

KAP1 controls endogenous retroviruses in embryonic stem cells

Helen M. Rowe¹, Johan Jakobsson^{1†}, Daniel Mesnard¹, Jacques Rougemont¹, Séverine Reynard¹, Tugce Aktas², Pierre V. Maillard¹, Hillary Layard-Liesching¹, Sonia Verp¹, Julien Marquis¹, François Spitz², Daniel B. Constam¹ & Didier Trono¹

More than forty per cent of the mammalian genome is derived from retroelements, of which about one-quarter are endogenous retroviruses (ERVs)¹. Some are still active, notably in mice the highly polymorphic early transposon (ETn)/MusD and intracisternal A-type particles (IAP)^{2,3}. ERVs are transcriptionally silenced during early embryogenesis by histone and DNA methylation^{4–6} (and reviewed in ref. 7), although the initiators of this process, which is essential to protect genome integrity⁸, remain largely unknown. KAP1 (KRAB-associated protein 1, also known as tripartite motif-containing protein 28, TRIM28) represses genes by recruiting the histone methyltransferase SETDB1, heterochromatin protein 1 (HP1) and the NuRD histone deacetylase complex⁹, but few of its physiological targets are known. Two lines of evidence suggest that KAP1-mediated repression could contribute to the control of ERVs: first, KAP1 can trigger permanent gene silencing during early embryogenesis¹⁰, and second, a KAP1 complex silences the retrovirus murine leukaemia virus in embryonic cells^{11–13}. Consistent with this hypothesis, here we show that KAP1 deletion leads to a marked upregulation of a range of ERVs, in particular IAP elements, in mouse embryonic stem (ES) cells and in early embryos. We further demonstrate that KAP1 acts synergistically with DNA methylation to silence IAP elements, and that it is enriched at the 5' untranslated region (5'UTR) of IAP genomes, where KAP1 deletion leads to the loss of histone 3 lysine 9 trimethylation (H3K9me3), a hallmark of KAP1-mediated repression. Correspondingly, IAP 5'UTR sequences can impose *in cis* KAP1-dependent repression on a heterologous promoter in ES cells. Our results establish that KAP1 controls endogenous retroelements during early embryonic development.

KAP1, a member of the RBCC (ring, B-box, coiled-coiled) or TRIM (tripartite motif) family of proteins, is recruited to genes by the tetrapod-specific, DNA sequence-specific KRAB-ZFPs (Krüppel-associated box domain-zinc finger proteins)¹⁴, which constitute the largest family of transcriptional regulators encoded by higher vertebrates. However, until now, few KAP1 target genes and their KRAB zinc finger intermediates have been identified¹⁵. To assess the potential role of KAP1 in ERV regulation we generated two conditional KAP1-knockout ES cell lines, in which the *Kap1* gene could be inactivated by a 4-hydroxytamoxifen (4-OHT)-inducible system. Both lines expressed pluripotency markers, and one was used to demonstrate contribution to mouse chimaeras after injection into blastocysts (Fig. 1, Supplementary Fig. 1 and data not shown). Control and KAP1-deleted ES cells (see Fig. 1a) were subjected to a combination of large-scale RNA sequencing and specific PCR with reverse transcription (RT-PCR) measurements. The results showed a modest increase in LINE1 (long interspersed nuclear elements 1) transcripts after KAP1 removal, but a marked upregulation of a range of ERVs, in particular IAP elements, which exhibited 15- and

66-fold overexpression in the two ES cell lines, respectively (Fig. 1b–d and Supplementary Figs 1 and 2). Because ES cells progressively lose self-renewal ability after KAP1 depletion (refs 16, 17 and data not shown), as a control we examined ES cells cultured in the absence of leukaemia inhibitory factor (LIF). Under these conditions we also observed a decrease in the stem cell markers NANOG and to a lesser extent OCT4 (also known as POU5F1), but IAP elements were not upregulated (Fig. 1a, c). Furthermore, partially restoring KAP1 levels by transduction with a KAP1-expressing lentiviral vector proportionately reduced the upregulation of IAP elements in KAP1-deleted ES cells (Supplementary Fig. 3). Notably, the stimulation of IAP transcription was accompanied by an increase in the IAP DNA load of KAP1-deleted compared to control ES cells, demonstrating that IAP genes were not only overexpressed but could also reverse transcribe and probably integrate into the genome (Fig. 1e). In contrast, IAP expression was not increased after KAP1 deletion in mouse embryonic fibroblasts (MEFs) (Supplementary Fig. 4), which supports a model in which ERV control in differentiated cells relies on more stable silencing mechanisms^{4,5}.

To investigate more thoroughly the mechanisms of IAP control during early embryogenesis, we treated ES cells with the DNA methyltransferase inhibitor 5-azacytidine (5-aza) and observed a similar upregulation of IAP elements. Notably, combining this drug with KAP1-knockout induced a synergistic effect on IAP overexpression (Fig. 1f). This suggests that DNA methylation and KAP1 repression act cooperatively to silence these elements. Moreover, validating the results of our ES cell-based experiments, we found that IAP elements were markedly increased in KAP1-depleted blastocysts cultured *ex vivo* (Supplementary Fig. 5), and, most importantly, were upregulated more than five-hundred times in KAP1-knockout embryos (Fig. 2).

These data indicate that KAP1 has a crucial role in controlling ERVs during early embryonic development. To determine whether this effect is direct, we performed chromatin immunoprecipitation (ChIP) studies. First, we found that KAP1 associates with the IAP genome in ES cells (Fig. 3a, b). Notably, KAP1 was significantly enriched over the IAP 5'UTR compared to the U3 ($P = 0.034$) or *gag* ($P = 0.026$) regions in control cells. Furthermore, KAP1 knockout correlated with a decrease in H3K9 trimethylation and an increase in H4 acetylation on the IAP genome (Fig. 3c, d). This combination of chromatin modifications is consistent with a loss of KAP1-mediated repression⁹. We also assessed the global level of IAP DNA methylation by Southern blot and more specifically by bisulphite sequencing, but did not detect a significant difference with or without KAP1 (not shown).

To evaluate the genetic diversity of IAP elements controlled by KAP1, we sequenced the 5'UTR of IAP complementary DNAs isolated from KAP1-depleted cells. The 33 sequences thereby analysed were highly diverse (Fig. 4a and Supplementary Fig. 7). Of interest

¹School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. ²EMBL Heidelberg, Meyerhofstrasse 1, 69117 Heidelberg, Germany. [†]Present address: Wallenberg Neuroscience Center, Department of Experimental Medical Sciences, Lund University, 22184 Lund, Sweden.

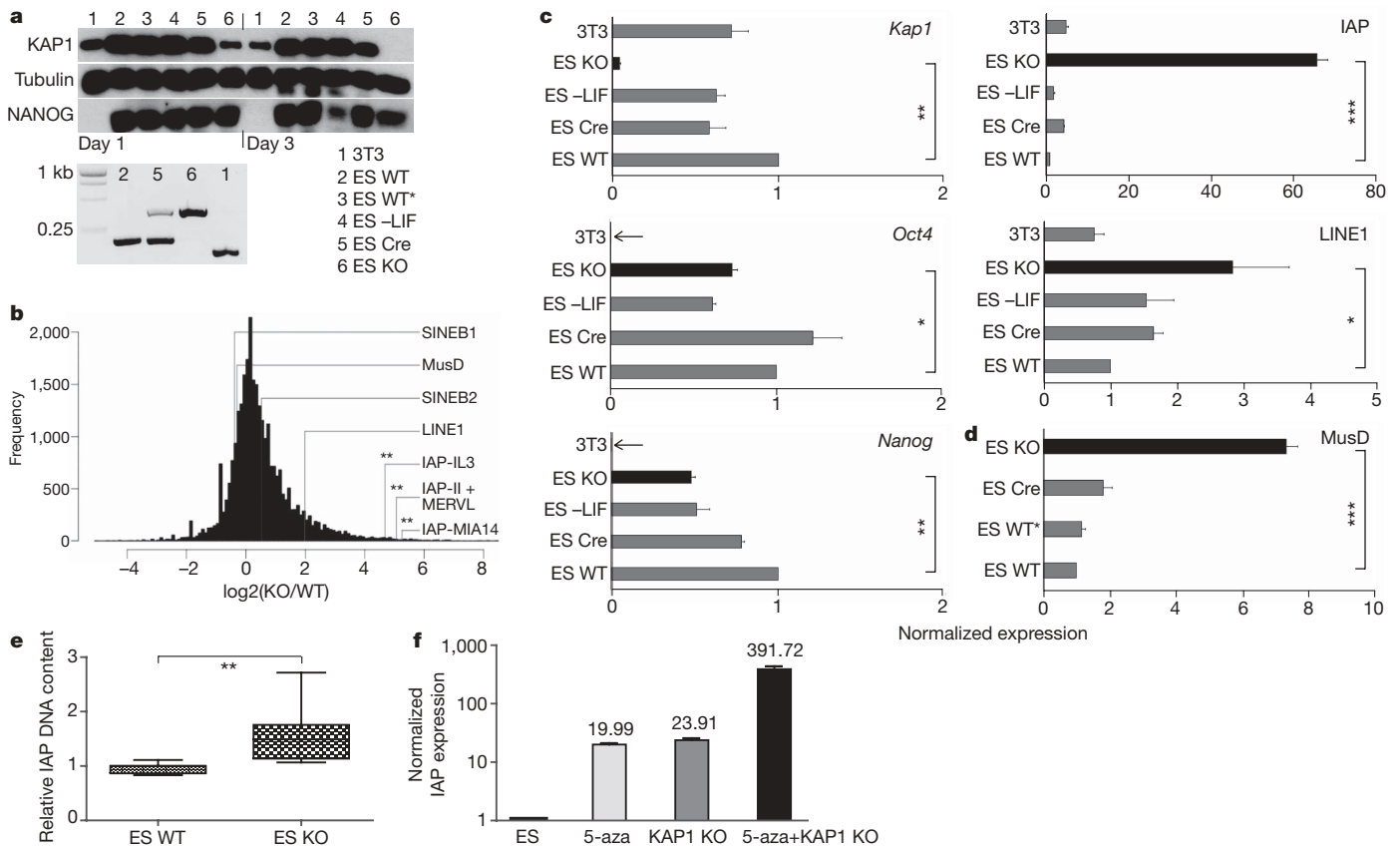


Figure 1 | IAP elements are upregulated in KAP1-depleted ES cells.

was their heterogeneity in the region coding for the primer-binding site (PBS)—the sequence complementary to the cellular transfer RNA that acts as primer for retroviral minus-strand DNA synthesis. Murine leukaemia virus (MLV) is indeed silenced in embryonic cells by the ZFP809-mediated recruitment of a KAP1-containing complex to its proline tRNA PBS (PBS Pro)^{11–13}. The PBS of IAP elements expressed in KAP1-depleted cells covered a range closely related to PBS Phe (Supplementary Fig. 8). To address whether these PBS variants were sufficient to confer KAP1-sensitivity to a heterologous promoter

in ES cells, we cloned them into a lentiviral vector expressing a green fluorescent protein (GFP) transgene from the MND (myeloproliferative sarcoma virus enhancer, negative control region deleted¹⁸) promoter. We then transduced KAP1-excisable ES cells with the resulting vectors and examined GFP expression. Although the MLV Pro sequence induced potent KAP1-dependent silencing compared to its functionally inactive B2 point mutant, none of the IAP PBS variants induced significant repression (Supplementary Fig. 8). This concurs with the results of a previous study in which the silencing activity of two

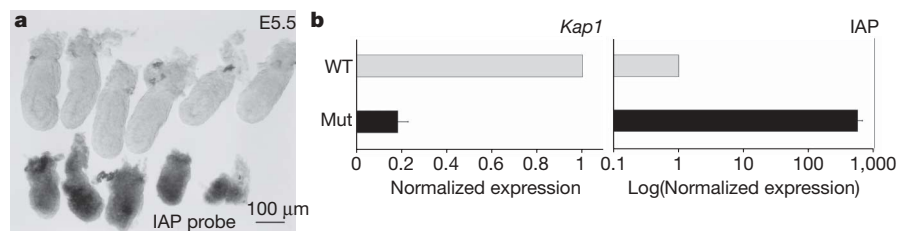


Figure 2 | KAP1 controls IAP elements in embryogenesis. **a**, Embryos from *Kap1*^{+/-} intercrosses were dissected and *in situ* hybridization performed with a 500-bp IAP probe (specific for a cDNA region spanning the IAP 5'UTR as shown in Fig. 4a). One litter is shown; in total, 23 out of 84 embryos (27%) stained positive for IAP elements. See also Supplementary

Figure 2. **b**, Mice were crossed as above and embryos dissected at E6.5, and two mutant versus two wild-type embryos selected and pooled based on morphology. Samples were analysed by qRT-PCR at three dilutions and results show the mean and s.d. One of two experiments is shown.

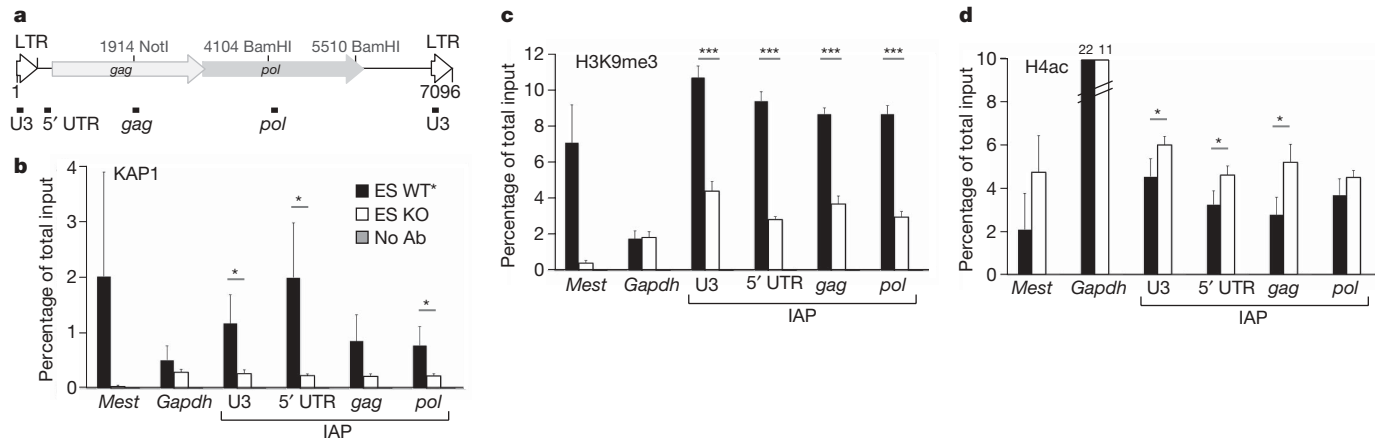


Figure 3 | KAP1 is enriched at the 5'UTR of IAP genomes and loss of KAP1 leads to loss of H3K9me3 and an increase in H4 acetylation. **a**, IAP map (based on IAP-MIA14; ref. 28) with positions of primers used for ChIPs. LTR, long terminal repeat. **b**, KAP1 ChIP results 4 days after 4-OHT addition to control (ES WT*) or Cre-expressing (ES KO) ES cells. Graphs show the mean enrichment in the immunoprecipitations ($n = 3$) relative to the total input samples and error bars show the s.d. A control with no antibody (Ab) gave background enrichment (mean 0.008%). All significant differences between wild-type and knockout for IAP primers are marked.

other IAP PBS sequences was evaluated¹⁹. However, when we cloned 500-base-pair-(bp)-long fragments overlapping the 5'UTR of IAP elements expressed in KAP1-null cells (shown in Fig. 4a) either upstream or downstream of the MND promoter, we could induce up to 53-fold KAP1-dependent silencing, which was comparable to the 50-fold repression induced by the MLV Pro sequence (Fig. 4b). Notably, the same region cloned from an IAP element expressed in control cells

Mest (mesoderm-specific transcript), a direct target of KAP1 in embryonic carcinoma (EC) cells²⁹, was a positive control and *Gapdh* a negative control. **c, d**, ChIPs as above but with an anti-H3K9me3 antibody (**c**), and with an antibody specific for acetylated H4 (**d**). Data are representative of 2–3 experiments. *P* values (by unpaired *t* tests) are *Kap1*: $P = 0.037$ (U3); $P = 0.036$ (5'UTR); $P = 0.047$ (*pol*), but summary values for the three experiments are $P < 0.0001$ for the U3 and 5'UTR, and $P < 0.003$ for *gag* and *pol*. H3K9me3: $P \leq 0.001$ (U3, 5'UTR, *gag* and *pol*). H4Ac: $P = 0.049$ (U3); $P = 0.037$ (5'UTR); $P = 0.022$ (*gag*).

(IAP1, which has multiple sequence differences, see Supplementary Fig. 10) failed to repress the MND–GFP reporter, thereby serving as a negative control.

Taken together, these data demonstrate that ERVs are repressed in murine embryonic stem cells by the recruitment of a KAP1-containing chromatin remodelling complex to their 5'UTRs, corroborating the previously noted marked enrichment for H3K9me3 at and near ERV

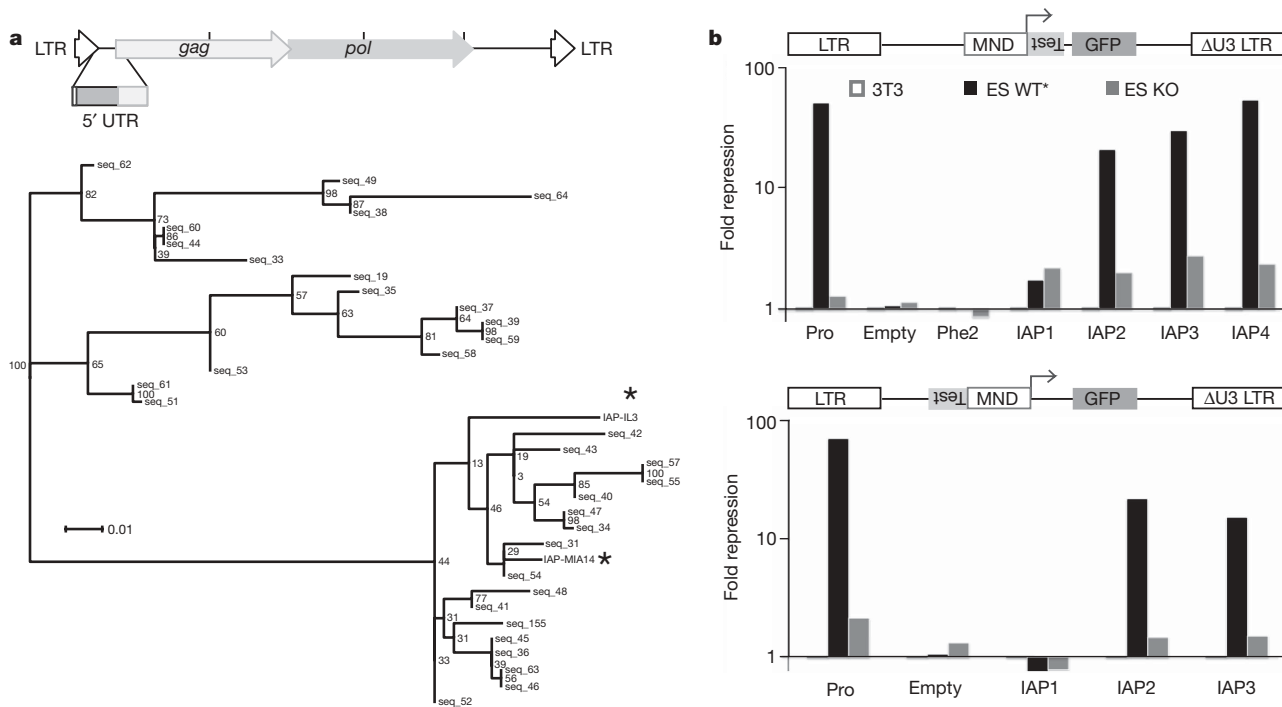


Figure 4 | IAP 5'UTR sequences expressed in KAP1-depleted ES cells are polymorphic and can repress a GFP reporter in a KAP1-dependent way. **a**, Map showing the 500-bp 5'UTR/5'gag region of IAP transcripts sequenced, and a phylogenetic tree of the 33 sequences obtained. Sequences for IAP-MIA14 and IAP-IL3 were included in the alignment and are marked by an asterisk. **b**, Lentiviral vector map with a GFP reporter and the test site where the following sequences were cloned antisense: IAP fragments 1–4 or PBS sequences Pro, B2 or Phe2 (see Supplementary Fig. 8). ES cells were transduced (1 day after 4-OHT treatment) with these vectors (or an empty

vector) and GFP was measured 3 days later in the SSEA-1^{hi} cell fraction (see Supplementary Fig. 9). Fold repression shows the ratio of expression between these vectors and the control B2 one, normalized to 3T3 cells where the ratio equals 1. The bottom panel shows the results for IAP fragments 1–3 or the PBS sequences, Pro or B2 cloned upstream of the promoter. All sequences are in antisense orientation except IAP2, which is in sense orientation. Results were normalized as above and are representative of 2–3 experiments. Results were also reproduced in E14 ES cells (see Supplementary Fig. 11).

DNA in these cells⁵. The described mechanism thus seems to represent a tetrapod-specific complement to the small-RNA-mediated retrotransposon silencing that is at play from plants to mammals^{20,21}.

By analogy with the demonstrated mechanism of repression of other targets including MLV, it is likely that KAP1 is tethered to IAP genomes by KRAB-ZFPs. The observed sequence diversity of KAP1-dependent ERVs further suggests a corresponding degree of polymorphism in the zinc fingers mediating their recognition. Our large-scale RNA sequencing analysis detected more than 250 KRAB-ZFP transcripts in ES cells, including 56 at levels higher than ZFP809 (data not shown). Phylogenetic studies further show that the DNA-binding domains of KRAB-ZFPs have been under strong positive selection during evolution, pointing to their participation in genetic conflicts²². Our data are consistent with a model in which rapidly mutating retroelements have been protagonists of these conflicts, exerting strong selective pressures on KRAB-ZFPs responsible for their control.

Finally, considering that epigenetic silencing can spread from repetitive elements to neighbouring genes^{5,23,24}, the work presented here opens new perspectives to explore ERV-mediated control of cellular genes in development and in adult tissues.

METHODS SUMMARY

ES cells. Two ES cell lines were derived from *Kap1^{loxP/loxP}* mice (gift from F. Cammas), karyotyped and cultured as described²⁵. KAP1-knockout cells were analysed 4 days after treatment with 4-OHT (used at 1 μ M overnight, from Sigma). Differentiation was monitored using an anti-SSEA-1 antibody (BD Pharmingen, MC480). Western blots were performed as described²⁶ using antibodies specific for KAP1 (Chemicon, MAB3662), NANOG (Abcam, 21603-100), OCT4 (Santa Cruz, sc5279) and α -tubulin (Sigma).

Lentiviral vectors. The LVCT¹⁰ vector was modified to express CAG-4-OHT-inducible Cre (from P. Chambon) and an SV40-puro cassette. An LV PGK-GFP vector was modified by substituting GFP with Cre for MEF experiments, and was also used to construct LV-silencing vectors by replacing PGK with MND¹⁸ and including test sequences upstream or downstream. Vectors were titrated on 3T3 cells.

RT-PCR. Total RNA was purified using a Trizol kit (Invitrogen), treated with DNase (Ambion), and 0.5 μ g was reverse transcribed using random primers and SuperScript II (Invitrogen). Alternatively, for low starting material, an RNeasy micro kit was used. Primers (see Supplementary Table 1) were used for SYBR green Q-PCR (Applied Biosystems) and their specificity confirmed with dissociation curves. All data are *Gapdh* normalized, although the actin gene gave similar results. IAP DNA PCR was normalized to the titin gene and results confirmed with *Gapdh*, *MusD* and major satellite primers.

ChIP. Chromatin was prepared according to the Upstate protocol, and starting material was normalized between wild-type and knockout samples. Triplicate immunoprecipitations were performed using protein A agarose beads (Millipore) and the following rabbit antibodies: KAP1 (ref. 9) (from D. Schultz), H3K9me3 (Abcam, ab8898) and acetyl H4 (Upstate, 06-866). Input and immunoprecipitation samples were analysed by SYBR green Q-PCR.

Bioinformatics. Sequences were aligned with Mafft and a phylogenetic tree generated using RAxML (Randomized Axelerated Maximum Likelihood). Illumina RNA-sequencing was analysed with MAQ 0.7.1 and reads mapped to the collection of mouse transcripts from RefSeq (version 36).

Received 24 September; accepted 11 November 2009.

1. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
2. Dewannieux, M., Dupressoir, A., Harper, F., Pierron, G. & Heidmann, T. Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells. *Nature Genet.* **36**, 534–539 (2004).
3. Zhang, Y., Maksakova, I. A., Gagnier, L., van de Lagemaat, L. N. & Mager, D. L. Genome-wide assessments reveal extremely high levels of polymorphism of two active families of mouse endogenous retroviral elements. *PLoS Genet.* **4**, e1000007 (2008).
4. Martens, J. H. et al. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24**, 800–812 (2005).
5. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
6. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genet.* **20**, 116–117 (1998).
7. Maksakova, I. A., Mager, D. L. & Reiss, D. Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. *Cell. Mol. Life Sci.* **65**, 3329–3347 (2008).

8. Morgan, H. D., Sutherland, H. G., Martin, D. I. & Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. *Nature Genet.* **23**, 314–318 (1999).
9. Sripathy, S. P., Stevens, J. & Schultz, D. C. 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 (2006).
10. Wiznerowicz, M. et al. The Kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *J. Biol. Chem.* **282**, 34535–34541 (2007).
11. Wolf, D., Cammas, F., Losson, R. & Goff, S. P. Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. *J. Virol.* **82**, 4675–4679 (2008).
12. Wolf, D. & Goff, S. P. TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell* **131**, 46–57 (2007).
13. Wolf, D. & Goff, S. P. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* **458**, 1201–1204 (2009).
14. Friedman, J. R. et al. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**, 2067–2078 (1996).
15. Krebs, C. J., Larkins, L. K., Khan, S. M. & Robins, D. M. Expansion and diversification of KRAB zinc-finger genes within a cluster including Regulator of sex-limitation 1 and 2. *Genomics* **85**, 752–761 (2005).
16. Fazio, T. G., Huff, J. T. & Panning, B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134**, 162–174 (2008).
17. Hu, G. et al. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev.* **23**, 837–848 (2009).
18. Haas, D. L. et al. The Moloney murine leukemia virus repressor binding site represses expression in murine and human hematopoietic stem cells. *J. Virol.* **77**, 9439–9450 (2003).
19. Modin, C., Lund, A. H., Schmitz, A., Duch, M. & Pedersen, F. S. Alleviation of murine leukemia virus repression in embryonic carcinoma cells by genetically engineered primer binding sites and artificial tRNA primers. *Virology* **278**, 368–379 (2000).
20. Tam, O. H. et al. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534–538 (2008).
21. Watanabe, T. et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539–543 (2008).
22. Emerson, R. O. & Thomas, J. H. Adaptive evolution in zinc finger transcription factors. *PLoS Genet.* **5**, e1000325 (2009).
23. Faulkner, G. J. et al. The regulated retrotransposon transcriptome of mammalian cells. *Nature Genet.* **41**, 563–571 (2009).
24. Whitelaw, E. & Martin, D. I. Retrotransposons as epigenetic mediators of phenotypic variation in mammals. *Nature Genet.* **27**, 361–365 (2001).
25. Bryja, V., Bonilla, S. & Arenas, E. Derivation of mouse embryonic stem cells. *Nature Protocols* **1**, 2082–2087 (2006).
26. Maillard, P. V., Reynard, S., Serhan, F., Turelli, P. & Trono, D. Interfering residues narrow the spectrum of MLV restriction by human TRIM5 α . *PLoS Pathog.* **3**, e200 (2007).
27. Cammas, F., Herzog, M., Lerouge, T., Chambon, P. & Losson, R. Association of the transcriptional corepressor TIF1 β with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes Dev.* **18**, 2147–2160 (2004).
28. Mietz, J. A., Grossman, Z., Lueders, K. K. & Kuff, E. L. Nucleotide sequence of a complete mouse intracisternal A-particle genome: relationship to known aspects of particle assembly and function. *J. Virol.* **61**, 3020–3029 (1987).
29. Riclet, R. et al. Disruption of the interaction between transcriptional intermediary factor 1 β and heterochromatin protein 1 leads to a switch from DNA hyper- to hypomethylation and H3K9 to H3K27 trimethylation on the MEST promoter correlating with gene reactivation. *Mol. Biol. Cell* **20**, 296–305 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. Turelli and I. Barde for advice, F. Cammas and R. Losson for the *Kap1^{loxP/loxP}* and *Kap1^{1+/-}* mice, D. Schultz for the KAP1-specific antibody, and K. Harshman at the genomics platform for the Illumina high throughput sequencing. This work was supported by grants from the Swiss National Science Foundation, the Infectigen Association and the Strauss Foundation to D.T., and by a post-doctoral fellowship from the Swedish Research Council to J.J.

Author Contributions H.M.R. performed experiments, with contributions from S.R., T.A., P.V.M., H.L.-L., S.V. and J.M.J.J. developed conditional KAP1-knockout ES cell lines, D.M. performed dissection and *in situ* hybridisation of embryos and J.R. did bioinformatics analyses. P.V.M., F.S., D.B.C., D.M. and J.J. contributed to experimental design. D.T. and H.M.R. designed the study, analysed the results and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.T. (didier.trono@epfl.ch).