De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET

Helen M. Rowe1, Marc Friedli1, Sandra Offner1, Sonia Verp1, Daniel Mesnard1, Julien Marquis1, Tugce Aktas2 and Didier Trono1,*

SUMMARY
Endogenous retroviruses (ERVs) undergo de novo DNA methylation during the first few days of mammalian embryogenesis, although the factors that control the targeting of this process are largely unknown. We asked whether KAP1 (KRAB-associated protein 1) is involved in this mechanism because of its previously defined role in maintaining the silencing of ERVs through the histone methyltransferase ESET and histone H3 lysine 9 trimethylation. Here, we demonstrate that introduced ERV sequences are sufficient to direct rapid de novo methylation of a flanked promoter in embryonic stem (ES) cells. This mechanism requires the presence of an ERV sequence-recognizing KRAB zinc-finger protein (ZFP) and both KAP1 and ESET. Furthermore, this process can also take place on a strong cellular promoter and leads to methylation signatures that are subsequently maintained in vivo throughout embryogenesis. Finally, we show that methylation of ERVs residing in the genome is affected by knockout of KAP1 in early embryos. KRAB-ZFPs, KAP1 and ESET are thus likely to be responsible for the early embryonic instatement of stable epigenetic marks at ERV-containing loci.

KEY WORDS: De novo DNA methylation, Endogenous retroviral silencing, KRAB-associated protein 1 (KAP1), TRIM28, TIF1β, KRAB zinc-finger protein (KRAB-ZFP), ESET (SETDB1)

INTRODUCTION
Endogenous retroviruses (ERVs) account for close to 10% of mammalian genomes (Waterston et al., 2002) and are both drivers of evolution and threats to genetic integrity because of their ability, first, to retrotranspose and, second, to alter the expression of neighbouring genes through cis-acting transcriptional influences. In mice, ERVs display residual retrotransposition activity, leading to polymorphic integrations and differential gene regulation between mouse strains (Qin et al., 2010; Takabatake et al., 2008; Zhang et al., 2008). Moreover, around 10% of spontaneous mutations in inbred mice are linked to ERVs (Maksakova et al., 2006), for example dactylaplasia that results from a MusD neo-insertion (Friedli et al., 2008; Kano et al., 2007). In humans, the non-LTR retroelement LINE1 (L1) retains some activity and, remarkably, individuals with Rett syndrome caused by mutations in the DNA methylation-binding protein MeCP2 are more prone to L1 retrotransposition (Muotri et al., 2010).

ERVs are duly inactivated during early embryogenesis by histone modifications and de novo DNA methylation (reviewed by Rowe and Trono, 2011). This process counteracts the prior genome-wide erasure of DNA methylation that begins at the zygote phase, leaving transposons in a low methylated state, including L1s and, to a lesser extent, intracisternal A-type particles (IAPs) (Feng et al., 2010; Lane et al., 2003). De novo DNA methylation is key to ERV repression from plants to mammals, and involves methylation of cytosine residues at CpG dinucleotides by the enzymes DNMT3A and DNMT3B, which act in conjunction with their catalytically inactive co-factor DNMT3L (Chedin et al., 2002; Okano et al., 1999; Suetake et al., 2004). Later in development and in adult tissues, levels of de novo methyltransferases are reduced compared with those in pre-implantation embryos (Carlson et al., 1992), yet pre-established DNA methylation patterns are perpetuated during DNA replication by the maintenance DNA methyltransferase DNMT1, which is recruited to hemi-methylated DNA by UHRF1 (Bestor et al., 1988; Bostick et al., 2007; Sharif et al., 2007). DNA methylation is particularly crucial for constraining the transcription of some ERVs, including IAPs, the expression of which become uncontrollable following inactivation of DNA methylation machinery (Bourc’his and Bestor, 2004; Chen et al., 2003; Gaudet et al., 2004; Walsh et al., 1998). However, IAPs are still silenced in DNMT1 knockout ES cells unless they are differentiated (Hutnick et al., 2010), highlighting the particular importance of DNA methylation later in development. Interestingly, alterations in ERV DNA methylation patterns impact on the expression of neighbouring genes (Duhl et al., 1994; Macfarlan et al., 2011; Macfarlan et al., 2012; Michaud et al., 1994; Rebollo et al., 2011).

DNA methylation is conditioned by the density and spacing of CpG dinucleotides, can be affected by DNA-binding proteins such as CTCF and REST, and is influenced by the histone code, with methylation of histone 3 at lysine 4 (H3K4) preventing the docking of the DNMT3A-DNMT3L complex (Glass et al., 2009; Jia et al., 2007; Lienert et al., 2011; Okitsu and Hsieh, 2007; Ooi et al., 2007; Stadler et al., 2011; Weber et al., 2007). The factors and pathways that recruit de novo DNA methyltransferases to specific genomic targets, notably ERVs, remain largely undefined, although it has recently been shown that DNMT3L and the lysine methyltransferase G9a are factors required for the initiation of proviral de novo DNA methylation (Leung et al., 2011; Ooi et al., 2010).

Here, we have investigated the role played in this process by KAP1 (KRAB-associated protein 1, also known as tripartite motif finger protein (KRAB-ZFP), ESET (SETDB1))
MATERIALS AND METHODS

Lentiviral vectors

All MND vectors have been described previously (Rowe et al., 2010). For ERV-directed repression of a cellular promoter, the transfer vector pRRLSIN.cPPT.PK.GFP.WPRE (available from Addgene) was employed, which contains a human phosphoglycerate kinase 1 (PGK) promoter to distinguish it from the endogenous mouse promoter during methylation analysis. Pro and B2 sequences were cloned into this vector by annealing primers upstream (into the XhoI site) or downstream (between the BamHI and AgeI sites) of the promoter in the antisense orientation. Or IAP1 or IAP4 [described previously (Rowe et al., 2010)] were cloned upstream of the promoter into the XhoI site in the antisense orientation. Cloning was verified by sequencing (Microsynth). Vectors were produced by transient transfection of 293T cells with the transfer vector, packaging and VSVG envelope plasmids (Barde et al., 2010), and titrated on T3T fibroblasts.

Cell culture and flow cytometry

Mouse ES cells were cultured in Glasgow Minimum Essential Media (GMEM, Sigma: G5154) with sodium pyruvate (1 mM, Sigma S8636), MEM non-essential amino acids (1×, Gibco 11140-035), L-glutamine (2 mM; Gibco 25300-024), 2-mercaptoethanol (0.1 mM, Sigma), ES cell tested FBS (Gibco 16141-079) and leukemia inhibitory factor (LIF, 50 units/500 ml; Chemicon ESG11107). Cells were grown on 0.2% gelatin (Sigma: 48723-500G-F)-coated plates and split every 2 days. ES cell lines used were used were two KAP1loxP/loxP lines called ES3 and ES6, and their derived KAP1-conditioned knockout cell lines transduced with tamoxifen (4-OHT)-inducible cre (Rowe et al., 2010). Knockout cells were collected 3-4 days after treatment with 4-OHT (overnight at 1 μM; Sigma H7904) as stated. G9a parental or stable knockout ES cells and Esets-inducible knockout ES cells were from Yoichi Shinkai (Dong et al., 2008; Matsui et al., 2010; Tachibana et al., 2008) (The RIKEN Institute, Japan). Dnmt1−/−, Dnmt3a−/−, Dnmt3b−/− (triple knockout or TKO) ES cells were from Masaki Okano (Tsumura et al., 2006) (The RIKEN Institute, Japan). F9 EC (embryonic carcinoma) cells, primary mouse embryonic fibroblasts (MEFs), T3T fibroblasts and 293T cells were also used where stated. Vector titres and GFP repression were measured by flow cytometry and for KAP1 and ESET knockout experiments, cells were stained with an anti-SSEA-1 PE-conjugated antibody or isotype control (BD Pharmingen: 560142 and 555584) and SSEA-1 high-expressing cells gated for undifferentiated cells.

Quantitative bisulfite pyrosequencing

Genomic DNA was converted (200 ng/sample) using an Epitexit Bisulfite kit (Qiagen: 59104) and used for PCR (primers, one tagged with biotin, were designed using PyroMark Assay Design Software 2.0). PCR products were checked by Sanger sequencing and verified on agarose gels for each experiment before immobilizing on 96-well plates using a Vacuum Prep Workstation and pyrosequencing using PyroMark Gold Reagents (Qiagen 972804; Center for Integrative Genomics, University of Lausanne, Switzerland). Results were analysed using Pyro Q-CpG Software. Primer sequences are in supplementary material Table S1.

Immunoblotting

Cells were washed with ice-cold PBS and resuspended in radioimmunoprecipitation buffer to prepare total cell extracts. Protein was quantified by BCA Protein Assay Reagents (Pierce) and normalized for loading on a 10% denaturing SDS-polyacrylamide gel. Wet transfer was performed and primary antibodies used were: anti-KAP1 (Mouse mAb, MAB3662, Chemicon), anti-IAP GAG [a kind gift from Bryan Cullen, Duke University, NC, USA (Bogerd et al., 2006)] and anti-PCNA (mouse mAb, clone PC10, catalog number NA03, Calbiochem).

293T cell transfection

Cells were seeded at 1×105 per well in a 12-well plate. Then, 100 μl transfection mix was prepared per well [containing 6 μl FuGene 6 (Promega) and 1 μg DNA in serum-free media] and incubated at room temperature for 15 minutes before adding drop-wise to cells.

Quantification of RNA and DNA copy number

RNA was extracted and qRT-PCR performed as described previously (Rowe et al., 2010). Primer specificity was confirmed by dissociation curves and samples normalized to Titin, Gapdh or Eef1a1 in Fig. 2D because it reacts with both human and mouse Eef1a1. For in vivo experiments, genomic DNA was extracted from embryos to measure DNA copy numbers by performing Taqman qPCR for HIV Gag, GFP and Titin as a normalizer, in comparison with a titration curve of plasmid containing sequences of HIV Gag, GFP and Titin. A cell line with a known number of vector copies was used as a control. Results presented are means of
values obtained for HIV Gag and GFP primers. See supplementary material Table S1 for primer sequences.

**Lentiviral transgenesis**

Lentiviral vectors for transgenesis were prepared using Episerf medium (Invitrogen 10732-022); the particle concentration was obtained by p24 ELISA (Perkin-Elmer: NEK050B001KT) and the infectious titre determined on HCT116 cells using GFP flow cytometry. Vectors for each transgenesis experiment were produced and titred in parallel in order to be in the same range: Fig. 4 titres, 2.9 x 10^9 (IAP2 MND) and 4.2 x 10^9 (PGK); supplementary material Fig. S4B titres, 2.3 x 10^9 (IAP4 PGK) and 2.4 x 10^9 (IAP1 PGK).

Transgenesis was performed by perivitelline injection of vectors into fertilized oocytes that were transferred to foster mothers (strain B6D2F1/J) and then recovered at E13. Photographs were taken using the same saturation, gain and exposure settings for all embryos, and GFP results displayed all represent the same image processing settings.

**Statistics**

GraphPad Prism version 4.00 (www.graphpad.com) was used for all statistical analyses. For all DNA methylation analyses, two groups (as detailed in the figure legends) were compared across multiple CpG positions using paired two-tailed t-tests (except where stated). Any additional statistical analyses are specified in the figure legends.

**RESULTS**

**Introduced ERV sequences can induce rapid de novo DNA methylation in ES cells**

As repeat-derived sequences are largely methylated in ES cells, we set up a method to measure the de novo DNA methylation directed at ERV-specific elements by cloning these into lentiviral vectors, a gene transfer system itself not prone to spontaneous methylation, compared with MLV-based vectors (Wang et al., 1998). To quantify levels of de novo DNA methylation, we used bisulfite pyrosequencing and verified this technique with primers for endogenous ERVs and the Oct4 promoter; **P = 0.0022. Points show individual CpG positions. (B) ERV-directed repression of a GFP reporter in ES cells. (Left) Flow cytometry 4 days post-transduction. All vectors had an internal ‘MND’ promoter (Rowe et al., 2010), which was empty or contained an ERV sequence (Pro, B2, IAP1, IAP2, IAP3 or IAP4) either upstream (when written before MND) or downstream (when written after MND) of the promoter. Both E14 ES and ES3 ES cell lines showed significant ERV repression (**P = 0.0260 and ***P = 0.0311, respectively, unpaired one-tailed t-tests) compared with control vectors (green). Repression is normalized to the MND B2 control vector and to expression in 3T3 control cells. (Right) GFP mRNA quantification (normalized to titin) for one representative cell line 6 days post-transduction; repressed vectors showed significantly less mRNA than controls (green), *P = 0.0042 (unpaired one-tailed t-test). (C) Day 4 analysis of de novo methylation at the MND promoter. n=10 (CpG positions – see map). Here, the methylation profile is also shown for the MND Pro vector in 3T3 control cells. Significant differences in ES cells: the MND Pro versus MND B2 vector (**P = 0.0006); IAP2 MND (**P = 0.001) or IAP3 MND or MND IAP2 or MND IAP4 (**P ≤ 0.0001 for all) all versus the IAP1 MND control vector.
A time-course study showed that this process was already under way by day 3 and then progressively increased, with mean levels across the promoter reaching around 75% at day 7 (supplementary material Fig. S1B). Control vectors, whether they contained no KAP1-targeted retroviral sequence, the point-mutated B2 PBS Pro derivative or the previously identified KAP1-insensitive IAP1 element (Rowe et al., 2010), underwent neither repression nor DNA methylation. Moreover, in 3T3 fibroblasts, no ERV-directed promoter methylation was detected, suggesting this process may be restricted to ES cells (Fig. 1C).

**ERV DNA methylation patterns are conditioned by KRAB-ZFP expression profiles**

KAP1 has previously been shown to repress ERVs in ES cells and MEFs based on mRNA-sequencing data. RPKM indicates the number of sequencing reads normalized by gene length and to the total read number. The Pro sequence is repressed in MEFs where ZFP809 expression is retained. The MND Pro vector is repressed in MEFs compared with the empty (MND) or IAP2 MND vector, whereas all vectors are highly expressed in 3T3 cells (bottom). Mock indicates nontransduced. The Pro sequence directs de novo promoter methylation in MEFs. DNA methylation was measured at days 4 and 12 post-transduction with the stated vectors. Promoter methylation of the MND Pro vector was also measured in 3T3 cells at day 12 as a control. P-values: **P=0.0052 (day 4) and ***P=0.0001 (day 12).**

293T cells were complemented with ZFP809 (or lacZ as a control) and expression compared with endogenous levels in the stated cells by qRT-PCR. Data are mean+s.e.m. of triplicate transfections (293T cells), two independent cell lines (MEFs) or individual cell lines (EC and ES cells). ZFP809 is responsible for Pro-sequence-directed repression and DNA methylation. 293T cells were transduced with the stated vectors (or nontransduced, Mock) and then transfected with either ZFP809 or lacZ. GFP repression (E) and promoter methylation (F) were measured 5 days post-transfection. **P=0.0042.

KAP1 is itself still present in MEFs. We therefore used mRNA-sequencing data from ES cells and MEFs to compare expression levels of KRAB-ZFPs between these cell types. Although most KRAB-ZFPs were enriched in ES cells compared with MEFs (although there is no evidence that any of them repress ERVs) (with mean expression levels at 3.421 and 1.4242, respectively), ZFP809 displayed similar levels in the two cell types (Fig. 2A; data not shown). Correspondingly, when we compared the silencing of the PBS Pro and IAP vectors in MEFs, we found the former but not the latter underwent repression in these cells (Fig. 2B), consistent with the lack of activation of endogenous IAPs in MEFs upon KAP1 removal (Rowe et al., 2010). Indeed, we propose that IAPs are already stably methylated in MEFs so it is not necessary for their cognate KRAB-ZFP to be maintained. PBS Pro-induced repression was accompanied by de novo DNA methylation in MEFs, albeit at lower levels than in ES cells (Fig. 2C). In order to demonstrate that this mechanism is determined by the presence of the KRAB-ZFP, we complemented human 293T cells with mouse ZFP809, leading to repression but also to significant PBS Pro-dependent DNA methylation of the adjacent promoter (Fig. 2D-F). This level of DNA methylation, however, is not sufficient to maintain silencing, which is reversible (supplementary material Fig. S2).
KAP1 and ESET are required for de novo DNA methylation of ERVs

In order to determine the role of KAP1 in ERV de novo DNA methylation, we used a previously described conditional knockout ES cell line (Rowe et al., 2010). As Kap1 deletion is ultimately lethal in early embryos and in ES cells (Rowe et al., 2010) (supplementary material Fig. S3), experiments were performed within 3 days of inducing this process (a time-point established from supplementary material Fig. S1B). At that point, both repression and DNA methylation of the lentivirally introduced, IAP- or Pro PBS-flanked promoter could be documented in control cells, whereas both processes were abrogated by KAP1 removal (Fig. 3A-C). Similar experiments in Eset-conditioned knockout ES cells demonstrated that this histone methyltransferase is also required for ERV-directed repression and de novo DNA methylation (Fig. 3D-F). Both KAP1 and ESET are thus necessary not only for the maintenance of proviral silencing (Leung et al., 2011; Matsui et al., 2010; Rowe et al., 2010), but also required initially for de novo DNA methylation of introduced ERV sequences.

ERV-guided KAP1-induced DNA methylation is recapitulated in embryogenesis

De novo DNA methylation patterns, which are established early in embryogenesis, are maintained later in development through the action of DNMT1 (Gaudet et al., 2004; Walsh et al., 1998). We investigated whether DNA methylation marks promoted by KAP1 in ES cells could also be induced in vivo, and whether in this setting they were perpetuated until late in embryogenesis. To achieve this, control and KAP1-sensitive IAP sequence-containing MND-GFP lentiviral vectors were injected into fertilized mouse oocytes, and the resulting embryos were collected at E13 to assess integration, repression and promoter methylation. Although all embryos containing the control vector were bright or dull green, none of the IAP LV-harbouring embryos expressed GFP to detectable levels, whether examined by microscopy or by flow cytometry (Fig. 4A,B). Of note, MEFs transgenic for the non-repressed promoter displayed variability in GFP intensity within the population, likely owing to KAP1-independent silencing mechanisms, including those relating to integration site differences (Barklis et al., 1986; Ellis, 2005). Strikingly, DNA methylation at E13 mirrored data obtained in ES cells, with high promoter methylation for all IAP vector-harbouring embryos, whereas DNA methylation of the lentivirally introduced promoter was lower and more variable in the absence of an adjacent KAP1-tethering element (Fig. 4C). These results suggest that in early embryos, KAPI machinery acts as a lock to ensure active DNA methylation and thus robust ERV silencing throughout development.
**ERV sequences can induce repression and DNA methylation of a cellular promoter**

As cellular promoters can lie in close proximity to ERVs in the genome and ERV methylation status can influence cellular gene expression (Duhl et al., 1994; Macfarlan et al., 2011; Macfarlan et al., 2012; Michaud et al., 1994; Rebollo et al., 2011), we assessed whether our KAP1-dependent ERV sequences could also direct *de novo* DNA methylation of a cellular promoter. PBS Pro could induce some repression of a juxtaposed PGK (phosphoglycerate) promoter, albeit less efficiently than for MND, and with only minimal, if any, DNA methylation at day 3 post-transduction (Fig. 5A). However, by day 6, repression was clear (around sevenfold, supplementary material Fig. S4) so we used this time-point to demonstrate that both PBS Pro and IAP sequences could direct *de novo* DNA methylation of the cellular promoter with levels further increasing after 12 days (Fig. 5B). Furthermore, PBS Pro or a KAP1-sensitive IAP sequence induced repression and DNA methylation of the adjacent PGK promoter when lentiviral vectors containing these elements were used to generate transgenic embryos (Fig. 5C; supplementary material Fig. S4B).

**KAP1 shapes DNA methylation of endogenous retroviruses in ES cells and embryos**

Our Southern blot analyses previously failed to detect significant global loss of DNA methylation at ERVs upon *Kap1* deletion (Rowe et al., 2010) but, here, we revisited this issue using the more-sensitive bisulfite pyrosequencing technique to examine ERVs in ES cells and early embryos. As expected, IAPs became highly overexpressed following KAP1 removal in ES cells, which here we record to translate as a marked accumulation of IAP GAG protein (Fig. 6A). We then measured IAP methylation, using the *Oct4* promoter as a control (Fig. 6B); whereas global IAP LTRs were unaffected in *Kap1*-knockout MEFs, they displayed a mild but significant decrease in knockout ES cells in line with the minimal decrease previously observed at ERVs in *Eset*-knockout ES cells (Matsui et al., 2010). A comparatively greater loss of IAP DNA methylation was measured in an ES cell line stably deleted for the G9a histone methyltransferase (Dong et al., 2008; Tachibana et al., 2008). However, this cell line grows continuously, allowing cumulative decreases in DNA methylation, whereas *Kap1*-deleted cells divide only a few times before dying. The overexpression of IAPs in *Kap1*-deleted cells is most probably a consequence of the removal of histone repressive marks such as H3K9me3 (Matsui et al., 2010; Rowe et al., 2010), especially as we confirmed with conventional bisulphite sequencing at one IAP locus that no molecules are completely or even largely demethylated (supplementary material Fig. S5). However, it is possible that such a small decrease in DNA methylation contributes to the phenotype, particularly as we find that 5-Aza treatment of MEFs induces fairly conservative decreases in IAP DNA methylation (supplementary material Fig. S6), yet a massive upregulation of their transcription (Fig. 7B). Interestingly, we also found, using two different primer sets, that DNA methylation levels at ERVs were mildly but significantly lower in E5.5 *Kap1* knockout embryos compared with wild-type littermates (Fig. 6C; supplementary material Fig. S7). Although only 1 out of 23 further knockout embryos was obtained (at E4.5), probably owing to the high lethality of this mutation, this embryo also displayed less DNA methylation at IAPs than its two control littermates (**P**=0.0002 and 0.0014, respectively, data not shown). Altogether, these data support a role for KAP1 in genome-wide ERV DNA methylation during early embryogenesis.
DNA methylation becomes crucial for ERV silencing later in development

Finally, in order to address the relevance of DNA methylation of ERV sequences and endogenous ERVs in development, we derived MEFs from transgenic embryos in which our ERV-derived proviruses were integrated. Treatment with the DNA methyltransferase inhibitor 5-aza (but not the HDAC inhibitor, TSA) was sufficient to relieve silencing and this was true for both ERV sequences newly introduced within the context of lentiviral vectors and endogenous IAPs that were upregulated by up to around 700-fold (Fig. 7). To determine whether DNA methylation already played a role in silencing ERVs at an earlier developmental stage, we also examined DNMT triple knockout (TKO) ES cells (Tsumura et al., 2006), in which all three DNA methyltransferases are deleted (supplementary material Fig. S8). Levels of IAP GAG protein and of IAP and MERV-L RNA were significantly higher in these than in control ES cells (supplementary material Fig. S8A,B), consistent with a previous study (Matsui et al., 2010). It was recently reported that in Dnmt1 KO ES cells and in another ES cell line where all DNMT proteins were depleted (by knockout of Dnmt3a/3b and knockdown of Dnmt1) (Meissner et al., 2005), IAP protein was detected only upon LIF (leukemia inhibitory factor) withdrawal for six or more days, a time point at which OCT4 protein was undetected in differentiated cultures by immunofluorescence (Hutnick et al., 2010). However, the TKO ES cells we have used here were maintained in standard ES cell conditions (see Materials and methods) and were largely undifferentiated as indicated by comparable levels of OCT4 protein to wild type (supplementary material Fig. S8A) as well as high surface expression of SSEA1 (supplementary material Fig. S8C), in contrast to ES cells cultured without LIF, which downregulated this marker (supplementary material Fig. S8D). In addition, these cells have previously been found to conserve their self-renewal ability (Tsumura et al., 2006). We still cannot exclude the possibility that some differentiating cells contribute more prominently to the observed IAP upregulation. Importantly, the overexpression of these elements in ES cells (around 40-fold) is modest compared with that seen in MEFs (Fig. 7), supporting a crucial role for DNA methylation in ERV silencing later in development (Gaudet et al., 2004; Walsh et al., 1998), in contrast to in ES cells where KAP1-dependent histone methylation is present (Hutnick et al., 2010; Karimi et al., 2011; Matsui et al., 2010; Rowe et al., 2010) (reviewed by Leung and Lorincz, 2012; Rowe and Trono, 2011). Indeed, both 5-aza treatment and Kap1 knockout need to be performed in order to induce such a dramatic upregulation of IAPs in ES cells (around 400-fold) (Rowe et al., 2010).
Development 140 (3)

Fig. 6. KAP1 shapes DNA methylation of endogenous retroviruses in ES cells and embryos. (A) KAP1 knockout in ES cells leads to an accumulation of IAP GAG p73 expression. ES cells with loxp-flanked KAP1 untreant (1), treated with 4-OHT (2), or transduced with 4-OHT-inducible Cre, without (3) or with (4) 4-OHT. Cells were harvested 4 days post-4-OHT treatment. (B) KAP1 wild-type and knockout ES cells were analysed for DNA methylation of global IAP LTRs (bottom). Methylation of the Oct4 promoter (top), G9a KO ES cells and KAP1 KO MEFs were used as controls. P-values for the IAP LTR: KO MEFs were no different from wild-type MEFs (P=0.2186); KAP1 knockout ES cells were significantly different from wild-type (P=0.0008) and G9a knockout ES cells were significantly different from wild type (P=0.0002). (C) KAP1 heterozygous mice (C57BL/6) were crossed and embryos dissected at ES 5.5 to measure DNA methylation of endogenous IAPs using primers either in the 5'UTR (top) or LTR (bottom). Two knockouts and one wild-type embryo were analysed. Top: ***P=0.0059 and ***P=0.0006, respectively. Bottom: P=0.1613 (not significant) and *P=0.0493, respectively (paired one-tailed t-tests). (D) Summary model: KRAB-ZFPs, KAP1 and ESET are necessary for de novo methylation of ERVs, occurring within 48 hours. This process is sequence specific and takes place in ES cells and during embryogenesis.

DISCUSSION

The present work demonstrates the involvement of sequence-specific KRAB-ZFPs, KAP1 and ESET in establishing early embryonic DNA methylation patterns at ERV sequences that are subsequently maintained through embryogenesis. This suggests that these factors are required during early development first for the histone-based silencing and then for the de novo methylation of genomic ERVs, even though the evidence provided here is limited to newly introduced ERV sequences. Our results support a model (Fig. 6D) whereby sequences located in the 5' region of ERV genomes recruit specific KRAB-ZFPs, KAP1, ESET and most likely other chromatin modifiers, the conjugated action of which leads to transcriptional repression and de novo promoter methylation. This phenomenon implies the secondary recruitment of de novo DNA methyltransferases, consistent with the documented immunoprecipitation of KAP1 with DNMT3A and DNMT3B (Li et al., 2008; Quenneville et al., 2011; Zuo et al., 2012), and the reported interaction of DNMT3A with the B-Box coiled coil domain of KAP1 (Zuo et al., 2012). As ERV DNA methylation requires ESET, it could be that this enzyme directly recruits DNMT3AB and DNMT3L. In that respect, histone methyltransferases have previously been revealed to act as docking sites for DNA methylation machinery (Dong et al., 2008; Tachibana et al., 2008) and, importantly, ESET has been proposed to bind directly to DNMT3A and tether it to promoters, leading to their DNA methylation-mediated silencing (Li et al., 2006). Alternatively, de novo DNA DNA methylation may result indirectly from a lack of transcription and active marks such as H3K4me3 (Jia et al., 2007; Ooi et al., 2007), owing to the recruitment of H3K9me3 and histone deacetylation at ERVs.

We reveal here that KAP1-recruiting ERV sequences are able to trigger repression and de novo DNA methylation of a strong cellular promoter, as well as of a retroviral promoter, at least within the context of newly introduced lentiviral vectors. The slower kinetics observed with the PGK promoter, compared with its MND counterpart, may be due to partial protection of its CpG-dense core by transcription factors, as recently proposed (Lienert et al., 2011; Stadler et al., 2011). These data indicate that KAP1-mediated silencing of ERVs may dampen expression from nearby cellular transcription units by inducing promoter CpG methylation. This is supported by our observation that the methylation status of CpG islands correlates with their immediate proximity to KAP1-recruiting sites in ES cells (Quenneville et al., 2012). Relevant to ERVs, in a recent study, polymorphic copies of ERVs across mouse strains were employed to determine whether histone and DNA methylation marks could spread from ERVs to cellular promoters.
Results showed that, whereas H3K9me3 marks could spread several kilobases, it was infrequent for DNA methylation to spread from an IAP to neighbouring gene promoters, except where genes lie within 500 bp of an IAP (Rebollo et al., 2011). Our results now implicate a KAP1/ESET complex in this mechanism. As DNA methylation rarely spreads from ERVs as far as nearby promoters in the genome, in future work it will be interesting to investigate the possible presence of barrier elements protecting cellular genes from the spread of repressive marks, including DNA methylation, which is nucleated at ERV sequences.

Here, we observed significant DNA methylation of a PBS Pro-, retroviral promoter-containing lentiviral vector in MEFs, which express similar levels of ZFP809 to ES cells, as well as in 293T cells engineered to express high levels of this KRAB-ZFP. By contrast, a KAP1-restricted IAP sequence induced neither repression nor de novo DNA methylation of the MND promoter in MEFs, suggesting that the cognate KRAB-ZFP is not expressed in these targets. The matching KRAB-ZFP(s) could be any of those enriched in ES cells except ZFP568, which has already been shown not to be responsible for repressing IAPs in vivo (Shibata et al., 2011). The induction of only low methylation in MEFs and 293T cells, compared with ES cells, by ZFP809 explains why the repression is reversible, as previously noted when KRAB is artificially tethered to a promoter in cell lines (Barde et al., 2009; Groner et al., 2010; Quenneville et al., 2012; Wiznerowicz and Trono, 2003).

Surprisingly, we found very little decrease in DNA methylation levels at resident ERVs in KAP1 knockout embryos, despite the role we have demonstrated for KAP1 in shaping DNA methylation patterns at ERV sequences in embryogenesis. It could be that the loss of DNA methylation that follows the loss of repressive histone marks cannot be measured accurately owing to the earlier lethality of knockout embryos (at or before E5.5) (Cammas et al., 2000) or that the majority of DNA methylation at genomic ERVs is KAP1 independent at this stage. Along these lines, as IAPs are relatively resistant to genome-wide DNA demethylation in preimplantation embryos (Lane et al., 2003), methylation may be maintained without requiring de novo methyltransferase activity. Importantly, this and previous studies (Matsui et al., 2010; Rowe et al., 2010) indicate that histone modifications, rather than DNA methylation, are primarily responsible for silencing ERVs in ES cells, as IAPs become immediately and markedly upregulated upon KAP1 or ESET removal in ES cells, in the absence of any significant change in their DNA methylation status. It is noteworthy that ERVs exhibit a range of DNA methylation rates in ES cells, suggesting that some, including many IAPs, may escape demethylation, whereas others may become demethylated but rapidly re-methylated following their KRAB/KAP1 recognition. Moreover, the impact of DNA methylation per se on transcription, at least during this period, needs to be further investigated, as blocking histone-based repression of IAPs suffices to induce their upregulation. It is possible that subtle alterations at specific positions, which are crucial to block transcription, have so far escaped analyses, or that DNA methylation does not have the same consequences in ES and somatic cells: for example, if some factors important for repressing transcription at DNA methylated loci are expressed in differentiated but not in undifferentiated cells.

Our in vivo experiments with vectors reveal that ERV methylation patterns are faithfully copied late into development, leading to extremely high DNA methylation levels (upwards of 80%) that mimic natural levels at IAP LTRs resident in the genome. The self-perpetuating nature of DNA methylation alleviates the need for the continuous expression of cognate ERV-binding KRAB-ZFP repressors. It also explains why the maintenance of DNA methylation at ERV’s throughout development is crucial for suppressing their transcription, notably for IAPs (Walsh et al., 1998). In summary, the present work implies that ERV sequence recognition by KRAB-ZFPs/KAP1 and the lysine methyltransferase ESET is necessary to target ERV-containing loci for rapid de novo DNA methylation in early development, leading to stable site-specific DNA methylation signatures across the genome.

Acknowledgements

We thank P. V. Maillard for advice; A. Raymond and J. Chrest (from the Center for Integrative Genomics, University of Lausanne, Switzerland) for use of their pyrosequencer; and Carine Gubelmann and Andrea Corsinotti for the ZFP809 plasmid. We also thank D. Schuebler for supplying us with the G9a knockout and parental ES cells, originally from Y. Shinkai (The Riken Institute, Japan), who provided us with the Eset-inducible knockout ES cells; and B. Cullen (Duke University) for the anti-IAP GAG antibody. We thank F. Spitz (EMBL, Heidelberg), in whose lab T.A. performed the experiment depicted in supplementary material Fig. S5.

Funding

This work was supported by grants from the Swiss National Science Foundation and the European Research Council to D.T.

Competing interests statement

The authors declare no competing financial interests.

Author contributions

H.M.R. conceived the study, designed and performed the experiments, analyzed the data and wrote the manuscript. S.O., S.V., M.F., D.M. and T.A. contributed to experiments. J.M. provided help with pyrosequencing setup and D.T. conceived the study, designed experiments and wrote the manuscript.

Supplementary material

Supplementary material available online at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.087585/-/DC1

References


methyltransferase in E5 cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J.* 27, 2691-2701.


Gonczy, A. C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. and Reik, Emerson, R. O. and Thomas, J. H.

Groner, A. C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. and Reik, Emerson, R. O. and Thomas, J. H.


**Fig. S1.** ERV sequences can induce rapid de novo DNA methylation in ES cells. This relates to Fig. 1. (A) Genomic DNA samples from Fig. 1B,C (time-point day 4) were used to measure relative vector copies by Q-PCR, to verify that repressed vectors did not represent non-integrated samples. (B) De novo methylation and repression are rapid and progressive thereafter. Left: ERV-containing (Pro or IAP2) vectors or a control vector (B2) were measured for their levels of promoter methylation at 3, 4 or 7 days post-transduction. (Right) In parallel, fold repression was recorded at days 3 and 7 post-transduction. Left P values: day 3, MND B2 versus MND Pro (0.0016) or versus IAP2 MND (0.0009); day 4, MND B2 versus MND Pro (0.0007) or versus IAP2 MND (0.0011); day 7, MND B2 versus MND Pro (0.0016) or versus IAP2 MND (0.0001).
**Fig. S2.** ERV DNA methylation patterns are conditioned by KRAB-ZFP expression profiles. This relates to Fig. 2. Reversibility of MND Pro silencing in MEFs and 293t cells where levels of DNA methylation induced are very low. (A) MEFs shown in Fig. 2B transduced with either the MND or MND Pro vector were cultured for 2 weeks and then transduced with shRNA vectors that were Puromycin selected (either against KAP1 or with an empty vector control) as stated. One week later, GFP was analysed by flow cytometry. (B) 293t cells from Fig. 2E transduced with either the MND or MND Pro vectors and transfected with ZFP809 and then washed and maintained for 5 days longer to allow loss of ZFP809 expression and reversibility of Pro repression.
Fig. S3. KAP1 and ESET are required for de novo DNA methylation of ERVs. This relates to Fig. 3. KAP1 knockout is lethal in ES cells around 4-5 days post-Kap1 excision. Cells were therefore assessed for DNA methylation at 3 days post-Kap1 excision (see Fig. 3). Here, Kap1 LoxP-flanked cells were transduced with a 4-OHT-inducible Cre vector in the absence (KAP1 Cre ES) or presence (KAP1 KO ES) of 4-OHT. Two cell lines are shown (ES3 and ES6). Four days post Kap1- excision, cells were harvested, stained and assessed for cell death by flow cytometry. Note that cell death is underestimated as some cells die and detach before harvesting.
**Fig. S4. ERV sequences can induce repression and DNA methylation of a cellular promoter.** This relates to Fig. 5. (A) Side-by-side comparison of repression of the MND promoter versus the cellular promoter PGK at day 3 and day 6 post-vector transduction of F9 EC cells and ES cells. Results were normalized to expression in 3T3 cells and fold repression of PGK vectors was normalized to the PGK B2 control vector, whereas MND vectors were normalized to the MND B2 control vector. Bars show means and s.d. of triplicate infections. *P* values (unpaired two-tailed *t* tests) for the PGK Pro vector versus the PGK control: EC day 3, *P*=0.0492; ES day 3, *P*=0.0406; EC day 6, *P*=0.0016; ES day 6, *P*=0.0012. For the MND Pro vector versus the MND control: EC day 3, *P*=0.0119; ES day 3, *P*=0.0009; EC day 6, *P*=0.0125; ES day 6, *P*=0.0026. (B) ERV sequences can direct de novo methylation of the PGK promoter in vivo. Lentiviral transgenesis was performed with either the IAP4 PGK vector or a control vector, IAP1 (see Fig. 1) PGK that escapes repression. Five embryos per group were assessed for DNA methylation of the PGK promoter. *P*=0.0270.
Fig. S5. KAP1 shapes DNA methylation of endogenous retroviruses in ES cells and embryos. This relates to Fig. 6. Bisulphite sequencing with TOPO cloning showing the difference in DNA methylation status between molecules at one IAP locus. KAP1 knockout (ko) ES cell samples shown in red and wild-type (wt) ones in black. Sequences are ordered depending on methylation density. Unfilled and filled lollipops represent unmethylated and methylated CpGs, respectively.
**Fig. S6.** DNA methylation of introduced ERV sequences and resident ERVs becomes crucial late in development. This relates to Fig. 7. MEFs from Fig. 7A,B (line PGK Pro 1) were used for DNA methylation analysis by bisulphite pyrosequencing at endogenous IAPs in parallel to the expression analysis shown in Fig. 7B. DNA methylation was reduced in the 5-Aza-treated group.

**Fig. S7.** KAP1 shapes DNA methylation of endogenous retroviruses in ES cells and embryos. This relates to Fig. 6. *Kap1* heterozygous mice were crossed and embryos dissected at E5.5 to measure DNA methylation of endogenous IAPs. Here, the results of the genotyping that was carried out by PCR with a mix of three primers (see Fig. 3) are shown. The 171 and 390 bp products represent *loxP*-flanked or excised *Kap1*, respectively. Embryos selected for analysis (two knockouts and one wild-type embryo) are labelled. The image shows two parts of the same gel but all at the same exposure and analysis settings.
Fig. S8. DNA methylation of introduced ERV sequences and resident ERVs becomes crucial late in development. This relates to Fig. 7. (A) Accumulation of IAP GAG p73 in DNA methyltransferase triple knockout ES cells (DNMT TKO). 3T3 cells were a positive control as they overexpress IAP GAG. Global OCT4 protein levels were comparable between wild-type and TKO ES cells. (B) qRT-PCR showing upregulation of IAP and MERVL transcripts in DNMT TKO cells. Samples were normalized to Gapdh, and G9a levels were also similar between samples whereas DNMT1 and DNMT3B were verified to be absent in DNMT TKO cells. (C) Wild-type and DNMT TKO ES cells were stained with an anti-SSEA1 antibody or left unstained, and histogram results were overlaid to verify that cells were largely undifferentiated. (D) In another experiment, ES cells were cultured for 6 days in the presence or absence of LIF, and then stained as in C to verify that cells downregulate SSEA1 upon differentiation in order to validate its use as a marker of undifferentiation in C.
Table S1. Primer sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>qRT-PCR primers</strong></td>
<td></td>
</tr>
<tr>
<td>GFP_F</td>
<td>CTGCTGCCGACAACCAC</td>
</tr>
<tr>
<td>GFP_R</td>
<td>ACCATGTGATCAGCCTTTC</td>
</tr>
<tr>
<td>GFP_probe</td>
<td>CCAGTCCGCCCTGAGCAAAGACC</td>
</tr>
<tr>
<td>Titin_F</td>
<td>TTCAGTGTGTAGCCTAGGC</td>
</tr>
<tr>
<td>Titin_R</td>
<td>AAAACGAGCGTGACGGTGAGC</td>
</tr>
<tr>
<td>Titin_probe</td>
<td>TGCACCGGAAGCGTCTCAGTCTCAGTC</td>
</tr>
<tr>
<td>Actin_F</td>
<td>TAGGCACCAGGGGTGTGATGG</td>
</tr>
<tr>
<td>Actin_R</td>
<td>CATGGCTGGGTTGGAAGG</td>
</tr>
<tr>
<td>KAP1_F</td>
<td>CGGAAATGTGAGCGTGTTTCT</td>
</tr>
<tr>
<td>KAP1_R</td>
<td>CGGTAGCCAGCTGATGCAA</td>
</tr>
<tr>
<td>Gapdh_F</td>
<td>TCCATGACAACCTTGGCCATTG</td>
</tr>
<tr>
<td>Gapdh_R</td>
<td>CAGTCTTCTGGGTTGGAAGTG</td>
</tr>
<tr>
<td>G9A_F</td>
<td>AGACAGCCGGTGGTGAA</td>
</tr>
<tr>
<td>G9A_R</td>
<td>CCCTCGGAGGCTCTCGTT</td>
</tr>
<tr>
<td>DNMT1_F</td>
<td>CCAGGCATTTTGCTGGA</td>
</tr>
<tr>
<td>DNMT1_R</td>
<td>CTTGCGATCTGCTGTGAACA</td>
</tr>
<tr>
<td>DNMT3b_F</td>
<td>AACTCCATAGACACCGGCAA</td>
</tr>
<tr>
<td>DNMT3b_R</td>
<td>GTCCTTGCCATTGACAGTCACTAC</td>
</tr>
<tr>
<td>ESET_F</td>
<td>TGGCAACAGCGGTCTGAGA</td>
</tr>
<tr>
<td>ESET_R</td>
<td>CAGAAGTTATCAGAGCTGTCATCA</td>
</tr>
<tr>
<td>Zfp809_F</td>
<td>AATTTGGAGCGTGGATTTGG</td>
</tr>
<tr>
<td>Zfp809_R</td>
<td>GGGAGGCTCTGTGTTGAAG</td>
</tr>
<tr>
<td>EeF1a1_F</td>
<td>AGCAAAAATGACCCACCATA</td>
</tr>
<tr>
<td>EeF1a1_R</td>
<td>GGCCTGGATGGTTCAGGATA</td>
</tr>
<tr>
<td>EeF1a1_probe</td>
<td>CACCTGAGCAGTGAAAGCCAGCTGCTT</td>
</tr>
</tbody>
</table>
### qPCR copy number primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV_GAG_F</td>
<td>GGAGCTAGAAGCATTCGCAGTTA</td>
</tr>
<tr>
<td>HIV_GAG_R</td>
<td>GGTGTAAGCTGTCACGATATTGTC</td>
</tr>
<tr>
<td>HIV_GAG_probe</td>
<td>ACAGCCTTTCTGTGTTTCTAACAGGCAGG</td>
</tr>
<tr>
<td>GFP_F</td>
<td>CTGCTGCCCAGAACACCAC</td>
</tr>
<tr>
<td>GFP_R</td>
<td>ACCATGTAATCGTCCTGTCCTGTC</td>
</tr>
<tr>
<td>GFP_probe</td>
<td>CCACTCCGCCCCTGACGAAGACC</td>
</tr>
<tr>
<td>Titin_F</td>
<td>TTCAGTCATGCTGCTAGCGC</td>
</tr>
<tr>
<td>Titin_R</td>
<td>AAAACGACAGTGACGTAGGC</td>
</tr>
<tr>
<td>Titin_probe</td>
<td>TGCACGGAAGCCTCTGCTCAAGTC</td>
</tr>
</tbody>
</table>

### Bisulphite pyrosequencing primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct4_Promoter_F</td>
<td>AGGGTGAAGGATTTTGAA</td>
</tr>
<tr>
<td>Oct4_Promoter_biotin_R</td>
<td>CCACCCTCTAACCCTAACCCT</td>
</tr>
<tr>
<td>Oct4_Promoter_seq</td>
<td>GGTGAAAATGAAGGTTT</td>
</tr>
<tr>
<td>IAP LTR_biotin_F</td>
<td>GGTGGTGAATGAGGGATTTT</td>
</tr>
<tr>
<td>IAP LTR_R</td>
<td>CTCTACTCCATATACTCTACTCT</td>
</tr>
<tr>
<td>IAP LTR_seq</td>
<td>ATACTCTACCTCCCT</td>
</tr>
<tr>
<td>IAP 5’UTR_F</td>
<td>GGGTTGTAGGTAATATGAGGAGTGA</td>
</tr>
<tr>
<td>IAP 5’UTR_biotin_R</td>
<td>ACAATTAAATCCTTTCTAAACATCTACTT</td>
</tr>
<tr>
<td>IAP 5’UTR_seq</td>
<td>ATTTTGTTGGTGTGT</td>
</tr>
<tr>
<td>MND_F</td>
<td>TTAGATGTTTATGAGGTGTTAAGGA</td>
</tr>
<tr>
<td>MND_biotin_R</td>
<td>TCACTCAAAAAAAACCTCCCCAAAAA</td>
</tr>
<tr>
<td>MND_seq</td>
<td>TGATTTGGTTTTATTAATGAAATTAA</td>
</tr>
<tr>
<td>hPGK_F</td>
<td>GGTGGGTTGAGGTTTTTTTAAGGTA</td>
</tr>
<tr>
<td>hPGK_biotin_R</td>
<td>CTAAACCAACCCCTATTAACCACACCAT</td>
</tr>
</tbody>
</table>
hPGK_seq

GTTTTTAAGTAGGGAAGGTTTT