The role of Smarcad1 in retroviral repression in mouse embryonic stem cells

Igor Bren[1], Carmit Strauss[1], Sharon Schlesinger[1*]

[1]Department of animal science, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem.

*** Correspondence:** Dr. Sharon Schlesinger: sharon.shle@mail.huji.ac.il

Keywords: H3.3, Smarcad1, Trim28, Retrovirus**,** Embryonic stem cells, Epigenetic silencing

Running title: Smarcad1's Role in Suppressing Retroviruses

Abstract

The replication of Moloney murine leukemia virus (MLV) is suppressed in mouse embryonic stem cells (ESCs) by the Trim28-SETDB1 complex. The chromatin remodeler Smarcad1 interacts with Trim28 and was suggested to enable the deposition of histone variant H3.3. However, the role of Trim28, H3.3, and Smarcad1 in the repression of MLV in ESCs is not fully understood.

In this study, we investigate the role of Smarcad1 in the silencing of incoming MLV retroviruses, specifically focusing on the role of the chromatin structure. We found that Smarcad1 binds to proviral DNA and that Smarcad1 depletion leads to up-regulation of retroviral expression. Additionally, our findings demonstrate that Smarcad1 is necessary to properly recruit Trim28 and deposition of H3.3 on the MLV provirus and other Trim28-repressed genomic loci. Finally, we also show that the mutual depletion of Smarcad1 and Trim28 results in enhanced de-repression of the MLV provirus, indicating that these two proteins may also function independently to maintain repressive chromatin states.

Our results provide evidence for the crucial role of Smarcad1 in the silencing of retroviral elements in embryonic stem cells. Further research is needed to fully understand how Smarcad1 and Trim28 cooperate and their implications for gene expression and genomic stability.

Take home:

- Depletion of Smarcad1 impairs retroviral repression
- Smarcad1 is necessary for proper recruitment of Trim28 and H3.3 deposition of H3.3
- Depletion of both Smarcad1 and Trim28 results in an enhanced de-repression of the MLV provirus

Introduction

The replication of the murine leukemia virus (MLV) is restricted in mouse pluripotent cells, namely embryonic stem cells (ESC) (Barklis et al., 1986; Teich et al., 1977). Transcriptional silencing of the long-terminal repeat (LTR) viral promoter is facilitated by a silencing complex composed of multiple factors that work together to establish and maintain a repressive chromatin state in proviral DNA. Trim28 (tripartite motif-containing 28, also known as Kap1 or Tif1b) mediates the recruitment of chromatin modifiers to the proviral DNA (Wolf and Goff, 2007). Trim28 is recruited into the provirus by the zinc finger DNA binding protein ZFP809, which recognizes and binds a specific sequence in the provirus called the proline primer binding site (PBSpro)(Wolf and Goff, 2009). Once Trim28 is bound to the proviral DNA, it recruits several factors involved in transcriptional silencing and heterochromatin formation (Pannell et al., 2000), including H3K9 methyltransferase SETDB1 (ESET), which then methylate histone H3 lysine 9 (H3K9) at the proviral DNA, leading to the formation of heterochromatin (Matsui et al., 2010; Schultz et al., 2002). This methylation marks the provirus as a repressive chromatin state, preventing the transcriptional machinery from accessing the viral promoter in the LTR (Schlesinger and Goff, 2013). Other components of the silencing complex, such as the YY1 cofactor (Schlesinger et al., 2013) or the heterochromatin protein (Iyengar and Farnham, 2011; Sripathy et al., 2006; Wolf et al., 2008), bind to the viral DNA or chromatin and help maintain the repressive chromatin state, preventing the transcriptional machinery from accessing the proviral DNA. The expression of most endogenous retrovirus (ERV) repeats is also limited through a similar ensemble (Karimi et al., 2011; Matsui et al., 2010; Rowe et al., 2010). Interestingly, the histone 3 variant H3.3 was also suggested to play a role in establishing heterochromatin in ERV sequences, as H3.3 depletion leads to reduced H3K9me3 marking, resulting in depression of ERVs and adjacent genes (Elsasser et al., 2015; Navarro et al., 2020). Histone variant H3.3 exhibits a dynamic exchange with the soluble pool of nucleoplasmic histones (Ahmad and Henikoff, 2002; Loyola and Almouzni, 2007), which is enhanced in ESCs (Melcer and Meshorer, 2010; Meshorer and Misteli, 2006; Schlesinger et al., 2017). H3.3 is important for maintaining ESC pluripotency by regulating important gene expression programs for lineage specification (Goldberg et al., 2010; Ha et al., 2014; Wong et al., 2009). Following integration, hyperdynamic H3.3 exchange on the MLV provirus was observed, together with transcriptional repression, suggesting a functional role for H3.3 in the establishment and maintenance of the silencing (Ray-Gallet and Almouzni, 2021). Another important component of the silencing complex is the nucleosome remodeler Smarcad1.

Smarcad1 (SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1) is a member of the SWI/SNF family of ATP-dependent chromatin remodeling enzymes, required for genome integrity and the establishment of repressive chromatin (Rowbotham et al., 2011). Smarcad1 has been shown to induce nucleosome disassembly and reassembly, suggesting that it plays a role in the dynamic regulation of chromatin structure (Markert et al., 2021). Smarcad1 is also an important factor in the regulation of transposable elements (TE) and, more specifically, ERVs regulation in embryonic stem cells, as the depletion of Smarcad1 leads to their derepression (Sachs et al., 2019). Furthermore, Smarcad1 is known to interact directly with Trim28 through its CUE1 (CUE domain containing 1) and RBCC protein domains (ring finger and B-box type 2 and coiled coil domain), respectively (Ding et al., 2018). Furthermore, Smarcad1 was suggested to evict nucleosomes, generating accessible DNA necessary for proper recruitment of Trim28 and deposition of H3.3 in Trim28-repressed retroviral sequences (Navarro et al., 2020).

Here, we aimed to study the specific role of Smarcad1 in the silencing of incoming MLV in embryonic stem cells. We show that Smarcad1 is localized to the provirus after MLV infection and demonstrate its role in the establishment and maintenance of MLV repression in mouse embryonic stem cells. This is because proper recruitment of Trim28 into the provirus requires Smarcad1, as well as H3.3 deposition after MLV infection. Interestingly, mutual depletion of Smarcad1 and Trim28 results in enhanced MLV de-repression. These data suggest that Smarcad1 is closely intertwined in the mechanism by which the silencing complex represses the transcription of the provirus, possibly responsible for long-term silencing through dynamic regulation of the structure of the chromatin. However, further research is needed to fully understand the mechanisms and potential implications for other cellular processes such as gene expression and genomic stability.

Materials and methods

Cell lines and cell culture

HEK293T, a Human Embryonic Kidney-derived, and NIH3T3, a mouse embryonic fibroblast-derived immortalized cell lines were grown on non-coated 100mm x 20mm tissue culture plates (Corning, CLS430167). Their growth medium was composed of Dulbecco's modified Eagle medium high glucose (DMEM, BI, 01-055-1A), supplemented with Fetal Bovine Serum 10% (FBS, Rhenium, lots: 2289265 & 210430K), L-glutamine 2mM (BI, 03-020-1A), penicillin 100 units/ml (BI, 03-031-1B), and streptomycin 100 μg/ml (BI, 03-031-1B). The mouse embryonic stem cell (mESC) lines R1 and

KH2 were grown on feeder-free and gelatin-coated tissue culture plates. Gelatin was added to promote cell adhesion. Their growth medium was composed of DMEM high glucose, FBS 15%, L-glutamine 2mM, penicillin 100 units/ml, streptomycin 100 µg/ml, MEM non-essential amino acids 200mM (Rhenium, 11140-035), sodium pyruvate 1mM (Rhenium,11360039), and β-mercaptoethanol 0.12mM. In addition, this medium contained Leukemia inhibitory factor (LIF) 1000 units / ml, PD0325091 1µM (PeproTech, PD 0325901) and CHIR99021 3µM (PeproTech, 2520691). The combination of the last three components was termed 2i+LIF and was used to maintain naive pluripotency. KH2 cells express a single doxycycline-inducible H3.3 copy. This cassette also contains a gene that confers resistance to hygromycin. Therefore, to maintain this cassette in KH2 genome, their growth medium was supplemented with hygromycin B 140µg/ml (ENZO,ALX-380-309-G001). To induce H3.3-HA transcription, KH2 medium was supplemented with doxycycline 10µg/ml. For antibiotic selection, the mediums were supplemented with puromycin 2.5µg/ml or G418 [neomycin] 500µg/ml (Rhenium , 11811031). The puromycin and neomycin selection processes required 2 and 5 days to complete, respectively. All cells were grown in a humidified incubator at 37 ° C and CO2 5%. The cells were tested for mycoplasma (Hylabs, KI 5034I) fortnightly. Mycoplasma-negative cells were cryopreserved in liquid nitrogen to establish a working cell bank. If mycoplasma was detected, these cells were discarded, and a new mycoplasma-free vial was thawed.

Cell passaging, cryopreservation, and thawing

All cell lines were passaged every 3-4 days. Briefly, cells were washed with Phosphate buffered saline (PBS, BI, 02231A). Next, 2ml of trypsin EDTA solution C (BI, 03-053-1A) were added at a final concentration of 0.05% and 0.02% for trypsin and EDTA, respectively. The cells were trypsinized for 5 minutes at 37°C to generate single-cell suspensions. Then, trypsin activity was inhibited by adding a double volume (4ml) of complete medium to plates. Subsequently, 5µl of each cell suspension was mixed with 5µl trypan blue (T8154-100ML), and the combined volume of 10µl was loaded on a cell counting slide (BIO-RAD, #1450011). These slides were inserted into an automated cell counter (BIO-RAD, TC20) to count the cells in each sample. The number of cells seeded on new tissue culture plates following cell counts ranged between 100,000 to 500,000 depending on purpose. For cryopreservation, cells were trypsinized and counted as described above. Next, these cell suspensions were centrifuged (5 minutes, 400g) and the supernatants aspirated. The pellets were resuspended in a freezing medium composed of FBS and DMSO (Sigma, W387520) in a 9: 1 ratio. The cells resuspended in freezing medium were then transferred to cell-freezing tubes. Each

tube contained ~1,000,000 cells resuspended in 1ml of freezing medium. Finally, the tubes were placed in Mr. FrostyTM freezing containers (thermo Fisher Scientific, 5100-0001) and transferred to -80 ° C. 24 hours later, the tubes were transferred to liquid nitrogen for long-term storage. When necessary, the cells were thawed. Briefly, a vial stored in liquid nitrogen was transferred to a warm bath (37 ° C) for one minute. Next, the interior of this vial was transferred to a new 15ml tube containing 8ml of pre-warmed complete medium. The combined 9ml volume was well mixed and the tube was centrifuged (5 minutes, 400g). The supernatant was aspirated to eliminate DMSO remnants. These pellets were resuspended in complete medium and seeded in new tissue culture plates. Finally, these plates were transferred to the incubator. The water inside incubators and water baths was supplemented with Aquaguard-1 (BI, 01-867-1B) and Aquaguard-2 (BI, 01-916-1E), respectively, to prevent contamination.

shRNA design and cloning

pLKO.1 (Addgene plasmid #13425) is a 3rd generation lentiviral backbone used for cloning and expression of shRNA sequences in target cells. This vector contains a 4114bp sequence flanked by long terminal repeats (LTR). When transcribed, this 4114bp sequence serves as a genetically engineered lentiviral ssRNA genome that can be packaged in a lentiviral particle and used for the infection of target cells. This lentiviral genome contains a neomycin resistance gene driven by hPGK promoter, and a multiple cloning site (MSC) used for the insertion of shRNAs downstream of the U6 promoter. Following lentiviral infection and neomycin administration, the antibiotic resistance is used to purify all non-infected cells and confirm that only resistant and therefore shRNA-expressing cells remain in the population. The U6 promoter is specifically used to drive the transcription of shRNAs as it recruits RNA pol III. Although the transcription products of RNA pol II are G-capped, poly-adenylated, spliced, and finally translated by ribosomes, the RNA products transcribed by RNA pol III are processed differently. They fold and form a functional secondary structure. Therefore, following RNA pol III-mediated transcription, the shRNA forms a dsRNA that is processed by the RNAi machinery and leads to the degradation of a target endogenous mRNA. To confirm that our pLKO.1 vector contained all the sequences and restriction sites required for the insertion of shRNA, a 672bp region was amplified that included the U6 promoter and the MSC. Sanger sequencing of this amplicon confirmed the presence of all regions required for the cloning process. The vector pLKO.1 was simultaneously cleaved with the restriction enzymes AgeI (NEB, R3552S) and EcoRI (NEB, R3101S), generating a linearized vector and a ~24bp fragment representing the physical

distance between the two restrictions sites. Afterwards, the products were separated on 1% agarose gel and the linearized pLKO.1 vector was purified from the gel. Then, it was ligated with the ~54bp dsDNA fragments that served as shRNA by T4 ligase (NEB, M0202L) in a ratio of 3:1, insert:vector. shRNA sequences were taken from (ref to Jackline's paper from 2019). The ligation products were transformed to cells competent for DH5 E. coli (Invitrogen, 18265017) by heat shock and the cells were transferred to LB-containing tubes. These were transferred to the incubator for 1 hour recovery at 37°C, 180rpm. Next, the bacteria were seeded on agar-coated plates containing ampicillin 100 µg/ml (Sigma, A0166-25G), and the plates were moved to 37°C overnight. The next day, colonies were observed. These colonies were used as a template for PCR with primers that target the MSC. The PCR products were separated on agarose gel to examine shRNA insertion events relative to a purified and non-shRNA-containing pLKO.1 vector. Positive colonies were identified and seeded in 250 ml of liquid LB for overnight expansion at 37 ° C, 180 rpm. On the next day, the shRNA-containing pLKO.1 vectors were extracted by Midi Prep (Invitrogen, K210005) from the overnight DH5α cultures. Sanger sequencing of these amplicons confirmed the presence and successful insertion of the shRNAs into the pLKO.1 vectors.

Transfection and production of lentiviruses/retroviruses

HEK293T cells were used in this study as a factory for the production of lentiviruses and retroviruses. For this purpose, shRNA-containing pLKO.1 plasmids were co-transfected with psPAX2 (Addgene Plasmid #12260) and pMD2.G (Addgene Plasmid #12259) vectors to HEK293T cells. Briefly, when expressed, the psPAX2 backbone encodes structural capsid gag proteins, and the reverse transcriptase and integrase enzymes used to reverse transcription of the lentiviral ssRNA genome to dsDNA and its integration to the host cell genome, respectively. The pMD2.G backbone encodes lentiviral envelope proteins. When all three elements mentioned above, including the lentiviral genome encoded by the pLKO.1 vector, are simultaneously expressed, they are sufficient for the assembly of functional lentiviruses. Following assembly, the lentiviral particles exit the host cells and accumulate in the medium. This conditioned medium is then collected and used to infect target cells. In practice, 24 hours before transfection, 3 million HEK293T cells were seeded in a 100mm x 20mm tissue culture plate (Corning, CLS430167). The next day, in one tube, the shRNA-containing pLKO.1 vector (10µg) was mixed with psPAX2 (9µg) and pMD2.G (1µg). This mixture was supplemented with Opti-MEM (Gibco, 31985-047) to a final volume of 500µl. In a second tube, Opti-MEM was mixed with the transfection reagent polyethyleneimine (PEI) to a final concentration

of 40µg/ml and volume of 500µl. The mixtures in both tubes were then combined to a final volume of 1ml and incubated at room temperature for 20 minutes to allow the formation of PEI-DNA aggregates. In the meantime, the medium of HEK293T cells was replaced by 6ml of low-serum medium (DMEM, FBS 2%, L-glutamine 2mM, penicillin 100 units/ml, and streptomycin 100 µg/ml). After the 20-minute incubation was completed, the interior of the incubating tube (transfection reagent and the three plasmids) was added to the HEK293T cell plate. The plate was then gently rotated to allow an even distribution of the PEI-DNA aggregates. This process increases the probability of interaction between cells and aggregates, and therefore their internalization frequency. Finally, the cells were moved back to the incubator. On the next day, the medium was removed and substituted by 5.5ml of low-serum medium to wash the remaining PEI. From the next day and for the following 48 hours (48 hours after transfection and more), the conditioned medium containing functional lentiviruses was collected every 24 hours. Upon collection, it was filtered through a 0.45µm filter and supplemented with HEPES buffer 10mM (BI, 03-025-1B) and polybrene 12.5µg/ml (Merck, TR-1003-G) to adjust the pH and increase viruses' infection efficiency, respectively. Finally, it was aliquoted (750µl per vial), frozen and stored in -80°C. The production of retroviruses was very similar to that of lentiviruses. The only modification was the plasmids used for the transfection process. In this case, the plasmids pLKO.1, psPAX2 and pMD2.G plasmids were replaced by the plasmids PBS-P GFP (10µg), eco2-gag-pol (9µg), and VSVG (1µg) plasmids. The cotransfection of these vectors to HEK293T cells resulted in the accumulation of Murine Leukemia Viruses (MLV) in the medium. These retroviruses were harvested and stored like the lentiviruses.

Retroviral infection

To first estimate the infection efficiency of a batch of retroviruses and its multiplicity of infection (MOI), NIH3T3 cells were infected with the GFP reporter MLV. In practice, 24 hours prior to infection, NIH3T3 cells were trypsinized and counted as described above. 100,000 cells were seeded on a required number of wells of a 6-well plate. On the next day, a vial containing retroviruses was thawed at 37°C and diluted 1:3, 1:9, and 1:27. Then, the medium in which the cells were grown was discarded, and the dilutions were added for infection. 24 hours after infection, the conditioned medium was removed and replaced by a complete medium. 48 hours after the conditioned medium was changed, all cells were trypsinized and analyzed by flow cytometry. The measurements obtained were then placed on a graph where the GFP signal is located on the Y-axis and viral dilution on the X-axis. The curve was used to quantitatively estimate retroviral infection efficiency and determine

MOI. These data were obtained for each batch of retroviruses produced in this study. Based on these measurements, ESCs were infected with the GFP-reporter MLV as described above at MOIs that ranged between 0.5 and 1. For ESC infections, LIF was added to the conditioned medium to prevent differentiation caused by the absence of 2i+LIF conditions. All retroviral infections lasted 24 hours at 37 ° C. At the end of this period, the conditioned medium was replaced by a complete medium. ESCs infected with MLV were analyzed by flow cytometry two days after infection (2dai) and on. Upon infection, reverse transcription of the MLV ssRNA genome to dsDNA and its integration to cells' genome, the provirus encodes only GFP. Therefore, these proviruses are incapable of retrotransposition or propagation in the genome.

Lentiviral infection

To estimate the efficiency of a batch of lentiviruses, ESCs were infected as described above without prior dilutions. 24 hours prior to infection, the cells were trypsinized and counted. 100,000 cells were seeded on a required number of wells of a 6-well plate. On the next day, the medium in which the cells were grown was removed and substituted with 750µl of freshly-thawed lentivirus, supplemented with LIF. 24 hours after infection, the conditioned medium was discarded and replaced by a complete medium. 48 hours after infection, an antibiotic selection was applied as described in the Cell Lines and cell culture method. When the selection process was completed, RNA was extracted from the cells and used as a template for cDNA synthesis. This cDNA was analyzed by quantitative real-time PCR to examine whether the lentiviral infection resulted in depletion of a target endogenous mRNA. An efficient batch of lentiviruses was marked as ready-to-use, while inefficient batches were discarded. All lentiviral infections were 24-hour long at 37°C. Upon infection, reverse transcription of the lentiviral ssRNA genome into dsDNA, and its integration into the cell genome, the provirus encodes only shRNA and a gene that confers antibiotic resistance. Therefore, these proviruses are incapable of transposition or propagation in the genome.

RNA extraction, cDNA synthesis, and real-time quantitative PCR (qPCR)

Total RNA was extracted from cells using TRI reagent (Sigma, T9424) according to the manufacturer's instructions. RNA concentration and purity parameters (260/280 & 260/230) were determined by a NanoDrop spectrophotometer. 1µg RNA was reverse transcribed using the qScript cDNA Synthesis Kit (Quantabio, 95047-100). To control for genomic DNA impurities in the RNA samples, reactions without reverse transcriptase (-RT) were performed simultaneously. Following reverse transcription, the cDNA samples were diluted 1: 5 to a final volume of 100µl. Sample

preparation for real-time quantitative PCR (qPCR) was carried out in MicroAmpTM Fast Optical 96-Well Reaction Plates (Rhenium, 4346906). Each well contained a final reaction volume of 20µl: 3µl of diluted cDNA and 17µl of reaction mix consisting of a pair of primers (working stock concentration 10µM), Fast SYBR™ Green Master Mix (ABI, AB-4385614), and water. Housekeeping genes such as Ubiquitin C (UBC) or GAPDH were used to calculate the relative expression level of genes of interest by the ΔΔCT method. Each sample was tested as a triplicate. All primers used in this study for the quantification of relative gene expression have been previously tested and found agreeable with standard curve evaluation. In addition, all primers generated single-peak melting profiles. To further confirm that the measurements obtained by qPCR originate from mRNA and represent the true expression level of the examined genes, one of the primers in each pair targeted an exon-exon junction region that is not found in genomic DNA. All primers used in this study are listed in Supplementary Table S1. All qPCR reactions were performed on a StepOnePlus™ Real-Time PCR system. To examine the expression level of endogenous retroviruses (ERVs), the procedure described above was performed with a slight modification. In this case, the RNA was treated with TURBO™ DNase (2U/µL) (ThermoFisher, AM2238) prior to cDNA synthesis. -RT controls were included to all assays in which ERV expression level was examined.

Flow cytometry

In this study, cells were infected with GFP-reporter MLV retroviruses and then analyzed via flow cytometry to quantify the GFP level in a given sample. For this purpose, cells were trypsinized as described above and filtered through a 35µm nylon filter into BIOFLOAT™ 96-well U-bottom plates (Sarstedt, 82.1582). Flow analysis was performed using a CytoFLEX flow cytometer equipped with a 488 nm laser. A minimum of 100,000 cells were examined per sample. Data generated during the flow analysis were exported as FSC files that were further analyzed using FlowJo V.10 software. The FSC-A/SSC-A plot was used to determine the relevant cell population and exclude cell debris. Next, the FSC-A/FSC-H plot was used to exclude doublets and larger cell-aggregates. Finally, the FITC-A/PE-A plot was used to quantify the proportion of GFP-positive cells. Non-infected cells were first analyzed to set the GFP-positive gate.

Chromatin immunoprecipitation (ChIP)

The number of cells used for an immunoprecipitation ranged from 2 to 3 million.

Cells were trypsinized, filtered through a 70µm strainer, and counted as described in the Cell passaging, cryopreservation, and thawing method. Next, the cell suspensions were supplemented

with formaldehyde 1% (Sigma, F8775) for 10 minutes at room temperature to covalently cross-link DNA and proteins. The crosslinking process was quenched by adding a 120 mM glycine solution 120mM (Sigma, G8898). The samples were incubated at room temperature for 5 minutes and then centrifuged for 5 minutes (1,500rpm, 4°C). The supernatants were aspirated and the pellets were resuspended in 10 ml of cold PBS. The samples were centrifuged again under the same conditions, and the supernatants were aspirated. The pellets were resuspended in 2 ml of PBS supplemented with protease inhibitor 1:25 (PI, Sigma, 11836170001). Samples were centrifuged once more as described above and pellets were resuspended in 200µl ChIP lysis buffer, supplemented with PI (1:25). The samples were incubated on ice for 15 minutes and transferred to tubes suitable for sonication (Evergreen, 830-00249). The water in the sonicator chamber were pre-cooled to 4°C. Then, the samples were sonicated for ~22 minutes (Pulse: 15sec ON, 45sec Off, amplitude 70%) using a Qsonica Q800R2 sonicator. To confirm that chromatin was properly fragmented and a variety of DNA fragments suitable for immunoprecipitation (200 - 700bp) was generated, 10µl of sonicate were mixed with proteinase K 100µg/ml (PK, Sigma, P2308). These mixtures were transferred to a thermal cycler (50 ° C for 10 minutes followed by 65 ° C for 50 m minutes) to reverse the crosslinks. Next, these samples were stained with loading dye, loaded on 1% agarose gel without prior DNA extraction, and visualized via gel electrophoresis (110V, 30 minutes). If no additional sonication was required, the sonicates were centrifuged for 10 minutes (8,000rpm, 4°C). At this stage, 5% of each sample was collected and stored at 4°C overnight to be further used as input. The remaining supernatant (95%) was transferred to new tubes and diluted (1:10) in ChIP dilution buffer, supplemented with PI (1:25). 20µl Magna ChIP™ Protein A Magnetic Beads (Sigma, 16-661) and antibodies (1µg per 1 million-cell chromatin, all antibodies are listed in Supplementary Table S2) were added to tubes, which were then transferred to 4°C for an overnight incubation, shaking at 10 loops/minute. On the next day, the samples were immunoprecipitated using a DynaMag™ 2 magnetic rack (ThermoFisher) and washed with low-salt, high-salt, LiCl, and TE buffers by that specific order. All buffers were supplemented with PI (1:25). Each washing step included a 2-minute incubation on ice in the presence of buffer, followed by the subsequent immunoprecipitation. After the final washing step was completed, the samples were immunoprecipitated, the supernatants aspirated, and the pellets (as well as the input fractions) were resuspended in 100µl ChIP elution buffer. All samples (immunoprecipitated and input) were then transferred to 62°C for 6 hours, shaking at 300rpm to reverse the cross links. Finally, the samples were immunoprecipitated to

exclude the magnetic beads and the supernatants were used for extraction of the DNA fragments using QIAquick PCR Purification Kit (QIAGEN). Elution buffer volumes used for the immunoprecipitated and input fractions were 100µl and 200µl, respectively. These samples were then used for qPCR without any further processing to examine enrichment over input. The stock concentrations of all the reagents used in this assay and the composition of all ChIP buffers are described in Supplementary Table S3.

H3.3 turnover rate quantification

To examine the H3.3 turnover rate, KH2 cells were first infected with the MLV GFP-reporter as described in the retroviral infection method. Next, these cells were trypsinized and counted. 300,000 cells were seeded on a 100mm x 20mm tissue culture plate and kept for 3 days to reach high cell numbers. Then, the cells' growth medium was supplemented with doxycycline (DOX) 10µg/ml to activate the transcription of H3.3-HA for 1 or 8 hours. Subsequently, ChIP was performed as described above using an anti-HA antibody (Abcam, ab9110). The DNA fragments immunoprecipitated by the anti-HA antibody after 1 and 8 hours of DOX administration were purified and used as template for qPCR. The primers included in this assay targeted four different regions of the MLV provirus after integration into cells' genome (40nt, PBS, MLV TSS, and GFP). The qPCR values obtained after 1 hour of DOX treatment were normalized to those obtained after 8 hours of DOX treatment to quantify the percentage of saturation of the H3.3-HA signal observed within the first hour of DOX administration. The logic behind this approach is that genomic regions for which a high percentage is calculated exhibit rapid H3.3 turnover, while those for which a low percentage is calculated exhibit low H3.3 turnover.

Whole cell extraction and Western blotting (WB)

For whole cell extraction, the cells were trypsinized and counted as described above. Next, 10 million cells (per extract) were transferred to a new tube and centrifuged for 10 minutes (3,000g, 4°C). They were washed with ice-cold PBS and centrifuged once more under the same conditions. The supernatant was aspirated and the pellet was resuspended in hypotonic lysis buffer composed of ice cold Tris pH 7.4, EDTA 0.2mM, DTT 0.5mM, protease inhibitor (PI, 1:25) and $NaVO_4$ 1mM. These samples were incubated on ice for 10 minutes. Then, the cells within the samples were mechanically lysed by pipetting 10 times through a SGE autosampler syringes (Merck). The samples were incubated on ice for 15 minutes. Subsequently, a comparable volume of high salt buffer composed of Tris pH 7.4, EDTA 0.2mM, DTT 0.5mM, and NaCl 1M was added to the samples and they were

centrifuged for 30 minutes (13,000g, 4°C). The supernatants containing the extracts were collected and diluted 1:1 in SDS sample buffer. The pellets were discarded. The samples were denaturated at 95 ° C for 10 minutes and loaded on Bolt™ 4 to 12%, Bis-Tris, 1.0 mm, Mini Protein Gels (ThermoFisher, NW04120BOX). 20X Bolt™ MES SDS Running Buffer (ThermoFisher, B0002) was used for a high resolution separation of proteins smaller than 110 KDa, while 20X Bolt™ MOPS SDS Running Buffer (ThermoFisher, B0001) was used for a high-resolution separation of proteins larger than 110KDa. After separation, the proteins were transferred to a nitrocellulose membrane at 20V, in the presence of transfer buffer composed of Bolt™ Transfer Buffer (20X) 50ml/L, (ThermoFisher, BT0006), 100 ml / L methanol and water. The membrane was blocked for 25 minutes using 5% skim milk in TBST. Next, it was incubated in the presence of primary antibodies (listed in Table 2) and 1% skim milk in TBST at 4°C overnight. The next day, the membrane was washed three times with TBST, 15 minutes per wash. Then it was incubated in the presence of horseradish peroxidase (HRP)-conjugated secondary antibodies for one hour at room temperature. The membrane was washed three times under the same conditions. The detection was performed by Pierce™ ECL Western Blotting Substrate (ThermoFisher). Images were analyzed using Image-Lab software.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 9.5.0 software. Data are presented as mean values ± SEM. The number of biological replicates per experiment is provided in figure legends. Statistical significance was determined using the Student's t-test. And considered statistically significant at $p < 0.05$. Significance levels are $* \; p < 0.05$; $** \; p < 0.01$; $** * \; p < 0.001$.

Results

Smarcad1 is localized to the provirus after MLV infection.

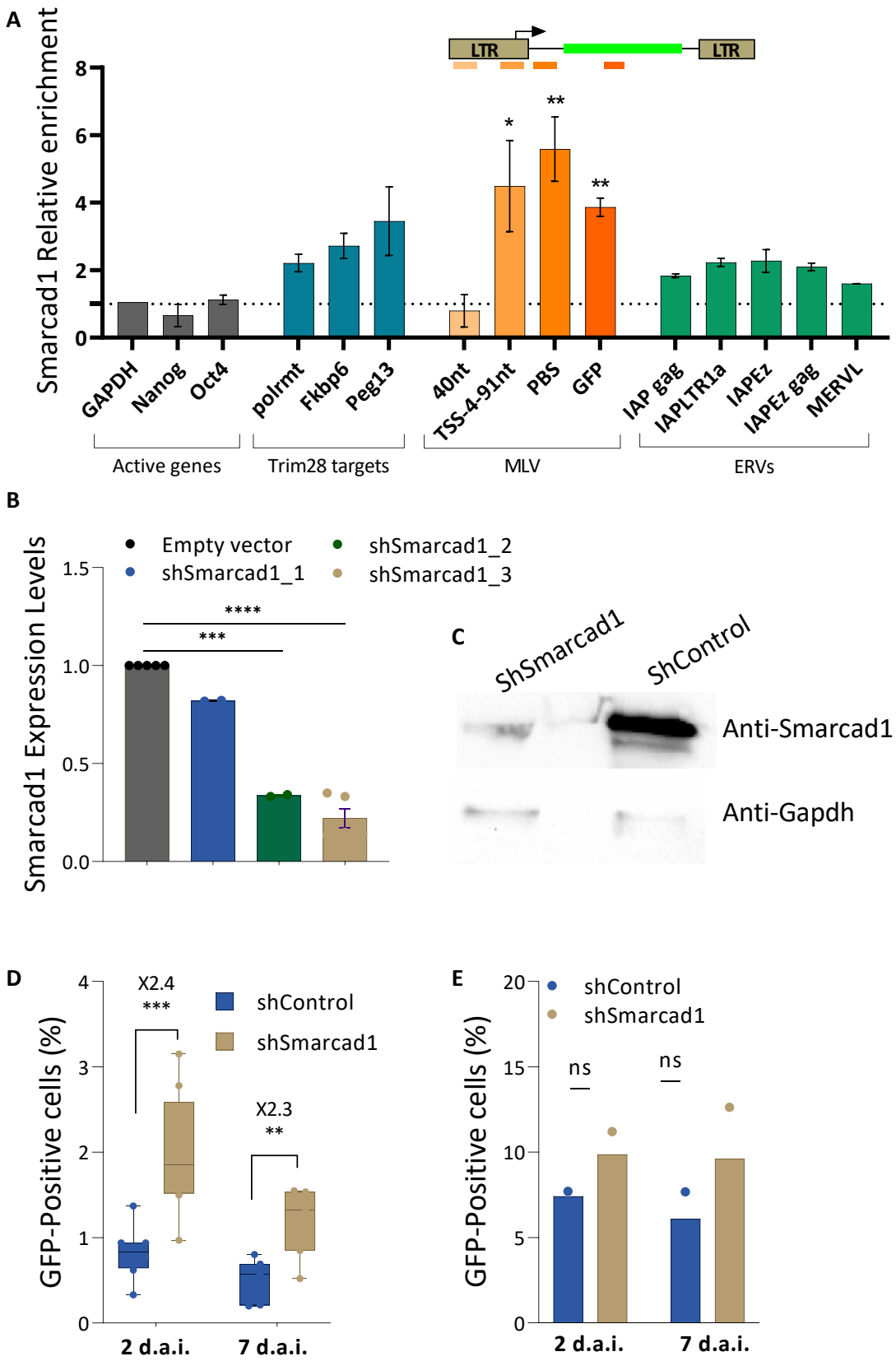
To examine the hypothesis that Smarcad1 plays a role in MLV repression, we transduced ESCs with MLV-like vector carrying GFP reporter (MLV-GFP) controlled by the LTR (as in (Schlesinger and Goff, 2013)). ChIP-qPCR performed two days after infection (2dai) showed eminent enrichment of Smarcad1 in some MLV regions: PBS, TSS, and coding regions (Figure 1A). Smarcad1 occupancy was not observed in the 5'UTR of the MLV LTR (the 40nt region), near the negative control region (NCR) of the provirus, suggesting that Smarcad1 is located around the PBS together with other members of the Trim28-dependent retroviral silencing complex. Importantly, Smarcad1 enrichment levels in MLV sequences were higher than those observed for ERV and other Trim28 genomic targets that

were previously shown to be bound by Smarcad1 (Ding et al., 2018; Sachs et al., 2019). Therefore, these data suggest that the recruitment of Smarcad1 to the provirus is mediated by Trim28.

Smarcad1 plays a role in the establishment and maintenance stages of MLV repression in mouse embryonic stem cells.

To examine whether Smarcad1 enrichment around the retroviral TSS has a functional meaning for the transcriptional regulation of the reporter, we depleted Smarcad1 using lentivirus-mediated delivery of short-hairpin RNAs (shRNAs sequences taken from (Sachs et al., 2019)). KD efficiency was assessed using RT-qPCR (Figure 1B) and verified using Western blot (Figure 1C and Supplementary Figure S1A). For further experiments, we used Sh_3 (ShSmarcad1) and sh_1 as control (ShControl) after we validated the stability of shRNA depletion and normal expression levels of pluripotency-related genes and other key factors related to retroviral epigenetic silencing (Supplementary Figure S1B). Next, we infected the two lines with MLV-GFP and followed the population fluorescence by flow cytometry two and seven days after infection (Figure 1D). Smarcad1 deletion results in 2.4-fold expression upregulation immediately after the proviral integration (2dai), which is maintained for several weeks (Supplementary figure S1C). The infection and integration efficiency were not affected by Smarcad1 depletion, as there is no change in the number of integrated genomic proviral copies (Supplementary figure S1D). Interestingly, no change in GFP expression is found in the less firmly repressed MLV virus, which carries the PBSgln and is therefore not as bound by Trim28 (Figure 1E). Therefore, these data suggest that Smarcad1 plays a role in the onset of MLV repression, probably through its binding to Trim28. To examine whether Smarcad1 is also required for maintaining MLV repression in ESCs, we first infected KH2 with MLV-GFP and only then depleted Smarcad1. A stable elevated GFP signal was observed following Smarcad1 depletion, demonstrating that it is also required for the maintenance stage of MLV repression (Supplementary figure S1E).

Figure 1. Smarcad1 is recruited to newly integrated retroviruses and plays a role in their repression. *(A) ChIP-qPCR was performed using Smarcad1 antibody, 2 days after MLV infection. Primers for open and repressed chromatin loci were used, values shown are % input normalized to non-enriched sites (Gapdh and Oct4). N=2, Data are the mean ± s.e.m, asterisk shows significant difference from the negative control genes (B) Smarcad1 expression levels were measured by RT-qPCR, normalized to UBC control gene. N=2-5, data are the mean ± s.e.m (C) Immunoblotting of ShSmarcad1 (Sh_3) and ShControl ESC extract using Smarcad1 antibody, with anti Gapdh as loading controls. (D) The percentage of GFP-positive cells after infection by PBSpro or (E) PBSgln virus in the wt and Smarcad1 depleted ESCs. N=2-6, data are the mean ± s.e.m. In all panels, P value was calculated using two-tailed unpaired Student's t-test, *P<0.05, **P<0.01, ***P<0.001.*


A
Smarcad1 Relative enrichment
LTR
LTR
GAPDH
Nanog
Oct4
polrmt
Fkbp6
Peg13
40nt
TSS-4-91nt
PBS
GFP
IAP gag
IAPLTR1a
IAPEz
IAPEz gag
MERVL
Active genes
Trim28 targets
MLV
ERVs
B
Smarcad1 Expression Levels
Empty vector
shSmarcad1_1
shSmarcad1_2
shSmarcad1_3
****
***
C
ShSmarcad1
ShControl
Anti-Smarcad1
Anti-Gapdh
D
GFP-Positive cells (%)
X2.4
***
X2.3
**
shControl
shSmarcad1
2 d.a.i.
7 d.a.i.
E
GFP-Positive cells (%)
shControl
shSmarcad1
ns
ns
2 d.a.i.
7 d.a.i.

Smarcad1 is required for proper Trim28 recruitment and H3.3 deposition in nucleosomes that wrap the MLV provirus after infection.

Smarcad1 is an ATP-dependent chromatin remodeling enzyme shown to regulate DNA accessibility at Trim28-binding ERV, possibly by nucleosome eviction (Navarro et al., 2020). Therefore, we hypothesized that in the absence of Smarcad1, the protein complex required for MLV repression cannot be properly recruited to or assembled at the provirus. To further decipher how Smarcad1 depletion disrupts retroviral repression, we first applied ChIP using Trim28 antibody on WT and Smarcad1 KD cells 2 and 7 days after infection. The enrichment of Trim28 in the PBS region of the MLV provirus immediately after infection was found to be independent of Smarcad1 (Figure 2A). However, 7 days after infection, Trim28 enrichment has increased significantly in the wt ESC but not in the Smarcad1 KD cells (Figure 2A). Therefore, Smarcad1 is not required for the initial recruitment to the provirus, but it is important for stabilizing and strengthening the binding properly over time. Genomic Trim28 targets, like the promoter region of Polrmt and Fkbp6 and the ERVK subfamily IAPEz, also lost Trim28 enrichment following Smarcad1 depletion (Figure 2B). Next, as the removal of Smarcad1 resulted in decreased accumulation of the histone variant H3.3 from IAPEz sequences (Navarro et al., 2020), we used KH2 ESCs (Beard et al., 2006) expressing a single copy of H3.3a-HA controlled by Doxycyclin (Dox) (Jin et al., 2009; Kraushaar et al., 2013; Yildirim et al., 2014) for our H3.3-HA ChIP assays. Exploring the dynamic accumulation of H3.3-HA in the proviral genome 48 hours after infection and 8 hours after Dox induction allowed us to focus on the onset of retroviral silencing immediately after integration. Surprisingly, no H3.3 enrichment was observed in the provirus (Figure 2C), although Trim28 is already bound there at that time point (Figure 2A). The accumulation of H3.3 is only observed several days after the integration of MLV and depends on Smarcad1 (Figure 2C). A similar effect of H3.3 eviction following Smarcad1 depletion was observed on Trim28 target sequences like Polrmt and imprinting control regions like Peg13, as well as the IAPEz subfamily (Figure 2D). As Trim28 binding is required for SetDB1 recruitment and the establishment of the H3K9me3 chromatin mark, similar effects are observed in anti-H3K9me3 ChIP for MLV provirus (Figure 2E) and genomic loci (Figure 2F). These data show that although Trim28 recruitment to the PBS is not dependent on Smarcad1, Smarcad1 is important for reinforcing Trim28 attachment to the area and is essential for H3.3 accumulation in retroviral sequences. These findings are also true for most ERV subfamilies examined here, in which Trim28 binding is barely

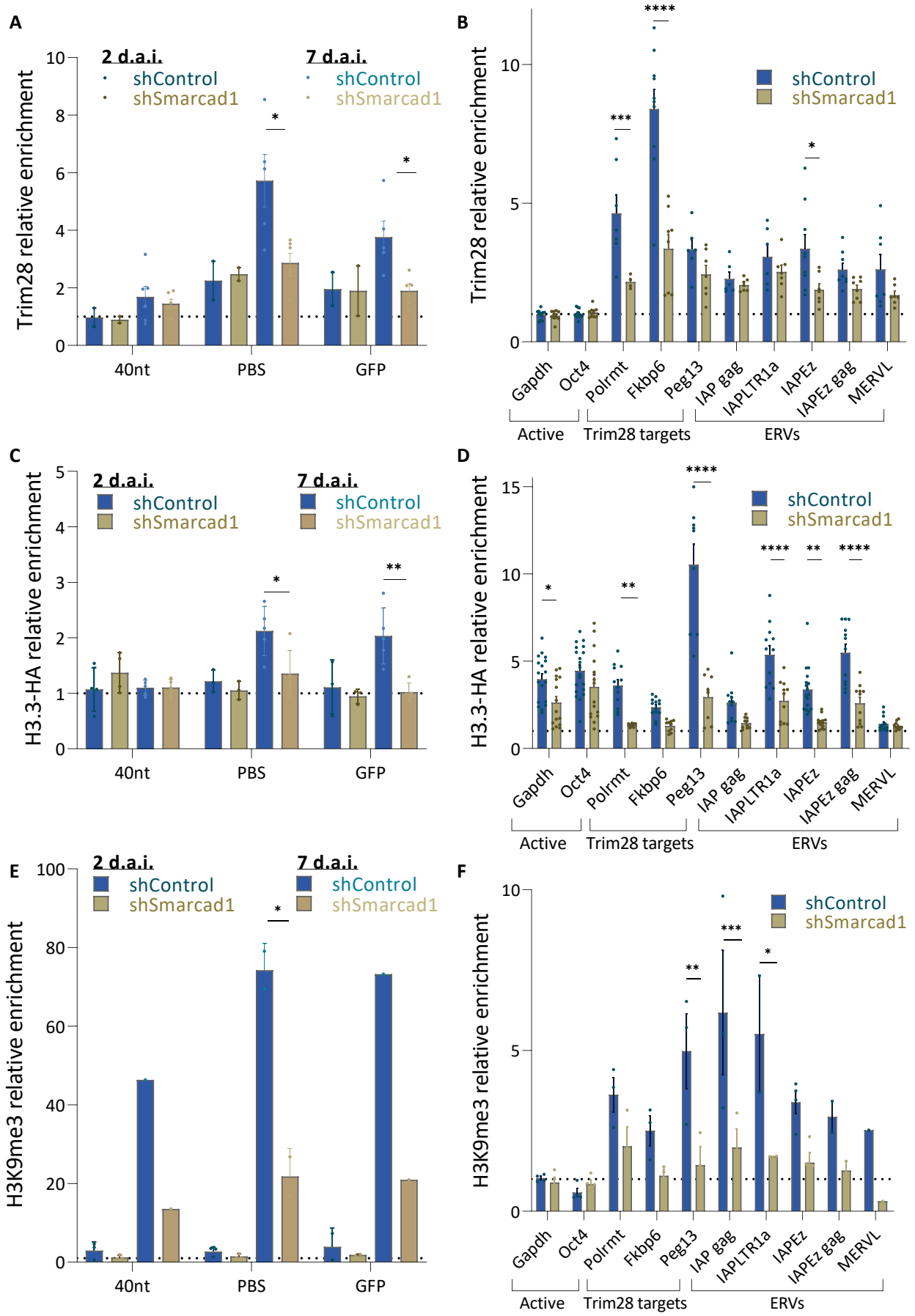

A
2 d.a.i.
shControl
shSmarcad1
7 d.a.i.
shControl
shSmarcad1
Trim28 relative enrichment
40nt
PBS
GFP
B
shControl
shSmarcad1
Trim28 relative enrichment
Gapdh
Oct4
Polrmt
Fkbp6
Peg13
IAP gag
IAPLTR1a
IAPEz
IAPEz gag
MERVL
Active
Trim28 targets
ERVs
C
2 d.a.i.
shControl
shSmarcad1
7 d.a.i.
shControl
shSmarcad1
H3.3-HA relative enrichment
40nt
PBS
GFP
D
shControl
shSmarcad1
H3.3-HA relative enrichment
Gapdh
Oct4
Polrmt
Fkbp6
Peg13
IAP gag
IAPLTR1a
IAPEz
IAPEz gag
MERVL
Active
Trim28 targets
ERVs
E
2 d.a.i.
shControl
shSmarcad1
7 d.a.i.
shControl
shSmarcad1
H3K9me3 relative enrichment
40nt
PBS
GFP
F
shControl
shSmarcad1
H3K9me3 relative enrichment
Gapdh
Oct4
Polrmt
Fkbp6
Peg13
IAP gag
IAPLTR1a
IAPEz
IAPEz gag
MERVL
Active
Trim28 targets
ERVs

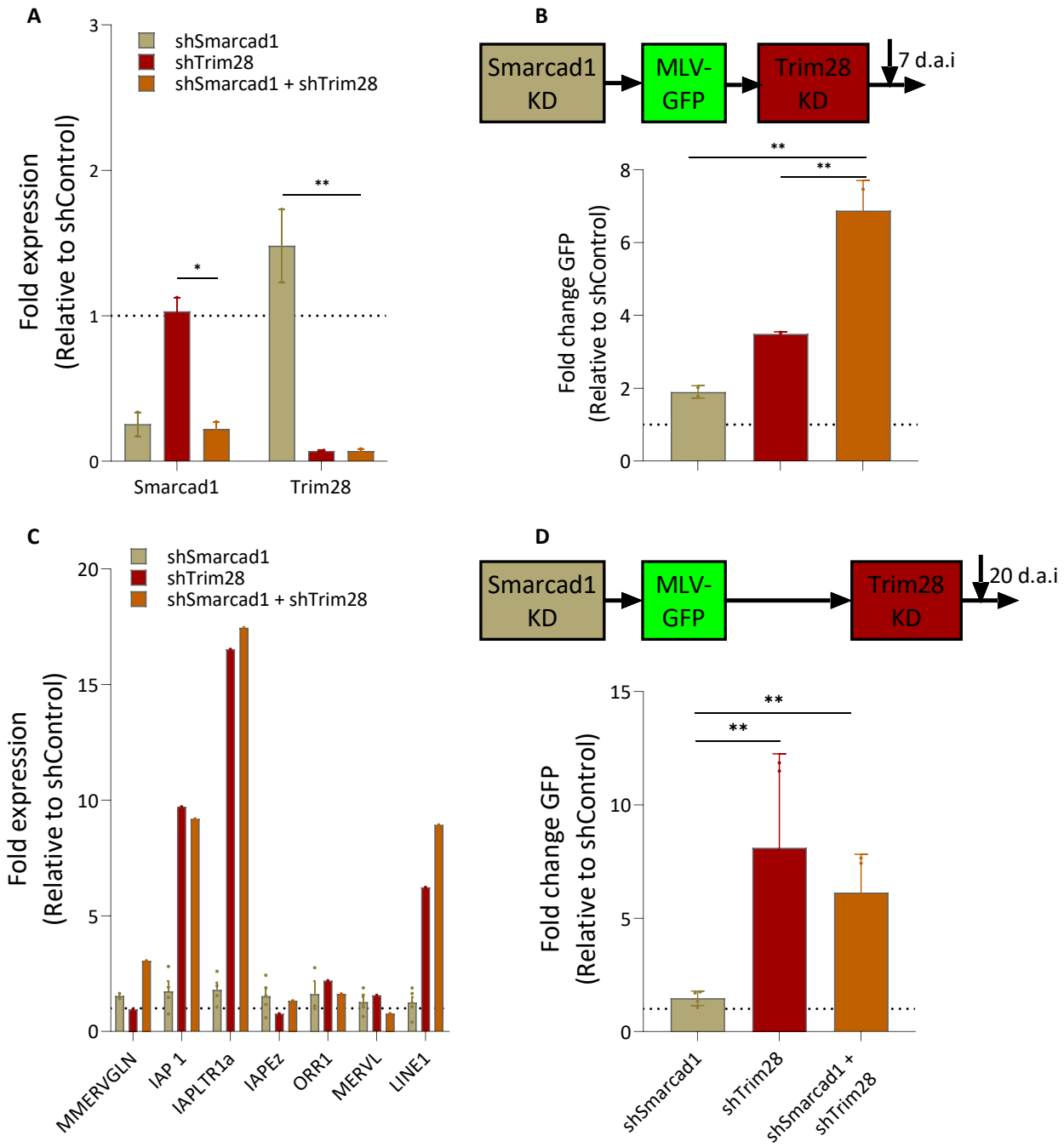
Figure 2. Smarcad1 reinforce Trim28 binding and allow H3.3 accumulation and heterochromatinization of proviral sequences. *(A) ChIP-qPCR was performed using Trim28 antibody at day 2 (N=2) and 7 (N=6) after MLV infection, with primers for MLV provirus specific regions and (B) for open and repressed genomic chromatin loci, including ERVs (N=12). Values shown are % input normalized to non-enriched sites (Gapdh and Oct4). (C) ChIP-qPCR was performed using HA antibody, at day 2 (N=3) and 7 (N=5) after MLV infection, with primers for MLV provirus specific regions and (D) for open and repressed genomic chromatin loci, including ERVs (N=12). Values shown are % input normalized to non-enriched sites (b-Globin). (E) ChIP-qPCR was performed using H3K9me3 antibody, at day 2 (N=2) and 7 (N=3) after MLV infection, with primers for MLV provirus specific regions and (F) for open and repressed genomic chromatin loci, including ERVs (N=2-4). Values shown are % input normalized to non-enriched sites (Gapdh and Aprt). All data are the mean ± s.e.m, in all panels P value was calculated using two-tailed unpaired Student's t-test, *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.*

affected by Smarcad1 depletion, while H3.3 enrichment and heterochromatinization are significantly reduced. Interestingly, the effect of Smarcad1 depletion on ERV class II expression is mild (Supplementary Figure S2A). This could be due to other epigenetic repression mechanisms applied to those sequences (Groh et al., 2021; Schlesinger et al., 2020).

Smarcad1 and Trim28 mutual depletion results in enhanced MLV de-repression.

Up to this point, we have established that Smarcad1 is required for the repression of MLV proviruses, selected class II ERVs and non-ERV Trim28-repressed genomic loci not repressed with Trim28 of ERV. Mechanically, we show that Smarcad1 improves the recruitment of Trim28 to these sites, allows the proper deposition of H3.3 into the nucleosomes that wrap DNA in these genomic regions and promotes their further marking with H3K9me3 repressive histone marks. To further explore the mechanism leading to MLV repression, we target Trim28 from the newly MLV-infected (2d.a.i.) Smarcad1 KD cells using Trim28 shRNA as in (Schlesinger et al., 2013) and 4 days of G418 selection, and verified the double KD by RT-qPCR at day 7 after MLV infection (Figure 3A). Importantly, double KD did not have an effect on the expression of other chromatin modifiers and key pluripotency genes (Supplementary Figure S2B). Next, we analyze the cells via flow cytometry for MLV-GFP expression. As expected, Smarcad1 KD and Trim28 KD resulted in a ~2-fold and ~4-fold increase in the GFP signal relative to shControl, respectively (Figure 3B). Interestingly, mutual depletion of both Smarcad1 and Trim28 resulted in a ~7-fold increase in GFP signal relative to shControl, demonstrating an additive effect. These observations were consistent in two independent biological replicates, suggesting that the mutual depletion of both regulators results in enhanced derepression of MLV. In addition, the change in expression level of selected ERVs was examined following the depletions of Smarcad1, Trim28 and the double KD. Up-regulation of some ERV was comparable

Figure 3. Smarcad1 and Trim28 depletion has an additive effect on MLV upregulation *(A) Double KD was achieved by infection with a lentiviral vector with Smarcad1 shRNA, puromycin selection and then Trim28 shRNA lentiviral vector with G418 selection. Depletion was verified by RT-qPCR, normalized to UBC control gene. Data are the mean ± s.e.m (n=2) (B) The percentage of GFP-positive cells 14 days after shSmarcad1 infection, 7 days after infection by MLV-GFP_PBSpro, and 5 days after shTrim28 infection. Data are the mean ± s.e.m (n=2) (C) Expression levels of selected ERVs following Smarcad1 and/or Trim28 depletion were measured by RT-qPCR. Data are the mean ± s.e.m (n=2-5) (D) The percentage of GFP-positive cells 27 days after shSmarcad1 infection, 20 days after infection by MLV-GFP_PBSpro and 5 days after shTrim28 infection. Data are the mean ± s.e.m (n=2). In all panels, the P value was calculated using a two-tailed unpaired Student's t-test, *P<0.05, **P<0.01.*



between shTrim28 and double KD cells, while others appeared to be more responsive to Smarcad1 depletion (Figure 3C). However, no additive effect is not observed for ERV repression. Therefore, we hypothesize a discrepancy in the role of Smarcad1 in silencing endogenous and exogenous retroviral sequences. To further test this hypothesis, we repeated the double KD assay in different order: first we depleted Smarcad1, second, we infected the cells with MLV and maintained them in culture for 16 days, third we depleted Trim28 and finally analyzed the cells via flow cytometry and qPCR after 4 days of selection. As before, we measured a ~2-fold and ~5-fold increase in thein GFP signal after depletions of Smarcad1 and Trim28, respectively (Figure 3D), except that this time the GFP signal observed in double KD cells was comparable to that of shTrim28. The corresponding RNA data confirmed Smarcad1 and Trim28 depletion (Supplementary Figure S2C). These observations suggest that the additive effect of mutual Smarcad1 and Trim28 depletion reduces over time and imply that the mutual role they play is more significant in the establishment of MLV repression. However, the decrease in Oct4 and Nanog levels in shTrim28 and double KD cells might suggest that these cells were in the process of losing pluripotency, a known effect of lengthy Trim28 depletion (Seki et al., 2010).

Discussion

We studied the role of Smarcad1 chromatin remodeler in retroviral repression in mouse ESC. The novelty of our approach is in the addition of the time dimension by following the transcriptional and epigenetic changes over time since the integration of the new proviral sequences into the genome. This, together with the ability to follow H3.3 accumulation and play with the timing of the depletion of two key factors, has allowed us to examine retroviral silencing by Smarcad1 from a whole new angel, focusing on change over time.

Our study has explored the Smarcad1 contribution of the chromatin remodeler to the repression of incoming retroviruses in mouse embryonic stem cells (ESCs). We focused on establishing the timeline of the transcriptional and epigenetic changes that occur after the integration of new proviral sequences into the genome. By doing so, we were able to gain new insights into the mechanisms of retroviral silencing by Smarcad1. We monitored the accumulation of H3.3, and controlled the timing of depletion of two key factors, to get a novel and comprehensive understanding of how the repression of retroviruses changes over time.

We begin at the point of MLV transduction to the cells. From this point, in about 24 hour, the newly incoming retroviral sequence is being integrated into the cells' genome (Goff, 2021). we examine the recruitment of Smarcad1 to the incoming MLV retroviral sequence. It was previously shown that binding of Smarcad1 and Trim28 to ERVs was not significantly altered in SetDB1 KD cells (Sachs et al., 2019), in agreement with our findings that these two factors already bound to the newly integrated provirus two days after infection (Figure 1A and 2A). It is interesting to note that both the variant H3K9me3 mark and the H3.3 are not enriched on viral chromatin at this stage yet (Figure 2B, C). Already in this early stage, immediately after integration, Smarcad1 is essential for complete silencing (Figure 1D), although Trim28 binding is not affected by its depletion (Figure 2A). Although retroviral silencing is already prominent (Figure 1D) and (Schlesinger and Goff, 2013), no enrichment of H3.3 and H3K9me3 has yet observed (Figure 2C, E). This is the first evidence that Smarcad1 and Trim28 might have an independent role in the retroviral repression. However, Trim28 seems to have a role in the recruitment of Smarcad1 to the proviral genome, as we do not detect the effect to Smarcad1 KD on the PBSgln provirus (Figure 1E), to which Trim28 is not bound at this stage (Schlesinger et al., 2013). Thus, the establishment of the silencing involves the binding of both Trim28 and Smarcad1 to the provirus, which then act both to repress its transcription.

Next, we examine the maintenance phase of the repression, 7 days after infection (similar results were also observed at later time points, data not shown). Here, the up-regulation after Smarcad1 depletion was stably maintained (supplementary Figure S1C). this effect was also evident when the Smarcad1 was depleted after the MLV integration (supplementary Figure S1E), as is the case for endogenous retroviruses (Supplementary Figure S2A and (Sachs et al., 2019)). Similarity in the chromatin structure of the proviral DNAs (Figure 2, all panels) indicate that similar epigenetic silencing mechanisms are applied to incoming viruses as well as to some ERVs. As we previously suggested (Schlesinger and Goff, 2013) we see that at this late time point after infection, the repression machinery overlaps for the silencing of exogenous and endogenous retroviruses. Another noteworthy observation is the correlation between Smarcad1 binding and H3.3 deposition on the provirus as well as some silenced genomic sites and ERVs. Once Smarcad1 is depleted H3.3 accumulation pause, in agreement with the idea suggested by (Navarro et al., 2020) that Smarcad1 nucleosome remodeling action is required for H3.3 deposition.

As both Trim28 and Smarcad1 binding is needed for H3.3 accumulation on retroviral sites, the question of their co-dependance strongly emerge. The ChIP described in Figures 2A and B is not

conclusive in this regard: although Smarcad1 KD cells show depleted Trim28 binding, Trim28 is still there, enriched over background more than 2 times. Similar results are presented by (Sachs et al., 2019) for IAPs, in which Smarcad1 depletion results in minor but evident Trim28 binding reduction. We hypothesize that Smarcad1 plays a role in the reinforcement and stabilization of Trim28 binding to the MLV provirus, as well as to genomic sites. To examine whether this effect on Trim28 binding is the only role of Smarcad1 in silencing, we sequentially depleted both proteins. Surprisingly, the MLV-GFP expression levels in the double KD cells were higher than in each KD (Figure 3A), indicating an additive effect of those factor. Therefore, the role of Smarcad1 in the silencing, at least in the early stages, is not only to support Trim28 binding. This is not true for ERVs and proviral silencing maintenance (Figure 3C, D), supporting the idea that at the determining point of silencing establishment Smarcad1 plays a Trim28 independent role.

Our findings suggest that Smarcad1 plays a critical role in the establishment and maintenance of MLV repression, as well as other Trim28-targeted genomic loci. The study found that Smarcad1 is recruited to the MLV provirus and that its presence around the provirus is required for proper Trim28 binding and H3.3 deposition to the provirus. The study further showed that Smarcad1 depletion leads to a 2.4-fold upregulation of GFP expression immediately after proviral integration and that Smarcad1 is also required for maintaining the repression. Furthermore, the results indicate that Smarcad1 is important for stabilizing and strengthening the binding of Trim28 to the provirus over time. The findings suggest that Smarcad1 is an important player in the regulation of MLV repression in ESCs and provide insights into the mechanisms underlying retroviral epigenetic silencing.

In conclusion, the study provides evidence that Smarcad1 is a critical component in the repression of MLV in mouse embryonic stem cells. Smarcad1 plays a role in both the onset and the maintenance of MLV repression, and its depletion leads to increased expression of the MLV-GFP reporter, suggesting that Smarcad1 is necessary for proper MLV repression in embryonic stem cells.

Declarations

Acknowledgments

We thank Dr. Jacqueline Mermoud for help with the design and cloning of the Smarcad1 shRNA, for ChIP advice, and for many insightful and encouraging discussions along the way. We wish to thank Prof. Eran Meshorer for providing the H3.3-HA KH2 cell line. We thank all members of the Schlesinger lab for helpful discussions and support.

Ethical Approval

Not applicable

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

I.B. and C.S conducted the research. Conceptualization was done by I.B. and S.S.; Methodology was developed by I.B. and C.S.; I.B. drafted the figures; S.S. wrote the main manuscript text and prepared the figures; S.S. was responsible for the acquisition of the financial support, management and coordination of the research. All authors reviewed the manuscript.

Funding

This work was funded by Israel Science Foundation grant number 761/17

Availability of data and materials

Not applicable

References

Ahmad, K., and Henikoff, S. (2002). Epigenetic consequences of nucleosome dynamics. Cell *111*, 281-284.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391-399.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. genesis *44*, 23-28.

Ding, D., Bergmaier, P., Sachs, P., Klangwart, M., Ruckert, T., Bartels, N., Demmers, J., Dekker, M., Poot, R.A., and Mermoud, J.E. (2018). The CUE1 domain of the SNF2-like chromatin remodeler SMARCAD1 mediates its association with KRAB-associated protein 1 (KAP1) and KAP1 target genes. Journal of Biological Chemistry *293*, 2711-2724.

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*, 240-244.

Goff, S.P. (2021). Silencing of Unintegrated Retroviral DNAs. Viruses *13*.

Goldberg, A.D., Banaszynski, L.A., Noh, K.-M.M., Lewis, P.W., Elsaesser, S.J., 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.

Groh, S., Milton, A.V., Marinelli, L.K., Sickinger, C.V., Russo, A., Bollig, H., de Almeida, G.P., Schmidt, A., Forne, I., Imhof, A.*, et al.* (2021). Morc3 silences endogenous retroviruses by enabling Daxx-mediated histone H3.3 incorporation. Nat Commun *12*, 5996.

Ha, M., Kraushaar, D.C., and Zhao, K. (2014). Genome-wide analysis of H3.3 dissociation reveals high nucleosome turnover at distal regulatory regions of embryonic stem cells. Epigenetics & chromatin *7*, 38.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. J Biol Chem *286*, 26267-26276.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet *41*, 941-945.

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

Kraushaar, D.C., Jin, W., Maunakea, A., Abraham, B., Ha, M., and Zhao, K. (2013). Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3.3. Genome Biol *14*, R121.

Loyola, A., and Almouzni, G. (2007). Marking histone H3 variants: how, when and why? Trends Biochem Sci *32*, 425-433.

Markert, J., Zhou, K.D., and Luger, K. (2021). SMARCAD1 is an ATP-dependent histone octamer exchange factor with de novo nucleosome assembly activity. Science Advances *7*.

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

Melcer, S., and Meshorer, E. (2010). Chromatin plasticity in pluripotent cells. Essays Biochem *48*, 245-262.

Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. Nature reviews Molecular cell biology *7*, 540-546.

Navarro, C., Lyu, J., Katsori, A.M., Caridha, R., and Elsasser, S.J. (2020). An embryonic stem cell-specific heterochromatin state promotes core histone exchange in the absence of DNA accessibility. Nat Commun *11*, 5095.
Pannell, D., Osborne, C., Yao, S., Sukonnik, T., Pasceri, P., Karaiskakis, A., Okano, M., Li, E., Lipshitz, H., and Ellis, J. (2000). Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. EMBO J *19*, 5884-5894.
Ray-Gallet, D., and Almouzni, G. (2021). The Histone H3 Family and Its Deposition Pathways. Adv Exp Med Biol *1283*, 17-42.
Rowbotham, S.P., Barki, L., Neves-Costa, A., Santos, F., Dean, W., Hawkes, N., Choudhary, P., Will, W.R., Webster, J., Oxley, D.*, et al.* (2011). Maintenance of Silent Chromatin through Replication Requires SWI/SNF-like Chromatin Remodeler SMARCAD1. Mol Cell *42*, 285-296.
Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J.*, et al.* (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463*, 237-240.
Sachs, P., Ding, D., Bergmaier, P., Lamp, B., Schlagheck, C., Finkernagel, F., Nist, A., Stiewe, T., and Mermoud, J.E. (2019). SMARCAD1 ATPase activity is required to silence endogenous retroviruses in embryonic stem cells. Nature Communications *10*, 1-16.
Schlesinger, S., Branco, M., and Mermoud, J. (2020). Roles and regulation of endogenous retroviruses in pluripotency and early development. In Stem Cell Epigenetics, T.G. Meshorer E, ed. (Academic Press), p. 464.
Schlesinger, S., and Goff, S.P. (2013). Silencing of proviruses in embryonic cells: efficiency, stability and chromatin modifications. EMBO reports *14*, 73-79.
Schlesinger, S., Kaffe, B., Melcer, S., Aguilera, J.D., Sivaraman, D.M., Kaplan, T., and Meshorer, E. (2017). A hyperdynamic H3.3 nucleosome marks promoter regions in pluripotent embryonic stem cells. Nucleic Acids Res *45*, 12181-12194.
Schlesinger, S., Lee, A., Wang, G., Green, L., and Goff, S. (2013). Proviral silencing in embryonic cells is regulated by yin yang 1. Cell reports *4*, 50-58.
Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes & Development *16*, 919-932.
Seki, Y., Kurisaki, A., Watanabe-Susaki, K., Nakajima, Y., Nakanishi, M., Arai, Y., Shiota, K., Sugino, H., and Asashima, M. (2010). TIF1 beta regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner (vol 107, pg 10926, 2010). Proc Natl Acad Sci USA *107*, 13556-13556.
Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol Cell Biol *26*, 8623-8638.
Teich, N.M., Weiss, R.A., Martin, G.R., and Lowy, D.R. (1977). Virus infection of murine teratocarcinoma stem cell lines. Cell *12*, 973-982.
Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. Journal of Virology *82*, 4675-4679.
Wolf, D., and Goff, S. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46-57.
Wolf, D., and Goff, S.P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature *458*, 1201-1204.
Wong, L.H., Ren, H., Williams, E., McGhie, J., Ahn, S., Sim, M., Tam, A., Earle, E., Anderson, M.A.,

Mann, J., *et al.* (2009). Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. Genome Res *19*, 404-414.
Yildirim, O., Hung, J.H., Cedeno, R.J., Weng, Z., Lengner, C.J., and Rando, O.J. (2014). A system for genome-wide histone variant dynamics in ES cells reveals dynamic MacroH2A2 replacement at promoters. PLoS Genet *10*, e1004515.