Heterochromatic Regulation of Endogenous Retroviruses in Mouse Embryonic Stem Cells

Sophia Eleonora Groh

2022

Aus dem Lehrstuhl für Molekularbiologie
im Biomedizinischen Centrum (BMC)
Institut der Ludwig-Maximilians-Universität München
Vorstand: Prof. Dr. Peter B. Becker

Heterochromatic Regulation of Endogenous Retroviruses in Mouse Embryonic Stem Cells

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften *(Dr. rer. nat.)*
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von

Sophia Eleonora Groh

aus
Starnberg

2022

Gedruckt mit Genehmigung der Medizinischen Fakultät der Ludwig-Maximilians-Universität München

Betreuer:	Prof. Dr. Gunnar Schotta
Zweitgutachter:	Prof. Dr. Stefan Stricker
Dekan:	Prof. Dr. Thomas Gudermann
Tag der mündlichen Prüfung:	08. Februar 2023

Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Heterochromatic Regulation of Endogenous Retroviruses in Mouse Embryonic Stem Cells

selbstständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder teilweise übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 20. Juli 2022 ***Sophia Eleonora Groh***

Sophia Eleonora Groh

Publications

This cumulative thesis comprises **two** publications in peer-reviewed journals:

1. Sadic D., Schmidt K., **Groh S.**, Kondofersky I., Ellwart J., Fuchs C., Theis F. J., and Schotta G. (2015). Atrx promotes heterochromatin formation at retrotransposons. *EMBO Reports*, 16(7):836–50. DOI: 10.15252/embr.201439937

2. **Groh S.**, Milton A. V., Marinelli L. K., Sickinger C. V., Russo A., Bollig H., de Almeida G. P., Forné I., Schmidt A., Imhof A., and Schotta G. (2021). Morc3 silences endogenous retroviruses by enabling Daxx-mediated H3.3 incorporation. *Nature Communications* , 12(1):5996. DOI: 10.1038/s41467-021-26288-7

   I presented the Morc3 project, prior to its final publication, at the **EMBO Workshop The Mobile Genome: Genetic and Physiological Impacts of Transposable Elements** (August 29 to September 1, 2021) as an abstract selected short talk in the "Transposon Regulation" session.

Further I contributed to a review article in the field of ERV silencing and collaborated with colleagues to support other scientific projects, which led to the following publications:

Groh S. and Schotta G. (2017). Silencing of endogenous retroviruses by heterochromatin. *Cellular and Molecular Life Sciences*, 74(11):2055–2065. DOI: 10.1007/s00018-017-2454-8

Cernilogar F. M., Hasenöder S., Wang Z., Scheibner K., Burtscher I., Sterr M., Smialowski P., **Groh S.**, Evenroed I. M., Gilfillan G. D., Lickert H., and Schotta G. (2019). Pre-marked chromatin and transcription factor co-binding shape the pioneering activity of Foxa2. *Nucleic acids research*, 47(17):9069–9086. DOI: 10.1093/nar/gkz627

Abbreviations

5mC	5-methyl Cytosine
A	Adenine
ac	acetylation
AML	Acute Myeloid Leukaemia
ATP	Adenosine Triphosphate
ATRX	Alpha Thalassemia/mental Retardation syndrome X-linked
BLAST	Basic Local Alignment Search Tool
bp	base pair
C	Cytosine
ChIP	Chromatin Immunoprecipitation
CMT	Charcot-Marie-Tooth
Co-IP	Co-Immunoprecipitation
DAXX	Death domain-associated protein 6
DNA	Deoxyribonucleic acid
DNAme	DNA methylation
DNMT	DNA Methyltransferase
DSB	Double Strand Break
EGFP	Enhanced Green Fluorescent Protein
env	envelope
ERV	Endogenous Retro Virus
ES cells	Embryonic Stem cells
ESET	ERG-associated protein with SET domain
FACS	Fluorescence Activated Cell Sorting
G	Guanine
G4	G-quadruplex
gag	group specific antigen
GHKL	Gyrase, HSP90, Histidine Kinase, MutL
H3K9me3	Tri-methylation of histone H3 on lysine 9
HERV	Human Endogenous Retro Virus
HIRA	Histone regulator A
HP1	Heterochromatin protein 1
HSP90	Heat Shock Protein 90
HUSH	Human Silencing Hub
IAP	Intracisternal A Particle
IFN	Interferon
IP	Immunoprecipitation
IP-MS	Immunoprecipitation followed by Mass Spectrometry
K	Lysine
KAP1	KRAB-associated Protein 1
KI	Knock In

Kme	Lysine methylation
KMTs	Lysine Methyltransferases
KO	Knock Out
KRAB	Krüppel-associated Box
L1	Line1
LINEs	Long Interspersed Nuclear Elements
LTRs	Long Terminal Repeats
MEF	Mouse Embryonic Fibroblast
mES cells	Mouse Embryonic Stem cells
MORC	Microrchidia
mRNA	Messenger Ribonucleic acid
MS	Mass Spectrometry
NBs	Nuclear Bodies
NGS	Next Generation Sequencing
NP95	Nuclear Protein, 95 kDa
NuRD	Nucleosome Remodeling and Deacetylase
NDs	Nuclear Domains
NXP-2	Nuclear Matrix-binding Protein 2
PML	Promyelocytic Leukemia
pol	polymerase
pro	protease
PTMs	Posttranslational Modifications
RNA	Ribonucleic Acid
SETDB1	SET Domain Bifurcated Histone Lysine Methyltransferase 1
sgRNA	single guide RNA
SHIN	Short Heterochromatin Inducing
shRNA	small hairpin RNA
SIM	SUMO-Interacting Motif
SINEs	Short Interspersed Nuclear Elements
SUMO	Small Ubiquitin-like Modifier
T	Thymine
TEs	Transposable Elements
TFs	Transcription Factors
TIF1β	Transcriptional Intermediary Factor 1β
TRIM28	Tripartite motif-containing protein 28
UHRF1	Ubiquitin-like, containing PHD and RING Finger domains 1
WT	Wild Type
WB	Western Blot
XRV	Exogenous Retrovirus
ZFP	Zinc Finger Protein

Table of Contents

Zusammenfassung

Der Großteil des Säugetiergenoms liegt in Form von Heterochromatin vor, einer Struktur aus Desoxyribonukleinsäure (DNS) und Proteinen, die sich durch begrenzte Zugänglichkeit der Erbinformation, geringe Transkriptionsaktivität und Anreicherung von repetitiven Sequenzen auszeichnet. Die spezifische Regulierung und dichte Verpackung bestimmter DNS-Abschnitte ist entscheidend für die Entwicklung von Säugetieren, während Defekte zu Krebs und Zelltod führen können.
Die Unterdrückung von Retrotransposons aus der Gruppe der endogenen Retroviren (ERVs) durch Heterochromatisierung ist von wesentlicher Bedeutung für die Gewährleistung genomischer Stabilität und transkriptioneller Integrität. Die Zielmechanismen sowie die Akteure, die an der Etablierung und Aufrechterhaltung des heterochromatischen Zustands beteiligt sind, sind jedoch nicht vollständig bekannt.
Diese Doktorarbeit umfasst zwei Veröffentlichungen, in denen neue Beteiligte der Regulierung des Heterochromatins identifiziert und charakterisiert wurden, wobei die Mechanismen der ERV-Stilllegung in embryonalen Stammzellen der Maus als Modellsystem verwendet wurden.
Es wurde ein 160 Basenpaar (bp) langes Sequenzelement von Intracisternal A-Partikel (IAP) Retrotransposons identifiziert, welches die Bildung, Ausbreitung und Erhaltung von Heterochromatin auslöst. Diese kurze Heterochromatin induzierende (SHIN) Sequenz führt zu einer von den Proteinen SETDB1 und TRIM28 abhängigen Ablagerung von H3K9me3 (Trimethylierung von Lysin 9 des Histons 3), einer Histonmodifikation, die ein Kennzeichen des Heterochromatins darstellt. Ein SHIN Sequenz Reportersystem in Kombination mit genomweiten sh- und sgRNA-Screens identifizierte die neuen ERV Regulationsfaktoren ATRX und MORC3. Beide Proteine binden an IAP-Elemente und sind für eine effiziente Heterochromatinbildung und die Aufrechterhaltung von robustem Heterochromatin notwendig.
Als ein wesentliches Ergebnis konnte gezeigt werden, dass ein funktionsfähiger MORC3-ATPase-Zyklus und MORC3-SUMOylierung für die ERV-Chromatinregulation von besonderer Bedeutung sind. Proteomanalysen von mutierten MORC3-Proteinen zeigten eine beeinträchtigte Interaktion mit dem Histon-H3.3- Chaperon DAXX. Bedeutenderweise ist H3.3 an MORC3-Bindungsstellen in MORC3 ko und mutierten Zellen deutlich reduziert, was zeigt, dass MORC3 ein kritischer Regulator des DAXX vermittelten Histon H3.3-Einbaus in ERV-Regionen ist.
Zusammenfassend beschreibt diese Arbeit zwei neue Akteure der heterochromatischen ERV-Regulierung und gibt einen molekularen Einblick in den H3.3 abhängigen Stilllegungsmechanismus.

Summary

The majority of the mammalian genome is present as heterochromatin, a structure composed of deoxyribonucleic acid (DNA) and proteins, characterized by limited accessibility of the heredity material, low transcriptional activity, and enrichment of repetitive sequences. The specific regulation and dense packaging of certain DNA segments are crucial for mammalian development whereas defects can lead to cancer and cell death. Repression of retrotransposons of the so-called endogenous retroviruses (ERVs) by heterochromatization is essential to ensure genomic stability and transcriptional integrity. Yet, the targeting mechanisms as well as players involved in both, establishing and maintaining the heterochromatic state, are not completely understood.

This thesis comprises two publications that identified and characterized novel players involved in heterochromatin regulation using the mechanisms of ERV silencing in mouse embryonic stem cells as a model system.

A 160 base pair (bp) sequence element of Intracisternal A particle (IAP) retrotransposons was identified to trigger the formation, spreading, and maintenance of heterochromatin. This short heterochromatin inducing (SHIN) sequence leads to deposition, of H3K9me3 (trimethylation of lysine 9 on Histone 3), a histone modification that represents a hallmark of heterochromatin, in a SETDB1 and TRIM28 dependent manner. A SHIN sequence-reporter system in combination with genome-wide sh- and sgRNA screens identified the novel players ATRX and MORC3. Both factors bind to IAP elements and are necessary for efficient heterochromatin formation and maintenance of robust heterochromatin.

Of particular significance, a functional MORC3 ATPase cycle and MORC3 SUMOylation are important for ERV chromatin regulation. Proteomic analyses of MORC3 mutant proteins revealed compromised interaction with the histone H3.3 chaperone DAXX. Importantly, H3.3 was strongly reduced on MORC3 binding sites in MORC3 ko and mutant cells, indicating that MORC3 is a critical regulator of DAXX-mediated histone H3.3 incorporation on ERV regions.

In summary, this work describes two novel players of heterochromatic ERV regulation and provides molecular insight into the H3.3-dependent silencing mechanism.

Introduction

1.1 Genetic Information and Chromatin

This chapter will introduce fundamental principles of biology and answer essential questions such as: "How is heritable information carried in cells?", "How are details encrypted on top of the genetic code?" or "How does a cell remember what happened in the past?". Even though a broad range of explanations to these questions are nowadays considered general knowledge, they are nonetheless extremely fascinating and a good entry point for the investigations presented in this doctoral thesis.

1.1.1 Hereditary Material and the Molecule of Life

Probably one of the most intriguing questions, especially - but not only - for biologists, is: "What is life? And what makes it so special?", with a multitude of philosophical and spiritual starting points for discussions.
On the physical level, the frequently used keywords "self-organizing chemistry" emphasize the peculiar capacity of life to reproduce and pass on characteristics to the following generation. This particular feature is only realizable since it contains its own instructions within itself. Laws of how this information is transmitted to the next generation were already formulated in 1865 by Gregor Mendel. He described how traits of peas were inherited his experimental crosses (Mendel, 1866).
Yet the medium of heredity information transfer was unknown to him. About the same time Ernst Haeckel proposed the idea, that the components responsible for the inheritance of traits are localized in the nucleus of cells (Haeckel, 1866). However, for him and most of his contemporaries, proteins remained the prime suspects of interest. Only in 1944 Avery, MacLeod, and McCarty experimentally demonstrated that the peculiar molecule of deoxyribonucleic acid (DNA) comprises the hereditary material, when they succeeded in introducing new traits into bacteria, by insertion of foreign DNA (Avery et al., 1944). By now we know very well that DNA in fact encodes information of how to develop, grow, survive or reproduce. Therefore, it even acquired the honorable title "blueprint of life".
DNAs' discovery and the understanding of its role as the carrier of genetic information was only possible as a result of the collective and extended effort of many scientists. One of the pioneer steps for recognizing the chemical and physical nature of our hereditary material was when in 1869 Friedrich Miescher first isolated a new phosphorus-containing molecule. He initially called it "nuclein" as it was extracted from the cells' nuclei and was later renamed nucleic acid (Dahm, 2005).
Unveiling the composition of this nucleic substance represents another milestone, achieved by Albrecht Kossel, who was awarded the 1910 Nobel Price in Physiology or Medicine for his work. He discovered that nuclein is built up out of the four nitrogen-containing organic compounds adenine (A), guanine (G), thymine (T), and cytosine (C) (Jones, 1953).

Each of these building blocks, or nucleotides, is composed of three parts: A five-carbon sugar, a phosphate group, and a (nitrogen-containing) base. Phoebus Levene was the first to discover the order of these three components (phosphate-sugar-base) and the carbohydrate components of DNA and ribonucleic acid (RNA)(deoxyribose and ribose, respectively) (Pray, 2008). RNA is another important molecule generated in cells from a DNA template in a process called transcription. One of RNAs' critical functions is to transfer the information from DNA to proteins since it is a single-stranded copy of certain DNA sequences, and serves as direct instruction for the proteins to be built (Alberts et al., 2002).
The flow of genetic information predominantly occurs from DNA to RNA to proteins. However, the "central dogma of molecular biology" specified more precisely, that the transfer of information from nucleic acid to nucleic acid (e.g. from DNA to RNA, but also from RNA to DNA), as well as from nucleic acid to protein is possible, while transfer from protein to protein, or from protein to nucleic acid can not happen (Crick, 1970).
The complete set of DNA in an organism is called the genome and the succession of bases in any order is referred to as a sequence. The basic physical and functional units of heredity, encoding the instructions of all endogenous protein structures in the body, are called genes. The bases A, T, G, C together with uracil (U), which is present in the transcripts of RNA instead of T, are the alphabet of our genetic code. To translate this four letter information into the 20 amino acids of proteins, each group of three consecutive nucleotides, called a codon, in messenger RNA (mRNA) is used to specify either one amino acid or a stop signal during the translation process (Alberts et al., 2002).
While the bases are the bearers of information, enabling the DNA to encode, the sugar and phosphate backbone serve to connect these bases and form the structure. An important progress in understanding the complex organization of DNA was when Erwin Chargaff showed in 1949 that in any given type of cell, the amount of A approximately equals the amount of T, while the amount of C approximately equals the amount of G. This finding eventually led to the base pairing model, where A is paired with T and G with C, and a first indication on another crucial aspect for life: How the information could be faithfully copied to be passed on from one generation to the next? By having these ratios and each sequence being paired with the complementary sequence on a second strand of DNA, semiconservative replication is the answer to the challenge (Ekundayo and Bleichert, 2019).
Final steps of DNA structure characterization, obtained by Rosalind Franklin and Maurice Wilkins in the early 1950s, revealed great symmetry and consistency in the structure of DNA by X-ray diffraction patterns, and gave important clues about its dimensions (Rapoport, 2003). The X-rays showed that the two strands of DNA could not be parallel, suggesting a helical nature organization. The famous double helix structure of DNA was finally published in 1953 by James Watson and Francis Crick, and is still probably the most recognized symbol for life and life sciences (Watson and Crick, 1953). Even most non scientist have at some point heard of or seen an image depicting the famous double helix for which Watson, Crick and Wilkins received the 1962 Nobel Prize in Physiology or Medicine for their discovery (Kyle and Shampo, 1998).

1.1.2 Organization in the Shape of Chromatin

The spatial organization of DNA is essential for cells for several reasons. To begin with, eukaryotic cells need to overcome the striking difficulty to arrange the enormous amounts of genetic information inside the small nucleus. This is achieved by wrapping the DNA around protein complexes. The resulting structure, called chromatin, is capable of packaging the long nuclear DNA molecules tightly and leads to the formation of stainable structures (Alberts et al., 2002). It was a mile stone in understanding the organization of DNA when Walther Flemming, who described cells during their division by light microscopy, noticed this well stainable structure in the cells nucleii, and termed it chromatin from the Greek word *chroma* for color (Flemming, 1882). According to the observation of condensing chromatin during cell division the resulting structures were called chromosomes, which can be translated to "colored bodies". Only much later the chemical composition of this complex was resolved as a combination of DNA with several types of proteins.
Associated with that packaging comes another advantage for the cells: It provides the opportunity for an additional layer of information and gives the chance for further regulation, above the sequence of the DNA, in Greek *epi genetic*. Studies of heritable and regulatory effects above the level of DNA are therefore summarized in the field of epigenetics (Henikoff and Greally, 2016), from which some chromatin associated aspects will be highlighted in the following sections.
The fundamental structural units of chromatin are called the nucleosomes. Each nucleosome is composed of a segment of negatively charged DNA (ca. 147 base pairs), wrapped in 1.65 left-handed super helical turns around an octamer of positively charged proteins, termed histones. Each nucleosome core particle is constituted of two copies of each of the four highly conserved basic histone proteins H2A, H2B, H3 and H4 (Kornberg and Thomas, 1974; Luger et al., 1997). The primary level of chromatin structure, is the arrangement of the repeating nucleosome units which are aligned in arrays and sometimes referred to as "beads on a string" as they resemble the nucleosome and DNA structure under the electron microscope (Olins and Olins, 1974; Baldi et al., 2020).
However, there are several more layers of complexity that were discovered over the years. Nucleosomes are essentially globular in configuration, only the N-terminal ends protrude as unstructured "histone tails" which are dynamically modified (Luger et al., 1997). A large number of different covalent post-translational modifications (PTMs) (e.g. acetylation, phosphorylation, methylation) with a wide variety of regulatory functions can take place at this tails and contribute to nucleosome variability by changing their biophysical properties. Also histone variants, most prominently present in histone families H2A and H3, can substitute the usual histones and exhibit specialized functions, not only in genome packaging, but also in other processes such as DNA repair or meiotic recombination (Talbert and Henikoff, 2017). DNA itself can be modified as well, with the most studied example being methylation of cytosines, which contributes to condensed packaging of the nucleosomes (Moore et al., 2013). The linker histone H1 is not part of the core nucleosome, but arrays of nucleosomes can be further compacted by interaction of H1 with adjacent linker DNA, contributing to fiber formation of a higher structural organization (Thoma and Koller, 1977; Willcockson et al., 2021)
As indicated above, chromatin is not only spatially organizing DNA, but with its diverse features also confers information, which provides an opportunity for regulation

of transcription (Kouzarides, 2007). Thereby genome architecture reflects and influences genome function and, compaction of DNA contributes to regulated and controlled access to the genetic information (Berger, 2007; Duan and Blau, 2012). Histone PTMs, histone variants and DNA methylation (DNAme) correlate with either positive or negative transcriptional states which will be further detailed in the next section.

1.1.3 Transcription Regulation by Chromatin Conformation

Development of multi cellular organisms requires precise regulation of transcription to establish cell type-specific gene expression signatures. These programs then need to be remembered and maintained over many cell generations. In this way, complex organisms, even though their cells carry identical genetic material, can be composed of various distinct cell types with differing identities and gene activity patterns, depending on which regions of the genome are activated or repressed (Henikoff and Greally, 2016). Chromatin and its chemical modifications play a central role in this regulation.
In 1928, Emil Heitz coined the term heterochromatin to describe chromosome regions that remain condensed and visible under light microscopy in interphase nuclei (Heitz, 1928; Berger, 2019). Also in mitotic chromosomes heterochromatin is distinguished by the dark stainable patters compared to euchromatin which appeared as unstained areas. Roughly speaking, the organization of the genome can be divided into these two categories. Both are essential for eukaryotic cell function and have very special characteristics and typical histone modifications. Euchromatin consists of "open chromatin" which refers to a state with readily accessible DNA and less densely packed nucleosomes as a consequence of active PTM marks on the histones. Heterochromatin on the other hand is also termed "closed chromatin", indicating reduced access to the DNA through addition of repressive histone PTMs and binding of further proteins that pack nucleosomes into a compact structure (Liu et al., 2020). Heterochromatin can additionally be divided in two categories: constitutive heterochromatin, which are compacted genomic areas formed in many cell types at centromeres and telomeres; and facultative heterochromatin, which is more locus- or cell type-specific. If a gene is located in an euchromatic domain, it has the potential to be transcriptionally active. On the other hand, most genes within or adjacent to more densely compacted heterochromatic domains are usually associated with silencing (Duan and Blau, 2012). Nevertheless chromatin states are not rigid, they can be highly dynamic and influenced by numerous different factors, with long-range control of gene expression (Dixon et al., 2015). Although the whole picture of chromatin dynamics remains to be determined, so far several major pathways are well known to contribute to chromatin organization and thereby also regulation of gene expression. Molecular hallmarks of epigenetic control have been reviewed in more detail, for instance by Kouzarides (2007) or Allis and Jenuwein (2016). Some examples with particular interest for this thesis, such as DNA methylation, histone modifications, histone variants and chromatin remodeling, will be introduced in this section.
DNA methylation in the form of 5-methyl cytosine (5mC), in CpG dinucleotides is a typical mark for the compact and inactive heterochromatin and plays an important role for mammalian development (Greenberg and Bourc'his, 2019). It is mediated by DNA methyltransferases (DNMTs), a protein family which includes four active members in mammals: DNMT1, DNMT3A, DNMT3B, and DNMT3L (Saitou et al., 2012). DNMT1

acts together with the protein UHRF1 to maintain DNA methylation during replication, preferentially methylating hemi-methylated cytosines (Sharif et al., 2007). On the other hand DNMT3A and DNMT3B can act together with the regulatory cofactor DNMT3L to mediate *de novo* methylation of unmethylated cytosines (Okano et al., 1998, 1999; Ooi et al., 2007).

Histone modifications can support both active or repressive chromatin states and are mediated by specific enzymes (Zhou et al., 2011). The frequent lysine (K) methylation is established by a variation of histone lysine methyltransferases (KMTs) that can catalyze mono- (me1), di- (me2) and trimethylation (me3), leading to different epigenetic marks prevalent on tails of histones H3 and H4 (Zhang et al., 2021b). Histone tail methylations are in general associated with repression, as for instance the heterochromatin characteristic trimethylation of lysine 9 on Histone 3 (H3K9me3), H3K27me3 or H4K20me3. But there are also exceptions of methylations, which are markers of active chromatin, namely H3K4me and H3K36me, which are present on active genes and promoters (Pick et al., 2014). The precise modes of action are still in process of being understood where for example a recent study showed that enhancer-associated H3K4 methylation can facilitate the activation of these enhancers (Bleckwehl et al., 2021). Active chromatin regions, such as transcriptionally active genes, promoters and enhancers frequently display further characteristics. They are typically associated with the presence of different transcription-associated proteins including RNA polymerase II or transcription factors (TFs) and by acetylation of histone tails, with the key mark of H3K27 acetylation (ac), while histones in silenced genomic regions are characterized by general deacetylation (Kurdistani and Grunstein, 2003; Pick et al., 2014).

TFs as determinants of cell identity play central roles in differentiation, but many of them can only bind accessible DNA. The so called pioneer transcription factors have the ability to recognize DNA in the context of nucleosomes and their recent investigations revealed in more detail how essential they are for cell fate transitions during embryonic development (Cernilogar et al., 2019; Zaret, 2020). Especially for these rearrangements histones and their tail modifications can be seen as docking units, they not only influence the compaction of chromatin, but also recruitment of transcriptional or further compactional machinery (Bano et al., 2017). Thereby they promote the establishment of unique gene expression patterns that ultimately promote different biological outcomes. As this work is studying the mechanisms that lead to gene silencing and heterochromatin regulation, the introduction will further focus on repressive histone marks and their deposition mechanisms.

H3K9me3 is a central epigenetic modification that defines heterochromatin. Well characterized histone KMTs that catalyze H3K9 methylation in mammalian cells are SETDB1 (also known as ESET), G9a, SUV39H1 and SUV39H2, which share a catalytic SET domain (Rea et al., 2000; Dillon et al., 2005; Feldman et al., 2006). The different H3K9 methylation systems show a partial redundancy, however, knock out (ko) of all H3K9 KTMs, and resulting complete loss of H3K9 methylation, dissolves mouse heterochromatin organization (Montavon et al., 2021). The silencing mechanisms of these factors include H3K9me mediated condensation of chromatin by forming a binding site for heterochromatin protein 1 (HP1), as well as recruitment of DNMT3A and DNMT3B to *de novo* methylate DNA (Lachner et al., 2001). While SUV39H proteins are preferentially targeted to the pericentric heterochromatin (Lehnertz et al., 2003)

where they convert H3K9me1 into its di- and trimethylated states (Loyola et al., 2009), G9a also participates in the methylation of histone H3K9, silencing euchromatic genes (Feldman et al., 2006). Since this thesis is focused on the silencing mechanisms of retrotransposons, a process where H3K9me mediated by SETDB1 plays an essential role, it will be further introduced in the separated chapter 1.3.3 "SETDB1 marks ERVs with H3K9me3" from page 13.

Histone variants are distinguished from canonical histones by distinct protein sequences (Franklin and Zweidler, 1977) and have been extensively reviewed for example by Talbert and Henikoff (2017), Ferrand et al. (2020) or Martire and Banaszynski (2020). With their replication independent incorporation by designated chaperone systems and specific PTMs they form another level of epigenetic control. Among the important functions of histone variants are for example the role in determining centromere identity, by a H3 variant in mammals called CENP- A (Gambogi and Black, 2019) or regulation of DNA damage response where the variant H2A.X becomes phosphorylated as an early response to DNA double strand breaks (DSBs) (Fernandez-Capetillo et al., 2002). The most relevant histone variant for this thesis is H3.3, whose deposition and function will be described in more detail from page 15 in chapter 1.3.5 "DAXX deposits the Histone Variant H3.3".

Several protein factors are responsible for regulating and translating these epigenetic marks into specific gene expression signals, where the degree of packaging, also referred to as chromatin accessibility, plays a central role in modulating transcriptional output (recently reviewed by Klemm et al. (2019)). Repressed regions are usually less accessible by packaging of DNA into nucleosomes and their repressive marks. For example precise positioning can inhibit transcription or transcription factor binding by directly blocking binding sites (Thurman et al., 2012). Chromatin remodelers play an essential part in modulating the chromatin accessibility, for example by sliding or spacing nucleosomes, but also through eviction or exchange of histones with variants by the use of their ATP-ases activity (reviewed by Becker and Workman (2013)).

Thereby the ATP-dependent nucleosome remodeling factors help to create a dynamic chromatin environment by alteration of the nucleosome architecture, using the energy of ATP hydrolysis to destabilize, shift, restructure or remove nucleosomes (Becker and Hörz, 2002; Clapier and Cairns, 2009). The ATP-dependent chromatin remodeling complexes are categorized into four groups, SWI/SNF, ISWI, CHD, and INO80, depending on the sequence and structure of ATPases in the complexes. They function in a lot of cellular processes including transcriptional regulation, chromatin assembly, and DNA damage repair and are essential for normal development. If disrupted the effect on health can be disastrous, as they were shown to play key roles for many neurodevelopmental disorders (Mossink et al., 2021). This is also true for the ATRX protein investigated in the first study presented in this thesis. Mutation of this SWI/SNF family chromatin remodeler causes Alpha-Thalassemia X-Linked Mental Retardation (ATRX) Syndrome, among other characteristics which will be further introduced from page 14 in chapter 1.3.4 "ATRX stabilizes the Genome".

Besides gene regulation heterochromatin also exerts other critical functions. It is essential for chromosome segregation, genome integrity, and is sometimes referred to as the "guardian of the genome" (Allshire and Madhani, 2018; Janssen et al., 2018). Crucial heterochromatin targets, and the focus of the studies presented in this thesis are genomic

sequences belonging to the class of transposable elements. They are well suited models for heterochromatin establishment and maintenance.

1.2 Transposable elements

Transposable elements (TEs) are DNA sequences of ancient origin with the ability to move along the genome. Since they are of peculiar significance for the investigations of this, thesis they will be introduced in more detail in the following sections, beginning with their discovery as a milestone for the understanding of the function and regulation of our genome. Especially the remnants of retroviral germ line infections that occurred several million years ago, are the object of the presented studies. Due to their similarity to exogenous viruses, they carry the designation endogenous retroviruses (ERVs). Their structure and physical importance will be highlighted starting from page 8 with chapter 1.2.2 "Evolution and Structure of Endogenous Retroviruses", while their heterochromatic silencing mechanisms, as main interest of the presented studies, will be described in chapter 1.3 "Factors of Heterochromatic ERV Silencing" from page 11.

1.2.1 Discovery and Classification of TEs

The discovery of TEs is attributed to a prominent biologist, Barbara McClintock, who received the 1983 Nobel prize in Physiology or Medicine for the identification of mobile genetic elements or "jumping genes". She observed that different colour patterns in corn were caused by the movement of TEs, which inspired her assumption that they bear regulatory potential (McClintock, 1956). Since their initial discovery in corn, where they constitute more than 50% of the genome, TEs have been identified in a large variety of other organisms (Gogvadze and Buzdin, 2009).
The enormous extent of repetitive elements in mammalian genomes could only be fully appreciated with the extensive technical advances allowing the sequencing and therefore investigation of whole genomes. This huge milestone for our understanding of genes, as well as non-coding regions and gene regulation represents a critical achievement for life sciences (Behjati and Tarpey, 2013). At the beginning of the 2000s long lasting consortia, using steadily evolving sequencing techniques, unveiled the first reference genomes. The first version of the human reference genome was published in 2001 (Lander et al., 2001), and just one year later the first mouse genome followed (Waterston et al., 2002). However, these first references were still incomplete. While covering most of the protein coding genes, around 8% of the human genome remained unsolved, with especially long repetitive stretches being a problem for Next Generation Sequencing (NGS). The improvement of the reference sequences by constantly evolving technologies, as for example long-read sequencing, is an ongoing challenge. Just in spring 2022 the first truly complete reference assembly of the human genome was published by the "Telomere-to-Telomere" consortium, which for the first time provided the complete sequences between the chromosome ends (so called telomeres) (Nurk et al., 2022). While with improvements of the reference genome also the *de novo* repeat discovery leads to continuously improving repeat annotations (Hoyt et al., 2022), already the initial

sequences from two decades ago showed that in both mouse and human only about 2% of the genome is comprised of coding sequences. The huge remainder consists of non-protein-coding regions. Their functions are still under investigation, but it has been shown that they contribute to chromosomal structure, safeguard genomic integrity and regulate protein production for example through heterochromatin or non-coding RNAs (Allshire and Madhani, 2018; Statello et al., 2021). Very remarkably, in all mammalian genomes investigated to date, TEs account for significant portions of the genome, as in mouse and human, where they comprise between 40 and 60% (Lander et al., 2001; Waterston et al., 2002).

TEs are classified in two major groups by their mode of transposition (Wicker et al., 2007). DNA transposons, also called class II elements, only comprise a very small part of the mammalian genome. They move by a so-called "cut and paste" mechanism, where they are truly "jumping" DNA sequences from one location in the genome to another (Waterston et al., 2002; Wicker et al., 2007). Their movement is driven by the encoded transposase enzyme, without promoting accumulations in copy number (Feschotte and Pritham, 2007).

Class I elements, also referred to as retrotransposons, are the most abundant in both human and mouse as they move by a "copy and paste" mechanism. This involves reverse transcription of an RNA intermediate back into DNA which is shared between retroviruses and retrotransposons (Bannert and Kurth, 2006). This mode of activity is also indicated by their name, since the Latin word *retro* means backwards. The resulting DNA can subsequently reintegrate as an additional copy within the host genome (Coffin et al., 1997). By this mechanism retrotransposons were able to expand to a huge abundance that makes up about 90% of all TEs present in humans (Bannert and Kurth, 2004).

Retrotransposons are further classified by the presence or absence of flanking long terminal repeats (LTRs). Non-LTR retrotransposons include two subgroups of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) (Mager and Stoye, 2015; Deininger et al., 2003). The LTR containing retrotransposons, on the other hand, closely resemble the proviral-integrated form of infectious retroviruses, pointing out their origin from ancient retroviral integrations. This resemblance, lead to the synonymously used term endogenous retrovirus (ERV) (Stocking and Kozak, 2008). As this kind of TEs is the focus of the presented studies, they will be further introduced in the following section.

1.2.2 Evolution and Structure of Endogenous Retroviruses

Viruses are ancient companions of life and some of their integrations have been with our ancestors for millions of years, relying on host cellular systems for their multiplication. Endogenized retroviruses compose about 8% of the human genome and maintained the capacity to expand by their "copy and paste" mechanism. Their activity bears important potential of genomic variation for evolution, for example by genetic innovation through mutating or creating new protein coding sequences but also by altering genome organization and expression regulation (Friedli and Trono, 2015; Mager and Stoye, 2015; Ferrari et al., 2021).

ERVs originated from exogenous retrovirus (XRV) germ line infections which lead to transmission of the stably integrated XRV sequences, called provirus, in the genome of host cells over many host generations (Coffin et al., 1997). The provirus encodes typical viral proteins as capsid group specific antigen (gag), protease (pro), envelope (env) and polymerase (pol) flanked by LTRs which are essential for their replication (Jern and Coffin, 2008).
Most ERV integrations were highly mutated or partially deleted during the course of evolution, yet they still show homologies to the characteristic retroviral genome structure. While they might contain functional gag, pro and pol genes, the loss of a functional env sequence leaves them generally unable to generate active retroviral particles, and therefore rendering them incapable of horizontal transmission. Instead they still remain as "fossils" in the cells and continue to shape the genomes of their hosts from within (Coffin, 2004; Stoye, 2012).
A group of mouse specific ERVs, the intracisternal A-type particles (IAPs), are a prototype and largely studied model of mammalian retroelements that emerged from a still active retroviral progenitor by loss of their env sequence (Ribet et al., 2008). Even lacking an extracellular phase, IAPs can largely affect their host cells as their intracellular retrotransposition can lead to mutations and impairment of genomic stability. Consequently it is critical for host cells to regulate their activity by different measures. Post-transcriptional silencing, e.g. by RNA interference (RNAi) has been reported for IAPs (Ramírez et al., 2006) but is less understood, compared to transcriptional ERV silencing by heterochromatin formation which was investigated to a greater extent. The heterochromatic silencing mechanisms which have been investigated on ERVs and in particular IAPs, also as the focus of this thesis, will be described in chapter 1.3 "Factors of Heterochromatic ERV Silencing" from page 11. IAPs have been studied for a long time and their expression in early development and active transposition in tumors or germ line of some mouse strains has been of interest (Kuff and Lueders, 1988).
Also other ERVs movement and accumulation represent a major force shaping genomes and their activity shows the potential of gene regulation. ERVs impact on health and development is object to many studies and will be shortly highlighted in the next chapter.

1.2.3 ERVs Impact on health and development

Since their endogenization, many ERVs evolved to feature physiological roles in their hosts. Some gained new cellular functions, as for example coding of new proteins, like the syncytins. These proteins play crucial roles during placenta development and originated from ERV envelope proteins (Mi et al., 2000). But also through their capacity to cause mutations in existing genes and to alter the genomic architecture, ERVs are seen as drivers of evolution (Friedli and Trono, 2015). Their contribution to genome organization and gene regulation with dynamic roles in distinct developmental contexts, for example when silencing needs to be partially relieved, have been recognized for a while, as active ERVs were shown to enhance expression of neighboring genes (Jern and Coffin, 2008; Rowe and Trono, 2011; Rowe et al., 2013b). Especially these physiological roles in gene regulation during development gain increasing appreciation and have been observed manifold (Gifford et al., 2013; Durnaoglu et al., 2021), with examples as

driving species-specific germ line transcriptomes (Sakashita et al., 2020) or leading to characteristic dynamics during embryogenesis (Low et al., 2021). Already a decade ago it was recognized that many regulatory sequences had ERV origins (Jacques et al., 2013; Sundaram et al., 2014) and that ERVs can function as physiological regulators during normal development, e.g. by rewiring the core regulatory network of human ES cells (Kunarso et al., 2010) or regulating innate immune responses (Chuong et al., 2016; Grandi and Tramontano, 2018). Especially the LTRs bear regulatory potential as enhancers or promoters as they harbor different binding sites e.g. for transcription factors (TFs) and can act as regulatory elements to drive expression of nearby genes (Cohen et al., 2009).
However, this regulation needs to be very well titrated since ERV derepression can be associated with disease and genomic instability (Slotkin and Martienssen, 2007). Over the years a wide variety of pathological conditions have been observed in correlation with aberrant ERV regulation. The most prominent ones include cancer, neurological disorders and autoimmune disease (Küry et al., 2018; Jansz and Faulkner, 2021). As for example enhancers within ERVs were shown to convey oncogenic potential in acute myeloid leukaemia (AML) (Deniz et al., 2020) but also expression of oncogenes in other human cancers was driven by ERVs (Jang et al., 2019). This observations even made ERVs attractive potential therapeutic targets against cancer (Grabski et al., 2019). Activation of human ERVs (HERVs) was also shown to lead to neurotoxicity as it was compromising development and function of cortical neurons (Padmanabhan Nair et al., 2021) and an effect on neurodegeneration by ERV activation has been shown several times (Ochoa Thomas et al., 2020). The hematopoetic system is affected where ERV activation leads to apoptosis of pro-B cells (Pasquarella et al., 2016). Additionally also certain autoimmune disorders are linked to ERV expression. A recent example showed that antibodies against an ERV envelope protein activate neutrophils in systemic lupus (Tokuyama et al., 2021).
Besides the more studied aspects of evolution and gene regulation also more recent discoveries are made about ERV impact. One of them is for example, that ERVs function as means of our body to communicate with the exogenous microbiota as the skin microbiome was shown to promote ERV expression which then leads to increased immune responses and inflammation (Lima-Junior et al., 2021). Since ERVs share certain regulatory mechanisms with their exogenous homologs ongoing research keeps discovering how ERVs can drive both resistance and promotion of XRVs (Chiu and VandeWoude, 2021; Pluta et al., 2020).
Since regulation of ERV activity is key for health and development, cells needed to find a way to keep them in check. This demands for strict regulation and sophisticated silencing mechanism to control their activity. The precise role of ERVs as regulators of gene expression are still under investigation, however, the guarding role of heterochromatin to protect genome integrity against transposon activities has been observed for many years (Liu et al., 2020). As mentioned in chapter 1.1.3 “Transcription Regulation by Chromatin Conformation” heterochromatin plays an important role for the proper regulation of gene expression, which is key for functional and healthy development. To investigate heterochromatin formation in mouse ES cells, the publications summarized in this thesis used the particular ERV family of IAPs introduced in 1.2.2 “Evolution and Structure of Endogenous Retroviruses”as a model. The current understanding of

their heterochromatic silencing mechanisms will be outlined in the following sections.

1.3 Factors of Heterochromatic ERV Silencing

Silencing mechanisms that initiate and maintain the heterochromatin formation on target sequences and in particular the identification and mechanistic characterization of factors leading to the silencing of retrotransposons are the focus of this thesis.
Prominent heterochromatic modifications on ERVs in mouse ES cells include high levels of DNA methylation, H3K9me3 and H4K20me3 (Mikkelsen et al., 2007). Since H4K20me3 is not essential for transcriptional repression of IAPs (Matsui et al., 2010), the focus of the following chapter will lay on the introduction of DNAme and H3K9 and other factors known to contribute to ERV silencing in ES cells.
By analyzing IAP sub-fragments for their ability to induce transcriptional repression of an enhanced green fluorescent protein (EGFP) reporter led to the identification of a short heterochromatin inducing (SHIN) sequence of 160 base pairs (bp) in the central gag part of the IAP sequence (Sadic et al., 2015). The two publications presented in this thesis used screens on the basis of the SHIN sequence, and could confirm already well described silencing factors such as DNMT1/UHRF1, TRIM28 or SETDB1. But also new candidates were discovered, most significantly ATRX/DAXX and MORC3, whose functions were investigated in this work. For a better understanding, these main players in ERV silencing in ES cells will be introduced in more detail during the following chapter.

1.3.1 DNMT1 & UHRF1 maintain DNA Methylation

DNMT1 (DNA-methyltransferase 1) is the major maintenance DNA methyltransferase in mammals. Its loss causes widespread hypomethylation of the genome and embryonic lethality (Li et al., 1992). The protein UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1), also known as NP95 (nuclear protein, 95 kDa), forms a complex with DNMT1 and mediates its recruitment to replicating heterochromatic regions (Bostick et al., 2007; Sharif et al., 2007). UHRF1s critical role in maintenance of DNAme is highlighted as its inactivation shows a phenocopy of DNMT1-KO with hypomethylation and embryonic lethality (Bostick et al., 2007; Sharif et al., 2007). Interestingly, UHRF1 also binds methylated H3K9, in particular H3K9me3 and thereby links DNA methylation and H3K9 methylation, two major epigenetic marks associated with silencing and heterochromatin (Arita et al., 2012; Rothbart et al., 2013).
Contribution of DNAme to retrotransposon silencing was initially not implicated in the early stages of embryonic development. Previous studies report that the knockout of DNMT1, DNMT3a, DNMT3b or DNMT3L in Embryonic stem (ES) cells, derived from the undifferentiated inner cell mass of blastocysts, does not impair silencing of endogenous or exogenous retroviruses (Pannell et al., 2000; Hutnick et al., 2010; Quenneville et al., 2012).Yet DNMT1s role for the repression of ERVs, was demonstrated for later stages of embryonic development and in differentiated cells, where DNMT1 knockout embryos show upregulation of IAPs at embryonic day 8.5 (Walsh et al., 1998) and Mouse Embryonic Fibroblast (MEF) cells treated with DNAme inhibitor displayed

strong derepression of IAPs (Rowe et al., 2013a). Hypomethylation through impaired DNMT1/UHRF1 pathways can induce tumors in mice, which where shown to be partially caused by somatic ERV transposition leading to oncogenic mutations (Howard et al., 2008).
Just a few years ago it was shown that also ES cells show a transient activation of ERVs by acute depletion of DNMT1. However, this effect was not dependent on the loss of DNAme itself. Instead the increased presence of hemimethylated DNA prolonged the binding of UHRF1 which in turn disrupted the SETDB1-dependent H3K9me3 deposition (Sharif et al., 2016).
Apart from their discovery as strong hits in the silencing screen, which depicts the acute and transient effect mentioned above, the DNA methylation pathway components DNMT1 and UHRF1 were not further investigated in the presented studies. Yet another project in the Schotta laboratory dealt with the topic of DNAme and H3K9me dependent ERV silencing dynamics during extra embryonic differentiation, where DNA methylation has a dominant role over H3K9me3. The corresponding dissertation of Zeyang Wang can be accessed in the electronic thesis library of the LMU (Wang, 2020).

1.3.2 TRIM28 and KRAB-ZFPs recruit Silencing Factors

One of the key players of ERV silencing in ES cells is a protein called tripartite-motif-containing protein 28 (TRIM28), also known as Krüppel-associated box (KRAB)-associated protein 1 (KAP1) or transcriptional intermediary factor 1β (TIF1β) (recently reviewed by Randolph et al. (2022)). TRIM28 is a universal transcriptional co-regulator, which functions as a scaffold protein and recruitment factor for different chromatin modifiers including the nucleosome remodeling and deacetylase (NuRD) complex (Schultz et al., 2001), HP1 (Nielsen et al., 1999; P. et al., 2006), and the H3K9 methyl transferase SETDB1 (Schultz et al., 2002) regulating a broad range of physiologic processes and defense against foreign DNA (Seki et al., 2010; Cheng et al., 2014; Randolph et al., 2022). TRIM28 dependent heterochromatic features of HP1 and H3K9me3 establish persistent and heritable gene silencing by long-range spreading of repressive chromatin marks (Talbert and Henikoff, 2006; Groner et al., 2010). Both XRVs and ERVs are targeted for TRIM28 dependent silencing and the control of ERV based enhancers is crucial to preserve transcriptional dynamics in ES cells (Rowe et al., 2010, 2013b). The role of TRIM28 is most important during early embryogenesis or in ES cells, where its deletion reduces H3K9me3 and leads to ERV reactivation and cell death, while differentiated cells only showed mild effects (Rowe et al., 2010).
TRIM28 is mostly recruited by the Krüppel-associated box (KRAB) domain of certain zinc finger proteins (ZFPs) which recognize and bind specific DNA sequences to enable sequence-specific silencing (Friedman et al., 1996; Moosmann et al., 1996). These KRAB domain containing ZFPs are tetrapod-specific and the largest family of transcriptional regulators in higher vertebrates (Ecco et al., 2017). Their abundance and diversity in mammalian genomes was recently studied by mass spectronomy (MS) analysis, confirming that most KRAB-ZFPs indeed recruit TRIM28 and associated ERV silencing factors. However, a subset of evolutionary more ancient KRAB-ZFPs also interacted with factors related to other functions as genome architecture or RNA processing (Helleboid et al., 2019). By keeping ERV activity in check, the KRAB-ZFPs also participate in the

evolution of gene regulatory networks (Imbeault et al., 2017), especially during early embryogenesis when the majority of KRAB-ZFPs is expressed (Corsinotti et al., 2013). The ongoing expansion of KRAB-ZFPs is considered as an evolutionary arms race for the defense against ERVs (Jacobs et al., 2014; Bruno et al., 2019), where a variety of KRAB-ZFPs arises in co-evolution with expanding and mutating ERVs (Thomas and Schneider, 2011; Lukic et al., 2014). Especially the evolutionary younger KARB-ZFPs have been shown to bind active ERVs, including the IAP family (Wolf et al., 2020). As an example, ZFP809 has been shown to silence *de novo* integrated retroviral vectors as well as ERVs in ES and embryonic carcinoma cells, via binding to the proline tRNA primer-binding site (PBS-pro) of these sequences (Wolf and Goff, 2009; Wolf et al., 2015).
Investigations presented in the first paper of this thesis showed that the SHIN sequence, within the IAP gag region, harbors a strong sequence specific silencing capacity (Sadic et al., 2015). A major question arising from the identification of this sequence specific silencing mechanism is: Which is the recognition factor of the SHIN sequence? KRAB-ZFPs seemed likely candidates, and while ZFP819 had been shown to regulate IAPs through the U3, 5'-UTR, and pol regions, it does not recognize the gag part, where the SHIN sequence is located (Tan et al., 2013). Our experimental approaches could not identify the SHIN sequence recognition factor yet, neither by genome wide screens nor by investigating the role of 37 candidate KRAB-ZFPs with enriched expression in mouse ES cells compared to MEFs (Groh, unpublished data).
Besides the recruitment by KRAB-ZFPs also the addition of small ubiquitin-like modifiers (SUMO) to TRIM28, in a process called SUMOylation, plays an important role in the silencing of ERVs (Ivanov et al., 2007; Yang et al., 2015). SUMO proteins contribute especially to protein-protein interactions for example of TRIM28 with SETDB1 and also a role in the recruitment of TRIM28 to ERVs has been proposed (Geis and Goff, 2020). Still the precise mechanisms of TRIM28 location to the SHIN sequence is unknown and its further investigation bears potential for interesting findings.

1.3.3 SETDB1 marks ERVs with H3K9me3

The H3K9 methyltransferase SET Domain Bifurcated 1 (SETDB1, also known as ESET and KMT1E) was identified as a TRIM28 associated silencing factor already two decades ago (Schultz et al., 2002). Similar to the strong up regulation of ERVs and reduction of H3K9me3 observed in TRIM28 depleted ES cells (Rowe et al., 2010), also the loss of SETDB1 in early embryogenesis results in a strong de-repression of many ERV classes (Matsui et al., 2010; Karimi et al., 2011). Since the deletion of SETDB1 in MEFs, similar to TRIM28, did not lead to a strong up-regulation of ERV expression, their essential role for ERV silencing was believed to be most pronounced in the early phases of development and ES cells while DNA methylation gains importance in differentiated cells (Matsui et al., 2010; Rowe et al., 2010; Leung and Lorincz, 2012). However, more recent studies also support ERV silencing roles for SETDB1 and TRIM28 in somatic cells as for example B- or T-cells (L. et al., 2015; Pasquarella et al., 2016; Takikita et al., 2016; Fukuda and Shinkai, 2020). Importantly a critical role of SETDB1 has been reported in many cancers over the past years (recently reviewed by Lazaro-Camp et al. (2021)), making SETDB1 an attractive target for cancer therapy (Kang, 2018).

The heterochromatin mark H3K9me3 plays an important role for silencing euchromatic promoters and heterochromatin reprogramming, especially during embryo development (Wang et al., 2018; Nicetto and Zaret, 2019). As a recent study showed its not only the silencing potential of H3K9me3 that leads to the SETDB1-mediated silencing effects, but a more complex mechanism. One of the regulatory modes of SETDB1 is by controlling a dual heterochromatin state where SETDB1 dependent H3K9me3 together with H3K36me3, usually associated with transcription, affect gene expression profiles in mESCs by marking and repressing potential enhancers, which are also present on some ERVs. Removal of SETDB1 caused loss of both signatures and turned these regions into active enhancers causing up-regulation of several genes (Barral et al., 2022).
Not only SETDB1 itself, but also some of its interacting proteins have been found to contribute to provirus and ERV silencing, chromatin assembly factors CHAF1A/B, which as part of a histone chaperone complex, contribute to H3K9me3 on ERVs by inserting variants H3.1/2 during replication (Yang et al., 2015; Wang et al., 2018). Beyond that also the protein ATF7IP (also known as MCAF1 or AM), an important co-factor for SETDB1, is needed for efficient repression and stimulates SETDB1 enzymatic activity (Wang et al., 2003; Ichimura et al., 2005). A screening for ERV silencing factors confirmed ATF7IP contribution in ES cells (Fukuda et al., 2018) while mechanistically it was described that ATF7IP stabilizes SETDB1 and promotes its location to the nucleus (Timms et al., 2016; Tsusaka et al., 2020). From the clinical aspect also ATF7IP, together with SETDB1, has been proposed as a target for immunotherapy in cancer (Hu et al., 2021).
Similar to the DNA methylation pathway components DNMT1/UHRF1, also the SETDB1 and ATF7IP proteins were not further investigated in the present work as very detailed analysis were performed previously in the Schotta laboratory. The corresponding dissertations of Gustavo Pereira de Almeida studying ATF7IP and its ERV regulation mechanisms (de Almeida, 2018) and of Alessandra Pasquarella and Rui Fan who studied the role of SETDB1 during haematopoiesis (Pasquarella, 2015) and early embryonic development (Fan, 2015) can be found in the electronic theses library of the LMU.

1.3.4 ATRX stabilizes the Genome

ATRX belongs to the SWI/SNF family of chromatin remodelers (Picketts et al., 1996) and was identified as the gene mutated in the genetic disease of Alpha-Thalassemia X-Linked Mental Retardation (ATRX) Syndrome (Gibbons et al., 2003). Alterations of ATRX also contribute to different cancers (Dyer et al., 2017; Darmusey et al., 2021) where they for instance can indicate high-risk in neuroblastoma progression (Akter et al., 2021). On the molecular level ATRX has been shown to play a role in genome stabilization and transcriptional repression, especially at H3K9me3-containing repetitive regions. Most studied examples are telomeres, where it cooperates with the histone chaperone DAXX to install the histone variant H3.3 to maintain structural integrity (Wong et al., 2010; Lewis et al., 2010), and at pericentric heterochromatin, where it is localized by readout of histone H3 modifications (Eustermann et al., 2011; Fioriniello et al., 2020). Another target of ATRX are regulatory regions which have both H3K9me3 and H3K36me3, for example at the 3' exons of zinc finger genes, where it contributes to preserve SETDB1 mediated H3K9me3 enrichment (Valle-García et al., 2016).

ATRX loss-of-function can cause genomic instability by alternative lengthening of telomeres, replication stress as well as DNA damage through G-quadruplex (G4) and DNA secondary structures (Akter et al., 2021; Teng et al., 2021). Other functions of ATRX interestingly include RNA-binding of RepA/Xist to promote loading of PRC2 and contributing to X inactivation and other polycomb target genes by H3K27me3 deposition (Sarma et al., 2014; Ren et al., 2020). Recently the investigation of global chromatin accessibility and transcriptional changes in human cancer cells were associated with ATRX inactivation (Liang et al., 2020), supporting ATRX's important role for both genome stabilization and regulation.
The identification and characterization of ATRX as an ERV silencing factor is a novel finding of the first publication presented in this thesis (Sadic et al., 2015) and is connected to its well-characterized contribution to the depostion of H3.3 together with the chaperone DAXX which will be highlighted in the following section.

1.3.5 DAXX deposits the Histone Variant H3.3

The ATRX interaction partner DAXX (Death domain-associated protein 6) was originally identified to induce apoptosis (Yang et al., 1997). Nowadays multiple functionalities and implications in different biological processes as DNA repair, chromatin modification and defense to viral infections are known, with the most studied roles still being apoptosis and transcriptional repression (Tang et al., 2015; Mahmud and Liao, 2019). Proper DAXX function is indispensable for mammalian development, especially in early embryos (Bogolyubova and Bogolyubov, 2021; Ishiuchi et al., 2021). One of the main functions of DAXX is the formation of a histone chaperone complex, together with ATRX, which specifically mediates replication-independent chromatin assembly of the variant histone H3.3 (Drané et al., 2010; Lewis et al., 2010).
H3.3 differs from H3.1 and H3.2 by only four or five amino acids and is encoded by two genes (H3-3A and H3-3B), whose translations result in identical proteins (Martire and Banaszynski, 2020). Deposition of H3.3 into chromatin is meditated by two selective chaperones: While the HIRA complex deposits H3.3 in association with actively transcribed genes, and therefore is not of main interest for this thesis, the ATRX and DAXX complex incorporate H3.3 for heterochromatin formation and telomere stabilization (Shi et al., 2017; Loppin and Berger, 2020). This ATRX/DAXX mediated deposition has been shown to be required for heterochromatin formation in embryos (Santenard et al., 2010) and necessary for silencing ERVs (Elsasser et al., 2015). Additionally, H3.3 has been reported to silences retroviruses (Wolf et al., 2017) and also DAXX was identified as an HIV restriction factor with its antiviral function against HIV-1 replication by initiating epigenetic repression of retroviral infection and inhibiting reverse transcription (Shalginskikh et al., 2013; Dutrieux et al., 2015; Maillet et al., 2020).
Interestingly, DAXX has no reported own enzymatic activity, instead it acts as a scaffold to bridge different proteins. These interactions are based on SUMOylation, where DAXX itself is SUMOylated, but also harbors two SIM (SUMO-interacting Motif) domains which recruit and bind SUMO-conjugated proteins to regulate their function (Mahmud and Liao, 2019). Also the antiviral response is relying on SUMOylation dependent mechanisms and DAXX, ATRX as well as many others were found in the SUMO2 proteome during HSV-1 infection (Sloan et al., 2015). Several interaction partners of

DAXX and ATRX are composed in a multi-protein nuclear structure called PML bodies, which function as a platform for the establishment of heterochromatin (Chang et al., 2013; Delbarre and Janicki, 2021).
The contribution of DAXX to regulation of ERVs and other repetitive regions was, simultaneously to our study, supported by other publications (Elsasser et al., 2015; He et al., 2015). Also more recent investigations confirm that the loss of DAXX creates a chromatin state permissive to transcription, due to ERV up regulation and leads to impaired tissue regeneration for example in pancreatic cancers (Wasylishen et al., 2020). Very interestingly for this work it was recently shown that the H3.3 exchange and surrounding chromatin remodeling activities are highly dynamic with the protein SMARCAD1 evicting the nucleosomes (Navarro et al., 2020). Since not all factors and contributors necessary to replace evicted nucleosomes are known, the second study presented in this thesis will contribute to finding answers to this open question.

1.3.6 A Family of Silencing Factors: The MORC Proteins

In the quest to identify so far unknown silencing factors one family of proteins, the MORC proteins, were of particular interest. The first *Morc* gene, short for *Microrchidia*, was identified in mice over two decades ago. As so often, a serendipitous mutation paved the way for its initial discovery as it lead to a severe developmental phenotype. Insertion of a transgene, with deletion of flanking genomic regions, caused infertility and reduced testicle size of male mice by the complete arrest of spermatogenesis at an early meiotic stage. Since the medical condition of microorchidism describes abnormally small testes, the term *Microrchidia* was chosen to designate the mutation (Watson et al., 1998).
The subsequent characterization of the affected locus and gene product described the first member, MORC or later MORC1, of a conserved nuclear protein family present in a wide variety of organisms. Soon the human homolog was studied, and investigations of sequence similarities via BLAST (Basic Local Alignment Search Tool) and FASTA algorithms discovered conformity to functionally undescribed human genes, that were afterwards characterized as *MORC2* and *MORC3*. Additionally several matches to sequences of other species like the zebrafish *Danio rerio*, the nematode *Caenorhabditis elegans* or the model plant *Arabidopsis thaliana* were reported, which suggested early, that the discovered protein formed the first member of a whole new protein family with diverse roles across different kingdoms of life (Inoue et al., 1999).
Further comparative genomics showed similarity to *Morc* genes in the genomes of a wide range of organisms, from prokaryotes to eukaryotes. The evolutionary early appearance suggested a prokaryotic origin while on the functional side, fusions of different family members to several DNA- and peptide-binding domains implicated chromatin association (Iyer et al., 2007, 2008). The ATPase domain of MORC proteins in particular, attributes them as part of an even broader super family of chromatin associated proteins, the so called GHKL-ATPases. This protein family is termed after prototypical members such as the DNA Gyrase, the molecular chaperone HSP90, the Histidine Kinase, and the DNA repair enzyme MutL (Inoue et al., 1999; Dutta and Inouye, 2000). Conservation among MORC proteins, also between the plant and animal kingdom, is mainly present in this GHKL-type ATPase and its fusion to an S5 fold (a beta-alpha-beta fold, termed after domain 2 of the ribosomal protein S5) which together from the catalytically

active ATPase module (Koch et al., 2017). A more recent comparative analysis of *Morc* orthologs from plants and animals even lead to the definition of a "MORC domain" which was defined as a highly conserved motif composition, also including the GHKL-ATPase and S5 domains, spanning ~370 amino acids at the N-terminus (Dong et al., 2018).
Another characteristic attribute of many MORC proteins, both in plants and animals, are one or more C-terminal coiled-coil domains which are important for protein-protein interactions and specifically for the formation of homodimers, as seen for the human MORC3 (Inoue et al., 1999; Mimura et al., 2010). Additionally, human and mouse MORC proteins, unlike their plant homologs, harbor a CW type zinc finger domain. This structural component is termed after its conserved cysteine (one letter code: C) and tryptophan (one letter code: W) residues (Perry and Zhao, 2003) and was identified as a histone modification reader confirming the chromatin association of MORC proteins (He et al., 2010; Hoppmann et al., 2011).
While the mammalian MORC was first identified in the context of meiosis, the identification of plant MORCs was based on screens for viral resistance and plant immunity (Kang et al., 2008). The *Arabidopsis* MORCs were described to be required in different processes, such as pathogen-induced chromatin remodeling, heterochromatin condensation and gene as well as transposon silencing (Moissiard et al., 2012, 2014; Langen et al., 2014). In more recent years further characterizations emphasized their role in plant immunity, transcriptional gene silencing and plant growth by contribution to the plant-specific RNA-directed DNA methylation pathway (Galli et al., 2021; Xue et al., 2021).
The copy number of plant MORCs varies greatly between different lineages, where *Arabidopsis* with 7 *Morc* genes is in the middle of a range from a single copy in some algae up to 23 in common wheat (Dong et al., 2018). The duplications of *Morc* genes together with different expression patterns might indicate functional divergence and point at different roles as epigenetic regulators in various nuclear processes (Li et al., 2013; Dong et al., 2018).
In animals the copy number of *Morc* genes is more stable. Many species, as for example humans, contain 4 MORC protein family members (MORC1, MORC2, MORC3 and MORC4) with the highest number in mouse and rat, which have 5 MORC proteins (MORC1, MORC2a, MORC2b, MORC3 and MORC4).
An evolutionary analysis of the *Morc* gene family, confirmed an origin before the divergence of plants and animals as well as a further division of the branches within the plant and animal kingdoms. The animal branch could be subdivided by two, Animal-Group I and Animal-Group II, interestingly separating *Morc* genes from the same species. Even though both animal groups contained vertebrate and invertebrate species the human *MORC1* and *MORC2* were part of Animal-Group I while Animal-Group II contained human *MORC3* and *MORC4* (Dong et al., 2018).
Beside the structural analysis also diverse functions in a variety of organisms are getting elucidated in the recent years. As an example the MORC protein of the single cell, eukaryotic parasite *Toxoplasma gondii* was described as an upstream transcriptional repressor of sexual commitment, therefore playing a critical role in the parasites life cycle (Farhat et al., 2020). Another example is the increasingly detailed analysis of the *C. elegans* MORC, which provided insight into its role in trans generational maintenance of chromatin organization downstream of small RNAs (Weiser et al., 2017), as well as

compaction of DNA by forming DNA loops, where the *C. elegans* MORC has been shown to topologically entrap and condense DNA (Kim et al., 2019).
More relevant for this study however, are the functions of mammalian MORCs. Also here the research rapidly expands and several functions have already been investigated. The first member of the MORC family, MORC1, has been shown to be responsible for transposon repression in the germ line of male mice (Pastor et al., 2014). Even though MORC1 is mainly studied in spermatogenesis it also shows expression in the brain, where it was recently proposed as a potential new target in mood regulation, connected with early life stress and depression (Mundorf et al., 2021).
Early studies of the human MORC2 showed that it was highly expressed in cancer cells an had a repression potential (Wang et al., 2010), MORC2 further showed recruitment of histone deacetylases associated with transcriptional repression (Shao et al., 2010; Zhang et al., 2015) and chromatin remodeling in DNA damage response (Li et al., 2012a; Wang et al., 2015). MORC2s silencing potential was later confirmed by identification of its essential contribution to human silencing hub (HUSH) dependent silencing, which deposits Setdb1 mediated H3K9me3 (Tchasovnikarova et al., 2017). Mutations of MORC2 are associated with a neurological disorder, the Charcot-Marie-Tooth (CMT) disease, characterized by damage to the peripheral nerves (Sevilla et al., 2015).
Several mutations in the MORC2 ATPase domain have been implicated in CMT disease, and investigation of the most common mutation showed a function in hyperactivation of HUSH-mediated silencing in neuronal cells (Tchasovnikarova et al., 2017). MORC2 together with the HUSH complex binds evolutionary young Line1 (L1) transposons, the only active retrotransposon in human, and promotes its transcriptional silencing by mediating H3K9me3 despostion (Liu et al., 2017).
In rodents the *Morc2* gene underwent a duplication by retrotransposition, where the first copy *Morc2a*, gave rise to a second locus coding for the functional and germ cell-specific protein MORC2b, which is essential for fertility of both sexes by regulation of meiosis-specific genes (Shi et al., 2018).
MORC4 was for a long time a poorly studied member of the MORC family, apart from a role as a potential lymphoma biomarker due to its high expression in a subset of diffuse large B-cell lymphomas and roles in several kinds of inflammatory diseases like Crohn's disease or chronic pancreatitis (Liggins et al., 2007; Hong et al., 2016). Only recently a study reported the molecular mechanism of the MORC4 ATPase activation and its biological functions as it regulates the formation of nuclear bodies (NBs) and plays a role in the cell cycle S-phase progression (Tencer et al., 2020).
In summary MORC proteins have been associated with vital, physiological functions, that when impaired by *Morc* mutations or changed expression patterns result in severe developmental phenotypes including germ line sterility, immune system defects, neural disorders, and different kinds of cancer (Hong et al., 2016).
Also MORC3 has been shown to be an interesting factor for health. Since it is a key protein in the investigations of this thesis it will be introduced separately in more detail during the following section.

1.3.7 MORC3 - A promising candidate for ERV silencing

A special focus of this work lies on the protein MORC3. It presented itself among the highest hits and as the most promising novel factor discovered in the sgRNA SHIN silencing screen. At the outset of my research, MORC3 was a protein of largely unknown function, yet several of its described structures made it an appealing suspect for further investigation.

MORC3 is ubiquitously expressed in human (Takahashi et al., 2007) and also known as NXP-2 (Nuclear Matrix-binding Protein 2), KIAA0136, ZCWCC3 (Zinc finger CW-type coiled-coil domain protein 3) and ZCW5 (Li et al., 2013). Besides the identification and characterization of MORC family proteins mentioned in the previous section, human MORC3 was discovered independently as the nuclear protein NXP-2. As its name indicates it localizes to the nuclear matrix and has a RNA binding domain (Kimura et al., 2002). The curiosity of both these characteristics where highlighted in the recent study of another RNA binding, nuclear matrix protein, SAFB. SAFB modulates chromatin condensation and stabilizes heterochromatin foci in mouse cells by interaction with heterochromatin-associated repeat transcripts such as major satellite RNAs and promotion of phase separation (Huo et al., 2020). Even though the investigations of this thesis did not focus on MORC3 RNA- or nuclear matrix-binding properties, neither touch the aspect of phase separation, it should be mentioned that also MORC3, was recently shown to form nuclear condensates through phase separation (Zhang et al., 2019b), confirming earlier observations where overexpression of MORC3 lead to formation of „nuclear domains (NDs)" (Mimura et al., 2010) and potentially directing to additional functions beyond the results of this thesis. Especially the MORC3 RNA binding capacity should be further investigated since it was recently shown that RNA plays a crucial role in nuclear organization and promotes the formation of spatial compartments in the nucleus (Quinodoz et al., 2021; Zhang et al., 2021a).

MORC3, as NXP-2, has also been of clinical interest for over a decade as a frequent autoantigen. NXP-2 autoantibodies occur in a variety of juvenile dermatomyositis (Gunawardena et al., 2009), a disease characterized by inflammation of muscles and skin, and are correlated with increased cancer risk in dermatomyositis patients (Fiorentino et al., 2013). More recently the NXP2 autoantibodies are even used as a characterization tool and predictor of complication and decision of therapeutic strategy (Fujimoto et al., 2016). Besides the dermatomyositis, human MORC3 was mapped to chromosome 21 (Katsanis et al., 1997), while the mouse *Morc3* region is localized on chromosome 16 (Inoue et al., 2000), which is connected with the mouse model of Down syndrome (Ling et al., 2014) implicating a potential role of MORC3 in this genetic disorder. A study of partial trisomy of mouse chromosome 16, which leads to a neuropathology similar to the one observed in patients with Down syndrome, identified interferon-related signal transduction as the most significantly dysregulated pathway. *Morc3* was among the upregulated differentially expressed genes (DEG) but its functional contribution was not investigated (Ling et al., 2014). Fitting to the assumption of MORC3 involvement in interferon signaling, a recent data set on GEO describes MORC3 as a negative regulator of Type I Interferon, while to associated publication is not available yet the regulation of cell signaling pathways by MORC3 is intriguing (GEO Series GSE182755)

MORC3 plays an essential physiological role during development, since mutant mice show a peculiar phenotype: $Morc3^{-/-}$ mice die at, or within, a day after birth (Taka-

hashi et al., 2007) and also the heterozygous mice show a detectable phenotype with an altered haematopoietic stem cells niche and bone cell differentiation (Jadhav et al., 2016) even though the mechanisms behind are not yet understood.
On the molecular level MORC3 is a conserved protein with a length of 942 amino acids in mouse and 939 in human. Like the other MORC proteins, it contains the characteristic N-terminal GHKL-ATPase domain which undergoes conformational changes coupled with ATP turnover. The ATPase domain controls dimerisation, where the ATP unbound MORC3 is present as a monomer and forms a homodimer upon ATP binding (Mimura et al., 2010; Li et al., 2016) together with the coiled-coil domain at the carboxy-terminus, similar to the function of other GHKL ATPases which act as dimerized "molecular clamps" (Corbett and Berger, 2006). ATP hydrolysis and connected dissolution of the ATPase domain dimer requires DNA (Andrews et al., 2016).
The CW type zinc finger domain plays an important role for two aspects. On the one hand, it can interact with the ATPase domain, leading to negatively regulated ATPase activity. This inhibition can be relived by ligand binding to the CW-domain, resulting in higher ATPase activity (Andrews et al., 2016; Zhang et al., 2019c). Apart from its regulatory role on the ATPase cycle, the CW domain has also been shown to bind to histone H3 tails and several studies found MORC3 as an H3K4 binder (Li et al., 2012b; Eberl et al., 2013) with a preference for posttranslational modifications (PTMs) with the higher methylation state H3K4me3 (Andrews et al., 2016; Liu et al., 2016). MORC3 discovery as a reader of H3K4me3, is a very curious finding, since this active histone mark is associated with promoters and first ChIP-seq analyses described that MORC3 localizes to H3K4me3 marked chromatin (Li et al., 2016), while its previously described localization to promyelocytic leukemia-nuclear bodies (PML NBs) with regulation of the tumor suppressor p53 were reported (Takahashi et al., 2007). The localization to PML bodies is an interesting feature, as also the silencing factors DAXX, ATRX and SETDB1 have been found to associate with these structures in a SUMOylation depended manner (Lin et al., 2006). In fact the repression capacity of NXP2 was recognized even before the protein was studied as MORC3 in the context of the MORC family. After identification in a proteomic screen for SUMO2–binding proteins, silencing of a Gal4-reporter confirmed the SUMO mediated repression by NXP-2, but its targets were not investigated (Rosendorff et al., 2006). Later studies confirmed MORC3-SUMOylation and 5 Lysines were identified to play a crucial role in the regulation (Mimura et al., 2010) and potential SUMO dependent interactions.
Antiviral properties were described to restrict Herpes Simplex Virus 1 and Human Cytomegalovirus (Sloan et al., 2016). Viral proteins can also serve as ligands for the MORC3 CW domain by mimicking histones (Zhu and Qin, 2019) and thereby targeting MORC3 and its anti-viral function, as it was described for the NS1 protein the Influenza A Virus (Zhang et al., 2019a). Since the viruses aim at inhibiting MORC3, also the potential host cells need to counteract. A recent study proposed a model where the primary anti-viral function of MORC3 is self-guarded by a secondary interferon (IFN)-repressing function. Therefore, if a virus targets MORC3 to counteract its primary anti-viral function the secondary anti-viral IFN response will be unleashed (Gaidt et al., 2021). Also transposon silencing has recently been reported as a novel MIWI2 association partner and epigenetic regulator of piRNA dependent transposon silencing in male germ cells (Kojima-Kita et al., 2021).

Taken together MORC3 is as an interesting new silencing factor to investigate, since the molecular function of silencing in mESCs could not be explained yet. Its important role in ERV silencing was further strengthened by Desai et al. (2021), who also reported an ERV silencing function of MORC3 in mES cells shortly after the results of this thesis were published in Groh et al. (2021).

1.4 Aims of the Thesis

As laid out in the previous chapters of the introduction, the importance of ERV regulation for functional chromatin architecture in development and health are gaining increasing appreciation. Even though many studies are performed to investigate heterochromatin and silencing of transposons, the precise mechanisms of regulation are still not completely understood.
A crucial step to investigate heterochromatin and regulation of endogenous retroviruses was to identify an adequate model to study the silencing mechanisms. This goal was reached in the first publication comprised in this thesis, by the discovery of a specific and strong silencing recruitment sequence within the IAPEz family of ERVs. This significant achievement was driven by the first author of the publication, Dennis Sadic, who finished his PhD (Sadic, 2014) shortly after I joined the project. As a newly arrived student in the Schotta lab I was eager to contribute to the exciting project investigating ERV silencing and felt honored to be entrusted with the responsible role of conducting all the revision experiments for the paper "Atrx promotes heterochromatin formation at retrotransposons", printed in this thesis from page 23.
The objective of both presented publications was to understand the silencing mechanisms at this newly discovered reporter sequence by employment of different screening systems. They were constructed out of the SHIN sequence, and either based on initiation of silencing (publication 1, leading to the identification of ATRX) or maintenance of silencing (publication 2, identifying MORC3) in genome scale screening approaches to determine molecular functions and genes involved in mammalian heterochromatin formation.
After identification of novel hits from the screens, the goal was to validate and further investigate the proteins by genetic and biochemical experiments to uncover more about the underlying mechanisms.
The focus and key objective during my doctoral time was the functional characterization of MORC3 and eventually the identification of the molecular silencing mechanisms behind the observed phenotype. To get a clearer picture of how MORC3 acts in the context of chromatin, its location and binding sites in the genome were revealed by ChIP-seq of a FLAG-tagged MORC3 cell line.
The next aim was the descriptive analysis of the MORC3-KO in SHIN silencing assays, expression analyses by qPCR and RNA-seq as well as investigation of chromatin changes, as chromatin accessibility by ATAC-seq and change in characteristic histone marks as H3K9me3, H3K4me3 and H3K27ac.
The way for uncovering the molecular mechanism, by which MORC3 contributes to the silencing, was paved by the generation and investigation of different MORC3 mutants cell lines. To identify the most significant partners for MORC3 function, protein interaction

studies of MORC3 and its mutants by IP-MS were performed. Finally, H3.3 ChIP-seq experiments were conducted to confirm the newly discovered role of MORC3.
The results of this fascinating project are present in the second publication "Morc3 silences endogenous retroviruses by enabling Daxx mediated histone H3.3 incorporation", printed in this thesis from page 39.

Publication 1

Atrx promotes heterochromatin formation at retrotransposons

Dennis Sadic[1], Katharina Schmidt[1], Sophia Groh[1], Ivan Kondofersky[2,3], Joachim Ellwart[4], Christiane Fuchs[2,3], Fabian J. Theis[2,3] and Gunnar Schotta[1]

[1] Adolf-Butenandt-Institute, Ludwig Maximilians University and Munich Center for Integrated Protein Science (CiPSM), Munich, Germany [2] Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Computational Biology, Neuherberg, Germany [3] Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, Technische Universität München, Garching, Germany [4] Helmholtz Zentrum München, Institute of Molecular Immunology, Munich, Germany

---

This manuscripts is published in:
EMBO reports. DOI: 10.15252/embr.201439937

Author contributions:

I performed all revision experiments, consisting of FACS analysis of SHIN and RBS silencing as well as RT-qPCR analysis of ERV expression, which led to the discovery of the interesting phenotype of MusD upregulation.
The main work for this manuscript was performed by first author D.S., who performed most of the experiments and together with G.S. designed the experimental approach, wrote the manuscript. K.S. cloned and tested several IAP subsequeces and identified the SHIN sequence; J.E. performed the FACS sorting; I.K., C.F. and F.J.T. developed and carried out the statistical analyses.

Article

SOURCE DATA TRANSPARENT PROCESS

EMBO reports

Atrx promotes heterochromatin formation at retrotransposons

Dennis Sadic[1], Katharina Schmidt[1], Sophia Groh[1], Ivan Kondofersky[2,3], Joachim Ellwart[4], Christiane Fuchs[2,3], Fabian J Theis[2,3] & Gunnar Schotta[1,*]

Abstract

More than 50% of mammalian genomes consist of retrotransposon sequences. Silencing of retrotransposons by heterochromatin is essential to ensure genomic stability and transcriptional integrity. Here, we identified a short sequence element in intracisternal A particle (IAP) retrotransposons that is sufficient to trigger heterochromatin formation. We used this sequence in a genome-wide shRNA screen and identified the chromatin remodeler *Atrx* as a novel regulator of IAP silencing. Atrx binds to IAP elements and is necessary for efficient heterochromatin formation. In addition, Atrx facilitates a robust and largely inaccessible heterochromatin structure as *Atrx* knockout cells display increased chromatin accessibility at retrotransposons and non-repetitive heterochromatic loci. In summary, we demonstrate a direct role of Atrx in the establishment and robust maintenance of heterochromatin.

Keywords Atrx; Daxx; heterochromatin; histone H3.3; IAP retrotransposons
Subject Category Chromatin, Epigenetics, Genomics & Functional Genomics
DOI 10.15252/embr.201439937 | Received 1 December 2014 | Revised 7 April 2015 | Accepted 17 April 2015 | Published online 26 May 2015
EMBO Reports (2015) 16: 836–850

Introduction

Silencing of retrotransposons is essential for embryonic development. Firstly, novel retrotransposon insertions can impair genome stability [1], and secondly, aberrant activation of retrotransposon-based enhancers can confuse the transcriptome by affecting the expression of nearby genes [2,3]. In light of these findings, recent data have implicated retrotransposons in the development of human diseases, such as cancer [4]. Hence, it is important to understand how mammalian retrotransposons are targeted for silencing.

Silencing of retrotransposons involves the establishment of a heterochromatic domain covering the important regulatory regions of these elements. This heterochromatic structure shows a very similar modification signature to pericentric heterochromatin and features high levels of DNA methylation as well as H3K9me3 and H4K20me3 [5]. Chromatin modifications on retrotransposons are established by the same molecular machineries that act on pericentric heterochromatin, except for H3K9me3 which is not predominantly mediated by Suv39h enzymes but rather depends on the histone methyltransferase Setdb1 [6]. Targeting of the repressive machineries to retrotransposons is incompletely understood. In part, DNA methylation imprints of germ cells are used to re-establish heterochromatin during early embryonic development [7,8]. In addition, there is strong evidence for *de novo* pathways that recognize specific retrotransposon sequences [9–11]. These pathways are active in embryonic stem (ES) cells and can be utilized to study *de novo* heterochromatin establishment.

To identify novel players in heterochromatin establishment and maintenance on retrotransposons, we chose the abundant class of IAP retrotransposons as model system. In ES cells, IAP silencing depends on the transcriptional repressor Trim28 that interacts with Setdb1 and thus mediates the recruitment of H3K9me3. Trim28 and Setdb1 also regulate the establishment and turnover of DNA methylation at IAP sequences in ES cells [12]. Interestingly, in differentiated cells IAP silencing does not require the Trim28 pathway anymore and largely depends on DNA methylation, suggesting that Trim28 is mainly an important factor in the *de novo* pathway for heterochromatin establishment [12,13]. How IAP sequences specifically recruit Trim28 and Setdb1 in ES cells is largely unclear. Likely candidates for targeting Trim28 to IAPs are sequence-specific KRAB zinc finger proteins. Recruitment of Trim28 by KRAB zinc finger proteins is well described *in vitro* [14–17] and on few endogenous targets including MLV proviruses and imprinted regions [9,18,19]. This hypothesis predicts that retrotransposons contain nucleation sites for silencing, bound by KRAB zinc fingers. In accordance with this model, novel IAP insertions into the genome of mouse ES cells recruit Trim28-dependent silencing [10,13].

Here, we systematically tested sequence elements of IAP retrotransposons for their ability to induce heterochromatin formation and identified a small region of 160 bp (SHIN) that is sufficient to trigger silencing. Based on this sequence, we developed a shRNA screen and identified the SNF2-type chromatin remodeler Atrx as

1 Adolf-Butenandt-Institute, Ludwig Maximilians University and Munich Center for Integrated Protein Science (CiPS^{M}), Munich, Germany
2 Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Computational Biology, Neuherberg, Germany
3 Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, Technische Universität München, Garching, Germany
4 Helmholtz Zentrum München, Institute of Molecular Immunology, Munich, Germany
*Corresponding author. Tel: +49 89 2180 75 422; Fax: +49 89 2180 75 425; E-mail: gunnar.schotta@med.uni-muenchen.de

strong modifier of IAP silencing. Atrx was initially identified as the gene responsible for the X-linked alpha thalassemia/mental retardation (ATR-X) syndrome [20]. ATRX-deficient cells display numerous phenotypes connected to heterochromatic regions, such as defective sister chromatid cohesion and telomere dysfunction [21–23]. ATRX loss is further connected to alterations in DNA methylation patterns at imprinted regions, rDNA loci, telomeric repeats, and pericentromeric heterochromatin [24,25]. Recruitment to heterochromatic domains may be due to direct binding of the Atrx ADD domain to H3K9me3 in the context of unmethylated H3K4 [26,27]. Interaction with HP1, which is highly enriched at heterochromatin, has been shown to also contribute to stable recruitment of Atrx [26,27].

Recently, Atrx was found to interact with the histone H3.3-specific chaperone Daxx to facilitate H3.3 deposition in telomeric and pericentromeric regions [28–30]. Notably, this H3.3 deposition pathway is frequently mutated in pancreatic neuroendocrine tumors, glioblastoma, and pediatric neuroendocrine tumors [31–35]. These mutations are generally associated with an activation of the alternative lengthening of telomeres (ALT) pathway, a telomerase-independent mechanism to extend telomere length by recombination [32,36]. It is currently unclear if H3.3 deposition by Atrx/Daxx is really necessary to establish proper heterochromatin domains across telomeres and pericentromeric heterochromatin. Furthermore, it is not known if Atrx affects heterochromatin in other regions of the genome.

In this study, we show that Atrx is crucial for heterochromatin formation at IAP retrotransposons. Interestingly, heterochromatin establishment does not involve Daxx-mediated histone H3.3 incorporation. Instead, Atrx is important for generating an inaccessible chromatin structure on retrotransposons and non-repetitive regions in the genome which is necessary for robust maintenance of heterochromatin. Thus, our data reveal a novel role of Atrx in heterochromatin organization and may help to explain pathological phenotypes associated with ATRX syndrome or ATRX-deficient tumors.

Results

Identification of a short heterochromatin inducing sequence (SHIN)

To monitor the silencing potential of specific DNA sequence elements of mouse IAP retrotransposons, we developed a lentiviral hEF1α-EGFP reporter system (Fig 1A). When we combined the reporter with the well-characterized repressor binding site (RBS) of the MLV retrovirus [9,37], strong silencing of EGFP expression was observed 2 days post-transduction (Fig 1A). Similar effects were observed with the GAG region of IAP retrotransposons (Fig 1A). In order to more systematically test silencing potential of IAP sequence elements, we analyzed different IAP regions in our reporter assay (Fig 1B). A previously characterized region containing the UTR of IAP retrotransposons [10] displayed reporter silencing; however, the strongest effect was exerted by the ~2,000-bp-long IAP-GAG region (Fig 1B). As *de novo* silencing of IAP retrotransposons is restricted to ES cells [38], we did not detect silencing in mouse embryonic fibroblasts (MEF). To refine potential silencing initiating elements within the GAG region, we further dissected this element into 400-bp fragments (Supplementary Fig S1A) and could finally determine a region of 160 bp which was sufficient to induce strong reporter silencing (Fig 1C). This sequence could not be further shortened without compromising silencing potential (Fig 1C) and is highly conserved in more than 600 IAP elements, predominantly of the IAP-Ez subclass (Supplementary Fig S1C, Supplementary Table S1). As will be outlined below, this short sequence element is sufficient to trigger heterochromatin formation; thus, we termed this region SHIN for short heterochromatin inducing sequence. Notably, SHIN silencing is not restricted to the hEF1α promoter (Supplementary Fig S1B) and is orientation independent (Supplementary Fig S1B), suggesting that silencing is initiated by sequence-specific DNA binding factors.

The SHIN sequence induces heterochromatin

In ES cells, IAP silencing is mediated by heterochromatin that depends on Trim28 which, in turn, recruits the H3K9me3 methyltransferase Setdb1 [6,10,39]. Similar to pericentric heterochromatin, the H3K9me3-rich domains across IAP elements coincide with H4K20me3 and DNA methylation. To test if SHIN silencing is dependent on any of these chromatin modifications, we assessed the extent of SHIN repression in ES cell lines that are depleted for the major silencing machineries. Here, we measured the number of EGFP-positive cells relative to the control reporter that does not contain the SHIN sequence. Interestingly, cells deficient for Suv39h, Suv4-20h, Dnmt1, or Dnmt3ab can still effectively silence the SHIN reporter (Fig 2A). In contrast, Setdb1 knockout cells are severely compromised in SHIN silencing (Fig 2A). Also, lentiviral knockdown of Trim28, which did not even completely abolish Trim28 expression (Supplementary Fig S1D and E), resulted in strong silencing defects. These data demonstrate that SHIN silencing requires the major players that also regulate silencing of endogenous IAP elements.

Next, we asked if the SHIN sequence can directly trigger Trim28-/Setdb1-dependent heterochromatin formation. To address this question, we used a well-established recombinase-mediated cassette exchange (RMCE) cell line to integrate the SHIN sequence into a defined position of the ES cell genome (Supplementary Fig S1F) [40,41]. We combined the SHIN sequence with a Tet-inducible EGFP-T2A-zeocin reporter which is not transcribed in the absence of doxycycline (Fig 2B). We also generated a control reporter lacking the SHIN sequence (Fig 2B). RMCE of the two reporter constructs resulted in the HA36::SHIN and HA36::control cell lines (Fig 2B). ChIP-qPCR analysis revealed high levels of H3K9me3 across the RMCE locus in HA36::SHIN cells, comparable to endogenous IAP elements (Fig 2B). In contrast, only residual occupancy for H3K9me3 was detected in HA36::control cells (Fig 2B). To test if H3K9me3 across the SHIN reporter depends on Setdb1, we performed CrispR/Cas knockout of Setdb1 in HA36::SHIN cells. Prolonged knockout of Setdb1 is lethal for ES cells. Thus, we could not select for Setdb1 knockout cells and used a mixed population of Setdb1-deficient/Setdb-proficient cells for ChIP analyses 4 days post-sgRNA transduction (Fig 2C, Supplementary Fig S1G). H3K9me3 levels at major satellite repeats are dependent on Suv39h enzymes and therefore remained almost unchanged upon Setdb1 depletion (Fig 2C). Levels of H3K9me3 at

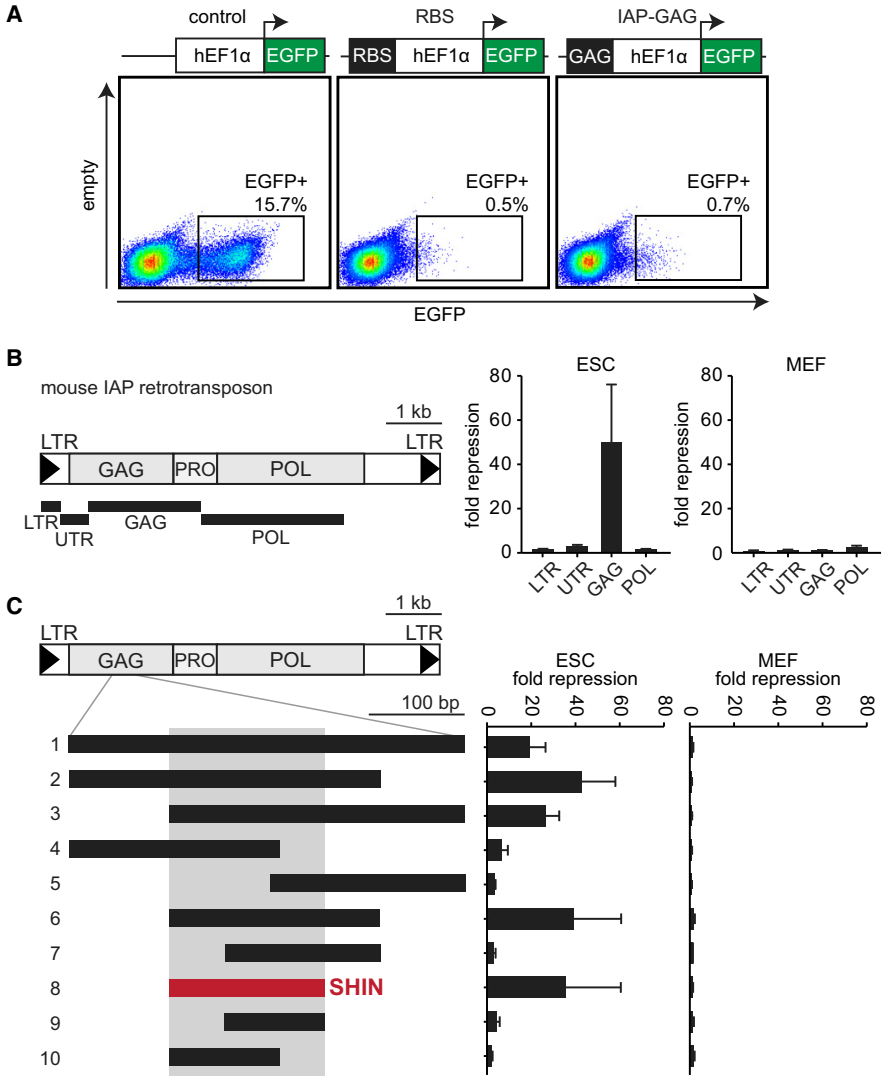


Figure 1. Identification of a short heterochromatin inducing sequence (SHIN) in mouse IAP retrotransposons.

A Reporter assay for retrotransposon silencing. Mouse ES cells were transduced with indicated lentiviral reporter vectors, and EGFP fluorescence was monitored 2 days later by FACS. PE was used as empty channel. RBS, repressor binding site of the MLV virus; IAP-GAG, GAG coding sequence of IAP retrotransposons.

B The GAG coding region of IAP retrotransposons leads to strong reporter silencing in mouse ES cells. Structure of IAP elements: LTR, long terminal repeats; UTR, 5′ untranslated region; GAG, GAG protein coding sequence; PRO, coding sequence of retroviral protease; and POL, coding sequence of retroviral polymerase. Indicated sequence elements of IAP retrotransposons were cloned upstream of the EGFP reporter cassette and tested for silencing activity in mouse ES cells and MEFs. Fold repression was calculated relative to the control EGFP vector at the same multiplicity of infection (MOI). Bar plots indicate the mean of four to seven biological replicates. Error bars indicate the standard deviation.

C Identification of the 160-bp SHIN sequence. Sub-fragments of the GAG coding region were tested for silencing activity as in (B). The smallest fragment which was still able to induce full repression was designated SHIN. Bar plots indicate the mean of two to six biological replicates. Error bars indicate the standard deviation.

Source data are available online for this figure.

endogenous IAP sequences were only mildly affected, indicating that this modification may have a low turnover on these sequences. Importantly, H3K9me3 was clearly reduced on a unique Setdb1 target region (pos control) and across the RMCE locus containing the SHIN sequence (Fig 2C). In summary, our data show that the SHIN sequence represents an initiation site for heterochromatin formation from which H3K9me3 can spread over the entire locus.

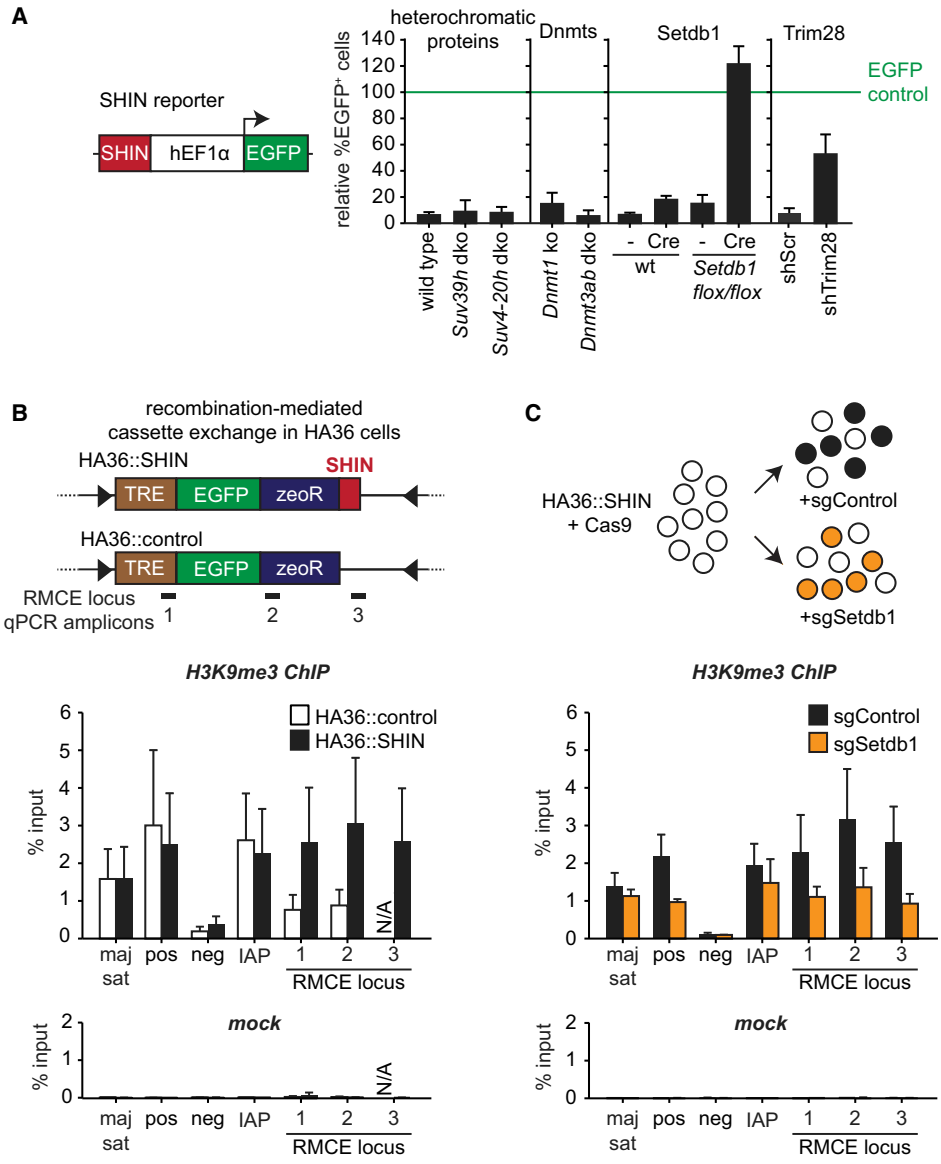


Figure 2. The SHIN sequence induces Trim28-/Setdb1-dependent heterochromatin formation.

A SHIN silencing depends on the Trim28/Setdb1 pathway. Mouse ES cells were transduced with the SHIN reporter or control vector, and EGFP fluorescence was monitored 2 days later by FACS. The percentage of EGFP-positive cells after SHIN reporter transduction is shown relative to the control transduction (relative % $EGFP^+$ cells). Setdb1 knockout cells were generated by the transduction of $Setdb1^{flox/flox}$ ES cells with Cre-expressing virus (Cre). Trim28 knockdown was performed by the transduction of a lentiviral shRNA vector (shTrim28 #1). Bar plots indicate the mean of three to six biological replicates. Error bars indicate the standard deviation.

B The SHIN sequence recruits H3K9me3. Recombination-mediated cassette exchange of the indicated SHIN reporter and control vector resulted in ES cell lines HA36::SHIN and HA36::control, respectively. ChIP-qPCR analysis of H3K9me3 across the RMCE locus in HA36::SHIN and HA36::control cells. RMCE PCR amplicons (1–3) are indicated. Control regions: maj sat, major satellite repeats; pos, Polrmt; neg, Tia1; and IAP, endogenous IAP elements. N/A = sequence not present in the HA36::control construct. Bar plots indicate the mean of three biological replicates. Error bars indicate the standard deviation.

C Setdb1 mediates SHIN-induced H3K9me3. HA36::SHIN cells were stably transduced with a Cas9 expression vector and then transduced with sgRNAs against Setdb1 or a control sequence, respectively. ChIP-qPCR for H3K9me3 was performed 4 days after sgRNA transduction. Since we could not select for Setdb1 knockout, the chromatin isolate for ChIP is composed of deleted and non-deleted cells. RMCE PCR amplicons (1–3) are indicated. Control regions: maj sat, major satellite repeats; pos, Polrmt; neg, Tia1; and IAP, endogenous IAP elements. Bar plots indicate the mean of three biological replicates. Error bars indicate the standard deviation.

Source data are available online for this figure.

SHIN silencing requires the Snf2-type chromatin remodeler Atrx

Heterochromatin formation across SHIN strictly depends on the Trim28/Setdb1 pathway and recapitulates silencing of endogenous IAP retrotransposons. Thus, we utilized our reporter assay to screen for additional factors that regulate retrotransposon silencing using a genome-wide shRNA screen (Fig 3A). We transduced wild-type ES cells with a pooled genome-wide shRNA library, selected shRNA expressing cells with puromycin and then transduced the SHIN reporter. Cells that escaped SHIN silencing were isolated based on their EGFP fluorescence, and shRNA sequences from these escaper cells were cloned into a plasmid library (Fig 3A). We tested 71 individual shRNA plasmids in a secondary screen. One shRNA which most strongly inhibited SHIN repression was found to target the SNF2-type chromatin remodeler Atrx (Fig 3B, Supplementary Table S2). This shRNA clearly reduces Atrx mRNA and protein levels in mouse ES cells (Supplementary Fig S2A and B). However, to rule out off-target effects, we generated four independent knockout ES cell lines (Supplementary Fig S2C) with loss of Atrx protein (Fig 3C). *Atrx* knockout (ko) cells grow normally and show no obvious alterations in their cell cycle profile (Supplementary Fig S2D). Importantly, SHIN silencing was defective in all *Atrx* knockout ES cell lines (Fig 3D), which confirmed the shRNA screening result.

We then asked if Atrx is important for the Trim28/Setdb1 pathway and performed SHIN silencing assays in wild-type versus *Atrx* ko cells, depleted for Setdb1 and Trim28, respectively (Supplementary Fig S2E and F). Although Setdb1 could be depleted by around 70% (Supplementary Fig S2E), SHIN silencing was only mildly impaired (Fig 3E). Setdb1 depletion in *Atrx* ko cells, in contrast, resulted in strong silencing defects (Fig 3E). This suggests that Atrx enhances the efficiency of Setdb1-dependent repression, although we cannot exclude the possibility that Atrx may act independently of Setdb1 in repressing the SHIN reporter, leading to synergistic effects when both Setdb1 and Atrx are depleted. Similar effects were observed for shRNA knockdown of Trim28 (Fig 3E).

To test if Atrx is required for Trim28/Setdb1-dependent repression outside of IAP retrotransposons, we performed silencing assays using a heterologous reporter system. The well-established repressor binding site of MLV retrotransposons (RBS) is recognized by the KRAB zinc finger protein Zfp809, which in turn recruits Trim28 and Setdb1 to induce silencing in a variety of cell types. In wild-type cells, the RBS-EGFP reporter was efficiently silenced (Fig 3F, Supplementary Fig S2G). However, RBS silencing was compromised upon knockdown of Atrx in MEFs and in *Atrx* ko ES cells (Fig 3F, Supplementary Fig S2G). Taken together, our data demonstrate Atrx as a crucial component of the Setdb1/Trim28 silencing pathway.

SHIN silencing is independent of Daxx-mediated H3.3 deposition

On telomeres and pericentric heterochromatin, Atrx recruits Daxx to mediate H3.3 deposition [28–30]. In order to test if the Daxx/H3.3 pathway plays a role in SHIN repression, we generated *Daxx* ko cells (Supplementary Fig S3A). We found that SHIN silencing is defective in *Daxx* ko cells (Supplementary Fig S3B), suggesting that Daxx is an important component for retrotransposon silencing. To address the question whether H3.3 deposition by Daxx is critical for SHIN silencing, we generated rescue cell lines expressing wild-type and mutant Daxx proteins (Supplementary Fig S3C). Re-expression of full-length Daxx protein in *Daxx* ko cells leads to a rescue in SHIN silencing (Supplementary Fig S3D). However, the expression of Daxx-mutant proteins with relaxed binding specificity for histone H3.3 (E231A) or impaired H3.3 interaction (R257A) also resulted in rescued SHIN silencing (Supplementary Fig S3D and E). In contrast, the expression of Daxx with a deletion of the C-terminal repressor domain [42] does not rescue SHIN silencing (Supplementary Fig S3D). These data point toward a H3.3-independent function of Daxx in SHIN repression. To confirm that H3.3 is really dispensable for retrotransposon silencing, we generated *H3.3* ko ES cell lines (Supplementary Fig S4A). Histone H3.3 is mainly transcribed from two loci in the mouse genome, H3f3a and H3f3b. Consequently, H3f3a/H3f3b double knockout cells have essentially lost H3.3 protein (Supplementary Fig S4B). Notably, SHIN silencing is completely intact in *H3.3* ko cells (Supplementary Fig S4C). Due to H3f3a duplications in the mouse genome that could eventually give rise to functional H3.3 transcripts (see blat search and alignment report in Supplementary Table S3), we performed H3f3a knockdown in H3f3b ko cells. These cells have lost H3.3 (Supplementary Fig S4D) but display robust SHIN silencing (Supplementary Fig S4E). Finally, we re-analyzed published H3.3 ChIP-seq datasets in control and *Atrx* ko ES cells [30]. We found only low enrichment of H3.3 in control ES cells (Supplementary Fig S4F). Surprisingly, upon deletion of Atrx, H3.3 becomes strongly enriched at IAP elements (Supplementary Fig S4F), suggesting that Atrx does not promote H3.3 incorporation at endogenous IAP elements, but rather inhibits excessive H3.3 deposition. Alternatively, higher chromatin turnover at IAP elements in the absence of Atrx may indirectly lead to more H3.3 incorporation. Together with our finding that H3.3 interaction mutants of Daxx can repress retrotransposon sequences and H3.3-depleted cells are not impaired in SHIN repression, our data demonstrate that SHIN silencing is independent of Daxx-mediated histone H3.3 incorporation.

Atrx is required for efficient heterochromatin formation

Next, we sought to investigate how Atrx mediates SHIN silencing on the molecular level. One possibility was that Atrx could mediate efficient heterochromatin formation across the SHIN locus. To test this hypothesis, we performed a re-silencing assay in HA36::SHIN ES cells in which we depleted Atrx by shRNA-mediated knockdown (Fig 4A). Strong induction of the Tet-inducible promoter using doxycycline alleviates SHIN silencing in a small population of cells, and constant selection pressure with zeocin allowed us to isolate a homogeneous population of cells with strong EGFP expression. Lack of promoter stimulation by doxycycline removal resulted in the rapid loss of EGFP expression within 4 days in both control (shScr) and Atrx-depleted (shAtrx) cells (Fig 4B, left panel). Loss of EGFP expression coincided with presence of H3K9me3 across the RMCE locus in shScr cells (Fig 4C, left panel, Supplementary Fig S5A). H3K9me3 could also be established in shAtrx cells; however, abundance of H3K9me3 appeared reduced in the test region 1 most distal to the SHIN sequence (Fig 4C, left panel), indicating reduced heterochromatin spreading. Surprisingly, re-silencing was very efficient in shScr cells despite continuous doxycycline induction (Fig 4B, right panel, Supplementary Fig S5A). A large proportion of these cells displayed strongly reduced EGFP expression and the establishment of H3K9me3 across the RMCE locus (Fig 4B and C, right panels).

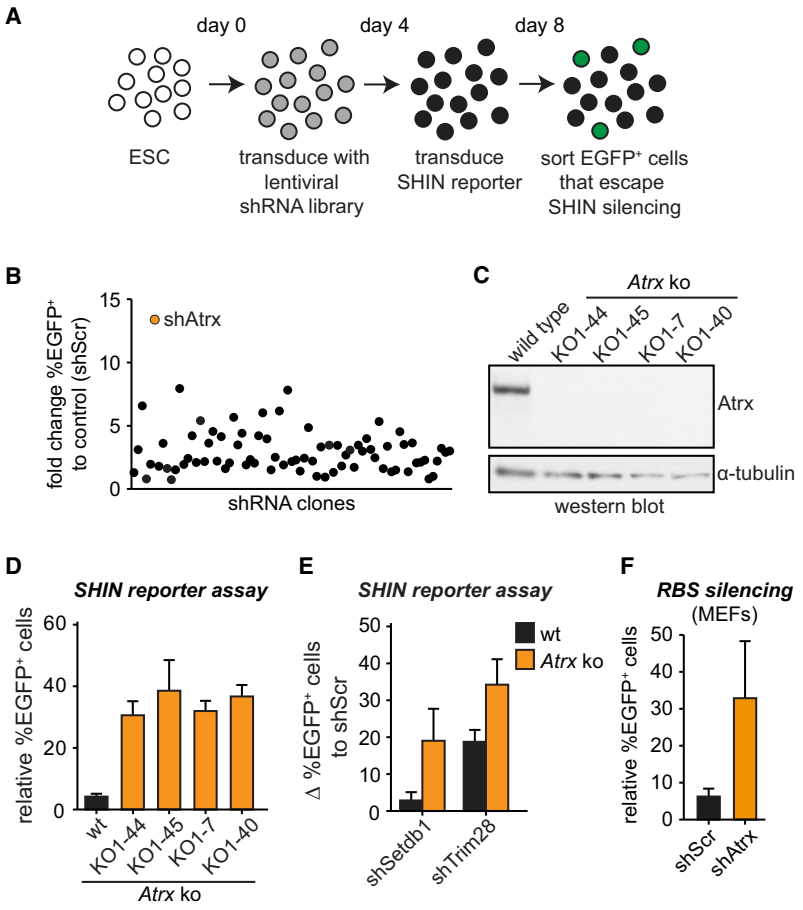


Figure 3. A shRNA screen identifies the chromatin remodeler Atrx as regulator of retrotransposon silencing.

A Scheme of the shRNA screen. ES cells were transduced with a genome-wide pooled shRNA library at a low MOI and then transduced with the SHIN reporter. Cells that escaped SHIN silencing were isolated by FACS sorting based on their high EGFP fluorescence. shRNA sequences of the EGFP$^+$ cells were amplified by PCR and cloned into a lentiviral vector.

B SHIN reporter assay in ES cells that were transduced with 71 random shRNA clones from the primary screen and scrambled control shRNAs. Dot plot displays the ratio of EGFP$^+$ cells in test shRNA versus control shRNA transductions. The test shRNA with strongest reduction in SHIN silencing was directed against Atrx (yellow dot). The top-scoring shRNAs of the secondary screen are listed in Supplementary Table S2.

C Western blot for Atrx in wild-type and four independent *Atrx* ko cell lines. Tubulin serves as a loading control.

D SHIN reporter silencing depends on Atrx. Wild-type and four independent *Atrx* ko ES cell lines were transduced with the SHIN reporter or control vector, and EGFP fluorescence was monitored 2 days later by FACS. The percentage of EGFP-positive cells after SHIN reporter transduction is shown relative to the control transduction (relative % EGFP$^+$ cells). The bar plot indicates the mean of three biological replicates. Error bars indicate the standard deviation.

E *Atrx* ko cells are sensitive to perturbations in the Setdb1/Trim28 silencing pathway. SHIN reporter silencing in wild-type and *Atrx* ko ES cells that were transduced with control (shScr), Setdb1, and Trim28 shRNAs. Bar diagram shows the difference in EGFP-positive cells upon knockdown of Setdb1 and Trim28 compared to control-treated cells. Bar plots indicate the mean of three to six biological replicates. Error bars indicate the standard deviation.

F Atrx is generally important for Setdb1-/Trim28-dependent silencing. Mouse embryonic fibroblasts were transduced with control and Atrx shRNAs. Silencing of an EGFP reporter construct containing the repressor binding site of the MLV retrovirus was monitored. The percentage of EGFP-positive cells after reporter transduction is shown relative to the control transduction (relative % EGFP$^+$ cells). The bar plot indicates the mean of three biological replicates. Error bars indicate the standard deviation.

Source data are available online for this figure.

In contrast, shAtrx cells are completely unable to induce heterochromatin formation across the RMCE locus in continuous presence of doxycycline (Fig 4B and C, right panels). In summary, these data indicate that Atrx is necessary for efficient establishment of heterochromatin. On transcriptionally inactive loci, heterochromatin can be formed, but spreading appears reduced. In addition, Atrx is strictly required when heterochromatin formation is challenged by transcriptional activity.

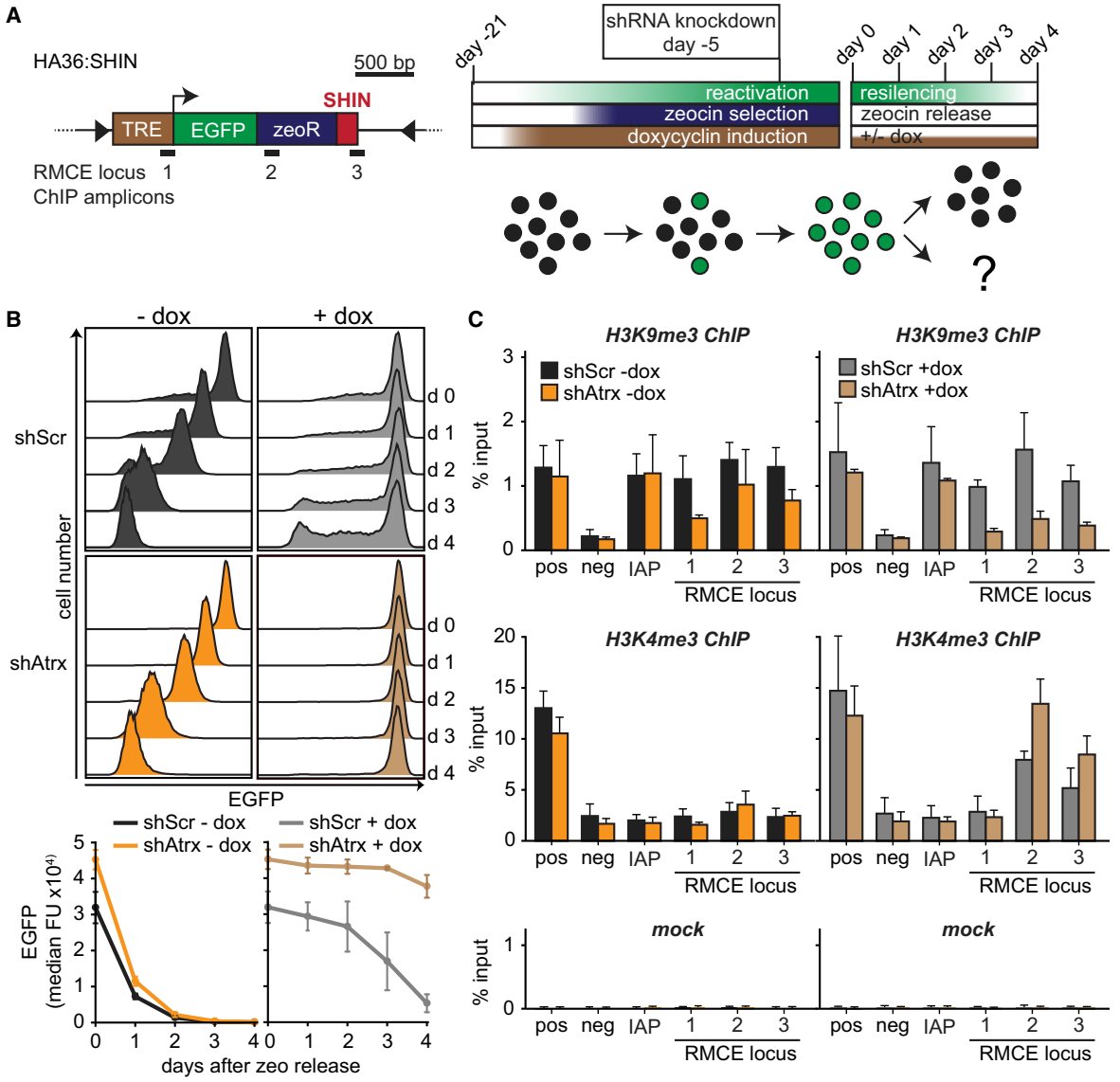


Figure 4. Atrx is required for efficient heterochromatin formation.

A Scheme of the re-silencing assay in HA36::SHIN cells. The silenced transgene in HA36::SHIN cells was reactivated by the addition of doxycycline which activates the Tet-responsive promoter (TRE). Cells in which the reporter was immediately activated (approx. 2–3%) were selected with zeocin. The EGFP$^+$ cells were then transduced with control or Atrx shRNAs. Five days later, re-silencing was triggered by the removal of the zeocin selection pressure (day 0), and EGFP expression was monitored in the presence or absence of doxycycline during the next 4 days (day 1–day 4).

B Atrx is necessary for efficient silencing in competition with transcriptional activity. Re-silencing experiments as outlined in (A) were performed. Top panel: histograms displaying EGFP fluorescence during the course of re-silencing. Removal of doxycycline (−dox) inactivates the TRE promoter, resulting in loss of EGFP expression. Despite maintained TRE promoter stimulation (+dox), a large number of control cells (shScr) can silence the reporter, while Atrx-depleted cells (shAtrx) are unable to induce silencing. Lower panel: Line plots indicating the average of the median EGFP fluorescence in three to four biological replicates. Error bars indicate the standard deviation.

C Atrx knockdown leads to impaired heterochromatin establishment. ChIP-qPCR analysis of H3K9me3 and H3K4me3 across the RMCE locus in HA36::SHIN cells at day 4 of the re-silencing assays as outlined in (A). Positions of the primer pairs at the RMCE locus are indicated in (A). IAP, endogenous IAP elements. H3K9me3 ChIP: pos, Polrmt; neg, Tia1; IAP, global endogenous IAPs. H3K4me3 ChIP: pos, Tia1; neg, Polrmt; IAP, global endogenous IAPs. Mock (beads only control): pos, Polrmt; neg, Tia1. Bar plots indicate the mean of three biological replicates. Error bars indicate the standard deviation.

Source data are available online for this figure.

Atrx is important for heterochromatin maintenance

Thus far, out data showed that Atrx is necessary for efficient *de novo* establishment of heterochromatin. However, in embryonic stem cells endogenous IAP elements are already covered by a heterochromatic structure which needs to be maintained. We thus wondered if Atrx is important for heterochromatin maintenance. When we analyzed published ChIP-seq datasets for Atrx, Setdb1, Trim28, and H3K9me3 in ES cells [43–45], we found strong enrichment of Atrx across IAP retrotransposons with highest occupancy over the SHIN region (Fig 5A). These data suggest that Atrx has a permanent role on these elements. ChIP-qPCR analyses for Atrx could confirm the enrichment on IAP elements and a non-repetitive control region (Fig 5B, top panel). However, heterochromatin marks, such as H3K9me3 and DNA methylation, were not altered in *Atrx* ko cells (Fig 5B, Supplementary Fig S5B and C).

The strong enrichment of Atrx, H3K9me3, Trim28, and Setdb1 on IAP sequences close to the SHIN region (Fig 5A) suggests that Atrx is recruited by H3K9me3 and/or Trim28/Setdb1. Thus, we tested if Atrx is *de novo* recruited to a Setdb1-dependent heterochromatin locus by analyzing HA36::SHIN cells. Enrichment of Atrx at the SHIN region within the reporter locus demonstrates that the recruitment of Atrx occurs in the context of Setdb1 and Trim28 (Fig 5C). Thus, we investigated whether Atrx generally associates with Setdb1/Trim28 target region in repetitive and non-repetitive regions of the genome. The overlap between Atrx, Trim28, and Setdb1 in non-repetitive regions is relatively small (Supplementary Fig S7A). Prominent enrichment was mainly observed on imprinted genes and few intragenic loci (Supplementary Fig S7B–D, Supplementary Tables S5 and S6). In contrast, our analyses revealed that Atrx co-occupies many ERV classes which are also enriched for Setdb1, Trim28, and H3K9me3, for example, IAP-Ez, MusD/ETn, and MMERVK10 elements (Fig 5D, Supplementary Fig S6, Supplementary Table S4). Thus, we sought to investigate if some of these ERV sequences become de-repressed in *Atrx* ko cells. Out of the three tested ERV classes with the enrichment of Setdb1/Trim28/Atrx, we detected the de-repression of MusD/ETn retrotransposons (Fig 5E), suggesting that Atrx function is rate limiting for the silencing of distinct ERV classes.

Challenges like DNA replication or transcription through intronic IAP repeats require heterochromatin to be constantly re-established. Based on our finding that Atrx is crucial for efficient establishment of heterochromatin, we wondered if heterochromatin maintenance is perturbed in *Atrx* ko cells when the Setdb1/Trim28 pathway has reduced activity. Moderate knockdown of Trim28 in wild-type cells resulted in slight up-regulation of endogenous IAP elements, demonstrating that heterochromatin maintenance is not fully ensured. However, when we performed Trim28 knockdown in *Atrx* ko cells, we observed severe derepression of IAP retrotransposons (Fig 5F). These findings indicate that Atrx is required and rate limiting for the repression of endogenous IAP elements when heterochromatin maintenance is challenged.

Another way of challenging heterochromatin is to provoke transcription across a heterochromatinized locus. In HA36::SHIN cells, the EGFP reporter gene is not transcribed and covered with H3K9me3 (Fig 2C). However, transcription can be induced by doxycycline that allows the recruitment of the reverse tetracycline transactivator (rtTA) to the Tet-responsive promoter of the reporter gene. Due to the heterochromatic state of the reporter gene in control cells (shScr + dox), doxycycline induction leads to a very low percentage of EGFP-positive cells (Fig 5G, Supplementary Fig S5D). In contrast, doxycycline induction upon Atrx knockdown (shAtrx + dox) results in a much higher percentage of EGFP-expressing cells (Fig 5G, Supplementary Fig S5D). These data demonstrate that heterochromatin is more vulnerable in *Atrx* ko cells.

Atrx regulates heterochromatin accessibility

Heterochromatin is generally characterized by a high compaction grade which makes it largely inaccessible to transcriptional activators and refractory to challenges. Our finding that transcription of a heterochromatic locus can be more easily stimulated in the absence of Atrx suggests that structural properties of heterochromatin like local nucleosome density or higher order folding are compromised when Atrx is lost. To test this hypothesis, we performed MNase accessibility assays in wild-type versus *Atrx* ko ES cells. Atrx is only binding to a small fraction of the mouse genome, and thus, we did not detect global alterations in MNase accessibility [43] (Fig 6A and B). To investigate Atrx

Figure 5. Atrx is important for heterochromatin maintenance.

A Atrx binds to IAP retrotransposons. Cumulative ChIP-seq coverage profiles across IAP elements for Atrx, Setdb1, Trim28, and H3K9me3. The structure of IAP elements is shown schematically; the position of the SHIN sequence is marked in dark gray. rpkm, reads per kilobase per million of reads.

B H3K9me3 is not altered in *Atrx* ko cells. ChIP-qPCR analysis for Atrx and H3K9me3 in wild-type (wt) and *Atrx* ko ES cells. pos, Polrmt; neg, Tia1; IAP, global endogenous IAPs; and IAP SHIN, SHIN sequence of IAP elements. Bar plots indicate the mean of three biological replicates. Error bars indicate the standard error of the mean.

C Atrx is recruited to newly formed heterochromatic sites. Atrx ChIP was performed in HA36::SHIN cells, in which the SHIN reporter locus was newly integrated into a defined locus by RMCE (see Fig 2B). Bar plots indicate the mean of three biological replicates. Error bars indicate the standard deviation.

D Setdb1, Atrx, H3K9me3, and Trim28 co-occupy distinct ERV elements. Binary heatmap showing ChIP-seq enrichment over input (> 1.5-fold) on all mouse ERV repeat classes. Selected ERV classes are indicated. Complete information is provided in Supplementary Fig S6 and Supplementary Table S4.

E MusD/ETn retrotransposons are de-repressed in *Atrx* ko cells. RT–qPCR analysis of selected retrotransposon classes in wild-type (wt) and *Atrx* ko ES cells. Bar plots indicate the mean of three to eleven biological replicates. Error bars indicate the standard deviation.

F Heterochromatin maintenance is compromised in *Atrx* ko ES cells when core heterochromatin proteins are depleted. Wild-type and *Atrx* ko ES cells were transduced with control or Trim28 shRNAs, and the expression of endogenous IAP elements was measured by qRT–PCR. Bar plots indicate the mean of three biological replicates. Error bars indicate the standard deviation.

G Heterochromatin is more vulnerable in *Atrx* ko cells. Addition of doxycycline to HA36::SHIN cells results in binding of the reverse Tet transactivator (rtTA) to the Tet-responsive promoter (TRE) in the SHIN reporter locus. Percentage of EGFP-expressing cells was monitored before and after the addition of doxycycline. Bar plots indicate the mean of three biological replicates. Error bars indicate the standard deviation.

Source data are available online for this figure.

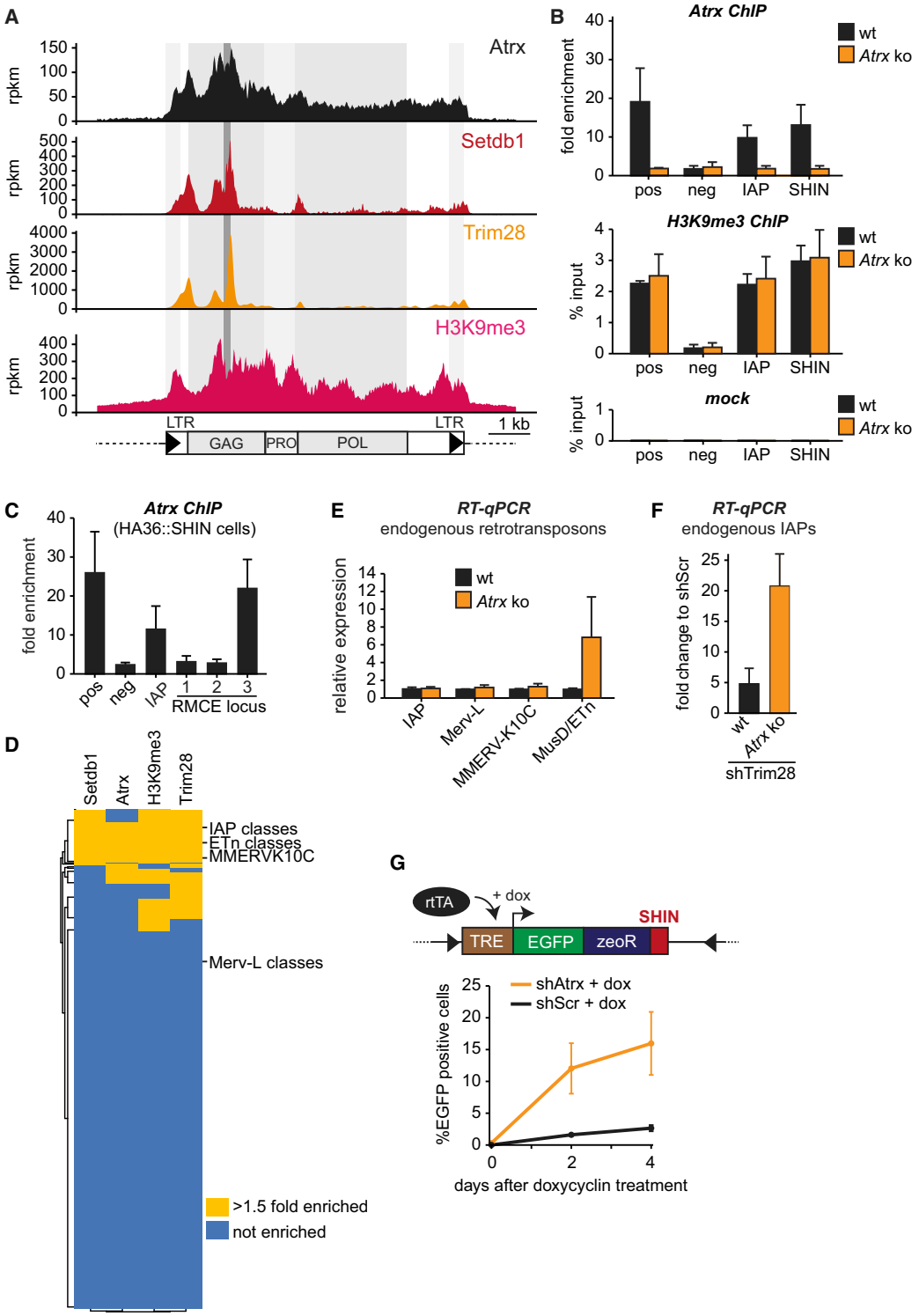


Figure 5.

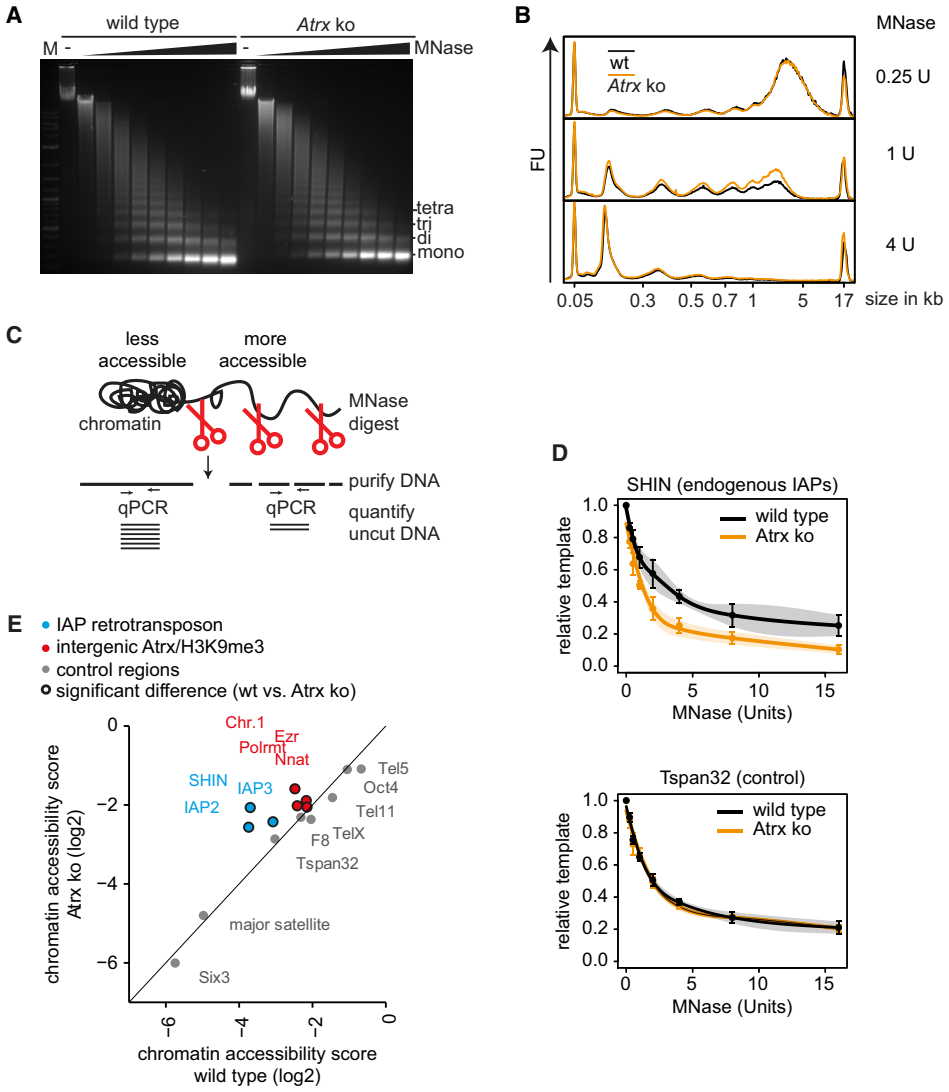


Figure 6. Atrx regulates heterochromatin compaction.

A Chromatin accessibility is not globally changed in *Atrx* ko cells. Nuclei of wild-type and *Atrx* ko ES cells were digested with increasing amounts of MNase (0–16 U), and DNA was analyzed on an agarose gel stained with ethidium bromide. M, size marker; "-", no MNase; size of mono-, di-, tri-, and tetra-nucleosomes is indicated.

B Overlay of DNA electropherograms obtained from Bioanalyzer runs of the MNase-digested DNA from wild-type and *Atrx* ko ES cells. The first and the last sharp peak represent the markers of the Agilent DNA 12000 kit.

C Schematic of the locus-specific chromatin accessibility assay. MNase cuts open chromatin faster than more compact, less accessible regions. The amount of uncut DNA of a specific locus thus correlates with accessibility of this region and can be measured by qPCR.

D Heterochromatin on IAP elements is less compact in *Atrx* ko cells. Locus-specific chromatin accessibility assays were performed in wild-type and *Atrx* ko ES cells using a range of MNase concentrations. Curves represent an example of a smoothing fit to the data points. Shaded areas demarcate a confidence interval based on ± one functional standard deviation.

E Generally increased heterochromatin accessibility in *Atrx* ko cells. Locus-specific chromatin accessibility assays were performed in wild-type and *Atrx* ko ES cells. Curve fitting to the data points resulted in a chromatin accessibility score which positively correlates with chromatin accessibility (see Supplementary Materials and Methods). The average chromatin accessibility score for indicated regions in wild-type and *Atrx* ko cells was plotted as dots from three biological replicates. The entire calculated curve fits from three biological replicates were used to assess statistically significant differences in digestion behavior (shown as dots with black border). Since the chromatin accessibility score only reflects the slope of the curve fit at 50% digestion degree, this score does not directly correlate with statistically significant differences in overall digestion behavior (also see Supplementary Materials and Methods).

Source data are available online for this figure.

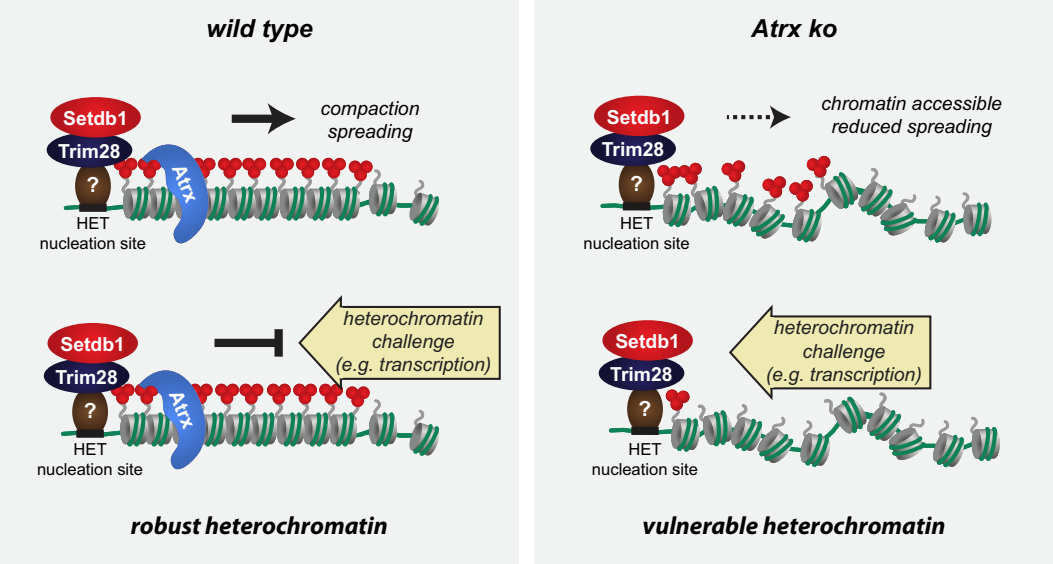


Figure 7. Atrx is important for the establishment and robust maintenance of heterochromatin.
Heterochromatin formation is thought to initiate on specific nucleation sites from which it can spread over large distances. Proteins that recognize these initiation sites and recruit Trim28/Setdb1 in the context of retrotransposons are currently unknown. Atrx is then necessary for the establishment and spreading of heterochromatin and mediates heterochromatin compaction. Thus, Atrx-deficient cells are characterized by more vulnerable heterochromatin which cannot easily be maintained upon challenges, such as transcription.

target regions, we employed a qPCR strategy to measure the MNase digestion degree on specific genomic loci with increasing MNase concentrations (Fig 6C). Curve fitting through these data points allows statistical assessment of the digestion degree between wild-type and Atrx-mutant cells (Supplementary Fig S7E). Further, we derive a chromatin accessibility score which represents the digestion rate at which 50% of the target locus is digested (Supplementary Fig S7E). MNase accessibility at endogenous IAP retrotransposons is significantly increased, whereas control regions, such as promoters of active and inactive genes, show no alterations (Fig 6D and E, Supplementary Fig S7F). Interestingly, chromatin accessibility was also increased on non-repetitive Setdb1/Trim28 targets (Fig 6E, Supplementary Fig S7E). These data demonstrate that Atrx renders Setdb1-dependent heterochromatin more inaccessible.

Discussion

In the mammalian genome, heterochromatin is very abundant across repetitive elements. Pericentric heterochromatin is mainly organized by the Suv39h–HP1–Suv4-20h pathway [46–48], while repetitive elements outside the pericentric compartment are controlled by the Trim28/Setdb1 pathway, with support by Suv39h enzymes [6,10,39,49]. A major characteristic of the heterochromatic state is its modification pattern of H3K9me3/H4K20me3/DNA methylation and the buildup of an inaccessible chromatin structure which does not allow strong transcriptional activity. Heterochromatin formation is thought to initiate on specific nucleation sites from which it can spread over large distances [50]. Initiation sites in pericentric heterochromatin may be binding sites for specific transcription factors [51] or involve other recruitment mechanisms for Suv39h/HP1 complexes [52]. In the context of retrotransposons, recruitment mechanisms may involve KRAB zinc finger proteins which could serve as targeting platform for Trim28/Setdb1 [9]. Our work identified a novel nucleation site in IAP retrotransposons (SHIN) that is sufficient to induce heterochromatin and transcriptional repression of strong promoters (Fig 7). SHIN silencing requires the Trim28/Setdb1 pathway for heterochromatin formation across an integrated reporter gene (Fig 2C). Cumulative coverage maps across IAP elements revealed strong enrichment of Setdb1 and Trim28 across the SHIN sequence; however, additional coverage was detected at the LTR/UTR region and in the beginning of the POL coding sequence (Fig 5A). These findings suggest that endogenous IAP elements may feature several independent nucleation sites for heterochromatin formation. In support of this hypothesis, previous analyses of the LTR/UTR sequence revealed significant silencing potential [10], although to a lower extent as compared to the SHIN sequence (Fig 1). Which proteins recognize the SHIN sequence is currently unclear. As silencing can only be initiated in ES cells, KRAB zinc finger proteins with selective expression in ES cells are likely candidates. As we could not further shorten the SHIN sequence to less than 160 bp without compromising silencing potential, it is quite likely that multiple DNA binding proteins are involved in Trim28/Setdb1 recruitment.

Identification of the SHIN sequence as a strong heterochromatin nucleation site allowed a pilot screen for regulators of heterochromatin establishment through which we identified the chromatin remodeler Atrx. Atrx co-localizes with Trim28/Setdb1 and

H3K9me3 on IAP elements, other ERV classes, and several single copy loci (Fig 5A, Supplementary Fig S7A–D). We found that Atrx is *de novo* recruited to Setdb1-dependent heterochromatin (Fig 5C), suggesting that Atrx could be directly recruited by Setdb1 or Trim28. However, as Atrx features domains which bind H3K9me2/3 and HP1 [26,27], a complex combinatorial nature of interactions leading to Atrx recruitment is likely. This may be further complicated by potential interactions with non-coding RNAs [53] that may affect heterochromatin formation on IAP elements [54].

We found that in the absence of Atrx reporter silencing is delayed and spreading of heterochromatin is reduced (Fig 4), which could be due to reduced activity of Setdb1 in establishing H3K9me3. Being a putative chromatin remodeler, we hypothesize that Atrx assists in generating a proper nucleosomal array which can then serve as efficient substrate for Setdb1. Consistent with this hypothesis, we found that reduced Setdb1 activity leads to defects in the establishment of silencing when Atrx is not present (Fig 3E). Secondly, heterochromatin spreading and stability rely on the formation of properly spaced nucleosomal arrays by HP1 molecules binding to H3K9me3 [55]. Atrx activity may be required to establish such arrays, and, as a consequence, loss of Atrx would lead to compromised spreading. Alternatively, Atrx may also promote heterochromatin spreading by supporting Suv39h activity at retrotransposons [49]. A third function of Atrx in heterochromatin establishment may be the recruitment of additional factors to assist in this process. On telomeres, Atrx interacts with Daxx to facilitate histone H3.3 deposition [28–30]. We found that H3.3 is dispensable for silencing retrotransposon sequences in reporter assays, suggesting that incorporation of this histone variant is not important for Trim28/Setdb1-dependent heterochromatin formation. Instead, we find that Daxx is important for this process (Supplementary Fig S3). Recently, Daxx was found to be involved in silencing of exogenous ASV retroviruses by H3K9me3 and DNA methylation [56]. Although it is unclear how silencing is triggered in this context and if Atrx, Trim28, or Setdb1 are involved in this process, these data suggest Daxx as a global regulator of retrovirus repression. Interestingly, we find that Daxx requires a functional SUMO interaction motif to contribute to Setdb1-/Trim28-dependent repression (Supplementary Fig S3), suggesting that the interaction with sumoylated proteins is important in this context [42]. This is consistent with the repressive activity of Daxx when recruited by sumoylated transcription factors [42]. As Setdb1 recruitment to Trim28 has been shown to be SUMO dependent, and because several heterochromatin proteins, like HP1, are sumoylated [57,58], it is likely that Daxx recruitment might rely on the sumoylation of Trim28 or other factors. More thorough genome-wide screens for retrotransposon silencing are necessary to identify such interaction partners and additional components of this silencing pathway.

Heterochromatin formation is the net result of a dynamic balance between factors that build up heterochromatic structures and processes which act antagonistically [59]. Thus, the efficiency of the heterochromatin buildup machinery is particularly critical in genomic regions with prominent activity of antagonistic processes, for example, transcription. Examples for such regions are retrotransposons which reside in intronic regions of highly expressed genes. Unfortunately, due to the lack of polymorphisms, it is not possible to directly examine heterochromatin formation on these elements. However, we could study heterochromatin establishment in competition with transcriptional activity on the RMCE reporter locus. In wild-type cells, this reporter locus could be efficiently silenced despite strong promoter activity. Atrx knockdown leads to impaired silencing. These data clearly demonstrate that in the absence of Atrx, the activity of the Trim28/Setdb1 pathway is strongly reduced. Importantly, the de-repression of MusD/ETn elements in *Atrx* ko cells represents a physiological situation in which the full activity of the Setdb1/Trim28 pathway is necessary for silencing (Fig 5E). The establishment of new heterochromatin domains, for example, in the context of differentiation, may also require Atrx for the full activity of the Setdb1/Trim28 pathway. The strong developmental phenotype of *Atrx* ko embryos [60] supports this hypothesis, although more analyses are necessary to identify Atrx target loci during developmental transitions.

Heterochromatin is generally characterized by low transcriptional activity. A likely explanation for this property is that binding of transcription factors and RNA polymerase is restricted by the largely inaccessible chromatin structure. We tested chromatin accessibility of different genomic regions in mouse ES cells and found that pericentric heterochromatin is highly inaccessible to MNase. A similarly low accessibility was only observed for a promoter region which is under control of the polycomb system (Six3, Fig 6E). IAP retrotransposons display higher chromatin accessibility as compared to pericentric heterochromatin, which may be due to a different organization of these regions. Pericentric heterochromatin is composed of large domains of major satellite repeats which cluster into higher order structures to form so-called chromocenters. As retrotransposons form much smaller domains of only several kilobases, higher order folding of the chromatin structure is probably limited. However, chromatin accessibility of IAP elements is still much lower as compared to active promoters and may protect these regions from various challenges. Thus, it is currently unclear if the observed differences in chromatin accessibility result from changes in large-scale chromatin compaction or from an altered local nucleosomal organization. In this study, we found that chromatin accessibility in *Atrx* ko cells is significantly increased on IAP elements, demonstrating that Atrx is important for proper heterochromatin organization. Interestingly, this more accessible heterochromatin architecture is indeed vulnerable to challenges. For example, strong transcriptional activators can more easily access their binding sites and lead to stronger transcriptional activation in *Atrx* ko cells. If Atrx-deficient heterochromatin is vulnerable to other challenges, for example, DNA damage, remains to be tested. Interestingly, chromatin architecture is not only altered on IAP retrotransposons. We also tested non-repetitive genomic regions with the enrichment of Atrx and H3K9me3 and consistently detected increased chromatin accessibility. Thus, our data provide strong evidence for a general role of Atrx in the establishment and robust maintenance of heterochromatin domains.

Materials and Methods

Reporter gene silencing assay

Reporter constructs were stably integrated into cells by lentiviral transduction, and the percentage of $EGFP^+$ cells was measured by FACS after 2–4 days. Lentiviral particles were generated using standard protocols, and virus titers were determined by titration in HeLa cells. Mouse ES cells were transduced on gelatinized multi-well dishes using spinoculation at a low multiplicity of

infection to ensure a linear relationship between virus titer and transduction rate. The ratio of the percentage of $EGFP^+$ cells generated by the reporter relative to the percentage of $EGFP^+$ cells generated by a control EGFP vector of the same virus titer was used to quantify reporter silencing (relative % $EGFP^+$ cells). In Fig 1 and Supplementary Fig S1, the reciprocal ratio (fold repression) is plotted.

shRNA screen

Feeder-independent ES cells were transduced with a genome-wide pooled shRNA library (Lentiplex, Sigma-Aldrich) at a low MOI, selected for shRNA expression using 0.5 μg/ml puromycin, and then transduced with the SHIN reporter. Cells that escaped SHIN silencing were isolated by FACS sorting based on their high EGFP fluorescence. shRNA sequences of the $EGFP^+$ cells were amplified by PCR and cloned into the pLKO1 backbone. For validation of hits, 71 random shRNA sequences of the primary screen were separately transduced into ES cells and the SHIN reporter assay was performed as indicated above. The number of EGFP-positive cells was compared to a non-silencing scrambled control. Details are given in the Supplementary Materials and Methods.

Resilencing assay

T86 cells (HA36::SHIN cells harboring a reverse Tet transactivator) were treated with 2 μg/ml doxycycline for 2 days, and the re-expression of the silenced reporter locus was selected with 50 μg/ml zeocin and 2 μg/ml doxycycline for at least 2 weeks. Reactivated reporter cells were then transduced with a lentiviral shRNA targeting either Atrx or a scrambled control sequence, and knockdown cells were selected by the addition of 1 μg/ml puromycin 2 days after transduction. After additional 3 days, zeocin selection was released and resilencing of the EGFP reporter was monitored daily in the presence or absence of doxycyline by FACS. The median EGFP fluorescence was calculated using FlowJo™ (TreeStar). For ChIP experiments, cells were harvested 4 days after zeocin release. Details are given in the Supplementary Materials and Methods.

Supplementary information for this article is available online: http://embor.embopress.org

Acknowledgements

Work in the lab of G.S. was funded by grants from the Deutsche Forschungsgemeinschaft (SFB 1064, SPP1356). F.J.T. and C.F. were supported by ERC grant "LatentCauses" and SPP1356. We thank Sandra Hake for providing antibodies. We are grateful to Tobias Straub and Felix Müller-Planitz for critical discussions. We thank Dirk Schübeler for generously providing HA36 ES cells. RMCE plasmids were provided by Matthew Lorincz and Dirk Schübeler.

Author contributions

GS and DS designed the experimental approach; DS, KS and SG carried out the bench experiments; JE performed the FACS sorting; IK, CF and FJT developed and carried out the statistical analyses; and GS and DS wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Bourc'his D, Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431: 96 – 99
2. Wilkins AS (2010) The enemy within: an epigenetic role of retrotransposons in cancer initiation. *BioEssays* 32: 856 – 865
3. Rowe HM, Kapopoulou A, Corsinotti A, Fasching L, Macfarlan TS, Tarabay Y, Viville S, Jakobsson J, Pfaff SL, Trono D (2013) TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. *Genome Res* 23: 452 – 461
4. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M (2014) Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res* 24: 1053 – 1063
5. 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
6. 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
7. Howlett SK, Reik W (1991) Methylation levels of maternal and paternal genomes during preimplantation development. *Development* 113: 119 – 127
8. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* 35: 88 – 93
9. Wolf D, Goff SP (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* 458: 1201 – 1204
10. 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
11. Turelli P, Castro-Diaz N, Marzetta F, Kapopoulou A, Raclot C, Duc J, Tieng V, Quenneville S, Trono D (2014) Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res* 24: 1260 – 1270
12. Leung D, Du T, Wagner U, Xie W, Lee AY, Goyal P, Li Y, Szulwach KE, Jin P, Lorincz MC *et al* (2014) Regulation of DNA methylation turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1. *Proc Natl Acad Sci USA* 111: 6690 – 6695
13. 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
14. 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
15. Kim SS, Chen YM, O'Leary E, Witzgall R, Vidal M, Bonventre JV (1996) A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. *Proc Natl Acad Sci USA* 93: 15299 – 15304
16. Le Douarin B, Nielsen AL, Garnier JM, Ichinose H, Jeanmougin F, Losson R, Chambon P (1996) A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. *EMBO J* 15: 6701 – 6715
17. Moosmann P, Georgiev O, Le Douarin B, Bourquin JP, Schaffner W (1996) Transcriptional repression by RING finger protein TIF1 beta that

interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res* 24: 4859 – 4867

18. 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
19. Iyengar S, Farnham PJ (2011) KAP1 protein: an enigmatic master regulator of the genome. *J Biol Chem* 286: 26267 – 26276
20. Gibbons RJ, Picketts DJ, Villard L, Higgs DR (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* 80: 837 – 845
21. De La Fuente R, Viveiros MM, Wigglesworth K, Eppig JJ (2004) ATRX, a member of the SNF2 family of helicase/ATPases, is required for chromosome alignment and meiotic spindle organization in metaphase II stage mouse oocytes. *Dev Biol* 272: 1 – 14
22. Ritchie K, Seah C, Moulin J, Isaac C, Dick F, Berube NG (2008) Loss of ATRX leads to chromosome cohesion and congression defects. *J Cell Biol* 180: 315 – 324
23. Wong LH, McGhie JD, Sim M, Anderson MA, Ahn S, Hannan RD, George AJ, Morgan KA, Mann JR, Choo KH (2010) ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res* 20: 351 – 360
24. Gibbons RJ, McDowell TL, Raman S, O'Rourke DM, Garrick D, Ayyub H, Higgs DR (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat Genet* 24: 368 – 371
25. Kernohan KD, Jiang Y, Tremblay DC, Bonvissuto AC, Eubanks JH, Mann MR, Berube NG (2010) ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. *Dev Cell* 18: 191 – 202
26. Eustermann S, Yang JC, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D *et al* (2011) Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* 18: 777 – 782
27. Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H *et al* (2011) ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* 18: 769 – 776
28. Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD (2010) Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci USA* 107: 14075 – 14080
29. Drane P, Ouararhni K, Depaux A, Shuaib M, Hamiche A (2010) The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev* 24: 1253 – 1265
30. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X *et al* (2010) Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140: 678 – 691
31. Cheung NK, Zhang J, Lu C, Parker M, Bahrami A, Tickoo SK, Heguy A, Pappo AS, Federico S, Dalton J *et al* (2012) Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *JAMA* 307: 1062 – 1071
32. Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S *et al* (2011) Altered telomeres in tumors with ATRX and DAXX mutations. *Science* 333: 425
33. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schulick RD, Tang LH, Wolfgang CL, Choti MA *et al* (2011) DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* 331: 1199 – 1203
34. Liu XY, Gerges N, Korshunov A, Sabha N, Khuong-Quang DA, Fontebasso AM, Fleming A, Hadjadj D, Schwartzentruber J, Majewski J *et al* (2012) Frequent ATRX mutations and loss of expression in adult diffuse astrocytic tumors carrying IDH1/IDH2 and TP53 mutations. *Acta Neuropathol* 124: 615 – 625
35. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M *et al* (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482: 226 – 231
36. Lovejoy CA, Li W, Reisenweber S, Thongthip S, Bruno J, de Lange T, De S, Petrini JH, Sung PA, Jasin M *et al* (2012) Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet* 8: e1002772
37. Wolf D, Goff SP (2007) TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell* 131: 46 – 57
38. Hutnick LK, Huang X, Loo TC, Ma Z, 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
39. 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
40. Baubec T, Ivanek R, Lienert F, Schubeler D (2013) Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* 153: 480 – 492
41. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D (2011) Identification of genetic elements that autonomously determine DNA methylation states. *Nat Genet* 43: 1091 – 1097
42. Lin DY, Huang YS, Jeng JC, Kuo HY, Chang CC, Chao TT, Ho CC, Chen YC, Lin TP, Fang HI *et al* (2006) Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Mol Cell* 24: 341 – 354
43. Law MJ, Lower KM, Voon HP, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A *et al* (2010) ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* 143: 367 – 378
44. 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
45. Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, Duc J, Jang SM, Turelli P, Trono D (2014) Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev* 28: 1397 – 1409
46. Hahn M, Dambacher S, Dulev S, Kuznetsova AY, Eck S, Worz S, Sadic D, Schulte M, Mallm JP, Maiser A *et al* (2013) Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev* 27: 859 – 872
47. Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A *et al* (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 107: 323 – 337

48. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA *et al* (2008) A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* 22: 2048 – 2061
49. Bulut-Karslioglu A, De La Rosa-Velazquez IA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, Galan C, Winter GE, Engist B, Gerle B *et al* (2014) Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol Cell* 55: 277 – 290
50. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR (2012) Dynamics and memory of heterochromatin in living cells. *Cell* 149: 1447 – 1460
51. Bulut-Karslioglu A, Perrera V, Scaranaro M, de la Rosa-Velazquez IA, van de Nobelen S, Shukeir N, Popow J, Gerle B, Opravil S, Pagani M *et al* (2012) A transcription factor-based mechanism for mouse heterochromatin formation. *Nat Struct Mol Biol* 19: 1023 – 1030
52. Muller-Ott K, Erdel F, Matveeva A, Mallm JP, Rademacher A, Hahn M, Bauer C, Zhang Q, Kaltofen S, Schotta G *et al* (2014) Specificity, propagation, and memory of pericentric heterochromatin. *Mol Syst Biol* 10: 746
53. Sarma K, Cifuentes-Rojas C, Ergun A, Del Rosario A, Jeon Y, White F, Sadreyev R, Lee JT (2014) ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell* 159: 869 – 883
54. Bierhoff H, Dammert MA, Brocks D, Dambacher S, Schotta G, Grummt I (2014) Quiescence-induced LncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol Cell* 54: 675 – 682
55. Canzio D, Chang EY, Shankar S, Kuchenbecker KM, Simon MD, Madhani HD, Narlikar GJ, Al-Sady B (2011) Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol Cell* 41: 67 – 81
56. Shalginskikh N, Poleshko A, Skalka AM, Katz RA (2013) Retroviral DNA methylation and epigenetic repression are mediated by the antiviral host protein Daxx. *J Virol* 87: 2137 – 2150
57. 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
58. Cubenas-Potts C, Matunis MJ (2013) SUMO: a multifaceted modifier of chromatin structure and function. *Dev Cell* 24: 1 – 12
59. Ebert A, Schotta G, Lein S, Kubicek S, Krauss V, Jenuwein T, Reuter G (2004) Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes Dev* 18: 2973 – 2983
60. Garrick D, Sharpe JA, Arkell R, Dobbie L, Smith AJ, Wood WG, Higgs DR, Gibbons RJ (2006) Loss of Atrx affects trophoblast development and the pattern of X-inactivation in extraembryonic tissues. *PLoS Genet* 2: e58

Publication 2

Morc3 silences endogenous retroviruses by enabling Daxx mediated histone H3.3 incorporation

Sophia Groh[1], Anna Viktoria Milton[1], Lisa Katherina Marinelli[1], Cara V. Sickinger[1], Angela Russo[1], Heike Bollig[1], Gustavo Pereira de Almeida[1], Andreas Schmidt[2], Ignasi Forné[2], Axel Imhof[2]and Gunnar Schotta[1]

[1] Division of Molecular Biology, Biomedical Center (BMC), Faculty of Medicine, Ludwig-Maximilians-University (LMU) Munich, Großhaderner Straße 9, 82152 Martinsried, Germany. [2] Protein Analysis Unit, Biomedical Center (BMC), Faculty of Medicine, Ludwig-Maximilians-University (LMU) Munich, Großhaderner Straße 9, 82152 Martinsried, Germany

This manuscripts is published in:
Nature communications. DOI: 10.1038/s41467-021-26288-7

Author contributions:

I shared the writing of the manuscript and design of the experimental approach with G.S. I performed most of the experiments, as the genome wide sgRNA screen, confirmed screening hits in SHIN initiation and maintenance silencing assays, generated wt26 based MORC3 KO, KI and some of the rescue cell lines, confirmed cell lines by western blot (WB) and FACS staining and also tested these cell lines in SHIN silencing assays. I further performed all RNA-seq, ChIP-seq (with exception of the SETDB1 ChIP) and ATAC-seq experiments and analyzed the resulting data. I performed ChIP for MS of the MORC3 deltaN mutant and contributed to the Co-IP of MORC3 mutants. A.V.M. performed ChIP for MS experiments with MORC3 and DAXX (WT and mutants). L.K.M. cloned and generated MORC3 rescue cell lines and performed SHIN initiation and maintenance assays. C.V.S. generated T90 based MORC3 KO cell lines and tested them in WB and SHIN assay. A.R. performed Co-IP WB of MORC3 and DAXX (WT and mutant). H.B. performed immunofluorescence staining microscopy. G.P.dA. performed SETDB1 ChIP-seq. A.S and I.F. performed MS experiments. A.I. contributed resources and competences for MS experiments. G.S. provided funding acquisition and project administration.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-26288-7 OPEN

Morc3 silences endogenous retroviruses by enabling Daxx-mediated histone H3.3 incorporation

Sophia Groh[1], Anna Viktoria Milton[1], Lisa Katherina Marinelli[1], Cara V. Sickinger[1], Angela Russo[1], Heike Bollig[1], Gustavo Pereira de Almeida[1], Andreas Schmidt[2], Ignasi Forné[2], Axel Imhof[2] & Gunnar Schotta[1✉]

Endogenous retroviruses (ERVs) comprise a significant portion of mammalian genomes. Although specific ERV loci feature regulatory roles for host gene expression, most ERV integrations are transcriptionally repressed by Setdb1-mediated H3K9me3 and DNA methylation. However, the protein network which regulates the deposition of these chromatin modifications is still incompletely understood. Here, we perform a genome-wide single guide RNA (sgRNA) screen for genes involved in ERV silencing and identify the GHKL ATPase protein Morc3 as a top-scoring hit. Morc3 knock-out (ko) cells display de-repression, reduced H3K9me3, and increased chromatin accessibility of distinct ERV families. We find that the Morc3 ATPase cycle and Morc3 SUMOylation are important for ERV chromatin regulation. Proteomic analyses reveal that Morc3 mutant proteins fail to interact with the histone H3.3 chaperone Daxx. This interaction depends on Morc3 SUMOylation and Daxx SUMO binding. Notably, in Morc3 ko cells, we observe strongly reduced histone H3.3 on Morc3 binding sites. Thus, our data demonstrate Morc3 as a critical regulator of Daxx-mediated histone H3.3 incorporation to ERV regions.

[1] Division of Molecular Biology, Biomedical Center (BMC), Faculty of Medicine, Ludwig-Maximilians-University (LMU) Munich, Großhaderner Straße 9, 82152 Martinsried, Germany. [2] Protein Analysis Unit, Biomedical Center (BMC), Faculty of Medicine, Ludwig-Maximilians-University (LMU) Munich, Großhaderner Straße 9, 82152 Martinsried, Germany. ✉email: gunnar.schotta@bmc.med.lmu.de

Endogenous retroviruses compose a significant portion of mammalian genomes. During evolution, most ERV integrations in mammals were highly mutated or partially deleted and are thus unable to generate functional retroviral particles. However, remnant ERVs harbor binding sites for transcription factors and can act as regulatory elements to drive host gene expression[1–6]. This physiological role for the regulation of normal development is contrasted by pathological effects of aberrant ERV regulation in neurological diseases[7] and cancer[8–10]. Many ERV families are thus transcriptionally repressed by a silencing pathway that involves the formation of H3K9me3 heterochromatin[11,12]. The current model for H3K9me3 establishment on ERVs is based on sequence-specific binding of KRAB-ZnF proteins, which recruit the co-repressor Trim28 and the histone methyltransferase Setdb1[12–14]. The establishment of DNA methylation is likely to happen during early embryogenesis, and maintenance of high methylation levels on ERVs is later ensured by Uhrf1/Dnmt1[15,16]. However, recent data indicate that Dnmt1/Uhrf1 can also induce de novo DNA methylation on ERVs[17]. Full establishment of H3K9me3 and DNA methylation, as well as reduced chromatin accessibility on ERVs, require activities, such as histone H3.3 deposition by Atrx/Daxx[18–21], chromatin remodeling by Smarcad1[22], and chromatin assembly by Chaf1a/b[23]. In ES cells, chromatin remodeling activities enforce a dynamic exchange of Histone H3.3[24]. In this context, histone H3.3 deposition is necessary to replace evicted nucleosomes and ensure low chromatin accessibility[24]. Currently, it is not clear how histone H3.3 turnover is regulated and coordinated with other chromatin-modifying activities to restrict chromatin accessibility and to mediate heterochromatin spreading on ERVs.

We have previously identified a small heterochromatin inducing sequence (SHIN) in Intracisternal A Type particle (IAP) elements[20]. The SHIN sequence is a 160 bp region from the GAG coding sequence in IAPEz elements. Genomic insertions of reporters containing the SHIN sequence attract Setdb1-dependent H3K9me3[20]. Using a genome-wide single-guide RNA (sgRNA) screen for genes involved in SHIN silencing we now identified Morc3 as a player in ERV silencing. Morc3 is a GHKL type ATPase protein that can form a closed dimer in the ATP-bound state[25]. It further contains a CW-type zinc finger domain that negatively regulates ATPase activity[26]. The binding of ligands, such as histone H3K4me3 or influenza virus protein NS1 peptides, to the CW domain, relieves suppression of ATPase activity and could regulate the turnover of ATPase domain-mediated dimerization and opening[27,28]. Although H3K4me3-mediated interaction of Morc3 with promoter regions in mouse ES cells was reported[25], the molecular roles of Morc3 in transcriptional regulation or chromatin organization are currently unclear. We demonstrate that Morc3 binds ERV sequences and that loss of Morc3 results in increased chromatin accessibility, reduced H3K9me3, and de-repression of ERVs. We detect an interaction of Morc3 with the histone H3.3 chaperone Daxx, which depends on the Morc3-ATPase cycle and SUMOylation. This interaction is crucial for Daxx-mediated histone H3.3 incorporation as Morc3 ko ES cells lose H3.3 on ERVs. Thus, our data indicate Morc3 as a regulator of Daxx-mediated histone H3.3 incorporation.

Results

Identification of Morc3 as an ERV silencing factor. To identify factors regulating IAPEz silencing we performed a genome-wide sgRNA screen based on the integration of an IAP SHIN reporter in mouse ES cells (T90 cells, Fig. 1a). This reporter contains a doxycycline-inducible promoter driving the expression of EGFP and a zeocin resistance gene. In the wild-type condition, doxycycline induction results in poor reporter activation due to its heterochromatic state, whereas impaired SHIN silencing allows doxycycline-induced reporter activity[20]. We used a pooled genome-wide lentiviral sgRNA library[29] to transduce T90 ES cells, followed by doxycycline induction to activate the reporter locus. Reporter activity also results in zeocin resistance. Therefore, we applied zeocin selection to detect cells with an activated reporter resulting from impaired heterochromatin. As a control, we collected a second pool of cells without selection pressure. We then identified the sgRNAs which were enriched in the zeocin selected cells (Fig. 1b, Supplementary Data 1). Top enriched sgRNAs targeted the major known ERV silencing factors, such as *Dnmt1*, *Uhrf1*, *Setdb1*, *Trim28*, and *Atrx*/*Daxx* (Fig. 1b). Another top hit was Morc3, a protein not previously implicated in ERV silencing. Due to the mode of selection with zeocin, we also identified genes that might be related to zeocin resistance, DNA damage repair, and apoptosis (Supplementary Data 1). We compared our data with a similar screen for IAP silencing factors which was recently published[30]. Overall, we found very little overlap between the two screens. Only the major silencing factors, such as Atrx, Daxx, Dnmt1, Uhrf1, Trim28, and our top hit Morc3 were present in both datasets (Supplementary Fig. S1a). Different silencing inducing sequences (SHIN sequence vs. 5′ LTR-UTR) used in the respective reporters could be responsible for the differences in the screening hits.

To validate selected screening hits, we performed SHIN silencing maintenance assays for candidate gene knock-outs (Fig. 1c). Control sgRNA treatment did not result in an activatable SHIN reporter, whereas sgRNA knock-out of Setdb1 and Morc3 resulted in significant SHIN de-repression (Fig. 1d, FACS gating strategy in Supplementary Fig. S1b). We further tested additional screening candidates in the SHIN maintenance assay and could partially confirm our screening results (Fig. 1e). We then complemented the functional testing with a SHIN initiation silencing assay in which we test whether a newly introduced SHIN reporter with a constitutively active promoter can be silenced in different genetic backgrounds (Fig. 1f). Here, the results differed from the maintenance assay as the DNA methylation pathway appeared less important for establishing silencing, while Setdb1/Trim28 and Atrx/Daxx were still critical for silencing (Fig. 1g). Since Morc3 was important in both maintenance and initiation of silencing we decided to functionally characterize this protein in more detail.

Morc3 binds different families of endogenous retroviruses. To assess if Morc3 directly regulates ERV silencing we performed ChIP-seq experiments using a knock-in cell line that expresses 3xFLAG tagged Morc3 at endogenous levels (Supplementary Fig. S2). We identified 2737 peaks that were shared between at least two out of three replicates (Supplementary Data 2). Annotation with genomic features revealed that most Morc3 peaks associate with ERVs (Fig. 2a). We then classified the ERV families to which Morc3 peaks associate and found a major enrichment with IAPEz, LTRIS2, and RLTR families (Fig. 2b). Due to the lack of polymorphisms, many ERV integrations cannot be precisely mapped. To better assess the association of Morc3 with ERV families we used RepEnrich2[31] to categorize all mappable Morc3 ChIP-seq reads into ERV families. This analysis confirmed the prominent association of Morc3 with IAP, RLTR, and LTRIS families (Fig. 2c, Supplementary Data 3). The most prominent ERV family bound by Morc3 are IAPEz elements, from which the SHIN sequence used in our screen originates. We generated cumulative coverage plots for IAPEz elements to determine whether Morc3 displays preferential binding with distinct IAPEz features. Interestingly, comparison of the Morc3 profile with ChIP-seq profiles for Setdb1, Trim28, and H3K9me3 (Fig. 2d) revealed prominent co-enrichment of the silencing factors with

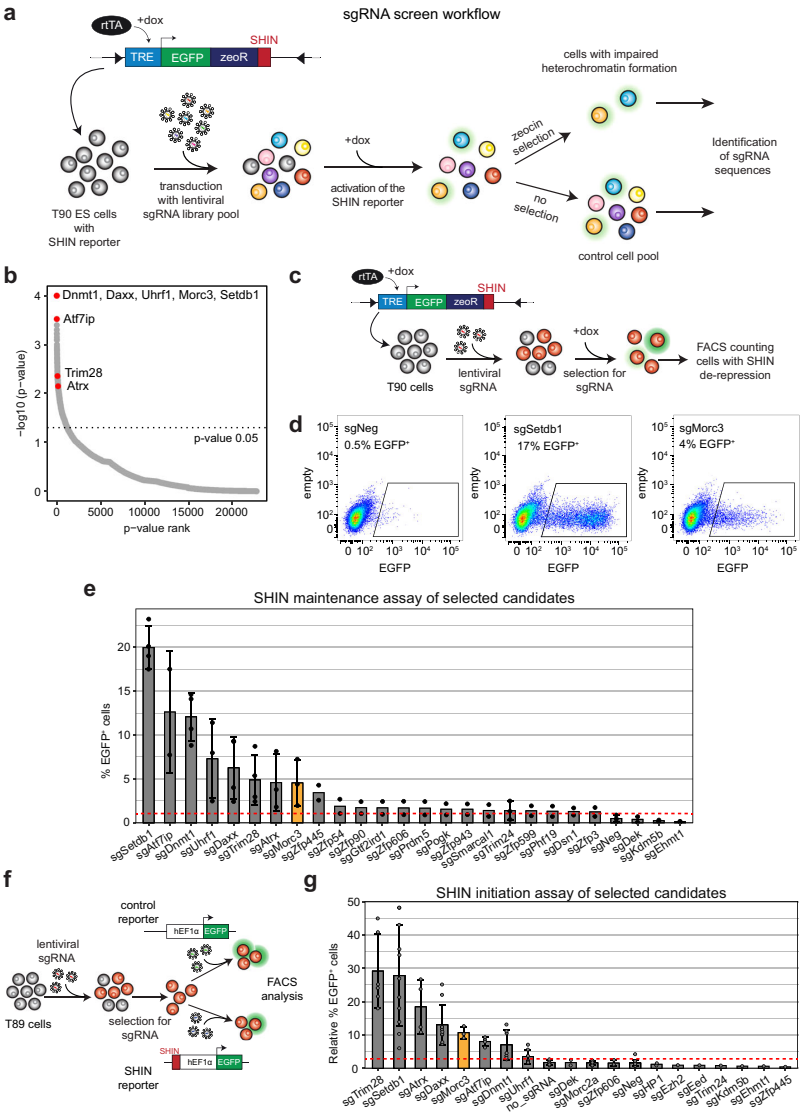


Fig. 1 A genome-wide sgRNA screen for SHIN silencing identifies Morc3. a Setup of the sgRNA screen. T90 ES cells containing an inducible SHIN reporter[20] were transduced with a genome-wide sgRNA library pool. Cells were selected for sgRNA vectors and the reporter was activated with doxycycline. Two independent cell batches were cultured to harvest non-selected cells representing the control, and zeocin-resistant cells representing cells with impaired heterochromatin on the SHIN reporter. **b** Dot plot showing the sgRNA screen results ordered by *p*-value rank (RIGER SecondBestRank scoring). Major ERV silencing factors are indicated. Morc3 represents a top hit in the screen. **c** Schematic of the SHIN silencing maintenance assay. T90 ES cells were transduced with sgRNAs for selected candidates and treated for integration with puromycin. Subsequently, doxycycline was added to the cells to induce reporter activity. EGFP expression was analyzed by FACS. **d** FACS plots depicting EGFP fluorescence in cells with activated SHIN reporter. Almost no activity was detected in cells transduced with a control sgRNA, demonstrating full SHIN silencing. sgRNAs targeting *Setdb1* or *Morc3* result in SHIN de-repression as indicated by cells showing EGFP expression. **e** Bar plot depicting the results of the SHIN maintenance silencing assay with selected candidate genes. The red dotted line indicates the background reporter activity. Data are presented as mean values +/− SD of biological replicates per sgRNA ($n = 2$–4, for details see "Statistics and reproducibility" section). **f** Schematic of the SHIN silencing initiation assay. T89 ES cells were transduced with sgRNAs for selected candidates and treated for integration with puromycin. Subsequently, the cells were transduced with a control virus without the SHIN sequence or the SHIN reporter with a constitutive promoter, respectively. EGFP expression was analyzed by FACS and the percentage of EGFP expressing cells, relative to the control reporter was calculated. **g** Bar plot depicting the results of the SHIN initiation silencing assay with selected candidate genes. The red dotted line indicates the background reporter activity. Data are presented as mean +/− SD of biological replicates per sgRNA ($n = 1$–12, for details see "Statistics and reproducibility" section).

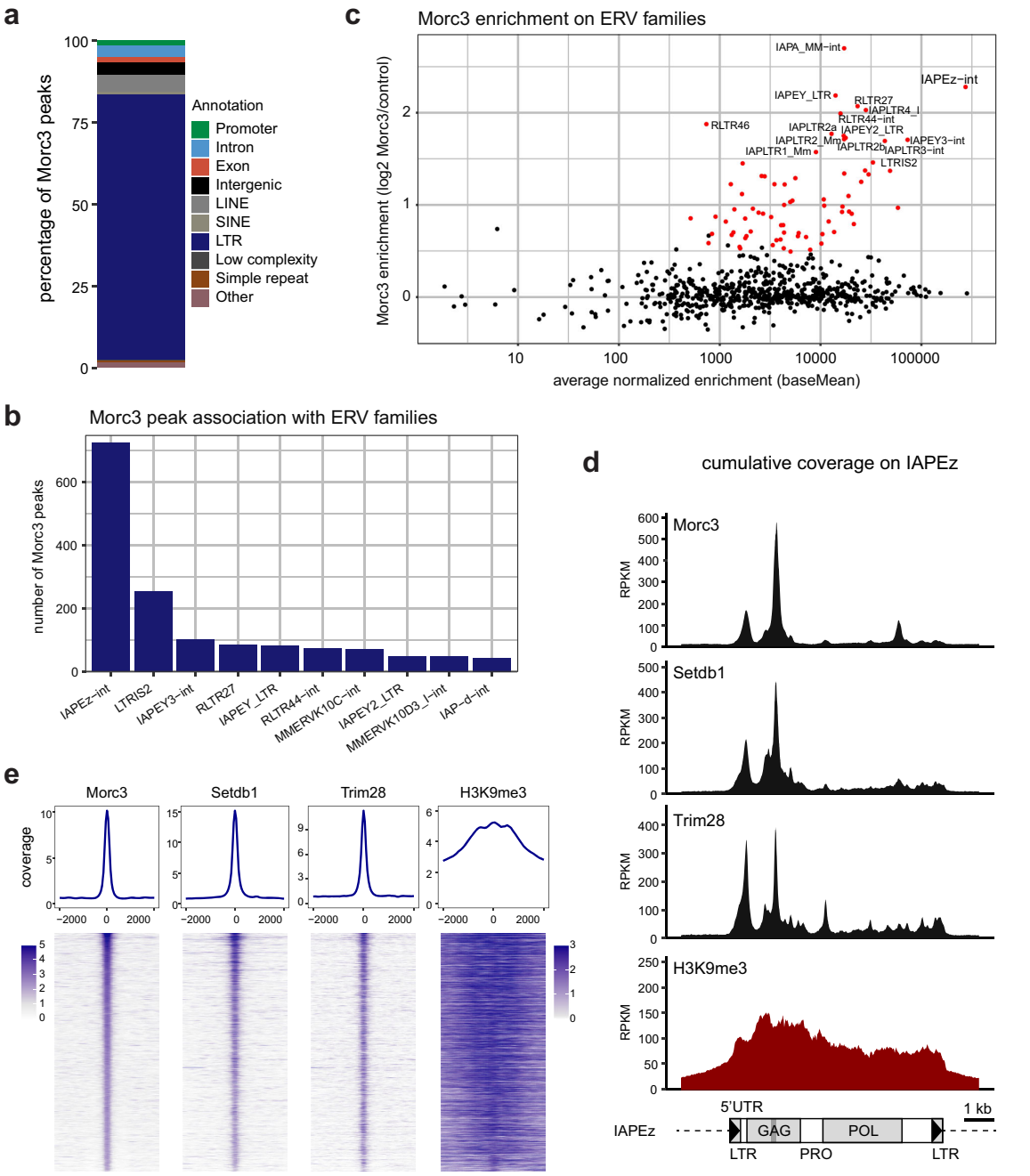


Fig. 2 Morc3 is enriched on different families of endogenous retroviruses. a Genomic features associated with the 2737 Morc3 peaks. **b** Morc3 peak association with ERV families. The bar plot shows the number of Morc3 peaks associating with the ten most enriched ERV families. **c** Dot plot showing Morc3 ChIP enrichment on ERV families. Red dots highlight ERV families with significant Morc3 over background enrichment. Selected ERV families are labeled. **d** Cumulative coverage plot of Morc3, Setdb1, Trim28, and H3K9me3 ChIP-seq profiles on IAPEz elements. Prominent enrichment is over the 5′ UTR and the GAG region. The position of the SHIN sequence is indicated as a dark gray bar. Setdb1 ChIP-seq is from a 3xFLAG knock-in ES cell line. Trim28 ChIP-seq data are from ref. [23]. **e** Morc3 peaks are associated with Setdb1 and Trim28 binding and feature high H3K9me3. Top panel: Density plot showing the average occupancy of Morc3, Setdb1, Trim28, and H3K9me3 on Morc3 peaks. Lower panel: Read-density heat map showing the normalized coverage of Morc3, Setdb1, Trim28, and H3K9me3 on Morc3 binding sites. Distance from the peak center is given in bp. Setdb1 ChIP-seq is from a 3xFLAG knock-in ES cell line. Trim28 ChIP-seq data are from ref. [23].

the IAP-GAG region containing the SHIN sequence and with the 5′UTR region that is also able to induce reporter silencing[14]. The similarity of the Morc3 binding pattern to Setdb1 and Trim28 suggests a functional interplay with the H3K9me3 pathway (Fig. 2d). To investigate whether Morc3 more generally associates with Setdb1 and Trim28 on its identified peaks, we generated a read-density heatmap of Setdb1, Trim28, and H3K9me3 on Morc3 peaks (Fig. 2e). This analysis also revealed prominent enrichment of Setdb1, and Trim28 on Morc3 peaks, further supporting Morc3 association with H3K9me3 repressed heterochromatin.

Morc3 knock-out results in ERV de-repression. To investigate the role of Morc3 in ERV silencing we generated Morc3 knock-out ES cells (Supplementary Fig. S3a). Transcriptome analysis by RNA-seq revealed significant dysregulation of 252 genes (adjusted p-value < 0.01) with a trend towards genes being upregulated (Fig. 3a, Supplementary Data 4). We then asked if Morc3 could be directly involved in the regulation of these genes by determining Morc3 peaks in the vicinity of their transcription start sites. We found that 64 upregulated and 18 downregulated genes had Morc3 peaks in <100 kb distance to their promoter (Supplementary Data 5). When we investigated the detailed peak annotation of these Morc3 peaks we found a strong association with ERV LTR sequences. The highest enriched ERV family was LRTIS2, which was found on 25 Morc3 peaks associated with upregulated genes (Supplementary Data 5). These data suggest that genes are indirectly regulated through de-repressed ERV LTRs which could act as enhancer elements for neighboring genes. Three prominent examples for upregulated genes are shown in Fig. 3B. *Irak3* has two Morc3 peaks in its gene body which associate with RLTR13B2 and ORR1A3, respectively. *Ube2l6* features Morc3 binding in an LTRIS2 element, overlapping the 3′UTR region. *Cd200* shows Morc3 binding with an LTRIS2 element downstream of its gene body. To attribute the effects on gene regulation to Morc3 function, we generated rescue ES cell lines by transducing the full-length Morc3 cDNA into Morc3 ko ES cells (Supplementary Fig. S3b). RNA-seq analysis of Morc3 rescue cell lines revealed largely normalized gene expression compared to Morc3 ko cells (Supplementary Fig. S4a). We validated selected Morc3 target genes by RT-qPCR in wild type, Morc3 ko, and rescue cells. All selected target genes displayed de-repression in Morc3 ko, whereas expression was reduced in rescue cells (Fig. 3c).

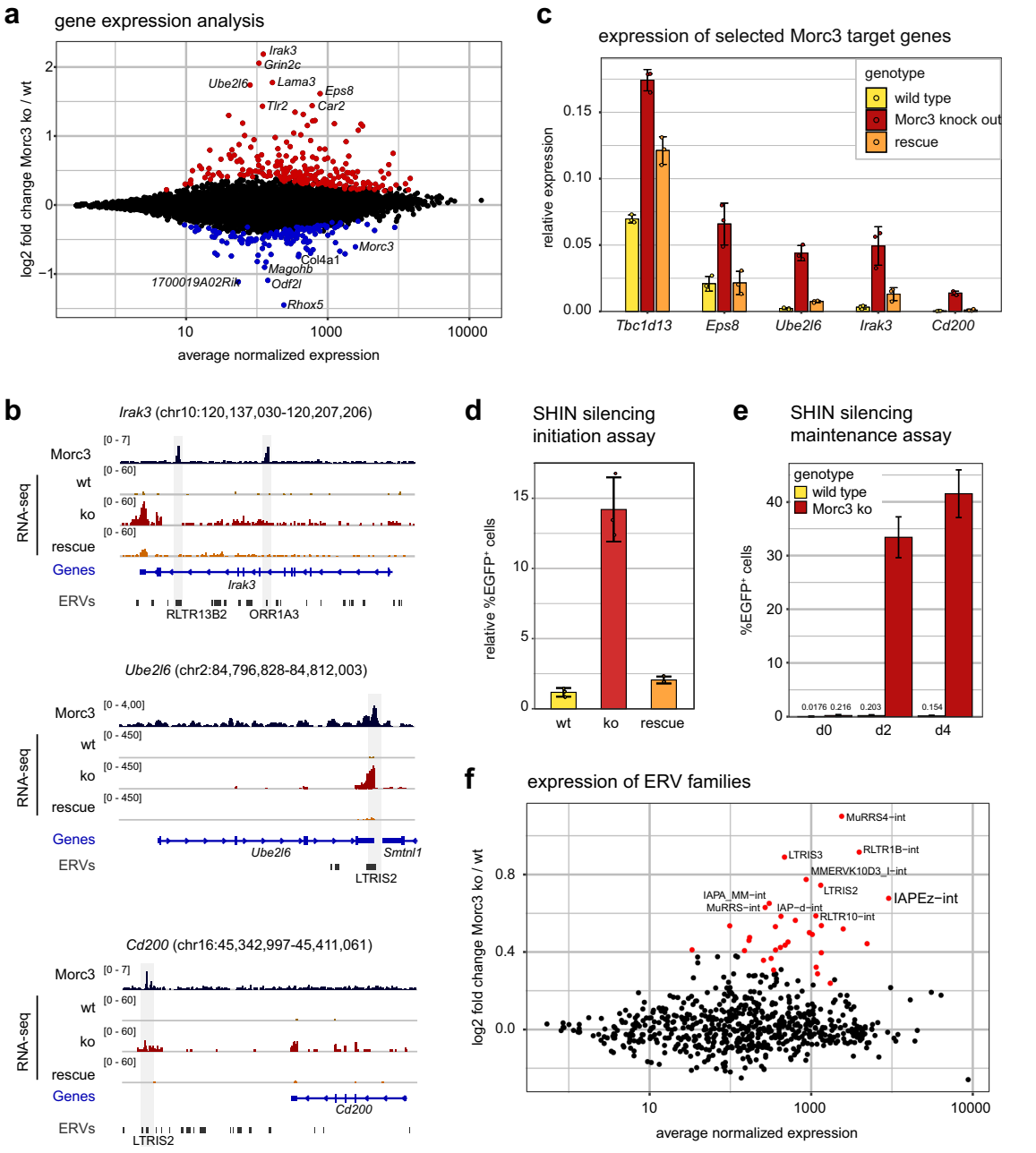
To determine Morc3 roles in ERV repression we performed SHIN silencing initiation assays. Morc3 knock-out cells displayed impaired SHIN silencing, whereas rescue cells can efficiently establish SHIN repression (Fig. 3d). We also generated Morc3 knock-out cells from T90 SHIN reporter ES cells to investigate Morc3-dependent silencing maintenance (Supplementary Fig. S3c). Without doxycycline induction, no EGFP expression could be detected in wild-type or Morc3 ko cells (Fig. 3e). We then monitored EGFP expression from the SHIN reporter with 2 and 4 days of doxycycline induction. Wild-type cells failed to activate the reporter and almost no EGFP positive cells could be detected (Fig. 3e). In contrast, silencing of the SHIN reporter was strongly impaired in Morc3 ko cells (Fig. 3e), demonstrating that Morc3 is important for the maintenance of SHIN silencing.

To determine if IAP and potentially other ERV families are de-repressed in Morc3 ko cells, we counted RNA-seq reads corresponding to ERVs with RepEnrich2 and calculated differentially expressed ERV families (Fig. 3f, Supplementary Data 6). We found that indeed, major Morc3 target families, such as IAP, RLTR, and LTRIS displayed significant de-repression, which can be restored in Morc3 rescue cells (Fig. 3f, Supplementary Fig. S4b). The extent of regulation is greatly reduced compared to Setdb1 knock-out where IAP de-repression is around 10–100-fold[13,32]. Therefore, our data suggest that Morc3 is a contributing factor for ERV silencing that acts redundantly with other mechanisms in the context of ERVs. However, individual ERV integrations, probably with less redundancy in silencing mechanisms, are strongly de-repressed in Morc3 ko ES cells and affect the expression of neighboring genes.

Morc3 knock-out leads to reduced H3K9me3 and increased chromatin accessibility. To characterize the chromatin changes coinciding with Morc3 loss, we measured H3K9me3 in wild type, Morc3 ko, and rescue cells. Visual inspection of H3K9me3 profiles on Morc3 peaks revealed different patterns of changes. Therefore, we clustered Morc3 peaks according to changes in H3K9me3 patterns and generated a read-density heatmap of Morc3 peak clusters with H3K9me3 pattern in wild-type cells, together with fold change representation in Morc3 ko and rescue cells (Fig. 4a). We found that H3K9me3 was generally reduced in cluster I. Clusters II and III displayed a biased reduction of H3K9me3 towards one side of the peak center. Clusters IV and V showed reduced H3K9me3 mainly in the peak center and cluster VI did not show appreciable changes in H3K9me3. Overall, the average H3K9me3 signal on Morc3 peak centers is significantly reduced in Morc3 ko cells and fully restored in Morc3 rescue cells (Fig. 4b). The genomic features underlying the different Morc3 peak clusters cannot easily explain the different patterns of change as all clusters display prominent enrichment of LTR elements (Supplementary Fig. S5a).

H3K9me3 heterochromatin is characterized by low chromatin accessibility which could prevent efficient access of transcriptional activators[33,34]. To test if chromatin accessibility changes in Morc3 ko cells, we performed ATAC-seq experiments in wild type, Morc3 ko, and rescue ES cells. Altogether we identified 444 peaks with significantly higher accessibility in Morc3 ko ES cells, while only 10 peaks were less accessible (Fig. 4c). Increased chromatin accessibility is directly related to Morc3 loss as we found reduced chromatin accessibility of these peaks in Morc3 rescue cells (Fig. 4c). Only a subset of the regulated ATAC-seq peaks was directly overlapping with Morc3 peaks, suggesting direct and indirect effects of Morc3 on chromatin accessibility (Fig. 4c).

We tested if enhanced chromatin accessibility in Morc3 ko could relate to activation of repressed enhancers by performing H3K27ac ChIP-seq. We observed a subset of Morc3 peaks which showed increased H3K27ac in Morc3 ko cells (Fig. 4d). These regions appeared to associate with low levels of H3K27ac already in wild-type cells (Supplementary Fig. S5b). We then asked for the overlap of H3K9me3 reduction, H3K27ac increase, and higher ATAC-seq accessibility in Morc3 ko cells. This analysis revealed that almost all peaks that displayed signs of enhancer activity (higher H3K27ac/ATAC-seq accessibility) displayed reduced H3K9me3 (Fig. 4e). The genomic features of these Morc3 peaks are characterized by ERV integrations and putative ENCODE enhancers (Supplementary Fig. S5c–e), supporting the view that Morc3 helps to repress distinct developmental enhancers. An example of a Morc3 repressed enhancer is shown in the genome browser screenshot for *Irak3* (Fig. 4f). This locus contains two Morc3 peaks in H3K9me3 chromatin. The peak distal to the TSS overlaps with an RLTR13B2 element and does not display chromatin changes in Morc3 ko cells. The second, promoter-proximal peak overlaps with an ORR1A3 element, but also with a region of developmental enhancers. This region shows strongly increased H3K27ac, ATAC-seq signals, and one-sided loss of H3K9me3 signal in Morc3 ko cells (Fig. 4f). Examples of



Morc3 repressed enhancers associated with ERV (LTRIS2) integrations are in the vicinity of Morc3 target genes *Cd200* and *Ube2l6* (Supplementary Fig. S6).

Since IAPEz elements represent the major target of Morc3 peaks, we calculated the difference in H3K9me3, ATAC, and H3K27ac signals comparing Morc3 ko vs. wild type and rescue vs. wild-type cells (Supplementary Fig. S5f). We found that H3K9me3 is decreased in the IAP-GAG region, concomitant with increased chromatin accessibility, while H3K27ac could not be detected (Supplementary Fig. S5f, top panels). These chromatin changes were reverted in Morc3 rescue cells (Supplementary Fig. S5f, lower panel). We extended this analysis using RepEnrich2 analysis and detected reduced H3K9me3 and increased chromatin accessibility on other ERV families (Supplementary Fig. S7a, b).

Together, our data demonstrate that Morc3 is critical for the full maintenance of H3K9me3 and low chromatin accessibility on distinct ERV families and distinct developmental enhancers.

Morc3-dependent heterochromatin requires a functional Morc3-ATPase cycle and -SUMOylation. Next, we thought to investigate the importance of conserved Morc3 domains and

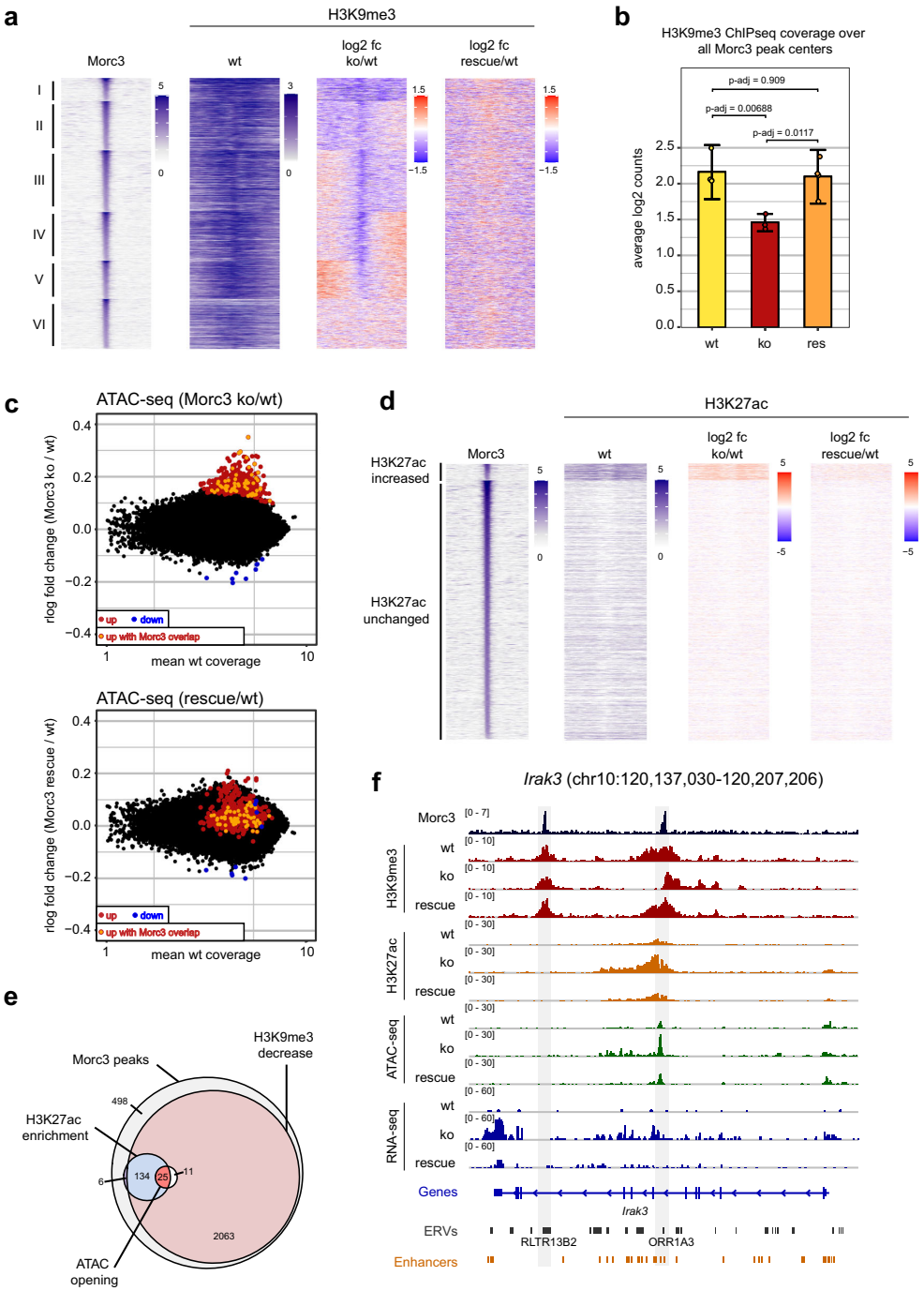
Fig. 3 Morc3 knock-out cells display de-repression of genes and distinct ERV families. a Dot plot showing average expression vs. log2-fold change of coding genes in wild type vs. Morc3 knock-out ES cells. Colored dots indicate genes with significantly changed expression (DESeq2 adjusted *p*-value < 0.05, $n = 3$ biological replicates for each condition). Positions of relevant genes are indicated. **b** Genome browser view of Morc3-dependent expression changes on selected target genes (*Irak3*, *Ube2l6*, and *Cd200*). Morc3 peaks are located on ERVs within or in close vicinity to these genes. Transcriptional upregulation of the target genes in Morc3 ko cells can be rescued by expression of wild-type Morc3. **c** RT-qPCR analysis of selected Morc3 target genes in wild type, Morc3 ko, and Morc3 rescue ES cells. Bar graph depicts mean relative expression to control genes (*Actin* and *Hprt*). Error bars indicate the standard deviation of replicate experiments ($n = 3$). Individual data points are shown as colored dots. **d** SHIN initiation silencing assay in wild type, Morc3 ko, and Morc3 rescue ES cells. Bar graph depicts the mean relative percentage of EGFP positive cells of SHIN-reporter transduced cells relative to control virus transduced cells. Error bars indicate the standard deviation of replicate experiments ($n = 3$). Individual data points are shown as colored dots. **e** SHIN maintenance silencing assay in wild type and Morc3 ko ES cells. Bar graph depicts the mean percentage of EGFP positive cells after doxycycline induction for 2 and 4 days, respectively. Error bars indicate the standard error of replicate experiments ($n = 3$). **f** Dot plot showing average expression vs. log2-fold change of ERV families in wild type vs. Morc3 knock-out ES cells. Colored dots indicate ERVs with significantly changed expression (DESeq2 adjusted *p*-value < 0.05, $n = 3$ for each condition). Positions of relevant ERV families are indicated.

SUMOylation for its role in ERV heterochromatin establishment. Morc3 has an N-terminal GHKL ATPase domain which undergoes conformational changes coupled with ATP binding and hydrolysis. The Morc3-ATPase domain is monomeric in the ATP unbound state and dimerizes upon ATP binding[25]. ATP hydrolysis and dissolution of the ATPase domain dimer are negatively regulated by the CW domain. Ligand binding of the CW domain relieves this inhibition and results in higher ATPase activity[28]. To test the function of the Morc3-ATPase cycle (Supplementary Fig. S8a) and SUMOylation for ERV silencing we generated distinct Morc3 mutant rescue cell lines (Fig. 5a). The ATP hydrolysis mutant (E35A) and the ATP binding mutant (G101A) should disrupt the ATPase cycle due to the inability to bind or hydrolyze ATP[35]. The ΔN mutant is unable to dimerize and displays reduced ATPase activity[28]. The CW ligand binding mutant (W419A) has a reduced ATPase activity and would result in a slowed-down ATPase cycle[28]. The 5KR mutant cannot be SUMOylated[35] and potentially affects SUMO-dependent protein interactions. The Morc3 mutant rescue cell lines display slight overexpression of Morc3 (Supplementary Fig. S8b), which is not directly related to the amount of Morc3 protein. In particular, the ATP hydrolysis mutant shows low Morc3 protein levels, which could suggest reduced stability of this mutant (Supplementary Fig. S8c). All mutant rescue cell lines display a similar nuclear distribution of Morc3 (Supplementary Fig. S8d). ChIP-seq analyses revealed concordant localization of most Morc3 mutant proteins with Morc3 peaks and IAPEz elements (Fig. 5b, Supplementary Fig. S8e). The ATP hydrolysis mutant failed to properly localize to Morc3 targets and the ATP binding mutant showed reduced binding to Morc3 binding sites (Fig. 5b, Supplementary Fig. S8e). In the ChIP-seq data of the rescue cells, we noticed subtle binding of Morc3 to H3K4me3 promoter regions which was reported before[25]. A more systematic analysis revealed that, although ChIP-seq of Morc3 in the 3xFLAG knock-in cells did not show association with H3K4me3, in wild type and some mutant rescue cell lines Morc3 could be detected on H3K4me3 promoters (Supplementary Fig. S9a). The CW mutant failed to localize to these regions, which would support H3K4me3-dependent recruitment of Morc3 through the CW domain[25]. Impaired H3K4me3 localization of the SUMO mutant suggests that SUMOylation contributes to H3K4me3 binding (Supplementary Fig. S9a). We did not observe changes in promoter H3K4me3 in Morc3 knock-out cells (Supplementary Fig. S9b) and we did not detect large-scale changes in transcription (see Fig. 3a). Therefore, we think that the major functional targets of Morc3 are ERV sequences and distinct enhancers.

We then functionally tested the Morc3 mutant rescue cells in SHIN initiation silencing assays and found that all mutant proteins failed to induce SHIN silencing (Fig. 5c). We also performed RT-qPCR analysis for selected Morc3 target genes and ERVs and observed de-repression, comparable to Morc3 knock-out cells (Fig. 5d, Supplementary Fig. S10a). To assess chromatin changes, we performed ATAC-seq and H3K9me3 ChIP-seq experiments in the mutant rescue cell lines. All mutants displayed increased accessibility on ATAC-seq peaks of ERV families (Fig. 5e). In line with increased chromatin accessibility, we detected reduced H3K9me3 on Morc3 peaks (Fig. 5f). Differences in cumulative ATAC-seq and H3K9me3 ChIP-seq coverage on IAPEz elements demonstrated compromised chromatin architecture in Morc3 mutant rescue cells (Supplementary Fig. S10b, c). Consistent with the de-repression of Morc3 target genes we detected increased chromatin accessibility on Morc3-associated peaks (Supplementary Figs. S10d and S11). Taken together our data demonstrate that a fully functional ATPase cycle and Morc3 SUMOylation are critical for Morc3 functionality.

The Morc3-ATPase cycle and SUMOylation are needed for its interaction with Daxx. Conformational changes during the Morc3-ATPase cycle, as well as SUMOylation, might affect interactions with other proteins that could be related to ERV silencing. To test this hypothesis, we determined the protein interaction context of Morc3 using ChIP-mass spectrometry analysis. This approach uses cross-linking and solubilization of chromatin through sonication. Therefore, proteins detected in this approach represent direct and stable protein interactions as well as low-affinity interactions which are more transient in nature, and proteins for which the interaction with Morc3 is mediated through DNA/RNA fragments. We found 489 proteins significantly enriched in the Morc3 ChIP-MS data (Fig. 6a, Supplementary Data 7). Importantly, we detected the major ERV silencing factors Setdb1, Trim28, Atrx/Daxx, and Dnmt1/Uhrf1. We also observed a strong association with SUMO proteins, corresponding to high SUMOylation of Morc3[35] and consistent with the role of SUMO in ERV repression[23]. The functional categorization of Morc3 interactors using Panther protein class enrichment analysis revealed that many proteins in the Morc3 context belong to chromatin and histone-modifying activities (Fig. 6b). Interestingly a large proportion of proteins related to RNA binding and RNA processing, suggesting that Morc3 may have roles in RNA metabolism.

Next, we investigated if the protein interaction context would change in the Morc3 mutant rescue cell lines. In this analysis, we did not include the ATP hydrolysis mutant (E35A) due to impaired recruitment to Morc3 peaks. Comparison between ChIP-MS results between wild type and mutant Morc3 proteins did not reveal large overall changes, suggesting that the interaction context is largely preserved in the mutants (Fig. 6c). However, focusing on the major ERV silencing factors, we found a selective reduction of Daxx in ChIP-MS data of most Morc3



mutant proteins (Fig. 6c). Only in the Morc3 CW mutant Daxx association was minimally impaired (Fig. 6c). Reduced association with Daxx cannot be attributed to its reduced abundance, since nuclear Daxx levels are maintained in the rescue cell lines (Supplementary Fig. S12). We validated interaction with Daxx by co-immunoprecipitation with wild-type Morc3 (Fig. 6d). In agreement with the ChIP-MS data, Morc3 mutant proteins failed to co-immunoprecipitate Daxx (Fig. 6d). These data indicate that a fully functional Morc3-ATPase cycle, as well as Morc3 SUMOylation, are needed for Daxx interaction.

Daxx is known to regulate ERV silencing through a C-terminal SUMO interaction motif (SIM)[20]. Since Morc3 is highly SUMOylated and Daxx failed to interact with the Morc3 SUMOylation mutant, we set out to determine if the interaction between Morc3 and Daxx is mediated through this C-terminal SIM. We performed ChIP-MS experiments comparing Daxx knock-out cell lines with re-

Fig. 4 Morc3 knock-out leads to reduced H3K9me3 and increased chromatin accessibility on target regions. a Changes in H3K9me3 on Morc3 peaks. Morc3 peaks were grouped in six clusters according to changes in H3K9me3 in Morc3 ko ES cells. Read-density heatmaps show the normalized coverage of Morc3 and H3K9me3 on Morc3 peaks, the log2-fold change in H3K9me3 signal between Morc3 ko and wild type ES cells, and the log2-fold change in H3K9me3 between Morc3 rescue and wild type cells. Depletion of H3K9me3 is colored in blue, increased H3K9me3 appears red. **b** H3K9me3 coverage on Morc3 peaks. Bar graph depicts mean normalized reads over all Morc3 peak centers (100 bp bin). Error bars indicate the standard error of replicate experiments ($n = 3$). Individual data points are shown as colored dots. *p*-values were calculated by a Tukey multiple comparisons of means Anova test with 95% family-wise confidence level. **c** Dot plot showing average coverage vs. log2-fold change of ATAC peaks in wild type vs. Morc3 knock-out ES cells (top panel) and wild type vs. Morc3 rescue ES cells (lower panel). Colored dots indicate ATAC peaks with significantly increased (red) or decreased (blue) coverage in Morc3 ko ES cells (DESeq2 adjusted *p*-value < 0.05, $n = 3$ for each condition), peaks with significantly increased ATAC coverage and overlapping with Morc3 peaks are marked in orange. **d** H3K27ac coverage on Morc3 peaks. Morc3 peaks were divided into 2 groups according to the change of H3K27ac in Morc3 ko compared to wt ES cells. Read-density heatmaps show the normalized coverage of Morc3 and H3K27ac on Morc3 peaks, the log2-fold change in H3K27ac signal between Morc3 ko and wild-type ES cells, and between Morc3 rescue and wild-type cells. Depletion of H3K27ac is colored in blue, increased H3K27ac appears red. **e** Venn diagram summarizing Morc3-dependent chromatin changes. **f** Genome browser view of Morc3-dependent chromatin changes on *Irak3*. Positions of two Morc3 peaks are indicated by gray boxes. The promoter-proximal Morc3 binding site displays selective loss of H3K9me3, increased H3K27ac, and increased chromatin accessibility, concomitant with elevated transcription.

expression of wild-type Daxx protein or Daxx$^{\Delta SIM}$. Lack of the C-terminal SIM reduced association with SUMO proteins and resulted in strongly reduced binding with Morc3 (Fig. 6e). In addition, wild-type Daxx could co-immunoprecipitate Morc3, whereas Daxx$^{\Delta SIM}$ failed to bind Morc3 (Fig. 6f). Together, our data indicate that SUMOylated Morc3 interacts with Daxx through the C-terminal SIM.

Morc3 is important for Daxx-mediated histone H3.3 incorporation. Recent studies have shown the role of Daxx-mediated H3.3 incorporation in ERV silencing and suggest a dynamic turnover of histones in ERV heterochromatin[18,24,36]. Therefore, we asked if the association of Morc3 with Daxx is critical for histone H3.3 incorporation on Morc3 target sites. We found histone H3.3 prominently enriched on Morc3 peak regions (Fig. 7a, Supplementary Fig. S13a). Histone H3.3 was largely lost in Morc3 ko cells and could be re-established in Morc3 wild-type rescue cells (Fig. 7a). Consistent with impaired Daxx interaction, all Morc3 mutants failed to re-establish histone H3.3 on Morc3 peak regions (Fig. 7a). On Morc3 target genes, we detected changed patterns of histone H3.3 which might be linked with the transcriptional activation of these loci (Fig. 7b, Supplementary Fig. S14). For example, *Irak3* showed slightly increased histone H3.3 incorporation in a broad region close to the promoter-proximal peak (Fig. 7b), which is associated with increased chromatin accessibility and H3K27ac (see Fig. 4f). The distal Morc3 peak, in contrast completely lost histone H3.3 association (Fig. 7b). Cumulative coverage of histone H3.3 confirmed that Morc3 peaks with repressed chromatin almost completely lost histone H3.3, whereas Morc3 peaks with gained H3K27ac signal could maintain significant levels of this histone variant (Fig. 7c), which might be linked with Daxx-independent histone H3.3 deposition in the context of enhancer activation[37,38].

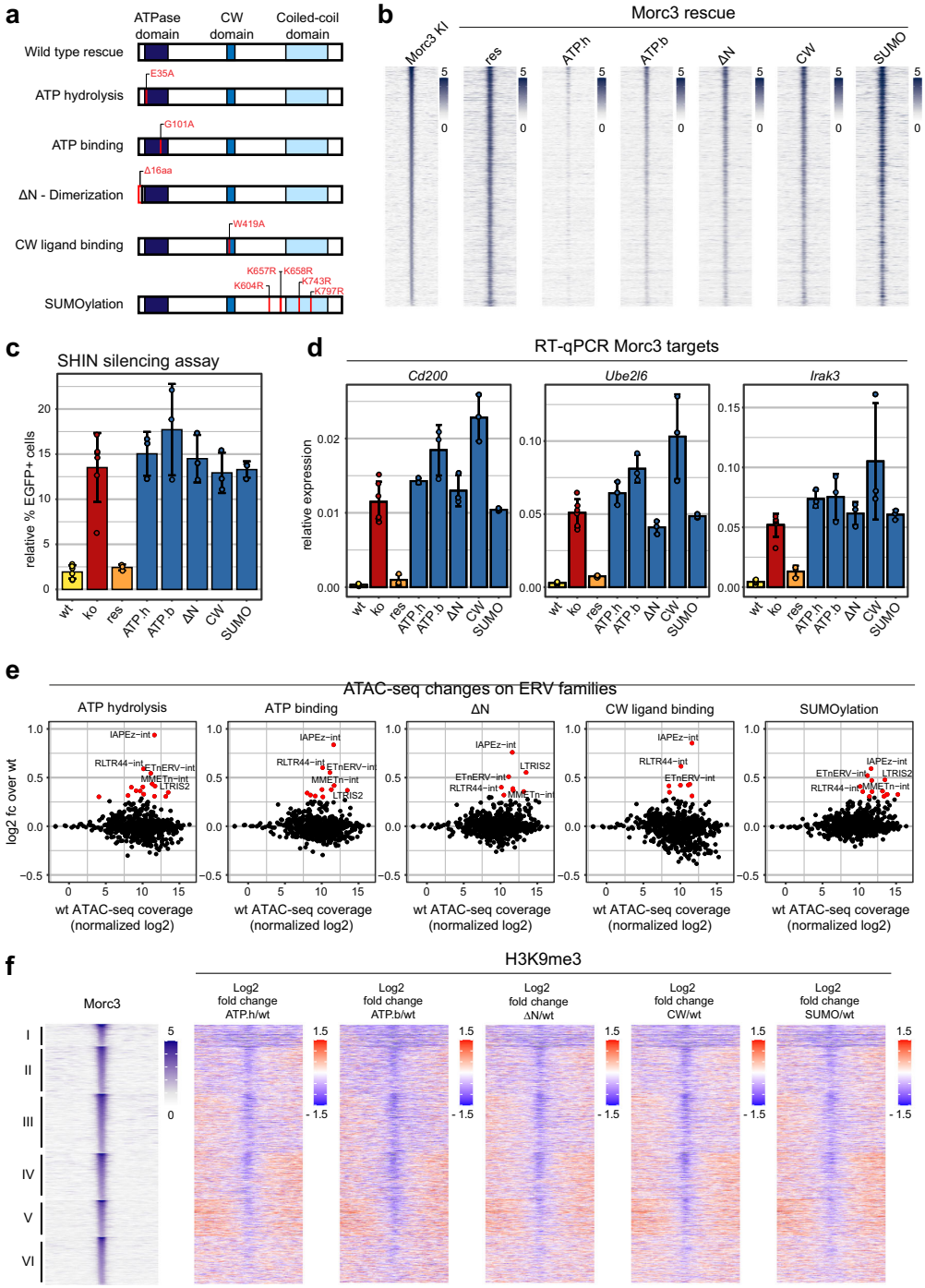
Then, we investigated histone H3.3 changes on ERV chromatin. We found that Morc3 target ERV families displayed strongly reduced histone H3.3 incorporation in Morc3 ko cells, which could be rescued with wild-type Morc3 (Fig. 7d). Cumulative coverage of histone H3.3 on IAPEz elements revealed a strong association in wild-type cells and almost complete loss in Morc3 knock-out cells (Fig. 7e). Re-expression of Morc3 could rescue histone H3.3 deposition on IAPEz elements (Fig. 7e), whereas Morc3 mutant proteins failed to rescue histone H3.3 deposition (Supplementary Fig. S13b).

Finally, we thought to compare the transcriptional changes observed in Morc3 knock-out cells with transcriptional changes in Daxx or histone H3.3 knock-out cells using published data for knock-out vs rescue cells[39]. Interestingly, a subset of Morc3 regulated ERV families displayed a trend towards being upregulated in Daxx and histone H3.3-ko cells (Supplementary Fig. S15A). In addition, several Morc3 target genes were upregulated in Daxx and histone H3.3-ko cells (Supplementary Figs. S15b and S16).

Together, our data demonstrate that Morc3 is critical for Daxx-mediated histone H3.3 deposition on distinct ERV families and enhancers.

Discussion

Our data show that Morc3 is a critical regulator of ERV chromatin in mouse ES cells. Other members of the MORC protein family have already been implicated in ERV regulation. Morc1 was found to regulate transposable elements in the mouse germline[40] and Morc2a regulates LINE1 repression in mouse ES cells[41]. In human cells, MORC2 was identified to influence HUSH complex silencing of HERVs[42,43]. Although the MORC family of proteins represent important regulators of ERV silencing in different cell types, their mechanism of action was largely obscure. We can show a specific function of Morc3 in regulating Daxx-mediated histone H3.3 incorporation to maintain ERV heterochromatin (Fig. 7F). We found that in absence of Morc3, ERV heterochromatin loses H3K9me3, shows increased accessibility and reduced levels of histone H3.3. A previous study has demonstrated high histone H3.3 turnover on ERV regions[24]. Smarcad1 evicts nucleosomes from ERV heterochromatin and, Daxx-dependent replacement by histone H3.3-containing nucleosomes is needed to maintain low chromatin accessibility. Loss of histone H3.3 or Daxx, therefore, leads to more accessible chromatin and less H3K9me3 on ERVs. We found that Morc3 knock-out cells display a very comparable phenotype, supporting the notion that Morc3 function is needed for proper Daxx activity. Based on our data we propose the following model for this process. High nucleosome turnover on ERV heterochromatin is induced by chromatin remodelers, such as Smarcad1. The re-establishment of these nucleosomes is mediated by Daxx, which contributes histone H3.3/H4 dimers. For this activity, Daxx needs to interact with SUMOylated Morc3 through its SUMO interaction motif. Morc3 undergoes an ATPase cycle with conformational changes that could be necessary for Daxx interaction and may function as “licensing step” for Daxx-mediated histone H3.3 deposition. In agreement with this model, we found that (I) Morc3 knock-out cells display impaired histone H3.3 deposition, (II) Daxx requires the C-terminal SIM domain to interact with Morc3, (III) the Morc3 mutant which cannot be SUMOylated failed to interact with Daxx (IV) Morc3 mutant proteins with impaired ATPase cycle fail to interact with Daxx. It is interesting to note that an efficient ATPase cycle is not only critical for



Morc3 function but also affects the silencing activity of MORC2[44].

The ATPase cycle is differently affected in the Morc3 mutant proteins. The Morc3-ATPase mutant protein is modeled after an Hsp90 mutant which binds ATP but fails to hydrolyze it[45]. We found that this mutant does not properly localize to Morc3 peak regions and, overall protein abundance is reduced which could suggest a higher degradation rate. The ATP binding mutant is expected to impair dimerization and ATP hydrolysis. This mutant localizes to Morc3 peaks, but with overall less intensity as compared to wild type. The ΔN mutant does not dimerize and displays reduced ATPase activity[28]. Localization to Morc3 peaks is largely intact, but failure to re-establish histone H3.3 and ERV heterochromatin demonstrates that Morc3 dimerization and/or

Fig. 5 Morc3-dependent heterochromatin requires a functional Morc3-ATPase cycle and SUMOylation. a Morc3 mutations which impair the ATPase cycle or Morc3 SUMOylation. Five mutant constructs were generated to impair ATP binding, ATP hydrolysis, dimerization, CW ligand interaction, and SUMOylation. **b** Most Morc3 mutant proteins bind to Morc3 peaks. Read-density heat map showing the normalized coverage of Morc3 wild type and mutant proteins on Morc3 binding sites. The Morc3 ATP hydrolysis mutant displays strongly reduced coverage on Morc3 peaks. **c** SHIN initiation silencing assay with Morc3 mutant rescue cell lines. Bar graph depicts mean relative percentage of EGFP positive cells of SHIN-reporter transduced cells relative to control virus transduced cells. Error bars indicate standard deviation of replicate experiments ($n = 3$). Individual data points are shown as colored dots. Data for wild type, ko, and rescue are from Fig. 3d for comparison. **d** RT-qPCR analysis of selected Morc3 target genes in Morc3 mutant rescue ES cells. Bar graphs depict mean relative expression to control genes (*Actin* and *Hprt*). Error bars indicate standard deviation of replicate experiments ($n = 3$). Individual data points are shown as colored dots. Data for wild type, ko, and rescue are from Fig. 3c for comparison. **e** Dot plot showing average wt ATAC-seq coverage vs. log2-fold change of ERV families in wild type vs. Morc3 mutant rescue ES cells (ATP hydrolysis mutant, ATP binding mutant, ΔN, CW ligand binding, SUMOylation mutant). Colored dots indicate ERV families with significantly increased (red dots) or decreased (blue dots) coverage in Morc3 mutant rescue ES cells (adjusted p-value < 0.05, $n = 2$ for each mutant condition and $n = 3$ for wt). **f** Morc3 mutant proteins fail to rescue H3K9me3 on Morc3 binding sites. Morc3 peaks group in six clusters according to changes in H3K9me3 in Morc3 ko ES cells (compare with Fig. 4a). Read-density heatmaps show the fold change in H3K9me3 between Morc3 mutant rescue and wild-type ES cells. Depletion of H3K9me3 is colored in blue, increased H3K9me3 appears red. All mutants fail to rescue the ko phenotype.

ATP hydrolysis are needed for Daxx-mediated histone H3.3 incorporation. The effect of the CW mutant is more complex. The CW domain acts as a negative regulator of ATPase activity and requires ligand binding to relieve auto-inhibition. Thus, Morc3 is expected to feature full ATPase activity only in the context of CW ligand binding. Mutation of the CW domain impairs ligand binding and therefore results in reduced ATPase activity, while dimerization of the ATPase domain is not affected. The loss-of-function phenotype of the Morc3 CW mutant would therefore suggest that ligand (H3K4me3) binding, and probably efficient ATPase activity are needed for Morc3 function in vivo.

All Morc3 mutant proteins fail to establish histone H3.3 on Morc3 binding sites. Although ΔN, CW, and SUMO mutants can properly localize to chromatin, they show impaired interaction with Daxx. This is most prominent in co-IP experiments, where only relatively stable interactions are captured. In ChIP-MS we observed differences between these mutants, the SUMO mutant completely impaired association, whereas ΔN showed strongly reduced and CW only slightly reduced Daxx association. This could suggest that Daxx can localize to Morc3 binding sites with ΔN and CW mutants, but stable interaction with Morc3 and histone H3.3 deposition might then require an efficient ATPase cycle. Unfortunately, we failed to ChIP Daxx and were unable to distinguish whether Daxx binding to Morc3 peaks, or activity on its binding sites, was impaired in Morc3 knock-out and mutant rescue cell lines.

Our data show that histone H3.3 deposition is not generally critical for ERV silencing. Several ERV families that are bound by Morc3 and lose H3.3 do not show strong transcriptional upregulation. However, individual ERV integrations and distinct developmental enhancers strongly respond to Morc3 loss and can affect the regulation of genes in their vicinity. It will thus be important to explore the role of Morc3 for histone H3.3 deposition during differentiation and development, to assess if other members of the MORC family also regulate Daxx function and to understand if the dysregulation of their targets may contribute to the different diseases associated with mutations in MORC family members[42,46–48].

Methods

Cell culture. Feeder-independent ES cells were cultured in ES cell medium (500 ml high glucose DMEM (Sigma, D6429), 91 ml (15%) fetal bovine serum (Sigma, F7524), 6.05 ml penicillin-streptomycin (Sigma, P4333), 6.05 ml MEM non-essential amino acid solution (Sigma, M7145), 1.2 ml 0.35% 2-mercaptoethanol (Sigma M7522), and 2.4 ml homemade LIF) on 10 and 15 cm gelatin-coated plates.

Lentivirus production. Lentiviral particles were produced to perform Daxx rescue, to carry reporter plasmids, to introduce sgRNAs for CRISPR-Cas9 knock-out, and to deliver the screening libraries. 293T cells were transfected with a mix of 24 µg plasmid DNA consisting of 8 µg lentiviral transfer vector, 8 µg of each of the packaging plasmids psPAX2 (#183) and pLP-eco-env (#811). All plasmids are listed in Supplementary Table S1. The DNA was mixed with 120 µl 2.5 M $CaCl_2$ and adjusted to a final volume of 1200 µl with H_2O. Subsequently, 1200 µl 2 × HBS solution (50 mM HEPES, 280 mM NaCl, 1.5 mM Na_2HPO_4, adjusted to pH 7.05 with NaOH) was added slowly and dropwise to the mix while vortexing. The transfection mix was added immediately to the 293T cells seeded 1 day ahead at a density of 4 million per 10 cm dish. Four to eight hours after transfection, the medium that contained precipitates of calcium phosphate and DNA was removed, cells were washed with 1 × PBS and fresh medium was added. The virus-containing culture supernatant was harvested 48 h after transfection.

Lentiviral transduction of mouse ES cells by spinoculation. For lentiviral transduction, 200 thousand ES cells per well were seeded on gelatinized 6-well dishes in ES medium containing 8 µg/ml polybrene (Sigma, #H9268). Viral supernatant was added to the cells and plates were spun in a prewarmed centrifuge at 1000 × g for 1 h at 37 °C to enhance viral transduction by improved viscosity for infection. After centrifugation, the medium was carefully replaced or diluted with fresh ES medium to reduce toxic side effects of polybrene.

sgRNA screen. The genome-wide sgRNA screen was conducted using the 2-vector system (lentiGuide-Puro) GeCKOv2 pooled library. The plasmid library was obtained from the Zhang lab (Addgene catalog #1000000053) and amplified in *E. coli*[29]. To validate the complexity of the library the purified plasmid pool was sequenced. Plasmids were used for lentivirus production by calcium phosphate transfection of 293T cells. To ensure complexity of the library ten parallel virus preparations with a total amount of 80 µg plasmid DNA were set up. Harvested supernatants were pooled. To identify optimal virus concentration for achieving a multiplicity of infection (MOI) of 0.3, virus titers were determined for each virus lot. Infectivity was tested by transducing ES cells with the GeCKOv2 library through spinoculation of 300 thousand cells per well in 6-wells with different volumes of virus supernatant (between 1 and 100 µl). After 48 h, cells were transferred to 15 cm plates and put under selection with 0.5 µg/ml puromycin until colonies were detected. The screening experiment was performed using T90 cells, which are ES cells carrying the TetO-EGFP-T2A-Zeo-GAG2.22 reporter at one specific integration[20]. T90 cells also carry the rtTA for doxycycline induction of the TRE promoter and a Cas9 transgene. For saturation, we aimed for infection of 100 cells per sgRNA. As the libraries carry 65.000 different sgRNAs, at least 6.5 million ES cells ought to be transduced. As the MOI should be below 30%, 30 million T90 cells were transduced with the lentiviral sgRNA library pool at an MOI between 27–28%. Transduced cells were selected by puromycin treatment after 2 days with 0.5 µg/ml puromycin. The SHIN reporter was activated 2 days after transduction with 0.1 µg/ml doxycycline. To determine the distribution of sgRNAs without selection pressure, the cell pool was divided into two groups. The control pool was cultured in medium containing 0.5 µg/ml puromycin and 0.1 µg/ml doxycycline to determine the baseline sgRNA distribution. The selection pool was treated with either 25 or 50 µg/ml zeocin from day 4 on. All cells were passaged every 2 days and harvested 8 days after transduction. For the control pool, 20 million cells, and for the zeocin-treated pool 400,000 cells were collected. Extraction of genomic DNA was performed using the DNeasy Blood and Tissue Kit Mini (Qiagen, #69504). DNA was sheared using a syringe and 27 × g needle, and DNA concentration was determined with the Q-bit fluorometer (Invitrogen). Quantitative PCR (qPCR) was used on genomic DNA extracted from ES cells in the library screen to determine the number of cycles necessary for amplification of the sgRNA sequences with oligonucleotides GS3369 and GS3371 (Supplementary Table S7). PCR was carried out with the Fast SYBR® Green Master Mix™ (Applied Biosystems) in a LightCycler480™ (Roche). The reactions were performed in a total volume of 20 µl in a 384-well plate (Sarstedt). Ct values were generated by the

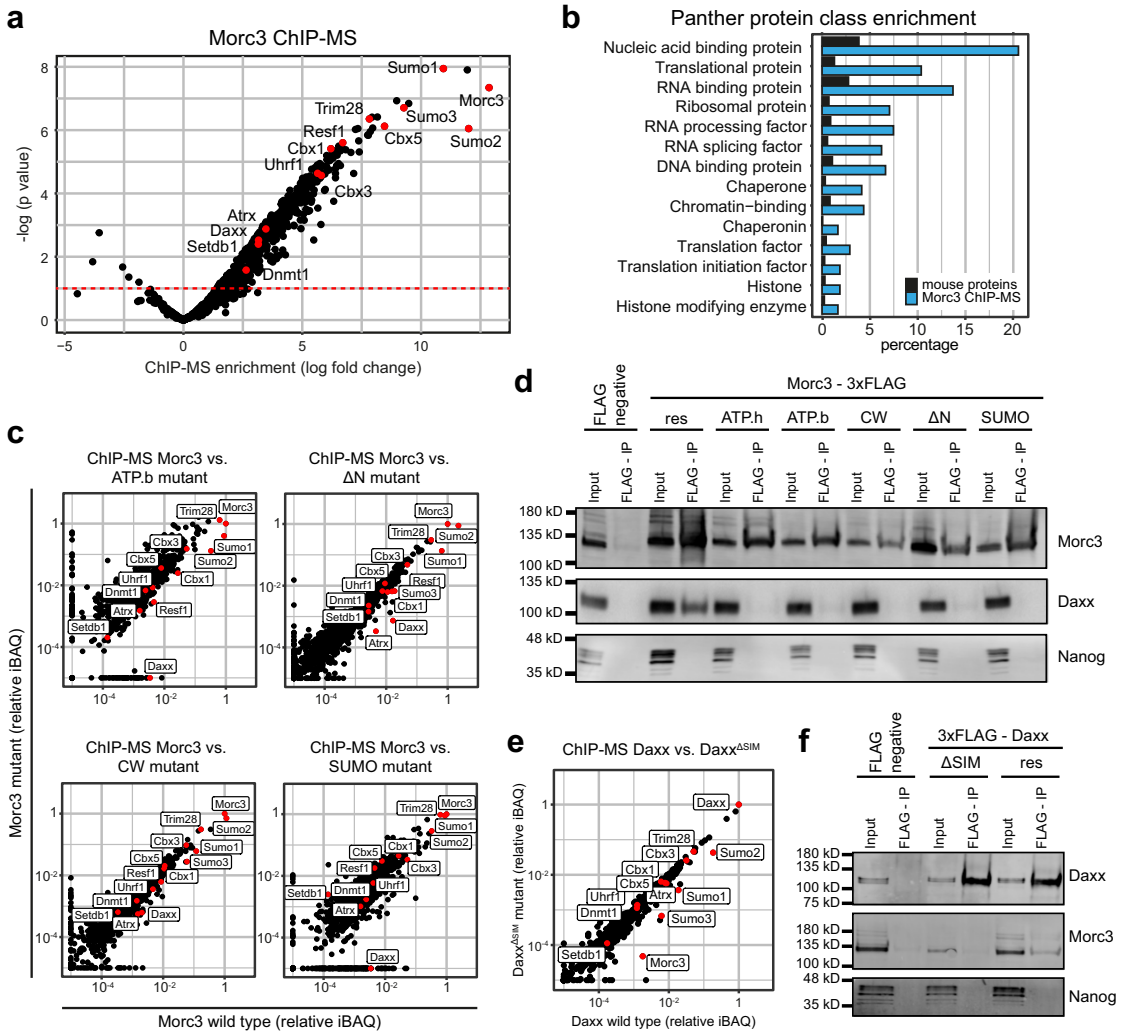
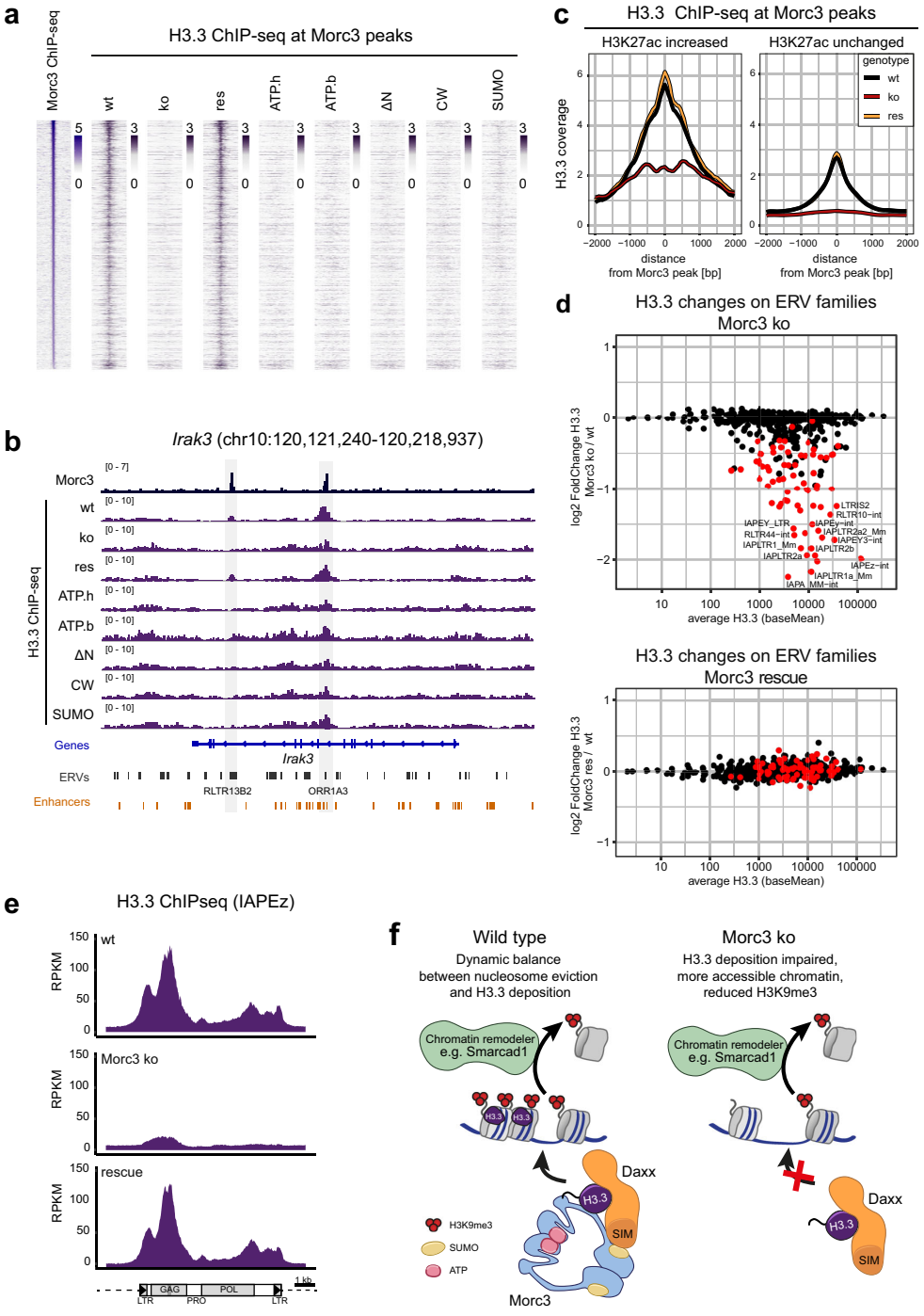


Fig. 6 Morc3-ATPase cycle and SUMOylation are needed for its interaction with Daxx. a ChIP-MS identification of Morc3-associated proteins. Morc3 ChIP-MS experiments ($n = 3$) were performed with Morc3-3xFLAG knock-in ES cells and wild-type ES cells using FLAG antibody. Dot plot shows log fold change enrichment vs. −log(p value) of proteins in Morc3-3xFLAG vs. wild type (background IP) cells. Positions of labeled proteins are indicated by red dots. **b** Panther protein class enrichment analysis of proteins which are significantly enriched in the Morc3 ChIP-MS. Bar graph shows the percentage of total mouse proteins (black) or Morc3 ChIP-MS proteins (blue) in significantly enriched Panther protein classes (Fisher's exact test, p value < 0.05). **c** ChIP-MS analysis of mutant Morc3 proteins. Dot plots show relative iBAQ values of proteins quantified in wild type vs. mutant Morc3 ChIP-MS data. Positions of ERV regulators are indicated. **d** Morc3 interaction with Daxx requires the ATPase cycle and SUMOylation. 3xFLAG tagged Morc3 wild type and mutant proteins were immunoprecipitated from nuclear extract of rescue cell lines using anti-FLAG magnetic beads. Wild-type cells without FLAG epitope served as negative control. Western blot analysis of Morc3, Daxx, and Nanog (negative control) revealed interaction of Morc3 wild type protein with Daxx. Mutant Morc3 proteins fail to co-immunoprecipitate Daxx. Uncropped blots in Source data. **e** ChIP-MS analysis of Daxx wild type vs. Daxx$^{\Delta SIM}$ mutant proteins. Dot plot shows relative iBAQ values of proteins quantified in Daxx rescue cell lines expressing wild-type Daxx or Daxx$^{\Delta SIM}$ mutant protein. Positions of ERV regulators are indicated by red dots. **f** Daxx requires the C-terminal SIM domain for Morc3 interaction. 3xFLAG tagged Daxx wild type and Daxx$^{\Delta SIM}$ proteins were immunoprecipitated from nuclear extract of rescue cell lines using anti-FLAG magnetic beads. Wild-type cells without FLAG epitope served as negative control. Western blot analysis of Daxx, Morc3, and Nanog (negative control) revealed interaction of Daxx wild type protein with Morc3. Mutant Daxx$^{\Delta SIM}$ failed to co-immunoprecipitate Morc3. Uncropped blots in Source data.

LightCycler480 software (Roche) using the 2nd derivative max function. The library preparation was performed in two steps of PCR reactions: For the first PCR, the amount of input genomic DNA (gDNA) for each sample was calculated to achieve 100x coverage over the GECKO library, which resulted in an input of at least 40 µg DNA per sample (assuming 6.6 µg of gDNA for 1 million cells). For the control sample, 46 separate 60 µl PCR reactions were performed with 1 µg gDNA in each reaction using Q5® High-Fidelity DNA Polymerase (New England Biolabs). For the zeocin-treated sample, the total amount of genomic DNA harvested was used. Oligonucleotides used to amplify lenti CRISPR sgRNAs for the first PCR were GS3367 and GS3368 (Supplementary Table S7). The first PCRs were pooled and 3 µl were used as a template for the second PCR. The second PCR served to attach Illumina adapters and to barcode samples and was done in a 60 µl reaction volume divided to two times 30 µl. Oligonucleotides used as primers for the second PCR include both a variable-length sequence to increase library complexity and an 8 bp barcode for multiplexing of different biological samples Amplification was carried out with 18 cycles for the first PCR and 30 cycles for the second PCR. Primers for



the second PCR for zeocin-treated samples were GS3369 and GS3371 (Index9) and for the control samples GS3369 and GS3370 (Index11) (Supplementary Table S7). The resulting amplicons of the second PCR were mixed, purified using AMPure XP beads (Beckman Coulter) in a volume ratio of 1:1 according to manufacturer's instructions. Before sequencing, the purified PCR product was quality controlled by Qubit fluorometer (Invitrogen) and Bioanalyzer (Agilent Technologies) measurement according to manufacturer's instructions. Sequencing was carried out by LAFUGA sequencing facility at the Gene Center of the LMU Munich using the Illumina HiSeq 1500 and a 50 bp single-end run.

SHIN silencing assays

Initiation assay. SHIN Reporter constructs (#940 EGFP only and #1074 EGFP SHIN) were stably integrated into cells by lentiviral transduction and the

Fig. 7 Morc3 is important for Daxx-mediated histone H3.3 incorporation. a Histone H3.3 is enriched on Morc3 peaks. Read-density heat map showing the normalized coverage of Morc3 and histone H3.3 in wild type, Morc3 ko, and rescue cell lines on Morc3 peaks. The Morc3 ko and all Morc3 mutant rescue cell lines lose histone H3.3 from Morc3 peak regions. **b** Genome browser view of Morc3-dependent histone H3.3 changes on an example target gene (*Irak3*). Positions of two Morc3 peaks are indicated by gray boxes. Both Morc3 binding sites display loss of histone H3.3. Histone H3.3 can be re-established in Morc3 wild type but not mutant rescue cell lines. **c** Density plot showing the average occupancy of histone H3.3 on Morc3 peaks with increased or unchanged H3K27ac in Morc3 ko ES cells. Morc3 peaks with increased H3K27ac signal maintain significant levels of this histone variant while Morc3 peaks with unchanged H3K27ac largely lose histone H3.3. **d** Dot plot showing average histone H3.3 coverage vs. log2-fold change on ERV families in wild type vs. Morc3 knock-out or Morc3 rescue ES cells. Colored dots indicate ERVs significantly enriched with Morc3 (see Fig. 2c). **e** Cumulative coverage plot of histone H3.3 ChIP-seq on IAPEz elements. In wild-type cells, prominent enrichment is over the 5′UTR and the GAG region. The position of the SHIN sequence is indicated as dark gray bar. Morc3 ko cells show a strong reduction of histone H3.3 coverage which is completely reverted in Morc3 rescue ES cells. **f** Model for the role of Morc3 in Daxx-mediated histone H3.3 deposition. In wild-type ES cells, Smarcad1 can evict nucleosomes from ERV chromatin[24]. Nucleosome reassembly requires Daxx-mediated H3.3 incorporation. Our data show that Morc3 interacts with Daxx through the SUMO interaction motif. An efficient Morc3-ATPase cycle and SUMOylation are important for Daxx interaction and assembly of histone H3.3 into ERV chromatin. In Morc3 ko ES cells, Daxx is inactive on ERV chromatin and evicted nucleosomes cannot be replaced. This results in accessible ERV chromatin and reduced H3K9me3.

percentage of EGFP$^+$ cells was measured by FACS after 2–4 days. Lentiviral particles were generated using standard protocols and virus titers were determined by titration in T37 cells. Mouse ES cells were transduced on gelatinized multi-well dishes using spinoculation at a low multiplicity of infection to ensure a linear relationship between virus titer and transduction rate. The ratio of the percentage of EGFP$^+$ cells generated by the reporter relative to the percentage of EGFP$^+$ cells generated by a control EGFP vector of the same virus titer was used to quantify reporter silencing (relative %EGFP$^+$ cells). Sequences of sgRNA oligonucleotides are summarized in Supplementary Table S2.

Maintenance assay. Cells containing the SHIN reporter (based on T90 cells) were incubated with 0.1 µg/ml doxycycline for 2 or 4 days and expression of the reporter locus was measured as the percentage of EGFP$^+$ cells. To test the effect of potential ERV silencing mutants, T90 cells were transduced with lentiviral sgRNA constructs. Two days after transduction, cells were selected for sgRNA expression with 1 µg/ml puromycin and SHIN reporter was induced with doxycycline.

Morc3-FLAG knock-in. The Morc3-3xFLAG knock-in cell line K14-E8 was generated by CRISPR-Cas9 mediated double-strand break induction close to the Morc3 STOP codon followed by homologous recombination providing a template of the same genomic region (#1529, Plasmids are summarized in Supplementary Table S1), including the 3xFLAG tag before the STOP and exchanging the sequence of the PAM from TGG to TAG so that the repaired sequence cannot be targeted (Supplementary Fig. S1). CRISPR and homology plasmids were transfected via jetPRIME (Polyplus-transfection) and single clones were analyzed for Morc3-3xFLAG expression by western blot. All cell lines used and generated are listed in Supplementary Table 3.

CRISPR-Cas9-mediated knock-out. Stable knock-out cell lines were generated by small-guide RNA (sgRNA) mediated Cas9 DNA cleavage using the pX330 plasmid (Addgene plasmid #42230). DNA oligonucleotides were hybridized and ligated into the BbsI digested pX330 to introduce the sgRNA sequence into the vector. Mouse ES cells were co-transfected with the pX330 plasmid harboring the sgRNA and a plasmid encoding a puromycin resistance gene (pLFIP) using jetPRIME (Polyplus-transfection) (#1500/#1501, Supplementary Table S1). After 2 days, transfected cells were selected by addition of 2 µg/ml puromycin to the medium. Puromycin selection was removed after 1 day, and individual cell clones were isolated after 4–6 days. Clonal cell lines were analyzed by western blotting or by PCR and Sanger sequencing.

Rescue cell lines. For the generation of Morc3 rescue cell lines, KO27-2 cells were transfected with the PiggyBac (PB) Transposon Vector System via jetPRIME (Polyplus-transfection). The system consists of two plasmids, one encoding the transposase (#1704) and a second one containing the respective rescue construct (Supplementary Table S1). After 2 days, transfected cells were selected by addition of 2 µg/ml puromycin to the medium. Individual cell clones were isolated while puromycin selection was maintained for 7–10 days.

For Daxx rescue cell lines, the Daxx KO cell line KO2-3 was transduced with lentivirus carrying the plasmids #1272 or #1273, respectively (Supplementary Table S1).

RT-qPCR. Total RNA was extracted using Trizol and the RNA Clean & Concentrator -25 Kit (Zymo Research, #R1017) including on-column DNAse digestion (Qiagen, #79254) according to manufacturer's instructions. For cDNA synthesis, 1 µg of RNA was used as input. The reaction was carried out using SuperScript III reverse transcriptase (Invitrogen, #18080044), Random Primer 6 (NEB, #S1230S), RNasin Ribonuclease inhibitor (Promega, #N2515). First, total RNA, random hexamer primers, and RNase-free water were mixed and incubated for 10 min at 70 °C, followed by 1 min incubation on ice. Next, SuperScript buffer, dNTPs, DTT, rRNasin, and SuperScript III reverse transcriptase were added and incubated at 25 °C for 8 min, followed by incubation at 50 °C for 50 min. Reaction was stopped by heat inactivation at 70 °C for 15 min. qPCR was carried out with the Fast SYBR® Green Master Mix (Applied Biosystems, 4385612) in a LightCycler480™ (Roche) according to the Fast SYBR Green Master Mix-protocol. Oligonucleotides used as primers for RT-qPCR are summarized in Supplementary Table S5. Every PCR-reaction was performed in a total volume of 10 µl in triplicates in a 384-well plate (Sarstedt). Two independent control genes (*Actin* and *Hprt*) were used as reference genes for RT-qPCR experiments and geometric mean of reference Ct values was used as normalization. Ct values were generated by the LightCycler480 software (Roche) using the 2nd derivative max function and fold changes were calculated using the $2^{-\Delta\Delta C}$ method.

ChIP-seq. For the standard ChIP protocol, 25 million ES cells were cross-linked in 7–8 ml pre-tempered ES medium containing 1% formaldehyde (Pierce #28906) for 10 min at 22 °C. Fixation was stopped by addition of 0.125 M final concentration of glycine, followed by two washing steps with PBS containing 10% Serum. Fixed cells were resuspended in 10 ml ice-cold buffer LB1 (50 mM Hepes-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.50% NP-40, 0.25% Triton X-100, 1 × Roche cOmplete Mini, EDTA-free Protease Inhibitor Cocktail (Roche, #04693159001), rocked at 4 °C for 10 min and after centrifugation resuspended in 10 ml ice-cold buffer LB2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1 × Roche cOmplete protease inhibitors), rocked at 4 °C for 5 min and centrifuged again. The pelleted nuclei were resuspended in 1 ml ice-cold shearing buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.1% SDS, 1 × Roche cOmplete protease inhibitors). Buffer compositions for LB1, LB2, and shearing buffer are based on ref. [49]. Exactly, 1 ml nuclei suspension was transferred to a 1 ml milliTUBE with AFA fiber (Covaris, #520130) and sonicated using a Covaris S220 device for a time of 15–20 min and the following settings: Peak power 140 W, Duty factor 20%, Cycles per burst 200, Temperature 4 °C. The sheared samples were added to a 1.5 ml tube with 110 µl of 10% Triton X-100 (final concentration 1%) and centrifuged at 18407 × g for 15 min at 4 °C. The soluble chromatin containing supernatant was divided in 110 µl aliquots (equivalent to 2–2.5 million cells) and for IP diluted with 890 µl complete Buffer A (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-Deoxycholate, 140 mM NaCl, 1 x Roche cOmplete protease inhibitors). Per IP 30 µl of Protein G Dynabeads (Invitrogen, #10004D) were used. For FLAG ChIP-seq experiments 1–2 µl of FLAG-M2 antibody (Sigma-Aldrich, #F3165) were used per IP and 4–5 IPs (equivalent to 10–12 million cells) were pooled to obtain enough material for library preparation. For H3K9me3 ChIP-seq experiments, 1 µl of H3K9me3 antibody (Active Motif, #39161) was used per IP. For H3K27ac ChIP-seq experiments, 2 µg of H3K27ac antibody (Diagenode, pAB-174-050) were used per IP. Beads were incubated with the corresponding antibody in complete Buffer A for 1.5 h prior to IP, followed by two washes with Buffer A (without Roche cOmplete protease inhibitors). Diluted Chromatin was added to the prebound beads and incubated for 4 h on a rotating wheel with 30 rpm at 4 °C. After IP beads were washed five times with Buffer A and 1 time with Buffer C (10 mM Tris-HCl pH 8.0, 10 mM EDTA) beads were resuspended in 100 µl elution buffer (10 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM EDTA, 0.5% SDS). RNA was degraded with 2 µl RNase A (10 mg/ml) for 30 min at 37 °C. Proteins were digested with 2 µl Proteinase K (10 mg/ml) at 55 °C for 1 h and cross-link reversal of immunoprecipitated DNA was carried out overnight at 65 °C. DNA was purified using the Agencourt AMPure XP beads (Beckman Coulter, #A63882).

For H3K9me3 and H3K4me3 ChIP-seq in wild type and Morc3 ko cells, a protocol was adapted from ref. [50]. Briefly, 2 million cells were lysed in 100 µl Buffer

B (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS, 1 x Roche cOmplete protease inhibitors) and sonicated in a microtube (Covaris; #520045) using the Covaris S220 device with the settings: Peak power 105 W, duty factor 2%, cycles per burst 200, Temperature 4 °C. The supernatant after centrifugation was diluted with 900 µl complete Buffer A and 150 µl chromatin (corresponding to ~300,000 cells) were used per IP with 10 µl Protein G Dynabeads and 1 µl of H3K9me3 antibody (Active Motif, #39161) or H3K4me3 antibody (Diagenode, CS-003-100) as described above.

Library preparation was performed using the Ultra II DNA Library prep kit for Illumina (NEB, #E7645S) according to manufacturer instructions. Sequencing was performed by LAFUGA on an Illumina Hiseq 1500 using 50 bp paired-end runs for FLAG, H3K9me3 and H3K4me3 ChIPs and a 50 bp single-end run for H3K27ac ChIP.

Histone H3.3 ChIP-seq. Histone H3.3 ChIP was performed based on Navarro et al.[24], with slight modifications. In detail, 10 million cells were cross-linked in 3 ml pre-tempered ES medium containing 1% formaldehyde (Pierce #28906) for 10 min at 22 °C. Fixation was stopped by addition of 0.125 M final concentration of glycine, followed by two washing steps with PBS containing 10% Serum. Fixed cells were resuspended in 400 µl of Sonication Buffer 1 (50 mM Tris-HCl (pH 8.0), 0.5% SDS, 1 x EDTA-free Protease Inhibitor Cocktail (Roche, #04693159001)) and sheared with the Bioruptor Pico (Diagenode) for 20 cycles (30 s on/30 s off) at 4 °C. To every 100 µl of sheared chromatin 500 µl lysis buffer (10 mM Tris-HCl (pH 8), 100 mM NaCl, 1% Triton X-100, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodiumdeoxycholate, 0.5% N-laurolsarcodine, 1 × Roche cOmplete protease inhibitors) were added and run full speed at 4 °C for 10 min on a tabletop centrifuge. Each IP used the supernatant corresponding to 3 million cells and 30 µl of Protein A Dynabeads (Invitrogen, #10002D) prebound to 5 µl of H3.3 antibody (Millipore, #09-838) for 1 h at RT followed by two washes with lysis buffer. Chromatin was added to the beads and incubated rotating overnight at 4 °C. The beads were washed twice each time, for 5 min with 1 ml RIPA (10 mM Tris-HCl (pH 8.0), 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 140 mM NaCl), 1 ml RIPA high salt (10 mM Tris-HCl (pH 8.0), 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 360 mM NaCl), 1 ml LiCl Buffer (10 mM Tris-HCl (pH 8.0), 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA) and a quick wash with 1 ml Buffer C (10 mM Tris-HCl pH 8.0, 10 mM EDTA). Elution and library preparation was performed as in the ChIP-seq description above with the 50 bp paired-end sequencing mode. All NGS files are listed in Supplementary Table 9.

ATAC-seq. The OmniATAC[51] transposition reaction was performed with 50 thousand ES cells using the Tagment DNA TDE1 Enzyme (Illumina, #20034197). DNA was purified using the PCR clean-up MinElute kit (Qiagen, #28006). The transposed DNA was subsequently amplified in 50 µl reactions with custom primers[52] listed in Supplementary Table 8. Libraries were purified and size selected for fragments <600 bp using the Agencourt AMPure XP beads (Beckman Coulter, #A63882). Sequencing was performed by LAFUGA on the Illumina Hiseq 1500 with 50 bp single-end reads.

RNA-seq. For samples GS271-GS276, total RNA was extracted using RNeasy Mini Kit (Qiagen, #74106) including on-column DNAse digestion (Qiagen, #79254). Ribosomal RNA was depleted using RNA Ribo-Zero rRNA Removal Kit (Illumina, #MRZH11124). For samples GS947-GS950, total RNA was extracted using Trizol and the RNA Clean & Concentrator -25 Kit (Zymo Research, #R1017) including on-column DNAse digestion (Qiagen, #79254). Due to the discontinued Ribo-Zero kit, ribosomal RNA was depleted with the NEBNext rRNA Depletion Kit (NEB, #E6350). Libraries were prepared with the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, #E7420S). Sequencing was performed by LAFUGA on an Illumina Hiseq 1500 using 50 bp paired-end runs.

ChIP-MS. Immunoprecipitation of bait proteins (Daxx or Morc3) was performed according to the "Rapid immunoprecipitation mass spectrometry of endogenous protein (RIME) for analysis of chromatin complexes" protocol[53]. Cells were harvested and fixed at 22 °C for 10 min in ES medium with 1% Formaldehyde (Pierce #28906) with a density of 4 million cells/ml. The fixation was stopped by 5 min incubation with glycine at a final concentration of 0.125 M. The cell pellet was washed twice with ice-cold PBS containing 10% serum and once with ice-cold PBS without serum. Cell pellets of 60 million cells were flash-frozen and stored at −80 °C for later use. All following steps were performed at 4 °C or on ice and centrifugation was performed at 2000 × g for 5 min if not specified. For each immunoprecipitation experiment, 60 million cells and 60 µl of Protein G Dynabeads (Invitrogen, #10004D) coupled with 6 µl of the FLAG-M2 antibody (Sigma-Aldrich, #F3165) were used. Beads were washed twice with LB3 (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-Lauroylsarcosine) and resuspended in 100 µl complete LB3 (LB3 containing 1 × Roche cOmplete Mini, EDTA-free Protease Inhibitor Cocktail (Roche, #04693159001)). The antibody was prebound to beads at 35 rpm on a rotating wheel for 2–4 h and the beads were washed 4 × 1 ml with LB3.

Cells were lysed by resuspending in complete LB1 (50 mM Hepes-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1 × Roche cOmplete Mini, EDTA-free Protease Inhibitor Cocktail (Roche, #04693159001)) and incubation on a rotating wheel at 35 rpm for 10 min followed by centrifugation. To obtain nuclei, the lysed cells were resuspended in LB2 (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, Roche cOmplete protease inhibitors) and incubated on a rotating wheel at 35 rpm for 5 min followed by centrifugation. The nuclei were resuspended in complete LB3 (complete LB3, 1x PhosStop EASYpack (Roche, #4906837001), 20 mM N-Ethylmaleimide (NEM) (Thermo Fisher, #23030)) at a maximum density of 20 million cells per 300 µl. Chromatin was sheared with the Bioruptor pico (Diagenode) for 10–12 cycles (30 s on/30 s off) at 4 °C. After shearing 1/10 volume of Triton X-100 was added followed by centrifugation at 20,000 × g for 10 min. The supernatant was added to the beads for chromatin-immunoprecipitation and incubated for 4 h at 35 rpm on a rotating wheel. The beads with the bound chromatin were washed with 4 × 1 ml of RIPA buffer (50 mM Hepes-KOH pH 7.6, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate) for 5 min at 20 rpm on a rotating wheel followed by the same washing process with 2 × 1 ml freshly prepared ice-cold 100 mM AMBIC (ammonium hydrogen carbonate) solution. On the last AMBIC wash the beads were transferred to a new tube. The beads were stored dry at −80 °C until MS analysis.

The sonicated chromatin was size checked with a 2100 Bioanalyzer (Agilent) using the DNA high sensitivity kit (Agilent, #5067-4626) or the DNA 7500 kit (Agilent, #5067-1506) following manufactures protocol. Shortly, Elution buffer (0.5% SDS, 300 mM NaCl, 5 mM EDTA, 10 mM Tris-HCl) was added 1:10 to the sonicated chromatin and incubated at 65 °C, shaking at 800 rpm overnight for de-cross-linking. The de-cross-linked chromatin was treated with 0.2 µg/ml of RNAse A and incubated at 37 °C, shaking at 700 rpm for 1 h. Next, the sample was treated with 0.2 µg/ml Proteinase K and incubated at 55 °C, shaking at 700 rpm for 2 h. The DNA was purified with the MinElute PCR purification kit (Qiagen, #28006) and concentration was measured with the Qubit dsDNA high sensitivity kit (Thermo Fisher, #Q32851) before loading the samples to the DNA HS/DNA 7500 chip.

Liquid chromatography–mass spectrometry (LC–MS). Beads were washed three times with 50 mM NH_4HCO_3 and incubated with 0.5 µg/µl Lys-C and 20 U benzonase in 6 M urea, 50 mM NH_4HCO_3 pH 7.5 for 90 min at 28 °C, washed with 50 mM NH_4HCO_3, and the combined supernatants were digested overnight with 0.2 µg/µl of trypsin in presence of 10 mM DTT. Digested peptides were alkylated with 30 mM IAA and desalted prior to LC–MS analysis.

For LC–MS/MS purposes, desalted peptides were injected in an Ultimate 3000 RSLCnano system (Thermo), separated in either a 15 cm analytical column (75 µm ID with ReproSil-Pur C18-AQ 2.4 µm from Dr. Maisch) with a 50 min gradient from 4 to 40% acetonitrile in 0.1% formic acid or in a 25 cm analytical column (75 µm ID, 1.6 µm C18, Aurora-IonOpticks) with a 50 min gradient from 2 to 35% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into a Qexactive HF (Thermo) operated in data-dependent mode to automatically switch between full-scan MS and MS/MS acquisition. Survey full-scan MS spectra (from m/z 375 to 1600) were acquired with resolution $R = 60{,}000$ at m/z 400 (AGC target of 3×10^6). The 10 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of 1×10^5 and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250 °C; ion selection threshold, 33,000 counts.

Nuclear extraction. Nuclei for IP and western blot analysis were isolated via ficoll gradient centrifugation. For each isolation 20 million cells were harvested and overlayed on a ficoll gradient consisting of 20% Ficoll (Lymphocyte Separation Medium1077, promo cell C-44010) and 80% NI-Stock (100 mM Tris/HCl pH 7.4, 10 mM $MgCl_2$, 10 mM $CaCl_2$l, 2% NP-40, 1.6% Triton X-100) with 0.1% DMSO. Centrifugation speed was stepwise increased at 4 °C from 37 to 150 × g in 8 min (30 s at 37 × g, 30 s at 58 × g, 30 s at 84 × g, 30 s at 114 × g, 6 min at 150 × g). The supernatant was carefully removed, and nuclei washed once in cold PBS.

Co-IP for western blot. Nuclei were isolated via ficoll gradient centrifugation and resuspended in complete 150 mM NaCl IP Buffer (50 mM HEPES ph 7.5, 150 mM NaCl, 0.05% NP-40, 20% glycerol and freshly added 1 × Roche complete, 1 × PhosphoSTOP, 25 mM NEM, Benzonase (Merck, #1016540001) (1 µl for each 400 µl of buffer) and 2.5 mM $MgCl_2$). Tubes were incubated 3 min at 37 °C for benzonase digestion, then salt concentration was adjusted to 300 mM by adding 5 M NaCl. After 30 min incubation on ice samples were centrifuged 30 min at 4 °C at maximum speed and the supernatant was used as input for IP. Since the bait proteins were FLAG-tagged, FLAG-M2 beads (Sigma, M8823-1ML) were used (80 µl beads for 80 million cells). IP was performed for 4 h followed by three washing steps, 5 min each with 1 ml 300 mM NaCl IP buffer (50 mM HEPES ph 7.5, 300 mM NaCl, 0.05% NP-40, 20% glycerol) on a rotating wheel at 4 °C. For elution of immunoprecipitated proteins, beads were first boiled in Lämmli buffer without β-mercaptoethanol (for Daxx elution) and then boiled in Lämmli buffer containing β-mercaptoethanol (for Morc3 elution). Morc3 western blot signals were detected by multiplexed fluorescent immunoblotting (LI-COR Biosciences). Development of Daxx western blots after co-IP was performed with the SuperSignal™ Western Blot Substrate Atto (Thermo Scientific, A45918).

Immunofluorescence analysis. Cells were seeded in 24-well plates containing 12 mm poly-L-lysine coated coverslips (50 thousand cells/well). Cells were fixed with

formaldehyde (500 µl/well, 3.7% formaldehyde in PBS for 10 min at RT). Coverslips were washed twice for 5 min in 1 ml PBS. Permeabilization was performed for 5 min in 500 µl/well in permeabilization solution (10 mM sodium citrate tribasic dihydrate, 200 µl Triton X-100). Two washing steps with 1 ml PBS followed by two washing steps with 1 ml washing solution I (2 l H_2O, 220 ml 10 x PBS, 2.2 ml Tween-20, 5.5 g BSA) were performed before samples were blocked for 30 min in 300 ml blocking solution I (50 ml washing solution I, 1.2 g BSA). Primary antibody (anti-Morc3, Rockland, 100-401-N96S) was diluted in 200 µl/well blocking solution I and incubated over night at 4 °C in the dark. Cells were then washed three times with washing solution I. Secondary antibody was diluted in 200 µl/well blocking solution II (0.5 ml blocking solution I, 0.5 ml serum (goat or donkey)) and incubated for 1 h at RT in the dark. Next, coverslips were washed three times 10 min with washing solution II (500 ml PBS, 500 µl Tween-20), followed by embedding coverslips with Vectashield containing DAPI (Vector Laboratories, Axxora - H-1200) and sealed with nail polish. Samples were stored at 4 °C in the dark until pictures were taken at inverted confocal microscope Leica SP5 (×64) and analyzed with ImageJ.

Statistics and reproducibility. Results shown in Fig. 1e, g are combinations of several experiments and the replicate number per sgRNA varies. Specifically in Fig. 1e $n = 4$ for sgDaxx, sgDnmt1, sgSetdb1 and Trim28; $n = 3$ for sgAtrx, sgMorc3, sgNeg, SgTrim24 and sgUhrf1, $n = 2$ for all remaining candidates. In Fig. 1g: $n = 12$ for sgNeg, $n = 11$ for sgSetdb1, $n = 10$ for sgDaxx, $n = 7$ for sgDnmt1 and sgUhrf1, $n = 6$ for no sgRNA control, sgMorc2a and sgTrim28, $n = 4$ for sgAtf7ip and sgAtrx, $n = 3$ for Morc3 and sgZfp606, $n = 2$ for sgDek and sg HP1, $n = 1$ for all remaining candidates.

All western blots are reproduced with $n \geq 2$ biologically independent experiments. All uncropped blots are present in the source data file.

Bioinformatics analysis

sgRNA screen data analysis. Fastq files were trimmed with Trimmomatic version 0.36 and options "CROP:43 HEADCROP:23" to contain only the unique sgRNA sequence. Reads were then mapped with bowtie to all sgRNA sequences of the Gecko V2 library. Cvs files of library sequences were downloaded from addgene (https://www.addgene.org/pooled-library/zhang-mouse-gecko-v2/). CSV files were converted to fasta format using R. Bowtie was used to generate the library index. Reads were then aligned with bowtie 1. The number of aligned reads to each sgRNA sequence was calculated with bedtools command "bedtools genomecov". The numbers of reads for each sgRNA per sample were normalized in R as follows: Normalized reads per sgRNA = (reads per sgRNA/total reads for all sgRNAs in sample) + 1. Hits were identified by conversion of sgRNA enrichment scores into gene rankings by statistical analysis using the SecondBestRank scoring method for RNAi gene enrichment ranking (RIGER) with the RigerJ tool.

ChIP-seq. Paired-end ChIP-seq reads were aligned to the mouse genome mm10 using Bowtie2 with default settings. The resulting BAM files were filtered to remove non-paired reads, low mapping quality, non-primary alignment, and PCR duplicates with "samtools view -b -f 2 -F 1280 -q 20". Homer tag directories were generated with "makeTagDirectory". BigWig files were generated from tag directories with "makeBigWig.pl mm10 -webdir. -url. -norm 1e7 -normLength 100 -fragLength 150 -update".

Morc3 peaks were identified using findPeaks with option "-style factor" over Input. The peak files of replicate experiments were merged using homer "mergePeaks" and, peaks common in two replicates were kept for the final peak list. The final Morc3 peak list was annotated using homer "annotatePeaks.pl", resulting in detailed peak annotation including association with repeat elements.

Morc3 association with repeats was determined using RepEnrich2. The fraction counts for each repeat family were normalized and repeat families with significant enrichment (padj < 0.05 & log2FoldChange > 0.5) in Morc3-FLAG ChIP-seq vs. wild type FLAG ChIP-seq (background) were calculated using DeSeq2. Enrichment of H3K9me3 on ERVs in wild type vs. Morc3 knock-out and Morc3 rescue cells was also calculated with RepEnrich2 and DeSeq2.

Heatmaps for ERV silencing factors Trim28, Setdb1, Morc3, and histone H3.3/H3K9me3 were plotted as log-transformed normalized coverage (calculated from Homer tag directories with annotatePeaks.pl).

Cumulative coverage plots on IAPEz elements were performed as described earlier[20]. Cumulative coverage plots on IAPEz of H3K9me3 mutant rescues were normalized based on non-Morc3-related H3K9me3 peaks in the wt sample.

Heatmaps of H3K9me3 and H3K27ac were plotted as log-transformed normalized coverage (calculated from Homer tag directories with annotatePeaks.pl). Heatmaps of H3K9me3 mutant rescues were normalized based on non-Morc3 related H3K9me3 peaks in the wt sample.

RNA-seq. Paired-end reads were aligned to the mouse genome version mm10 using STAR with default options "--runThreadN 32 --quantMode TranscriptomeSAM GeneCounts --outSAMtype BAM SortedByCoordinate". Read counts for all genes were normalized using DESeq2. Differentially expressed genes were determined using the DESeq2 results function (adjusted p-value < 0.01).

Expression of ERV families was calculated using RepEnrich2. Fraction counts were normalized and differentially expressed repeats were calculated using DeSeq2.

ATAC-seq. Single-end ATAC-seq reads were aligned to the mouse genome mm10 using Bowtie with options "-q -n 2 --best --chunkmbs 2000 -p 32 -S -m 1". Duplicated reads were subsequently removed using Picard. Homer tag directories were generated using makeTagDirectory and bigwig files were generated with makeBigWig.pl. ATAC peaks were identified using Homer findPeaks.pl with the option "-style factor" over Input. Peaks from all samples were merged using mergePeaks resulting in a unified Peak set. Raw ATAC coverage counts were calculated with annotatePeaks. Differential ATAC peaks were determined with the DESeq2 result function.

ATAC coverage of ERV families was calculated using RepEnrich2 on BAM files including multi-mapping reads. Fraction counts were normalized, and differentially accessible repeats were calculated using DeSeq2.

IGV Screenshots. IGV screenshots were taken from respective bigwig files using IGV 2.10.0.

ChIP-MS data analysis. MaxQuant (1.6.14.0) was used to identify proteins and quantify by iBAQ with the following parameters: Database, UP000000589_10090_Mmusculus_2020; MS tol, 10 ppm; MS/MS tol, 20 ppm Da; Peptide FDR, 0.1; Protein FDR, 0.01 Min. peptide Length, 7; Variable modifications, Oxidation (M); Fixed modifications, Carbamidomethyl (C); Peptides for protein quantitation, razor and unique; Min. peptides, 1; Min. ratio count, 2.

The final list of proteins found was filtered and statistically processed in R studio version 3.5.0 using the Linear Models for Microarray Data (LIMMA) R script Version 1.0.1 that is available on GitHub written by Wasim Aftab (https://github.com/wasimaftab/LIMMA-pipeline-proteomics).

For comparative analysis between Morc3 wild type and mutant proteins, iBAQ values for proteins identified with more than 2 peptides were normalized to Morc3 or Daxx, respectively, by dividing by the iBAQ value of the bait. Mean normalized iBAQ values from replicate experiments were plotted. Daxx was not detected in the Morc3 ATP binding and SUMOylation mutant ChIP-MS experiments.

Pathway enrichment analysis of significantly Morc3-associated proteins was performed using Panther.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding author upon reasonable request. ChIP-MS proteomics datasets generated in this study have been deposited to the ProteomeXchange database under the accession code PXD027368. ATAC-seq, ChIP-seq, and RNA-seq datasets generated in this study have been deposited to the GEO database under the accession code GSE159936. The published Trim28 ChIP-seq data used in this study are available in the GEO database under accession code GSM1819199. The published Daxx-ko and H3.3-ko RNA-seq data used in this study are available in the GEO database under accession code GSE102688. Source data are provided with this paper.

Received: 18 November 2020; Accepted: 22 September 2021;
Published online: 14 October 2021

References

1. Thompson, P. J., Macfarlan, T. S. & Lorincz, M. C. Long terminal repeats: from parasitic elements to building blocks of the transcriptional regulatory repertoire. *Mol. Cell* **62**, 766–776 (2016).
2. Chuong, E. B., Elde, N. C. & Feschotte, C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**, 1083–1087 (2016).
3. Fuentes, D. R., Swigut, T. & Wysocka, J. Systematic perturbation of retroviral LTRs reveals widespread long-range effects on human gene regulation. *Elife* **7**, https://doi.org/10.7554/eLife.35989 (2018).
4. Jacques, P. E., Jeyakani, J. & Bourque, G. The majority of primate-specific regulatory sequences are derived from transposable elements. *PLoS Genet.* **9**, e1003504 (2013).
5. Kunarso, G. et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634 (2010).
6. Sundaram, V. et al. Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res.* **24**, 1963–1976 (2014).
7. Kury, P. et al. Human endogenous retroviruses in neurological diseases. *Trends Mol. Med.* **24**, 379–394 (2018).
8. Deniz, O. et al. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat. Commun.* **11**, 3506 (2020).
9. Babaian, A. & Mager, D. L. Endogenous retroviral promoter exaptation in human cancer. *Mob. DNA* **7**, 24 (2016).

10. Jang, H. S. et al. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat. Genet.* **51**, 611–617 (2019).
11. Groh, S. & Schotta, G. Silencing of endogenous retroviruses by heterochromatin. *Cell Mol. Life Sci.* **74**, 2055–2065 (2017).
12. Geis, F. K. & Goff, S. P. Silencing and transcriptional regulation of endogenous retroviruses: an overview. *Viruses* **12**, https://doi.org/10.3390/v12080884 (2020).
13. Matsui, T. et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927–931 (2010).
14. Rowe, H. M. et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237–240 (2010).
15. Ramesh, V. et al. Loss of Uhrf1 in neural stem cells leads to activation of retroviral elements and delayed neurodegeneration. *Genes Dev.* **30**, 2199–2212 (2016).
16. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet* **20**, 116–117 (1998).
17. Haggerty, C. et al. Dnmt1 has de novo activity targeted to transposable elements. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-021-00603-8 (2021).
18. Elsasser, S. J., Noh, K. M., Diaz, N., Allis, C. D. & Banaszynski, L. A. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**, 240–244 (2015).
19. He, Q. et al. The Daxx/Atrx complex protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation. *Cell Stem Cell* **17**, 273–286 (2015).
20. Sadic, D. et al. Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep.* **16**, 836–850 (2015).
21. Wasylishen, A. R. et al. Daxx maintains endogenous retroviral silencing and restricts cellular plasticity in vivo. *Sci. Adv.* **6**, eaba8415 (2020).
22. Sachs, P. et al. SMARCAD1 ATPase activity is required to silence endogenous retroviruses in embryonic stem cells. *Nat. Commun.* **10**, 1335 (2019).
23. Yang, B. X. et al. Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell* **163**, 230–245 (2015).
24. Navarro, C., Lyu, J., Katsori, A. M., Caridha, R. & Elsasser, S. J. An embryonic stem cell-specific heterochromatin state promotes core histone exchange in the absence of DNA accessibility. *Nat. Commun.* **11**, 5095 (2020).
25. Li, S. et al. Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. *Proc. Natl Acad. Sci. USA* **113**, E5108–E5116 (2016).
26. Andrews, F. H. et al. Multivalent chromatin engagement and inter-domain crosstalk regulate MORC3 ATPase. *Cell Rep.* **16**, 3195–3207 (2016).
27. Zhang, Y. et al. MORC3 is a target of the influenza A viral protein NS1. *Structure* **27**, 1029–1033 e1023 (2019).
28. Zhang, Y. et al. Mechanism for autoinhibition and activation of the MORC3 ATPase. *Proc. Natl Acad. Sci. USA* **116**, 6111–6119 (2019).
29. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
30. Chelmicki, T. et al. m(6)A RNA methylation regulates the fate of endogenous retroviruses. *Nature* **591**, 312–316 (2021).
31. Criscione, S. W., Zhang, Y., Thompson, W., Sedivy, J. M. & Neretti, N. Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics* **15**, 583 (2014).
32. Karimi, M. M. et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8**, 676–687 (2011).
33. Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends Genet* **32**, 29–41 (2016).
34. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
35. Mimura, Y., Takahashi, K., Kawata, K., Akazawa, T. & Inoue, N. Two-step colocalization of MORC3 with PML nuclear bodies. *J. Cell Sci.* **123**, 2014–2024 (2010).
36. Wolf, G. et al. On the role of H3.3 in retroviral silencing. *Nature* **548**, E1–E3 (2017).
37. Chen, P. et al. H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. *Genes Dev.* **27**, 2109–2124 (2013).
38. Schwartz, B. E. & Ahmad, K. Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev.* **19**, 804–814 (2005).
39. Hoelper, D., Huang, H., Jain, A. Y., Patel, D. J. & Lewis, P. W. Structural and mechanistic insights into ATRX-dependent and -independent functions of the histone chaperone DAXX. *Nat. Commun.* **8**, 1193 (2017).
40. Pastor, W. A. et al. MORC1 represses transposable elements in the mouse male germline. *Nat. Commun.* **5**, 5795 (2014).
41. Fukuda, K., Okuda, A., Yusa, K. & Shinkai, Y. A CRISPR knockout screen identifies SETDB1-target retroelement silencing factors in embryonic stem cells. *Genome Res.* **28**, 846–858 (2018).
42. Tchasovnikarova, I. A. et al. Hyperactivation of HUSH complex function by Charcot-Marie-Tooth disease mutation in MORC2. *Nat. Genet.* https://doi.org/10.1038/ng.3878 (2017).
43. Liu, N. et al. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* https://doi.org/10.1038/nature25179 (2017).
44. Douse, C. H. et al. Neuropathic MORC2 mutations perturb GHKL ATPase dimerization dynamics and epigenetic silencing by multiple structural mechanisms. *Nat. Commun.* **9**, 651 (2018).
45. Panaretou, B. et al. ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo. *EMBO J.* **17**, 4829–4836 (1998).
46. Hong, G. et al. The emerging role of MORC family proteins in cancer development and bone homeostasis. *J. Cell Physiol.* **232**, 928–934 (2017).
47. Jadhav, G., Teguh, D., Kenny, J., Tickner, J. & Xu, J. Morc3 mutant mice exhibit reduced cortical area and thickness, accompanied by altered haematopoietic stem cells niche and bone cell differentiation. *Sci. Rep.* **6**, 25964 (2016).
48. Sloan, E., Orr, A. & Everett, R. D. MORC3, a component of PML nuclear bodies, has a role in restricting herpes simplex virus 1 and human cytomegalovirus. *J. Virol.* **90**, 8621–8633 (2016).
49. Boyer, L. A. et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
50. Dahl, J. A. & Collas, P. A rapid micro chromatin immunoprecipitation assay (microChIP). *Nat. Protoc.* **3**, 1032–1045 (2008).
51. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
52. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
53. Mohammed, H. et al. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326 (2016).

Acknowledgements

High throughput sequencing was performed by the Laboratory for Functional Genome Analysis (LAFUGA) of the Ludwig-Maximilian-University, Munich. We acknowledge the Core Facility Flow Cytometry and the Bioinformatics Core Unit at the Biomedical Center, Ludwig-Maximilian-Universität München, for providing equipment, service, and expertise. Mouse GeCKOv2 CRISPR knock-out pooled library was a gift from Feng Zhang (Addgene #1000000052). The NP95 antibody was a gift from Heinrich Leonhardt (Faculty of Biology, Ludwig-Maximilians-University (LMU) Munich, Großhaderner Straße 2, 82152 Martinsried, Germany). Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 213249687 – SFB 1064 TP3 and Projekt-ID 329628492 – SFB 1321 TP13) to G.S.

Author contributions

Conceptualization: S.G. and G.S. Investigation: S.G., A.V.M., L.K.M., C.V.S., A.R., H.B., G.P.dA., A.S., and I.F. Data curation: S.G., A.V.M., L.K.M., C.V.S., A.R., H.B., G.P.dA., A.S., and I.F. Formal analysis: S.G., A.V.M., A.S., I.F., and G.S. Funding acquisition: G.S. Methodology: S.G. and G.S. Project administration: S.G. and G.S. Resources: A.I. and G.S. Software: S.G., A.S., I.F., and G.S. Supervision: S.G. and G.S. Validation: S.G., A.V.M., L.K.M., C.V.S., A.R., H.B., and G.S. Visualization: S.G., H.B., C.V.S., A.V.M., A.R., and G.S. Writing—original draft: S.G., G.S. Writing—review & editing: S.G., A.V.M., L.K.M., C.V.S., A.R. H.B., G.P.dA., A.S., I.F., A.I., and G.S.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-26288-7.

Correspondence and requests for materials should be addressed to Gunnar Schotta.

Peer review information *Nature Communications* thanks Johan Jakobsson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Bibliography

Akter, J., Katai, Y., Sultana, P., Takenobu, H., Haruta, M., Sugino, R. P., Mukae, K., Satoh, S., Wada, T., Ohira, M., Ando, K., and Kamijo, T. (2021). Loss of p53 suppresses replication stress-induced DNA damage in ATRX-deficient neuroblastoma. *Oncogenesis*, 10(11):73.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular biology of the cell (4th edition)*. Garland Science, New York.

Allis, C. D. and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat Rev Genet*, 17(8):487–500.

Allshire, R. C. and Madhani, H. D. (2018). Ten principles of heterochromatin formation and function. *Nature Reviews Molecular Cell Biology*, 19(4):229–244.

Andrews, F. H., Tong, Q., Sullivan, K. D., Cornett, E. M., Zhang, Y., Ali, M., Ahn, J., Pandey, A., Guo, A. H., Strahl, B. D., Costello, J. C., Espinosa, J. M., Rothbart, S. B., and Kutateladze, T. G. (2016). Multivalent Chromatin Engagement and Inter-domain Crosstalk Regulate MORC3 ATPase. *Cell Rep*, 16(12):3195–3207.

Arita, K., Isogai, S., Oda, T., Unoki, M., Sugita, K., Sekiyama, N., Kuwata, K., Hamamoto, R., Tochio, H., Sato, M., Ariyoshi, M., and Shirakawa, M. (2012). Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proc Natl Acad Sci U S A*, 109(32):12950–12955.

Avery, O. T., MacLeod, C. M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. *Journal of Experimental Medicine*, 79(2):137–158.

Baldi, S., Korber, P., and Becker, P. B. (2020). Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nature Structural & Molecular Biology*, 27(2):109–118.

Bannert, N. and Kurth, R. (2004). Retroelements and the human genome: New perspectives on an old relation. *Proceedings of the National Academy of Sciences*, 101(suppl 2):14572–14579.

Bannert, N. and Kurth, R. (2006). The evolutionary dynamics of human endogenous retroviral families. *Annu Rev Genomics Hum Genet*, 7:149–173.

Bano, D., Piazzesi, A., Salomoni, P., and Nicotera, P. (2017). The histone variant H3.3 claims its place in the crowded scene of epigenetics. *Aging*, 9(3):602–614.

Barral, A., Pozo, G., Ducrot, L., Papadopoulos, G. L., Sauzet, S., Oldfield, A. J., Cavalli, G., and Déjardin, J. (2022). SETDB1/NSD-dependent H3K9me3/H3K36me3 dual heterochromatin maintains gene expression profiles by bookmarking poised enhancers. *Molecular cell*, 82(4):816–832.e12.

Becker, P. B. and Hörz, W. (2002). ATP-dependent nucleosome remodeling. *Annu Rev Biochem*, 71:247–273.

Becker, P. B. and Workman, J. L. (2013). Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol*, 5(9).

Behjati, S. and Tarpey, P. S. (2013). What is next generation sequencing? *Archives of Disease in Childhood - Education and Practice*, 98(6):236–238.

Berger, F. (2019). Emil Heitz, a true epigenetics pioneer. *Nature Reviews Molecular Cell Biology*, 20(10):572–572.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, 447(7143):407–412.

Bleckwehl, T., Crispatzu, G., Schaaf, K., Respuela, P., Bartusel, M., Benson, L., Clark, S. J., Dorighi, K. M., Barral, A., Laugsch, M., van IJcken, W. F. J., Manzanares, M., Wysocka, J., Reik, W., and Rada-Iglesias, Á. (2021). Enhancer-associated H3K4 methylation safeguards in vitro germline competence. *Nat Commun*, 12(1):5771.

Bogolyubova, I. and Bogolyubov, D. (2021). DAXX Is a Crucial Factor for Proper Development of Mammalian Oocytes and Early Embryos. *Int J Mol Sci*, 22(3).

Bostick, M., Kim, J. K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S. E. (2007). UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science*, 317(5845):1760–1764.

Bruno, M., Mahgoub, M., and Macfarlan, T. S. (2019). The Arms Race Between KRAB-Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *Annu Rev Genet*, 53:393–416.

Cernilogar, F. M., Hasenöder, S., Wang, Z., Scheibner, K., Burtscher, I., Sterr, M., Smialowski, P., Groh, S., Evenroed, I. M., Gilfillan, G. D., Lickert, H., and Schotta, G. (2019). Pre-marked chromatin and transcription factor co-binding shape the pioneering activity of Foxa2. *Nucleic acids research*, 47(17):9069–9086.

Chang, F. T. M., McGhie, J. D., Chan, F. L., Tang, M. C., Anderson, M. A., Mann, J. R., Andy Choo, K. H., and Wong, L. H. (2013). PML bodies provide an important platform for the maintenance of telomeric chromatin integrity in embryonic stem cells. *Nucleic Acids Res*, 41(8):4447–4458.

Cheng, B., Ren, X., and Kerppola, T. K. (2014). KAP1 represses differentiation-inducible genes in embryonic stem cells through cooperative binding with PRC1 and derepresses pluripotency-associated genes. *Molecular and cellular biology*, 34(11):2075–2091.

Chiu, E. S. and VandeWoude, S. (2021). Endogenous Retroviruses Drive Resistance and Promotion of Exogenous Retroviral Homologs. *Annu Rev Anim Biosci*, 9:225–248.

Chuong, E. B., Elde, N. C., and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science (New York, N.Y.)*, 351(6277):1083–1087.

Clapier, C. R. and Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annu Rev Biochem*, 78:273–304.

Coffin, J., Hughes, S., and Varmus, H. (1997). The Interactions of Retroviruses and their Hosts. *Cold Spring Harbor Laboratory Press*.

Coffin, J. M. (2004). Evolution of Retroviruses: Fossils in Our DNA. *Proceedings of the American Philosophical Society*, 148(3):264–280.

Cohen, C. J., Lock, W. M., and Mager, D. L. (2009). Endogenous retroviral LTRs as promoters for human genes: A critical assessment. *Gene*, 448(2):105–114. Genomic Impact of Eukaryotic Transposable Elements.

Corbett, K. D. and Berger, J. M. (2006). Structural basis for topoisomerase VI inhibition by the anti-Hsp90 drug radicicol. *Nucleic acids research*, 34(15):4269–4277.

Corsinotti, A., Kapopoulou, A., Gubelmann, C., Imbeault, M., Santoni de Sio, F. R., Rowe, H. M., Mouscaz, Y., Deplancke, B., and Trono, D. (2013). Global and Stage Specific Patterns of Krüppel-Associated-Box Zinc Finger Protein Gene Expression in Murine Early Embryonic Cells. *PLOS ONE*, 8(2):e56721–.

Crick, F. (1970). Central Dogma of Molecular Biology. *Nature*, 227(5258):561–563.

Dahm, R. (2005). Friedrich Miescher and the discovery of DNA. *Developmental Biology*, 278(2):274 – 288.

Darmusey, L., Pérot, G., Thébault, N., Le Guellec, S., Desplat, N., Gaston, L., Delespaul, L., Lesluyes, T., Darbo, E., Gomez-Brouchet, A., Richard, E., Baud, J., Leroy, L., Coindre, J.-M., Blay, J.-Y., and Chibon, F. (2021). ATRX Alteration Contributes to Tumor Growth and Immune Escape in Pleomorphic Sarcomas. *Cancers (Basel)*, 13(9).

de Almeida, G. P. (2018). Modulation of an essential histone methyltransferase in mouse embryonic stem cells. *Dissertation - Ludwig-Maximilians-Universität München*.

Deininger, P. L., Moran, J. V., Batzer, M. A., and Kazazian, H. H. J. (2003). Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev*, 13(6):651–658.

Delbarre, E. and Janicki, S. M. (2021). Modulation of H3.3 chromatin assembly by PML: A way to regulate epigenetic inheritance. *Bioessays*, 43(10):e2100038.

Deniz, Ö., Ahmed, M., Todd, C. D., Rio-Machin, A., Dawson, M. A., and Branco, M. R. (2020). Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat Commun*, 11(1):3506.

Desai, V. P., Chouaref, J., Wu, H., Pastor, W. A., Kan, R. L., Oey, H. M., Li, Z., Ho, J., Vonk, K. K. D., San Leon Granado, D., Christopher, M. A., Clark, A. T., Jacobsen, S. E., and Daxinger, L. (2021). The role of MORC3 in silencing transposable elements in mouse embryonic stem cells. *Epigenetics Chromatin*, 14(1):49.

Dillon, S. C., Zhang, X., Trievel, R. C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biology*, 6(8):227.

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., Diao, Y., Liang, J., Zhao, H., Lobanenkov, V. V., Ecker, J. R., Thomson, J. A., and Ren, B. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature*, 518(7539):331–336.

Dong, W., Vannozzi, A., Chen, F., Hu, Y., Chen, Z., and Zhang, L. (2018). MORC Domain Definition and Evolutionary Analysis of the MORC Gene Family in Green Plants. *Genome Biol Evol*, 10(7):1730–1744.

Drané, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev*, 24(12):1253–1265.

Duan, Z. and Blau, C. A. (2012). The genome in space and time: does form always follow function? How does the spatial and temporal organization of a eukaryotic genome reflect and influence its functions? *Bioessays*, 34(9):800–810.

Durnaoglu, S., Lee, S.-K., and Ahnn, J. (2021). Human Endogenous Retroviruses as Gene Expression Regulators: Insights from Animal Models into Human Diseases. *Mol Cells*, 44(12):861–878.

Dutrieux, J., Maarifi, G., Portilho, D. M., Arhel, N. J., Chelbi-Alix, M. K., and Nisole, S. (2015). PML/TRIM19-Dependent Inhibition of Retroviral Reverse-Transcription by Daxx. *PLOS Pathogens*, 11(11):1–22.

Dutta, R. and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. *Trends in Biochemical Sciences*, 25(1):24–28.

Dyer, M. A., Qadeer, Z. A., Valle-Garcia, D., and Bernstein, E. (2017). ATRX and DAXX: Mechanisms and Mutations. *Cold Spring Harbor perspectives in medicine*, 7(3):a026567.

Eberl, H. C., Spruijt, C. G., Kelstrup, C. D., Vermeulen, M., and Mann, M. (2013). A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol Cell*, 49(2):368–78.

Ecco, G., Imbeault, M., and Trono, D. (2017). KRAB zinc finger proteins. *Development*, 144(15):2719–2729.

Ekundayo, B. and Bleichert, F. (2019). Origins of DNA replication. *PLoS Genet*, 15(9):e1008320.

Elsasser, S. J., Noh, K. M., Diaz, N., Allis, C. D., and Banaszynski, L. A. (2015). Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*, 522(7555):240–4.

Eustermann, S., Yang, J.-C., Law, M. J., Amos, R., Chapman, L. M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R. J., Rhodes, D., Higgs, D. R., and Neuhaus, D. (2011). Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol*, 18(7):777–782.

Fan, R. (2015). Deletion of Setdb1 in Sox17 lineage cells impairs early embryonic development in the mouse. *Dissertation -Ludwig-Maximilians-Universität München.*

Farhat, D. C., Swale, C., Dard, C., Cannella, D., Ortet, P., Barakat, M., Sindikubwabo, F., Belmudes, L., De Bock, P.-J., Couté, Y., Bougdour, A., and Hakimi, M.-A. (2020). A MORC-driven transcriptional switch controls *Toxoplasma* developmental trajectories and sexual commitment. *Nat Microbiol*, 5(4):570–583.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nature Cell Biology*, 8(2):188–194.

Fernandez-Capetillo, O., Chen, H.-T., Celeste, A., Ward, I., Romanienko, P. J., Morales, J. C., Naka, K., Xia, Z., Camerini-Otero, R. D., Motoyama, N., Carpenter, P. B., Bonner, W. M., Chen, J., and Nussenzweig, A. (2002). DNA damage-induced G2–M checkpoint activation by histone H2AX and 53BP1. *Nature Cell Biology*, 4(12):993–997.

Ferrand, J., Rondinelli, B., and Polo, S. E. (2020). Histone Variants: Guardians of Genome Integrity. *Cells*, 9(11).

Ferrari, R., Grandi, N., Tramontano, E., and Dieci, G. (2021). Retrotransposons as Drivers of Mammalian Brain Evolution. *Life (Basel, Switzerland)*, 11(5):376.

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

Fiorentino, D. F., Chung, L. S., Christopher-Stine, L., Zaba, L., Li, S., Mammen, A. L., Rosen, A., and Casciola-Rosen, L. (2013). Most patients with cancer-associated dermatomyositis have antibodies to nuclear matrix protein NXP-2 or transcription intermediary factor 1gamma. *Arthritis Rheum*, 65(11):2954–62.

Fioriniello, S., Marano, D., Fiorillo, F., D'Esposito, M., and Della Ragione, F. (2020). Epigenetic Factors That Control Pericentric Heterochromatin Organization in Mammals. *Genes (Basel)*, 11(6).

Flemming, W. (1882). Zellsubstanz, Kern und Zelltheilung. *F.C.W. Vogel, Leipzig.*

Franklin, S. G. and Zweidler, A. (1977). Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature*, 266(5599):273–275.

Friedli, M. and Trono, D. (2015). The Developmental Control of Transposable Elements and the Evolution of Higher Species. *Annual Review of Cell and Developmental Biology*, 31(1):429–451.

Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G., and Rauscher, F. J. r. (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev*, 10(16):2067–2078.

Fujimoto, M., Watanabe, R., Ishitsuka, Y., and Okiyama, N. (2016). Recent advances in dermatomyositis-specific autoantibodies. *Current Opinion in Rheumatology*, 28(6).

Fukuda, K., Okuda, A., Yusa, K., and Shinkai, Y. (2018). A CRISPR knockout screen identifies SETDB1-target retroelement silencing factors in embryonic stem cells. *Genome Res*, 28(6):846–858.

Fukuda, K. and Shinkai, Y. (2020). SETDB1-Mediated Silencing of Retroelements. *Viruses*, 12(6).

Gaidt, M. M., Morrow, A., Fairgrieve, M. R., Karr, J. P., Yosef, N., and Vance, R. E. (2021). Self-guarding of MORC3 enables virulence factor-triggered immunity. *Nature*.

Galli, M., Martiny, E., Imani, J., Kumar, N., Koch, A., Steinbrenner, J., and Kogel, K.-H. (2021). CRISPR/SpCas9-mediated double knockout of barley Microrchidia MORC1 and MORC6a reveals their strong involvement in plant immunity, transcriptional gene silencing and plant growth. *Plant Biotechnol J*.

Gambogi, C. W. and Black, B. E. (2019). The nucleosomes that mark centromere location on chromosomes old and new. *Essays Biochem*, 63(1):15–27.

Geis, F. K. and Goff, S. P. (2020). Silencing and Transcriptional Regulation of Endogenous Retroviruses: An Overview. *Viruses*, 12(8):884.

Gibbons, R. J., Pellagatti, A., Garrick, D., Wood, W. G., Malik, N., Ayyub, H., Langford, C., Boultwood, J., Wainscoat, J. S., and Higgs, D. R. (2003). Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the alpha-thalassemia myelodysplasia syndrome (ATMDS). *Nat Genet*, 34(4):446–449.

Gifford, W. D., Pfaff, S. L., and Macfarlan, T. S. (2013). Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol*, 23(5):218–226.

Gogvadze, E. and Buzdin, A. (2009). Retroelements and their impact on genome evolution and functioning. *Cellular and Molecular Life Sciences*, 66(23):3727.

Grabski, D. F., Hu, Y., Sharma, M., and Rasmussen, S. K. (2019). Close to the Bedside: A Systematic Review of Endogenous Retroviruses and Their Impact in Oncology. *Journal of Surgical Research*, 240:145–155.

Grandi, N. and Tramontano, E. (2018). Human Endogenous Retroviruses Are Ancient Acquired Elements Still Shaping Innate Immune Responses. *Front Immunol*, 9:2039.

Greenberg, M. V. C. and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology*, 20(10):590–607.

Groh, S., Milton, A. V., Marinelli, L. K., Sickinger, C. V., Russo, A., Bollig, H., de Almeida, G. P., Schmidt, A., Forné, I., Imhof, A., and Schotta, G. (2021). Morc3 silences endogenous retroviruses by enabling Daxx-mediated histone H3.3 incorporation. *Nature Communications*, 12(1):5996.

Groner, A. C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, N., Bucher, P., and Trono, D. (2010). KRAB–Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading. *PLOS Genetics*, 6(3):e1000869–.

Gunawardena, H., Wedderburn, L. R., Chinoy, H., Betteridge, Z. E., North, J., Ollier, W. E., Cooper, R. G., Oddis, C. V., Ramanan, A. V., Davidson, J. E., McHugh, N. J., Juvenile Dermatomyositis Research Group, U. K., and Ireland (2009). Autoantibodies to a 140-kd protein in juvenile dermatomyositis are associated with calcinosis. *Arthritis Rheum*, 60(6):1807–14.

Haeckel, E. (1866). *Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie*. Bd. 2. G. Reimer.

He, F., Umehara, T., Saito, K., Harada, T., Watanabe, S., Yabuki, T., Kigawa, T., Takahashi, M., Kuwasako, K., Tsuda, K., Matsuda, T., Aoki, M., Seki, E., Kobayashi, N., Guntert, P., Yokoyama, S., and Muto, Y. (2010). Structural insight into the zinc finger CW domain as a histone modification reader. *Structure*, 18(9):1127–39.

He, Q., Kim, H., Huang, R., Lu, W., Tang, M., Shi, F., Yang, D., Zhang, X., Huang, J., Liu, D., and Songyang, Z. (2015). The Daxx/Atrx Complex Protects Tandem Repetitive Elements during DNA Hypomethylation by Promoting H3K9 Trimethylation. *Cell Stem Cell*, 17(3):273–286.

Heitz, E. (1928). *Das heterochromatin der moose*. Bornträger.

Helleboid, P.-Y., Heusel, M., Duc, J., Piot, C., Thorball, C. W., Coluccio, A., Pontis, J., Imbeault, M., Turelli, P., Aebersold, R., and Trono, D. (2019). The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. *EMBO J*, 38(18):e101220.

Henikoff, S. and Greally, J. M. (2016). Epigenetics, cellular memory and gene regulation. *Curr Biol*, 26(14):R644–8.

Hong, G., Qiu, H., Wang, C., Jadhav, G., Wang, H., Tickner, J., He, W., and Xu, J. (2016). The Emerging Role of MORC Family Proteins in Cancer Development and Bone Homeostasis. *Journal of Cellular Physiology*, 232(5):928–934.

Hoppmann, V., Thorstensen, T., Kristiansen, P. E., Veiseth, S. V., Rahman, M. A., Finne, K., Aalen, R. B., and Aasland, R. (2011). The CW domain, a new histone recognition module in chromatin proteins. *EMBO J*, 30(10):1939–52.

Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., and Eden, A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*, 27(3):404–408.

Hoyt, S. J., Storer, J. M., Hartley, G. A., Grady, P. G. S., Gershman, A., de Lima, L. G., Limouse, C., Halabian, R., Wojenski, L., Rodriguez, M., Altemose, N., Rhie, A., Core, L. J., Gerton, J. L., Makalowski, W., Olson, D., and et al., J. R. (2022). From telomere to telomere: The transcriptional and epigenetic state of human repeat elements. *Science*, 376(6588):eabk3112.

Hu, H., Khodadadi-Jamayran, A., Dolgalev, I., Cho, H., Badri, S., Chiriboga, L. A., Zeck, B., Lopez De Rodas Gregorio, M., Dowling, C. M., Labbe, K., Deng, J., Chen, T., Zhang, H., Zappile, P., Chen, Z., Ueberheide, B., and et al., K. (2021). Targeting the Atf7ip-Setdb1 Complex Augments Antitumor Immunity by Boosting Tumor Immunogenicity. *Cancer Immunol Res*, 9(11):1298–1315.

Huo, X., Ji, L., Zhang, Y., Lv, P., Cao, X., Wang, Q., Yan, Z., Dong, S., Du, D., Zhang, F., Wei, G., Liu, Y., and Wen, B. (2020). The Nuclear Matrix Protein SAFB Cooperates with Major Satellite RNAs to Stabilize Heterochromatin Architecture Partially through Phase Separation. *Mol Cell*, 77(2):368–383.

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(27):21082–21091.

Ichimura, T., Watanabe, S., Sakamoto, Y., Aoto, T., Fujita, N., and Nakao, M. (2005). Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins. *J Biol Chem*, 280(14):13928–13935.

Imbeault, M., Helleboid, P.-Y., and Trono, D. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature*, 543(7646):550–554.

Inoue, N., Hess, K. D., Moreadith, R. W., Richardson, L. L., Handel, M. A., Watson, M., and Zinn, A. R. (1999). New Gene Family Defined by MORC, a Nuclear Protein Required for Mouse Spermatogenesis. *Human Molecular Genetics*, 8(7):1201–1207.

Inoue, N., Wei, F., Seldin, M. F., Zinn, A. R., and Watson, M. L. (2000). Assignment of *microrchidia (Morc)* to mouse chromosome 16 by interspecific backcross linkage analysis and human chromosome 3q13 using somatic cell hybrids and in situ hybridization. *Cytogenetic and Genome Research*, 90(1-2):123–125.

Ishiuchi, T., Abe, S., Inoue, K., Yeung, W. K. A., Miki, Y., Ogura, A., and Sasaki, H. (2021). Reprogramming of the histone H3.3 landscape in the early mouse embryo. *Nature Structural & Molecular Biology*, 28(1):38–49.

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., Sadofsky, M. J., Zhou, M.-M., and Rauscher, F. J. r. (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol Cell*, 28(5):823–837.

Iyer, L. M., Abhiman, S., and Aravind, L. (2008). MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. *Biology direct*, 3:8–8.

Iyer, L. M., Anantharaman, V., Wolf, M. Y., and Aravind, L. (2007). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *Int J Parasitol*, 38(1):1–31.

Jacobs, F. M. J., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A. D., Katzman, S., Paten, B., Salama, S. R., and Haussler, D. (2014). An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, 516(7530):242–245.

Jacques, P.-É., Jeyakani, J., and Bourque, G. (2013). The majority of primate-specific regulatory sequences are derived from transposable elements. *PLoS Genet*, 9(5):e1003504.

Jadhav, G., Teguh, D., Kenny, J., Tickner, J., and Xu, J. (2016). Morc3 mutant mice exhibit reduced cortical area and thickness, accompanied by altered haematopoietic stem cells niche and bone cell differentiation. *Sci Rep*, 6:25964.

Jang, H. S., Shah, N. M., Du, A. Y., Dailey, Z. Z., Pehrsson, E. C., Godoy, P. M., Zhang, D., Li, D., Xing, X., Kim, S., O'Donnell, D., Gordon, J. I., and Wang, T. (2019). Transposable elements drive widespread expression of oncogenes in human cancers. *Nature Genetics*, 51(4):611–617.

Janssen, A., Colmenares, S. U., and Karpen, G. H. (2018). Heterochromatin: Guardian of the Genome. *Annu Rev Cell Dev Biol*, 34:265–288.

Jansz, N. and Faulkner, G. J. (2021). Endogenous retroviruses in the origins and treatment of cancer. *Genome biology*, 22(1):147–147.

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

Jones, M. E. (1953). Albrecht Kossel, a biographical sketch. *Yale J Biol Med*, 26(1):80–97.

Kang, H.-G., Kuhl, J. C., Kachroo, P., and Klessig, D. F. (2008). CRT1, an Arabidopsis ATPase that interacts with diverse resistance proteins and modulates disease resistance to turnip crinkle virus. *Cell Host Microbe*, 3(1):48–57.

Kang, Y.-K. (2018). Surveillance of Retroelement Expression and Nucleic-Acid Immunity by Histone Methyltransferase SETDB1. *Bioessays*, 40(9):e1800058.

Karimi, M. M., Goyal, P., Maksakova, I. A., Bilenky, M., Leung, D., Tang, J. X., Shinkai, Y., Mager, D. L., Jones, S., Hirst, M., and Lorincz, M. C. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell*, 8(6):676–687.

Katsanis, N., Beck, J. A., and Fisher, E. M. C. (1997). Mapping of a novel SH3 domain protein and two proteins of unknown function to human chromosome 21. *Human Genetics*, 100(3):477–480.

Kim, H., Yen, L., Wongpalee, S. P., Kirshner, J. A., Mehta, N., Xue, Y., Johnston, J. B., Burlingame, A. L., Kim, J. K., Loparo, J. J., and Jacobsen, S. E. (2019). The Gene-Silencing Protein MORC-1 Topologically Entraps DNA and Forms Multimeric Assemblies to Cause DNA Compaction. *Molecular Cell*, 75(4):700–710.e6.

Kimura, Y., Sakai, F., Nakano, O., Kisaki, O., Sugimoto, H., Sawamura, T., Sadano, H., and Osumi, T. (2002). The newly identified human nuclear protein NXP-2 possesses three distinct domains, the nuclear matrix-binding, RNA-binding, and coiled-coil domains. *J Biol Chem*, 277(23):20611–7.

Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet*, 20(4):207–220.

Koch, A., Kang, H.-G., Steinbrenner, J., Dempsey, D. A., Klessig, D. F., and Kogel, K.-H. (2017). MORC Proteins: Novel Players in Plant and Animal Health. *Frontiers in Plant Science*, 8:1720.

Kojima-Kita, K., Kuramochi-Miyagawa, S., Nakayama, M., Miyata, H., Jacobsen, S. E., Ikawa, M., Koseki, H., and Nakano, T. (2021). MORC3, a novel MIWI2 association partner, as an epigenetic regulator of piRNA dependent transposon silencing in male germ cells. *Sci Rep*, 11(1):20472.

Kornberg, R. D. and Thomas, J. O. (1974). Chromatin Structure: Oligomers of the Histones. *Science*, 184(4139):865–868.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, 128(4):693–705.

Kuff, E. L. and Lueders, K. K. (1988). The intracisternal A-particle gene family: structure and functional aspects. *Adv Cancer Res*, 51:183–276.

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(7):631–634.

Kurdistani, S. K. and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol*, 4(4):276–284.

Küry, P., Nath, A., Créange, A., Dolei, A., Marche, P., Gold, J., Giovannoni, G., Hartung, H.-P., and Perron, H. (2018). Human Endogenous Retroviruses in Neurological Diseases. *Trends Mol Med*, 24(4):379–394.

Kyle, R. A. and Shampo, M. A. (1998). Wilkins, Crick, and Watson: Nobel Prize for work on the structure of DNA. *Mayo Clin Proc*, 73(4):362.

L., C. P., E., K. K., Takeshi, E., Yoichi, S., and M., O. E. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proceedings of the National Academy of Sciences*, 112(27):8367–8372.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*, 410(6824):116–120.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., and et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921.

Langen, G., von Einem, S., Koch, A., Imani, J., Pai, S. B., Manohar, M., Ehlers, K., Choi, H. W., Claar, M., Schmidt, R., Mang, H.-G., Bordiya, Y., Kang, H.-G., Klessig, D. F., and Kogel, K.-H. (2014). The compromised recognition of turnip crinkle virus1 subfamily of microrchidia ATPases regulates disease resistance in barley to biotrophic and necrotrophic pathogens. *Plant Physiol*, 164(2):866–878.

Lazaro-Camp, V. J., Salari, K., Meng, X., and Yang, S. (2021). SETDB1 in cancer: overexpression and its therapeutic implications. *American journal of cancer research*, 11(5):1803–1827.

Lehnertz, B., Ueda, Y., Derijck, A. A. H. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. F. M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol*, 13(14):1192–1200.

Leung, D. C. and Lorincz, M. C. (2012). Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends in Biochemical Sciences*, 37(4):127–133.

Lewis, P. W., Elsaesser, S. J., Noh, K.-M., Stadler, S. C., and Allis, C. D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci U S A*, 107(32):14075–14080.

Li, D.-Q., Nair, S. S., and Kumar, R. (2013). The MORC family: new epigenetic regulators of transcription and DNA damage response. *Epigenetics*, 8(7):685–693. PMID: 23804034.

Li, D.-Q., Nair, S. S., Ohshiro, K., Kumar, A., Nair, V. S., Pakala, S. B., Reddy, S. D. N., Gajula, R. P., Eswaran, J., Aravind, L., and Kumar, R. (2012a). MORC2 signaling integrates phosphorylation-dependent, ATPase-coupled chromatin remodeling during the DNA damage response. *Cell Rep*, 2(6):1657–1669.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, 69(6):915–926.

Li, S., Yen, L., Pastor, W. A., Johnston, J. B., Du, J., Shew, C. J., Liu, W., Ho, J., Stender, B., Clark, A. T., Burlingame, A. L., Daxinger, L., Patel, D. J., and Jacobsen, S. E. (2016). Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. *Proceedings of the National Academy of Sciences*, 113(35):E5108–E5116.

Li, X., Foley, E. A., Molloy, K. R., Li, Y., Chait, B. T., and Kapoor, T. M. (2012b). Quantitative Chemical Proteomics Approach To Identify Post-translational Modification-Mediated Protein–Protein Interactions. *Journal of the American Chemical Society*, 134(4):1982–1985.

Liang, J., Liu, H., Li, G., Qian, J., Gao, R., Zhou, Y., and Wang, X. (2020). Global changes in chromatin accessibility and transcription following ATRX inactivation in human cancer cells. *FEBS Lett*, 594(1):67–78.

Liggins, A. P., Cooper, C. D. O., Lawrie, C. H., Brown, P. J., Collins, G. P., Hatton, C. S., Pulford, K., and Banham, A. H. (2007). MORC4, a novel member of the MORC family, is highly expressed in a subset of diffuse large B-cell lymphomas. *British Journal of Haematology*, 138(4):479–486.

Lima-Junior, D. S., Krishnamurthy, S. R., Bouladoux, N., Collins, N., Han, S.-J., Chen, E. Y., Constantinides, M. G., Link, V. M., Lim, A. I., Enamorado, M., Cataisson, C., Gil, L., Rao, I., Farley, T. K., Koroleva, G., Attig, J., Yuspa, S. H., Fischbach, M. A., Kassiotis, G., and Belkaid, Y. (2021). Endogenous retroviruses promote homeostatic and inflammatory responses to the microbiota. *Cell*, 184(14):3794–3811.

Lin, D.-Y., Huang, Y.-S., Jeng, J.-C., Kuo, H.-Y., Chang, C.-C., Chao, T.-T., Ho, C.-C., Chen, Y.-C., Lin, T.-P., Fang, H.-I., Hung, C.-C., Suen, C.-S., Hwang, M.-J., Chang, K.-S., Maul, G. G., and Shih, H.-M. (2006). Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Mol Cell*, 24(3):341–354.

Ling, K.-H., Hewitt, C. A., Tan, K.-L., Cheah, P.-S., Vidyadaran, S., Lai, M.-I., Lee, H.-C., Simpson, K., Hyde, L., Pritchard, M. A., Smyth, G. K., Thomas, T., and Scott, H. S. (2014). Functional transcriptome analysis of the postnatal brain of the Ts1Cje mouse model for Down syndrome reveals global disruption of interferon-related molecular networks. *BMC Genomics*, 15(1):624.

Liu, J., Ali, M., and Zhou, Q. (2020). Establishment and evolution of heterochromatin. *Ann N Y Acad Sci*, 1476(1):59–77.

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

Liu, Y., Tempel, W., Zhang, Q., Liang, X., Loppnau, P., Qin, S., and Min, J. (2016). Family-wide Characterization of Histone Binding Abilities of Human CW Domain-containing Proteins*. *Journal of Biological Chemistry*, 291(17):9000–9013.

Loppin, B. and Berger, F. (2020). Histone Variants: The Nexus of Developmental Decisions and Epigenetic Memory. *Annu Rev Genet*, 54:121–149.

Low, Y., Tan, D. E. K., Hu, Z., Tan, S. Y. X., and Tee, W.-W. (2021). Transposable Element Dynamics and Regulation during Zygotic Genome Activation in Mammalian Embryos and Embryonic Stem Cell Model Systems. *Stem Cells International*, 2021:1624669.

Loyola, A., Tagami, H., Bonaldi, T., Roche, D., Quivy, J. P., Imhof, A., Nakatani, Y., Dent, S. Y. R., and Almouzni, G. (2009). The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO Rep*, 10(7):769–775.

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Åresolution. *Nature*, 389(6648):251–260.

Lukic, S., Nicolas, J.-C., and Levine, A. J. (2014). The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. *Cell death and differentiation*, 21(3):381–387.

Mager, D. L. and Stoye, J. P. (2015). Mammalian Endogenous Retroviruses. *Microbiol Spectr*, 3(1):MDNA3–0009–2014.

Mahmud, I. and Liao, D. (2019). DAXX in cancer: phenomena, processes, mechanisms and regulation. *Nucleic Acids Res*, 47(15):7734–7752.

Maillet, S., Fernandez, J., Decourcelle, M., El Koulali, K., Blanchet, F. P., Arhel, N. J., Maarifi, G., and Nisole, S. (2020). Daxx Inhibits HIV-1 Reverse Transcription and Uncoating in a SUMO-Dependent Manner. *Viruses*, 12(6).

Martire, S. and Banaszynski, L. A. (2020). The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol*, 21(9):522–541.

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(7290):927–31.

McClintock, B. (1956). Controlling elements and the gene. *Cold Spring Harb Symp Quant Biol*, 21:197–216.

Mendel, G. (1866). Versuche Über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines zu Brünn*, 4:3–47.

Mi, S., Lee, X., Li, X.-p., Veldman, G. M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.-Y., Edouard, P., Howes, S., Keith, J. C., and McCoy, J. M. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*, 403(6771):785–789.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S., and Bernstein, B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448(7153):553–560.

Mimura, Y., Takahashi, K., Kawata, K., Akazawa, T., and Inoue, N. (2010). Two-step colocalization of MORC3 with PML nuclear bodies. *J Cell Sci*, 123(Pt 12):2014–24.

Moissiard, G., Bischof, S., Husmann, D., Pastor, W. A., Hale, C. J., Yen, L., Stroud, H., Papikian, A., Vashisht, A. A., Wohlschlegel, J. A., and Jacobsen, S. E. (2014). Transcriptional gene silencing by Arabidopsis microrchidia homologues involves the formation of heteromers. *Proceedings of the National Academy of Sciences*, 111(20):7474–7479.

Moissiard, G., Cokus, S. J., Cary, J., Feng, S., Billi, A. C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B. R., McCord, R. P., Hale, C. J., Feng, W., Michaels, S. D., Frand, A. R., Pellegrini, M., Dekker, J., Kim, J. K., and Jacobsen, S. E. (2012). MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing. *Science*, 336(6087):1448.

Montavon, T., Shukeir, N., Erikson, G., Engist, B., Onishi-Seebacher, M., Ryan, D., Musa, Y., Mittler, G., Meyer, A. G., Genoud, C., and Jenuwein, T. (2021). Complete loss of H3K9 methylation dissolves mouse heterochromatin organization. *Nature communications*, 12(1):4359–4359.

Moore, L. D., Le, T., and Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 38(1):23–38.

Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J. P., and Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res*, 24(24):4859–4867.

Mossink, B., Negwer, M., Schubert, D., and Nadif Kasri, N. (2021). The emerging role of chromatin remodelers in neurodevelopmental disorders: a developmental perspective. *Cellular and molecular life sciences : CMLS*, 78(6):2517–2563.

Mundorf, A., Koch, J., Kubitza, N., Wagner, S. C., Schmidt, M., Gass, P., and Freund, N. (2021). Morc1 as a potential new target gene in mood regulation: when and where to find in the brain. *Experimental Brain Research*.

Navarro, C., Lyu, J., Katsori, A.-M., Caridha, R., and Elsässer, S. J. (2020). An embryonic stem cell-specific heterochromatin state promotes core histone exchange in the absence of DNA accessibility. *Nature Communications*, 11(1):5095.

Nicetto, D. and Zaret, K. S. (2019). Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Current Opinion in Genetics & Development*, 55:1–10.

Nielsen, A. L., Ortiz, J. A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P., and Losson, R. (1999). Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO J*, 18(22):6385–6395.

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., and et al., G. G. B. (2022). The complete sequence of a human genome. *Science*, 376(6588):44–53.

Ochoa Thomas, E., Zuniga, G., Sun, W., and Frost, B. (2020). Awakening the dark side: retrotransposon activation in neurodegenerative disorders. *Curr Opin Neurobiol*, 61:65–72.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell*, 99(3):247–257.

Okano, M., Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, 19(3):219–220.

Olins, A. L. and Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science*, 183(4122):330–332.

Ooi, S. K. T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C. D., Cheng, X., and Bestor, T. H. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, 448(7154):714—717.

P., S. S., Jessica, S., and C., S. D. (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. *Molecular and Cellular Biology*, 26(22):8623–8638.

Padmanabhan Nair, V., Liu, H., Ciceri, G., Jungverdorben, J., Frishman, G., Tchieu, J., Cederquist, G. Y., Rothenaigner, I., Schorpp, K., Klepper, L., Walsh, R. M., Kim, T. W., Cornacchia, D., Ruepp, A., Mayer, J., Hadian, K., Frishman, D., Studer, L., and Vincendeau, M. (2021). Activation of HERV-K(HML-2) disrupts cortical patterning and neuronal differentiation by increasing NTRK3. *Cell Stem Cell*, 28(9):1566–1581.

Pannell, D., Osborne, C. S., Yao, S., Sukonnik, T., Pasceri, P., Karaiskakis, A., Okano, M., Li, E., Lipshitz, H. D., and Ellis, J. (2000). Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. *EMBO J*, 19(21):5884–5894.

Pasquarella, A. (2015). The role of repressive chromatin functions during haematopoiesis. *Dissertation - Ludwig-Maximilians-Universität München.*

Pasquarella, A., Ebert, A., Pereira de Almeida, G., Hinterberger, M., Kazerani, M., Nuber, A., Ellwart, J., Klein, L., Busslinger, M., and Schotta, G. (2016). Retrotransposon derepression leads to activation of the unfolded protein response and apoptosis in pro-B cells. *Development*, 143(10):1788–99.

Pastor, W. A., Stroud, H., Nee, K., Liu, W., Pezic, D., Manakov, S., Lee, S. A., Moissiard, G., Zamudio, N., Bourc'his, D., Aravin, A. A., Clark, A. T., and Jacobsen, S. E. (2014). MORC1 represses transposable elements in the mouse male germline. *Nat Commun*, 5:5795.

Perry, J. and Zhao, Y. (2003). The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. *Trends in Biochemical Sciences*, 28(11):576–580.

Pick, H., Kilic, S., and Fierz, B. (2014). Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function. *Biochim Biophys Acta*, 1839(8):644–656.

Picketts, D. J., Higgs, D. R., Bachoo, S., Blake, D. J., Quarrell, O. W., and Gibbons, R. J. (1996). ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum Mol Genet*, 5(12):1899–1907.

Pluta, A., Jaworski, J. P., and Douville, R. N. (2020). Regulation of Expression and Latency in BLV and HTLV. *Viruses*, 12(10):1079.

Pray, L. A. (2008). Discovery of DNA Structure and Function: Watson and Crick. *Nature Education*, 1:100.

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(4):766–773.

Quinodoz, S. A., Jachowicz, J. W., Bhat, P., Ollikainen, N., Banerjee, A. K., Goronzy, I. N., Blanco, M. R., Chovanec, P., Chow, A., Markaki, Y., Thai, J., Plath, K., and Guttman, M. (2021). RNA promotes the formation of spatial compartments in the nucleus. *Cell*, 184(23):5775–5790.e30.

Ramírez, M. A., Pericuesta, E., Fernandez-Gonzalez, R., Moreira, P., Pintado, B., and Gutierrez-Adan, A. (2006). Transcriptional and post-transcriptional regulation of retrotransposons IAP and MuERV-L affect pluripotency of mice ES cells. *Reprod Biol Endocrinol*, 4:55.

Randolph, K., Hyder, U., and D'Orso, I. (2022). KAP1/TRIM28: Transcriptional Activator and/or Repressor of Viral and Cellular Programs? *Front Cell Infect Microbiol*, 12:834636.

Rapoport, S. (2003). Rosalind Franklin: Unsung Hero of the DNA Revolution. *New York History*, 84(3):315–329.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406(6796):593–599.

Ren, W., Medeiros, N., Warneford-Thomson, R., Wulfridge, P., Yan, Q., Bian, J., Sidoli, S., Garcia, B. A., Skordalakes, E., Joyce, E., Bonasio, R., and Sarma, K. (2020). Disruption of ATRX-RNA interactions uncovers roles in ATRX localization and PRC2 function. *Nat Commun*, 11(1):2219.

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(4):597–609.

Rosendorff, A., Sakakibara, S., Lu, S., Kieff, E., Xuan, Y., DiBacco, A., Shi, Y., Shi, Y., and Gill, G. (2006). NXP-2 association with SUMO-2 depends on lysines required for transcriptional repression. *Proceedings of the National Academy of Sciences*, 103(14):5308–5313.

Rothbart, S. B., Dickson, B. M., Ong, M. S., Krajewski, K., Houliston, S., Kireev, D. B., Arrowsmith, C. H., and Strahl, B. D. (2013). Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev*, 27(11):1288–1298.

Rowe, H. M., Friedli, M., Offner, S., Verp, S., Mesnard, D., Marquis, J., Aktas, T., and Trono, D. (2013a). De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. *Development*, 140(3):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., Spitz, F., Constam, D. B., and Trono, D. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*, 463(7278):237–40.

Rowe, H. M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T. S., Tarabay, Y., Viville, S., Jakobsson, J., Pfaff, S. L., and Trono, D. (2013b). TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. *Genome research*, 23(3):452–461.

Rowe, H. M. and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. *Virology*, 411(2):273–287.

Sadic, D. (2014). Regulation of heterochromatic gene silencing in mouse. *Dissertation - Ludwig-Maximilians-Universität München*.

Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F. J., and Schotta, G. (2015). Atrx promotes heterochromatin formation at retrotransposons. *EMBO Reports*, 16(7):836–50.

Saitou, M., Kagiwada, S., and Kurimoto, K. (2012). Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development*, 139(1):15–31.

Sakashita, A., Maezawa, S., Takahashi, K., Alavattam, K. G., Yukawa, M., Hu, Y.-C., Kojima, S., Parrish, N. F., Barski, A., Pavlicev, M., and Namekawa, S. H. (2020). Endogenous retroviruses drive species-specific germline transcriptomes in mammals. *Nature Structural & Molecular Biology*, 27(10):967–977.

Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A. J., and Torres-Padilla, M.-E. (2010). Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nature Cell Biology*, 12(9):853–862.

Sarma, K., Cifuentes-Rojas, C., Ergun, A., Del Rosario, A., Jeon, Y., White, F., Sadreyev, R., and Lee, J. T. (2014). ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell*, 159(4):869–883.

Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G., and Rauscher, F. J. r. (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(8):919–932.

Schultz, D. C., Friedman, J. R., and Rauscher, F. J. r. (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-2alpha subunit of NuRD. *Genes Dev*, 15(4):428–443.

Seki, Y., Kurisaki, A., Watanabe-Susaki, K., Nakajima, Y., Nakanishi, M., Arai, Y., Shiota, K., Sugino, H., and Asashima, M. (2010). TIF1beta regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. *Proceedings of the National Academy of Sciences of the United States of America*, 107(24):10926–10931.

Sevilla, T., Lupo, V., Martínez-Rubio, D., Sancho, P., Sivera, R., Chumillas, M. J., García-Romero, M., Pascual-Pascual, S. I., Muelas, N., Dopazo, J., Vílchez, J. J., Palau, F., and Espinós, C. (2015). Mutations in the MORC2 gene cause axonal Charcot–Marie–Tooth disease . *Brain*, 139(1):62–72.

Shalginskikh, N., Poleshko, A., Skalka, A. M., and Katz, R. A. (2013). Retroviral DNA methylation and epigenetic repression are mediated by the antiviral host protein Daxx. *J Virol*, 87(4):2137–2150.

Shao, Y., Li, Y., Zhang, J., Liu, D., Liu, F., Zhao, Y., Shen, T., and Li, F. (2010). Involvement of histone deacetylation in MORC2-mediated down-regulation of carbonic anhydrase IX. *Nucleic Acids Res*, 38(9):2813–2824.

Sharif, J., Endo, T. A., Nakayama, M., Karimi, M. M., Shimada, M., Katsuyama, K., Goyal, P., Brind'Amour, J., Sun, M. A., Sun, Z., Ishikura, T., Mizutani-Koseki, Y., Ohara, O., Shinkai, Y., Nakanishi, M., Xie, H., Lorincz, M. C., and Koseki, H. (2016). Activation of Endogenous Retroviruses in Dnmt1(-/-) ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. *Cell Stem Cell*, 19(1):81–94.

Sharif, J., Muto, M., Takebayashi, S.-i., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M., and Koseki, H. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, 450(7171):908–912.

Shi, B., Xue, J., Zhou, J., Kasowitz, S. D., Zhang, Y., Liang, G., Guan, Y., Shi, Q., Liu, M., Sha, J., Huang, X., and Wang, P. J. (2018). MORC2B is essential for meiotic progression and fertility. *PLoS genetics*, 14(1):e1007175–e1007175.

Shi, L., Wen, H., and Shi, X. (2017). The Histone Variant H3.3 in Transcriptional Regulation and Human Disease. *Journal of molecular biology*, 429(13):1934–1945.

Sloan, E., Orr, A., and Everett, R. D. (2016). MORC3, a Component of PML Nuclear Bodies, Has a Role in Restricting Herpes Simplex Virus 1 and Human Cytomegalovirus. *Journal of virology*, 90(19):8621–8633.

Sloan, E., Tatham, M. H., Groslambert, M., Glass, M., Orr, A., Hay, R. T., and Everett, R. D. (2015). Analysis of the SUMO2 Proteome during HSV-1 Infection. *PLoS Pathog*, 11(7):e1005059.

Slotkin, R. K. and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*, 8(4):272–285.

Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*, 22(2):96–118.

Stocking, C. and Kozak, C. A. (2008). Murine endogenous retroviruses. *Cell Mol Life Sci*, 65(21):3383–3398.

Stoye, J. P. (2012). Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nature Reviews Microbiology*, 10(6):395–406.

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 research*, 24(12):1963–1976.

Takahashi, K., Yoshida, N., Murakami, N., Kawata, K., Ishizaki, H., Tanaka-Okamoto, M., Miyoshi, J., Zinn, A. R., Shime, H., and Inoue, N. (2007). Dynamic Regulation of p53 Subnuclear Localization and Senescence by MORC3. *Molecular Biology of the Cell*, 18(5):1701–1709. PMID: 17332504.

Takikita, S., Muro, R., Takai, T., Otsubo, T., Kawamura, Y. I., Dohi, T., Oda, H., Kitajima, M., Oshima, K., Hattori, M., Endo, T. A., Toyoda, T., Weis, J., Shinkai, Y., and Suzuki, H. (2016). A Histone Methyltransferase ESET Is Critical for T Cell Development. *J Immunol*, 197(6):2269–2279.

Talbert, P. B. and Henikoff, S. (2006). Spreading of silent chromatin: inaction at a distance. *Nature Reviews Genetics*, 7(10):793–803.

Talbert, P. B. and Henikoff, S. (2017). Histone variants on the move: substrates for chromatin dynamics. *Nature Reviews Molecular Cell Biology*, 18(2):115–126.

Tan, X., Xu, X., Elkenani, M., Smorag, L., Zechner, U., Nolte, J., Engel, W., and Pantakani, D. V. (2013). Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells. *Stem Cell Res*, 11(3):1045–59.

Tang, S.-Y., Wan, Y.-P., and Wu, Y.-M. (2015). Death domain associated protein (Daxx), a multi-functional protein. *Cell Mol Biol Lett*, 20(5):788–797.

Tchasovnikarova, I. A., Timms, R. T., Douse, C. H., Roberts, R. C., Dougan, G., Kingston, R. E., Modis, Y., and Lehner, P. J. (2017). Hyperactivation of HUSH complex function by Charcot–Marie–Tooth disease mutation in MORC2. *Nature Genetics*, 49(7):1035–1044.

Tencer, A. H., Cox, K. L., Wright, G. M., Zhang, Y., Petell, C. J., Klein, B. J., Strahl, B. D., Black, J. C., Poirier, M. G., and Kutateladze, T. G. (2020). Molecular mechanism of the MORC4 ATPase activation. *Nature Communications*, 11(1):5466.

Teng, Y.-C., Sundaresan, A., O'Hara, R., Gant, V. U., Li, M., Martire, S., Warshaw, J. N., Basu, A., and Banaszynski, L. A. (2021). ATRX promotes heterochromatin formation to protect cells from G-quadruplex DNA-mediated stress. *Nat Commun*, 12(1):3887.

Thoma, F. and Koller, T. (1977). Influence of histone H1 on chromatin structure. *Cell*, 12(1):101–107.

Thomas, J. H. and Schneider, S. (2011). Coevolution of retroelements and tandem zinc finger genes. *Genome Res*, 21(11):1800–1812.

Thurman, R. E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M. T., Haugen, E., Sheffield, N. C., Stergachis, A. B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., and et al. (2012). The accessible chromatin landscape of the human genome. *Nature*, 489(7414):75–82.

Timms, R. T., Tchasovnikarova, I. A., Antrobus, R., Dougan, G., and Lehner, P. J. (2016). ATF7IP-Mediated Stabilization of the Histone Methyltransferase SETDB1 Is Essential for Heterochromatin Formation by the HUSH Complex. *Cell Rep*, 17(3):653–659.

Tokuyama, M., Gunn, B. M., Venkataraman, A., Kong, Y., Kang, I., Rakib, T., Townsend, M. J., Costenbader, K. H., Alter, G., and Iwasaki, A. (2021). Antibodies against human endogenous retrovirus K102 envelope activate neutrophils in systemic lupus erythematosus. *The Journal of experimental medicine*, 218(7):e20191766.

Tsusaka, T., Fukuda, K., Shimura, C., Kato, M., and Shinkai, Y. (2020). The fibronectin type-III (FNIII) domain of ATF7IP contributes to efficient transcriptional silencing mediated by the SETDB1 complex. *Epigenetics & chromatin*, 13(1):52–52.

Valle-García, D., Qadeer, Z. A., McHugh, D. S., Ghiraldini, F. G., Chowdhury, A. H., Hasson, D., Dyer, M. A., Recillas-Targa, F., and Bernstein, E. (2016). ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment. *Epigenetics*, 11(6):398–414.

Walsh, C. P., Chaillet, J. R., and Bestor, T. H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genetics*, 20(2):116–117.

Wang, C., Liu, X., Gao, Y., Yang, L., Li, C., Liu, W., Chen, C., Kou, X., Zhao, Y., Chen, J., Wang, Y., Le, R., Wang, H., Duan, T., Zhang, Y., and Gao, S. (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nature Cell Biology*, 20(5):620–631.

Wang, G., Song, Y., Liu, T., Wang, C., Zhang, Q., Liu, F., Cai, X., Miao, Z., Xu, H., Xu, H., Cao, L., and Li, F. (2015). PAK1-mediated MORC2 phosphorylation promotes gastric tumorigenesis. *Oncotarget*, 6(12):9877–9886.

Wang, G.-L., Wang, C.-Y., Cai, X.-Z., Chen, W., Wang, X.-H., and Li, F. (2010). Identification and Expression Analysis of a Novel CW-Type Zinc Finger Protein MORC2 in Cancer Cells. *The Anatomical Record*, 293(6):1002–1009.

Wang, H., An, W., Cao, R., Xia, L., Erdjument-Bromage, H., Chatton, B., Tempst, P., Roeder, R. G., and Zhang, Y. (2003). mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. *Mol Cell*, 12(2):475–487.

Wang, Z. (2020). Dominant role of DNA methylation over H3K9me3 in ERV silencing during embryonic endoderm development. *Dissertation - Ludwig-Maximilians-Universität München*.

Wasylishen, A. R., Sun, C., Moyer, S. M., Qi, Y., Chau, G. P., Aryal, N. K., McAllister, F., Kim, M. P., Barton, M. C., Estrella, J. S., Su, X., and Lozano, G. (2020). Daxx maintains endogenous retroviral silencing and restricts cellular plasticity in vivo. *Sci Adv*, 6(32):eaba8415.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., and et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915):520–562.

Watson, J. D. and Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356):737–738.

Watson, M. L., Zinn, A. R., Inoue, N., Hess, K. D., Cobb, J., Handel, M. A., Halaban, R., Duchene, C. C., Albright, G. M., and Moreadith, R. W. (1998). Identification of morc (microrchidia), a mutation that results in arrest of spermatogenesis at an early meiotic stage in the mouse. *Proceedings of the National Academy of Sciences*, 95(24):14361–14366.

Weiser, N. E., Yang, D. X., Feng, S., Kalinava, N., Brown, K. C., Khanikar, J., Freeberg, M. A., Snyder, M. J., Csankovszki, G., Chan, R. C., Gu, S. G., Montgomery, T. A., Jacobsen, S. E., and Kim, J. K. (2017). MORC-1 Integrates Nuclear RNAi and Transgenerational Chromatin Architecture to Promote Germline Immortality. *Dev Cell*, 41(4):408–423 e7.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., and Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, 8(12):973–982.

Willcockson, M. A., Healton, S. E., Weiss, C. N., Bartholdy, B. A., Botbol, Y., Mishra, L. N., Sidhwani, D. S., Wilson, T. J., Pinto, H. B., Maron, M. I., Skalina, K. A., Toro, L. N., Zhao, J., Lee, C.-H., Hou, H., Yusufova, N., Meydan, C., Osunsade, A., David, Y., Cesarman, E., Melnick, A. M., Sidoli, S., Garcia, B. A., Edelmann, W., Macian, F., and Skoultchi, A. I. (2021). H1 histones control the epigenetic landscape by local chromatin compaction. *Nature*, 589(7841):293–298.

Wolf, D. and Goff, S. P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature*, 458(7242):1201–1204.

Wolf, G., de Iaco, A., Sun, M.-A., Bruno, M., Tinkham, M., Hoang, D., Mitra, A., Ralls, S., Trono, D., and Macfarlan, T. S. (2020). KRAB-zinc finger protein gene expansion in response to active retrotransposons in the murine lineage. *Elife*, 9.

Wolf, G., Rebollo, R., Karimi, M. M., Ewing, A. D., Kamada, R., Wu, W., Wu, B., Bachu, M., Ozato, K., Faulkner, G. J., Mager, D. L., Lorincz, M. C., and Macfarlan, T. S. (2017). On the role of H3.3 in retroviral silencing. *Nature*, 548(7665):E1–E3.

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(5):538–54.

Wong, L. H., McGhie, J. D., Sim, M., Anderson, M. A., Ahn, S., Hannan, R. D., George, A. J., Morgan, K. A., Mann, J. R., and Choo, K. H. A. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res*, 20(3):351–360.

Xue, Y., Zhong, Z., Harris, C. J., Gallego-Bartolomé, J., Wang, M., Picard, C., Cao, X., Hua, S., Kwok, I., Feng, S., Jami-Alahmadi, Y., Sha, J., Gardiner, J., Wohlschlegel, J., and Jacobsen, S. E. (2021). Arabidopsis MORC proteins function in the efficient establishment of RNA directed DNA methylation. *Nature communications*, 12(1):4292–4292.

Yang, B. X., El Farran, C. A., Guo, H. C., Yu, T., Fang, H. T., Wang, H. F., Schlesinger, S., Seah, Y. F., Goh, G. Y., Neo, S. P., Li, Y., Lorincz, M. C., Tergaonkar, V., Lim, T. M., Chen, L., Gunaratne, J., Collins, J. J., Goff, S. P., Daley, G. Q., Li, H., Bard, F. A., and Loh, Y. H. (2015). Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell*, 163(1):230–45.

Yang, X., Khosravi-Far, R., Chang, H. Y., and Baltimore, D. (1997). Daxx, a novel Fas-binding protein that activates JNK and apoptosis. *Cell*, 89(7):1067–1076.

Zaret, K. S. (2020). Pioneer Transcription Factors Initiating Gene Network Changes. *Annual review of genetics*, 54:367–385.

Zhang, Q., Song, Y., Chen, W., Wang, X., Miao, Z., Cao, L., Li, F., and Wang, G. (2015). By recruiting HDAC1, MORC2 suppresses p21 Waf1/Cip1 in gastric cancer. *Oncotarget*, 6(18):16461–16470.

Zhang, S., Übelmesser, N., Josipovic, N., Forte, G., Slotman, J. A., Chiang, M., Gothe, H. J., Gusmao, E. G., Becker, C., Altmüller, J., Houtsmuller, A. B., Roukos, V., Wendt, K. S., Marenduzzo, D., and Papantonis, A. (2021a). RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis. *Science Advances*, 7(43):eabg8205.

Zhang, Y., Ahn, J., Green, K. J., Vann, K. R., Black, J., Brooke, C. B., and Kutateladze, T. G. (2019a). MORC3 Is a Target of the Influenza A Viral Protein NS1. *Structure*, 27(6):1029–1033 e3.

Zhang, Y., Bertulat, B., Tencer, A. H., Ren, X., Wright, G. M., Black, J., Cardoso, M. C., and Kutateladze, T. G. (2019b). MORC3 Forms Nuclear Condensates through Phase Separation. *iScience*, 17:182–189.

Zhang, Y., Klein, B. J., Cox, K. L., Bertulat, B., Tencer, A. H., Holden, M. R., Wright, G. M., Black, J., Cardoso, M. C., Poirier, M. G., and Kutateladze, T. G. (2019c). Mechanism for autoinhibition and activation of the MORC3 ATPase. *Proceedings of the National Academy of Sciences*, 116(13):6111–6119.

Zhang, Y., Sun, Z., Jia, J., Du, T., Zhang, N., Tang, Y., Fang, Y., and Fang, D. (2021b). Overview of Histone Modification. *Adv Exp Med Biol*, 1283:1–16.

Zhou, V. W., Goren, A., and Bernstein, B. E. (2011). Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet*, 12(1):7–18.

Zhu, L. and Qin, J. (2019). A Viral Protein Mimics Histone to Hijack Host MORC3. *Structure*, 27(6):883–885.

Acknowledgements

First of all, I want to thank Prof. Dr. Gunnar Schotta for giving me the opportunity to dive into the exciting field of ERV silencing and to join his group for several professionally formative years.
I am also grateful for the funding of the SFB1064 “Chromatin Dynamics” and the support of the integrated training program (IRTG), particularly of the IRTG coordinator Dr. Elizabeth Schroeder-Reiter, who organized stimulating seminars, profitable workshops as well as networking events and retreats, which made the doctoral time a special experience also outside of the laboratory.
Furthermore, I want to express my gratitude to all members of the Molecular Biology Department, especially our head Prof. Dr. Peter Becker, for providing a stimulating and collaborative scientific environment.
Special thanks go to my TAC-committee members Prof. Dr. Axel Imhof and Prof. Dr. Felix Müller-Planitz for their valuable advice on my project.

My thanks further extend to my collaborators: In particular to the three excellent master students Lisa Marinelli, Anna Milton, and Cara Sickinger, who contributed greatly to this work through their outstanding dedication and significant participation in determining key findings of the MORC3 study. I also want to thank Angela Russo for supporting the revision experiments of the MORC3 manuscript and following the project after I left the group.
In addition, I am thankful for the collaboration with Dr. Andreas Schmidt and Dr. Ignasi Forné, who performed the Mass-Spectrometry experiments, and for the contribution of Dr. Gustavo Pereira de Almeida and Heike Bollig to the MORC3 study.
Moreover, I want to thank the whole Schotta group for the friendly and supportive atmosphere, in particular Dr. Dennis Sadic for his introduction to the ERV silencing project, and Dr. Filippo Cernilogar for his advice throughout the period of my doctoral project with a focus on NGS sequencing methods.
I am also grateful for the discussions with the BMC Computational Biology Club, especially to Dr. Tobias Straub and Dr. Tamas Schauer for their input and advice in terms of bioinformatic analyses.

Finally, credit goes to my family, whose constant support has made this work possible.