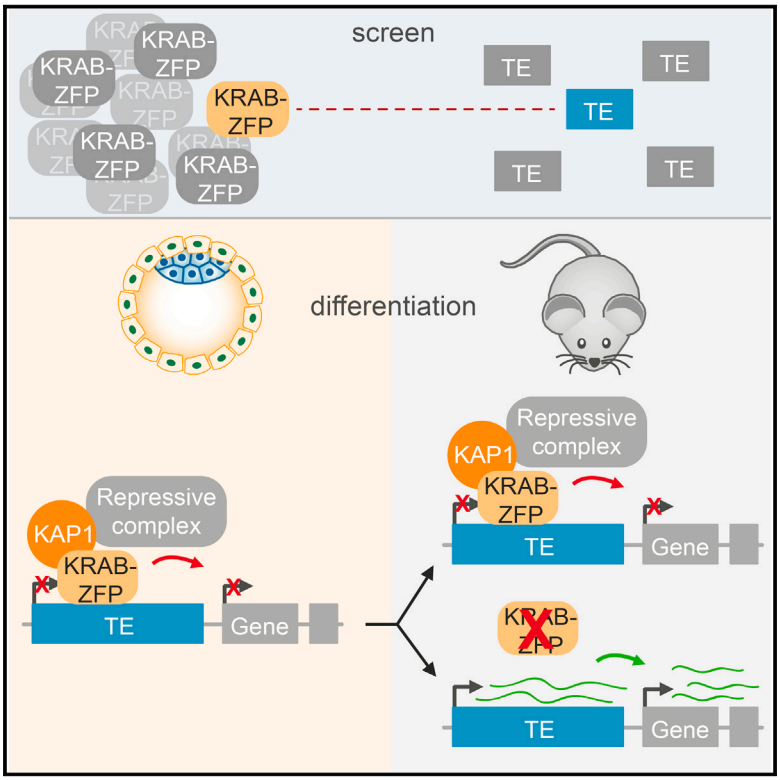
Article

Developmental Cell

Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues

Graphical Abstract



Authors

Gabriela Ecco, Marco Cassano, Annamaria Kauzlaric, ..., Helen M. Rowe, Priscilla Turelli, Didier Trono

Correspondence

didier.trono@epfl.ch

In Brief

KRAB-ZFPs and KAP1 are embryonic controllers of transposable elements (TEs) thought to irreversibly silence TEs and thus be dispensable in the adult. Ecco et al. demonstrate that these modulators continue to control TE expression in adult tissues, where they also act to control expression of neighboring cellular genes.

Highlights

- Large-scale functional screen matches KRAB-ZFPs to transposable element (TE) targets
- ZFP932 and its paralog Gm15446 regulate different endogenous retrovirus K subsets
- KRAB-ZFPs/KAP1 regulate TEs in somatic cells via histone modifications
- KRAB-ZFPs/KAP1 use TE-based platforms to regulate adult tissue cellular gene expression

Accession Numbers

GSE74278

Ecco et al., 2016, Developmental Cell *36*, 611–623
March 21, 2016
http://dx.doi.org/10.1016/j.devcel.2016.02.024

Developmental Cell
Article

Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues

Gabriela Ecco,[1] Marco Cassano,[1] Annamaria Kauzlaric,[1] Julien Duc,[1] Andrea Coluccio,[1] Sandra Offner,[1] Michaël Imbeault,[1] Helen M. Rowe,[1,2] Priscilla Turelli,[1] and Didier Trono[1,*]
[1]School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Station 19, 1015 Lausanne, Switzerland
[2]Present address: Centre for Medical Molecular Virology, Division of Infection and Immunity, University College London, 90 Gower Street, London WC1E 6BT, UK
*Correspondence: didier.trono@epfl.ch
http://dx.doi.org/10.1016/j.devcel.2016.02.024

SUMMARY

KRAB-containing zinc finger proteins (KRAB-ZFPs) are early embryonic controllers of transposable elements (TEs), which they repress with their cofactor KAP1 through histone and DNA methylation, a process thought to result in irreversible silencing. Using a target-centered functional screen, we matched murine TEs with their cognate KRAB-ZFP. We found the paralogs ZFP932 and Gm15446 to bind overlapping but distinguishable subsets of ERVK (endogenous retrovirus K), repress these elements in embryonic stem cells, and regulate secondarily the expression of neighboring genes. Most importantly, we uncovered that these KRAB-ZFPs and KAP1 control TEs in adult tissues, in cell culture and in vivo, where they partner up to modulate cellular genes. Therefore, TEs and KRAB-ZFPs establish transcriptional networks that likely regulate not only development but also many physiological events. Given the high degree of species specificity of TEs and KRAB-ZFPs, these results have important implications for understanding the biology of higher vertebrates, including humans.

INTRODUCTION

Transposable elements (TEs) account for more than half of the human and murine genomes (Lander et al., 2001; Waterston et al., 2002). Long considered as purely parasitic, they are now recognized as important motors of evolution, yet they also represent genomic threats requiring control from the earliest stages of development. Whether they are DNA transposons or retrotransposons (endogenous retroviruses [ERVS], LINEs, SINEs, and SVAs; reviewed in Friedli and Trono, 2015), TEs can disrupt genes, alter their transcription, or serve as ground for recombination and have been implicated in diseases such as cancer and diabetes (Hancks and Kazazian, 2012; Jern and Coffin, 2008). However, growing evidence indicates that TEs can be co-opted for the benefit of the host, with for instance expression of zygotic activation genes driven from the long terminal repeat (LTR) of murine endogenous retrovirus L (MERVL) in the mouse, and many binding sites for pluripotency factors residing within mobile DNA elements in the human genome (Bourque et al., 2008; Chuong, 2013; Dupressoir et al., 2012; Fort et al., 2014; Macfarlan et al., 2012).

TEs are repressed through RNA- and protein-based epigenetic mechanisms instated during the first days of embryogenesis. KRAB-containing zinc finger proteins (KRAB-ZFPs) constitute a large family of transcription factors implicated in this process. KRAB-ZFPs bind to specific DNA sequences through an array of zinc fingers, and recruit their cofactor KAP1, which serves as a scaffold for a heterochromatin-inducing complex encompassing histone methyltransferase, histone deacetylase, nucleosome remodeling, and DNA methyltransferase activities (reviewed in Rowe and Trono, 2011). Depletion of KAP1 or its partner histone methyltransferase SETDB1 in murine or human embryonic stem cells (ESCs) activates the expression of endogenous retroelements (EREs) (Matsui et al., 2010; Rowe et al., 2010; Turelli et al., 2014). This affects expression of nearby genes, as KAP1 and associated effectors control TE-originating promoter or enhancer effects (Rebollo et al., 2012; Rowe et al., 2013b; Wolf et al., 2015). Furthermore, a few individual KRAB-ZFPs have been confirmed to repress retroelements in pluripotent cells, such as ZFP809 for murine leukemia virus (MLV) and its endogenous relatives (Wolf and Goff, 2007, 2009; Wolf et al., 2015), or Gm6871 and ZNF93 for mouse and human LINEs, respectively (Castro-Diaz et al., 2014; Jacobs et al., 2014). Although recent findings indicate that many KRAB-ZFPs have EREs as their preferential genomic targets (Najafabadi et al., 2015) (and our unpublished results), detailed functional data are missing about most members of the family.

Encoded in the hundreds by the genomes of higher vertebrates, KRAB-ZFPs first emerged more than 350 million years ago, and the expansion of this gene family subsequently mirrored waves of retroviral invasion into the genomes of their host species (Thomas and Schneider, 2011). During this process, *Krab-zfp* genes underwent strong positive selection at positions encoding for amino acids predicted to determine DNA-binding specificity, consistent with a role in countering rapidly mutating genetic invaders (Emerson and Thomas, 2009). Furthermore, the study of a couple of KRAB-ZFP/TE target pairs suggests the parallel evolution of restriction factors

and TE mutants escaping their inhibition (Jacobs et al., 2014). This led to the suggestion of a host-invader arms race, where KRAB-ZFPs were primarily selected to silence TEs.

In ESCs, the KRAB-ZFP-mediated docking of KAP1 and associated epigenetic modifiers at TEs triggers the deposition of repressive chromatin marks such as trimethylation of histone 3 on lysine 9 (H3K9me3), and methylation of CpG dinucleotides by de novo DNA methyltransferases (Rowe and Trono, 2011). Once established, DNA methylation is perpetuated across cell divisions, and its establishment at TEs during early embryogenesis is thought to result in permanent silencing, without need for persistent expression of their cognate KRAB-ZFP repressors (Quenneville et al., 2012; Wolf et al., 2015). Cumulated evidence indicates that DNA methylation is indeed an important mechanism of TE control in somatic tissues (Hutnick et al., 2010; Jackson-Grusby et al., 2001; Rowe et al., 2013a; Walsh et al., 1998). However, we previously observed that a significant fraction of TEs bound by KAP1 in human ESCs still carries the corepressor in mature T lymphocytes (Turelli et al., 2014), and that *Kap1* deletion in neuronal progenitors activates some EREs (Fasching et al., 2015). A few ERVs are similarly induced in murine B lymphocytes and fibroblasts deleted for *Setdb1*, the histone methyltransferase associated with KAP1 (Collins et al., 2015; Wolf et al., 2015). Moreover, many KRAB-ZFPs are expressed not only in ESCs but also in a variety of tissues (Barde et al., 2013; Corsinotti et al., 2013; Lizio et al., 2015).

In order to investigate KRAB-ZFPs/TE interactions, we developed a functional screen to identify KRAB-ZFPs responsible for recognizing specific DNA sequences. This led us to characterize two members of the family, ZFP932 and its paralog Gm15446, which we found to regulate distinct subsets of endogenous retrovirus K (ERVK) in murine ESCs. Invalidating current models (Maksakova et al., 2008; Mikkelsen et al., 2007; Rowe and Trono, 2011; Wolf et al., 2015), we further determined that these two KRAB-ZFPs also regulate their TE targets in differentiated tissues, through histone-based mechanisms not always correlated with the DNA methylation status of these loci. Furthermore, the dynamic control of these TEs by their KRAB-ZFP repressors modulated the expression of cellular genes in several adult tissues examined, both in cell culture and in vivo. We conclude that TEs and their KRAB-ZFP controllers are broad regulators of cellular gene expression, likely engaged in influencing multiple aspects of the biology of higher species.

RESULTS

A Functional Screen Identifies KRAB-ZFPs Repressing Specific DNA Targets

In order to identify KRAB-ZFPs responsible for regulating specific TEs, we developed a large-scale functional screen based on the repression of a reporter cassette through the intermediate of nearby DNA baits (Figure 1A). Putative KRAB-ZFP-binding sequences were selected based on the coincidence of KAP1, SETDB1, and H3K9me3 peaks in chromatin immunoprecipitation/deep sequencing (ChIP-seq) studies of murine ESCs (Bilodeau et al., 2009; Rowe et al., 2013b). Sequences contained within TEs or zinc finger protein-coding genes were favored, and imprinting control regions (known for their recruitment of ZFP57) were excluded. We selected 19 such targets (Table S1), and added the $PBS^{Lys1,2}$ sequence, the primer binding site (PBS) sequence of many retroviruses, demonstrated to drive KAP1-mediated repression, but through an unknown KRAB-ZFP intermediate (Wolf et al., 2008). Lentivectors containing these baits upstream of a PGK-GFP cassette were first introduced into murine ESCs, which revealed that 10 of 20 of these TE-derived fragments induced KAP1-dependent repression of GFP expression (Figure S1). The corresponding vectors were used to engineer stable 293T cell lines, which were transfected in triplicates with individual members of a library of 211 murine KRAB-ZFPs (Table S2) and examined 6 days later for GFP expression with a fluorescence plate reader; the hits were then confirmed by fluorescence-activated cell sorting (FACS) analysis. A control cell line, transduced with a vector containing the MLV PBS^{Pro} sequence upstream of the PGK-GFP cassette, properly singled out ZFP809 as its cognate repressor, giving us confidence in our approach (Figure 1B). Of the ten tested DNA sequences, three allowed the identification of a specific KRAB-ZFP ligand (Figures 2A and S2A–S2C). The P9 bait, which comprises a SINE-B2 element, was repressed by Gm6871. The P5 and P8 baits, both derived from ERVKs (RLTR44-int for P5; RLTR9A3, MMERVK10D3_I-int, and RLTR6-int for P8), were repressed by ZFP932. To confirm the matches thereby identified, we introduced the DNA bait-PGK-GFP lentiviral vectors in murine ESCs depleted for the individual KRAB-ZFPs by RNAi. For P5, P8, or P9, depleting the corresponding KRAB-ZFP released GFP expression, while it had no effect on a vector containing the PBS^{Pro} control sequence (Figures 2B and S2D). Furthermore, repression could be restored by overexpressing an shRNA-resistant form of the specific KRAB-ZFP for P9 and P8 (significantly), and to a lesser extent for P5 (not significantly, possibly due to initial lower values). Finally, we could document the bindings of Gm6871 to the P9 sequence, and of ZFP932 to both of its genomic targets by chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR) (Figures 2C and S2E). Interestingly, Gm6871 also binds some LINE-1 elements, which it can repress in murine ESCs (Castro-Diaz et al., 2014). In Gm6871-specific ChIP-seq analyses, not only LINEs but also the SINE-containing P9 locus was detected (Figure S2B), indicating that a KRAB-ZFP can recognize TEs from different subgroups.

ZFP932 and Its Paralog, Gm15446, Bind to Distinct Subsets of ERVK Endogenous Retroelements

We then turned to the biological characterization of *Zfp932*. We realized that this mouse-specific KRAB-ZFP gene has a paralog, *Gm15446*, located in close proximity on chromosome 5 (Figure 3A) and also targeted by our *Zfp932*-directed siRNAs. When examining the predicted DNA-contacting amino acids within the zinc finger arrays (Liu et al., 2014) of ZFP932 and Gm15446, we saw that the two proteins differed by only two point mutations and the presence of one extra zinc finger in Gm15446. In 293T cells, overexpression of either protein led to repression of the same targets, albeit with some differences in efficiency (Figure S3A).

We next investigated by ChIP-seq the genomic targets of ZFP932 and Gm15446 in murine ESCs overexpressing HA-tagged derivatives of these proteins (Figure S3B). The intersection of duplicate ChIP-seq experiments yielded a total of 755 peaks for ZFP932 and 1,053 for Gm15446, among which 401

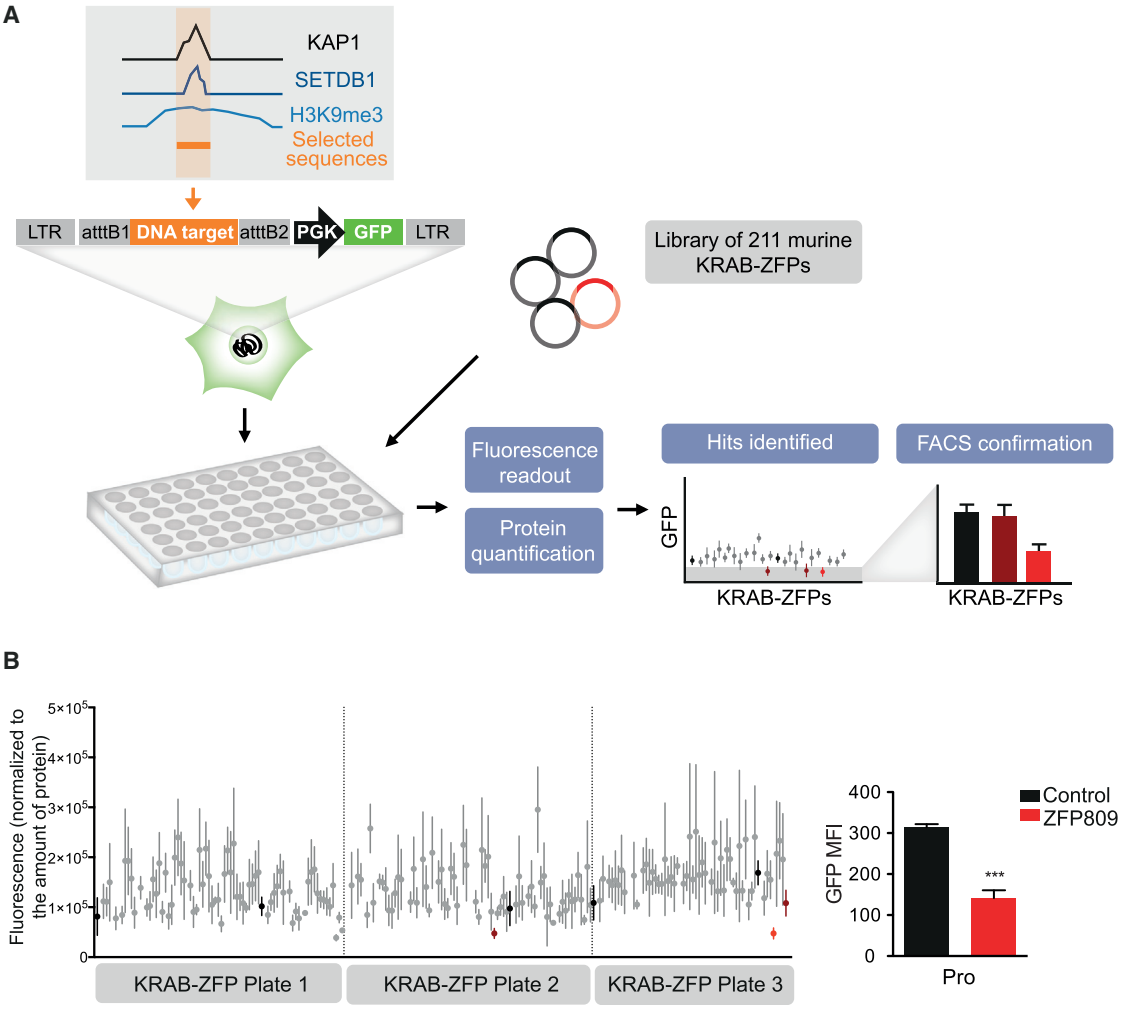


Figure 1. A Functional Screen to Match KRAB-ZFP Repressors with Their Genomic Targets

(A) Methodological outline. A library of murine KRAB-ZFP-expressing plasmids was transfected in 293T stable cell lines containing the DNA sequence of interest upstream a PGK-GFP cassette. Fluorescence readout in 96-well plates was normalized for protein content, and hits were tested by FACS for confirmation.

(B) Validation of the screen using the previously characterized ZFP809/PBS^{Pro} (Pro) pair. Fluorescence readout was performed (left) and identified hits were tested by FACS. The only hit confirmed by FACS was ZFP809 (right panel). Black, transfection control; red, hits identified by plate reader; light red, hits confirmed by FACS. Error bars represent SD, ***$p < 0.001$, Student's t test.

See also Figure S1, Table S1, and Table S2.

were shared (Figure 3B). Overlapping these data with genic regions and using our updated census of repeats in the mouse genome (see Supplemental Experimental Procedures for details), we observed that most of the KRAB-ZFPs peaks were on TEs (Figure 3C), either in or outside genes, the majority of which were ERVKs (68.7% of peaks for ZFP932 and 91.5% for Gm15446). To decrease the possibility of unspecific targets due to overexpression of the proteins, we overlapped the results of these ChIP-seqs with that of similar analysis performed on endogenous KAP1 in murine ESCs. In total, 226 peaks for ZFP932 and 448 for Gm15446 overlapped with KAP1 peaks, mostly at ERVKs, confirming that these KRAB-ZFPs recognize this ERE. Interestingly, some differences were noted between the two paralogs (Figure 3D). For instance, ZFP932 and Gm15446 were similarly enriched at RLTR44-int, IAP-d-int, and MMERVK10D3_I-int elements, but Gm15446 was more frequently found at MMERVK10C-int, IAPEy-int, and IAPEY3-int. Using ChIP-seq data, we identified a motif present in 80% of the ZFP932- and Gm15446-binding sites and 82.7% of ZFPs-enriched ERVKs, but absent from members of the ERV1 family, not targeted by these KRAB-ZFPs (Figure 3E). The underlying sequence is different from a ZFP932-recruiting motif identified within the *Ptch1* gene promoter in a limb mesenchymal cell line (He et al., 2011). However, we did not detect ZFP932/Gm15446 at this site. It could be that *Ptch1* binds ZFP932 only in particular cellular contexts, and via another protein.

We then mapped more precisely the ZFP932/Gm15446/KAP1 binding sites within the genome of their ERV targets. MLV and some IAPs (intracisternal A particle, a subtype of ERV) were previously found to recruit KAP1 via the proximal part of their provirus, consistent with the observed repression of their nearby 5′ LTR (Rowe et al., 2010; Wolf and Goff, 2009; Wolf et al., 2015).

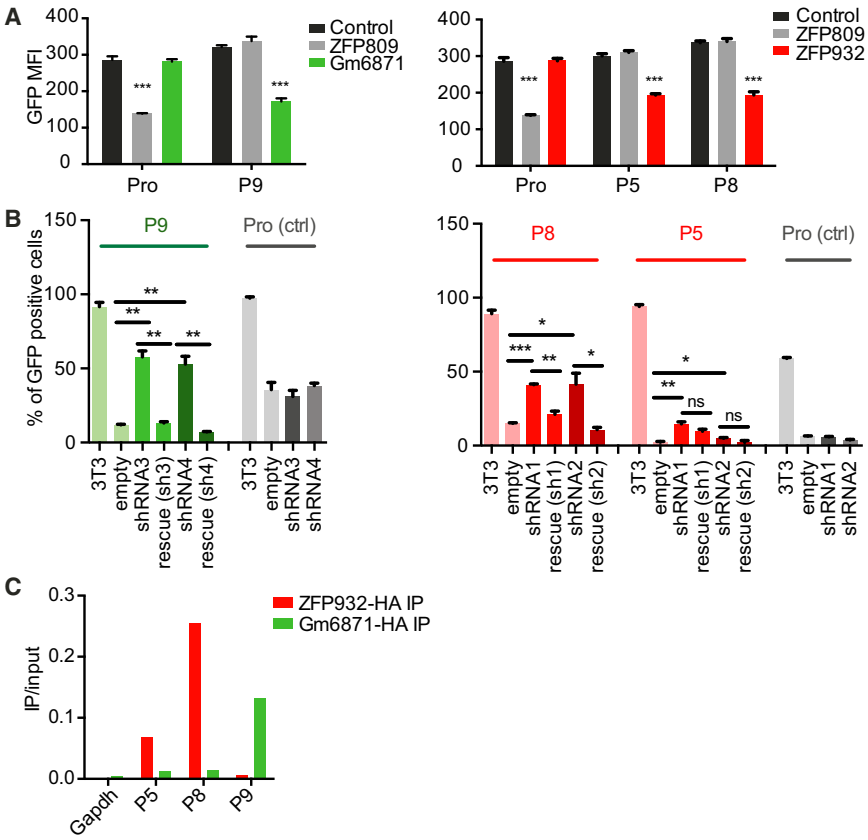


Figure 2. Identification of Gm6871 and ZFP932 as Ligands of KAP1-Repressed TE Sequences

(A) FACS confirmation of hits identified through screening of selected target sequences.

(B) GFP repression assay in control, ZFP932- or Gm6871-KD, or ZFP-complemented murine ESCs.

(C) HA ChIP-PCR of Gm6871 and ZFP932 in murine ESCs overexpressing the corresponding tagged protein (results representative of three independent experiments).

Error bars represent SD, *$p < 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ns, not significant, Student's t test. See also Figure S2.

Surprisingly, both ZFP932 and Gm15446 were instead enriched near the 3′ end of their target retroelements, together with KAP1 and H3K9me3 (Figure 3F). An analysis using multiple alignments of MMERVK10C-int elements revealed that the two paralogs and KAP1 bound a region situated just after *env*, upstream of the 3′ LTR, partly overlapping with the 3′ polypurine tract (PPT), an element important for reverse transcription (Figure 3G). Interestingly, KAP1 was also enriched at two 5′ sites in these retroelements, where it was likely recruited by other KRAB-ZFPs since neither of the two paralogs was found there. Finally we asked more broadly if KAP1 binds to multiple regions of retroelements other than ZFP932/Gm15446 targets, and we identified several KAP1 peaks on IAPEz, another subtype of ERVs (Figures S3C and S3D). These results demonstrate that KAP1 can bind to multiple locations of the same TE, including its 3′ end, via different KRAB-ZFPs.

ZFP932 and Gm15446 Regulate ERVK and Nearby Gene Transcription in Murine Embryonic Stem Cells

To gain insight on the transcriptional impact of these two KRAB-ZFPs, we next deleted *Zfp932* and *Gm15446* from murine ESCs by CRISPR/Cas9-mediated editing of the locus and defined the transcriptome of the resulting cells by RNA-seq (Figure 4A). Mainly members of the LTR/ERV class of TEs were upregulated, consistent with the ChIP-seq data. Overlapping these results with those of RNA-seq performed in *Kap1* knockout (KO) ESCs (Figure S4A) unveiled 141 ERV, 17 LINE, and 27 SINE integrants, the expression of which was significantly increased in both settings. ERVs that had been identified as ZFP932/Gm15446 binding targets were more upregulated in cells depleted for either protein or KAP1, compared with ERVs on which these KRAB-ZFPs and KAP1 had not been detected (Figure 4B). Of the bound ERVs with increased expression upon ZFPs depletion, many had only the 3′ PPT KAP1 peak, or had a 3′ peak markedly more pronounced than its 5′ counterparts.

It was previously demonstrated that the KRAB/KAP1-mediated control of endogenous retroelements preserves the transcriptional dynamics of ESCs by preventing TE-originating enhancer and promoter effects on nearby genes (Rowe et al., 2013b; Wolf et al., 2015). Seventy-one genes were significantly upregulated upon *Zfp932/Gm15446* KO, and 29 of those were significantly more expressed also in *Kap1* KO murine ESCs, six of which were within 50 kb of an ERV bound by both the KRAB-ZFPs and their cofactor. Among these genes, *Bglap3* stood out as one of the most upregulated genes in the *Krab-zfps* KO cells (fold change of 270.9, $p = 0.00000003$). *Bglap3* has three short and two long transcripts, the latter interestingly containing a ZFP932/Gm15446-bound IAP-d element (Figure 4C). By comparing FANTOM CAGE transcriptional start site (TSS) data with RNA-seq, we observed that both the 5′ and the 3′ LTRs of the IAP function as promoters of the *Bglap3* gene, with transcripts starting exactly at the beginning of the LTR R region. Both *Bglap3* and its IAP were more expressed in murine ESCs deleted for either *Zfp932/Gm15446* or *Kap1*, demonstrating that the retroelement partakes in the controlling unit of this gene (Figures 4C, S4B, and S4C). This effect was confirmed through ZFP932/Gm15446 knockdown (KD) experiments, properly controlled by restoration of the repression upon complementation with RNAi-resistant derivatives of the two KRAB-ZFPs (Figure S4D). The *Rgs20* locus provided another example of gene-TE pair regulated by ZFP932/Gm15446, in which no TSS overlapped with the TEs, suggesting that this gene is regulated through enhancer effects (Figures 4C and S4B).

We deepened our study of this phenomenon by examining the epigenetic status of the *Bglap3* locus (Figure 4D). Upon *Zfp932/Gm15446* deletion, ChIP-PCR analyses revealed loss of KAP1

and H3K9me3 and gain of the activation marks H3K27ac and H3K4me1 at the IAP, together with the accumulation of PolII at the predicted TSSs of *Bglap3*. Furthermore, pyrosequencing unveiled a marked drop in DNA methylation both underneath the KAP1 peak and at the 5′ end of the IAP. These data indicate that, in murine ESCs, ZFP932/Gm15446 epigenetically regulate not only the expression of their target ERVKs but also the transcriptional influence of these TEs on nearby genes, via both promoter and enhancer effects.

The KRAB/KAP1 System Regulates TEs in Differentiated Cells

To gain further insight into *Zfp932/Gm15446* functions, we assessed their expression in representative murine tissues, using CAGE datasets generated by the FANTOM consortium (Lizio et al., 2015) (Figure 5A). Both paralogs were broadly expressed in adult murine cells, as was *Zfp809*. In contrast, *Zfp459* and *Zfp819*, which we previously found to be highly transcribed in murine pluripotent stem cells (Corsinotti et al., 2013), were largely restricted to undifferentiated cells. Because ZFP932 and Gm15446 displayed a strong preference for ERVK family members as their genomic targets, we suspected that they might control these TEs in adult tissues as well. To probe this hypothesis, we first examined by RNA-seq the transcriptomes of hepatocytes harvested from control and liver-specific *Kap1* KO mice, wild-type and *Kap1*-deleted mouse embryonic fibroblasts (MEFs) (Figure S5A), and control and KAP1 KD C2C12 mouse myoblast cells. In all these settings, expression of ERVs was increased upon KAP1 depletion, even though induction was less pronounced than in *Kap1* KO ESCs (Figure 5B). We ranked the significantly upregulated families in each dataset and found that MERVK10C-int had the highest number of induced integrants in KAP1-depleted differentiated cells (Figure S5B). General MERVK10C-int primers confirmed by RT-qPCR the increased expression of these elements in liver and C2C12 depleted for KAP1, in spite of the absence of detectable changes in the DNA methylation status of these loci as measured by bisulfite sequencing (Figures S5C and S5D). Interestingly, there were differences in the subsets of ERVs deregulated in the various tissues following KAP1 removal (Figure 5B). For example, particular MMERGLN and ORR1A0 integrants located on chromosome 12 were upregulated in KAP1-depleted liver and ESCs, but not in their C2C12 counterparts (Figure S5C).

About 24% of ERVs bound by KAP1 in ESCs still bore the corepressor in C2C12 cells and MEFs (Figure 5C), confirming that the KRAB/KAP1 system is involved in the somatic control of TEs. Interestingly, for 30% of the ERV loci bound by KAP1 in C2C12 and 24% in MEFs no significant enrichment was found in ESCs. We also performed ZFP932- and Gm15446-specific ChIP-seq analyses in C2C12 cells expressing HA-tagged forms of these proteins (Figures 5D and S5E). We detected 169 ZFP932 and 209 Gm15446 binding sites, 98 of which were common, and the large majority of which were also bound by either one or both of these KRAB-ZFPs in ESCs. Most resided within TEs, namely ERVKs, and KAP1 was also detected at about a third of them, supporting a role for these ZFPs in regulating such elements in differentiated cells. We then asked if H3K9me3 marks were present at KAP1-bound ERVKs in C2C12. Using H3K9me3 ChIP-seq data in these cells, we observed a higher correlation of H3K9me3 at ERVKs with than without KAP1 (Figure 5E). General ERVK primers revealed upregulation of MERVK10C, IAP-d, and IAPEz in C2C12 cells depleted for SETDB1, suggesting that this histone methyltransferase partners with KAP1 also in differentiated cells to control TEs (Figure S5F). Finally, ERVKs that were bound by KAP1 in C2C12 or MEFs were significantly more expressed than their unbound counterparts when the corepressor was depleted (Figure 5F). Taken together, these results demonstrate a role for the KRAB/KAP1 system in the somatic control of TEs.

TEs and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues

We then asked if the KRAB/KAP1-mediated control of TEs also prevents the illegitimate transcription of nearby genes in adult cells. For this, we first analyzed our RNA-seq of KAP1-depleted cells, separating the genes in stable, up- or downregulated (deregulation criteria were 2-fold change cut-off and adjusted p value $\leq$ 0.05), and calculated their distance to closest KAP1-bound ERV. As expected, in ESCs, upregulated genes were significantly closer to a KAP1-bound ERV than their downregulated or stable counterparts (Figure 6A). In the case of C2C12 KAP1 KD, although the same trend was observed, upregulated genes were only significantly closer to KAP1-bound ERVs when compared with stable but not downregulated genes. Interestingly, for MEFs *Kap1* KO, we observed the same pattern as in ESCs, suggesting that, even if to a smaller extent, KAP1-regulated TEs can affect neighboring genes transcription also in differentiated cells.

We confirmed our hypothesis further by examining in more detail a few upregulated genes situated within 100 kb of a KAP1-bound ERV. At some loci, such as *Fbxw19*, KAP1 depletion resulted in activating both the ERV and its nearby gene in all cells examined (Figures 6B, S6A, and S6B). At others, such as *Gm13251*, removal of the regulator induced expression in liver and in ESCs but not in C2C12 cells (Figure S6B). We then examined the *Bglap3* locus, which we found enriched for ZFP932/Gm15446, KAP1, and H3K9me3 also in C2C12 cells (Figure S6C). Upon RNAi-mediated KD of these two KRAB-ZFPs, both the IAP and *Bglap3* were activated (Figure 6C), coincident with loss of H3K9me3 and KAP1, and mild but significant gain of H3K27ac at the IAP (Figures 6D and S6D). However, in spite of its transcriptional activation, this ERV remained DNA methylated (Figure 6E).

To expand these findings in vivo, we generated ZFP932/Gm15446 KD mice by transduction of fertilized oocytes with shRNA-expressing lentiviral vectors. We could not assess major phenotypic abnormalities in these animals, but a marked upregulation of *Bglap3* gene and the upstream IAP was measured in their liver, albeit again with no detectable loss of DNA methylation at this locus (Figures 6F and S6E). ChIP-PCR experiments in murine wild-type liver confirmed the enrichment of KAP1 and H3K9me3 at the IAP (Figure S6F). Interestingly, in the liver of conditional *Alb-Cre Kap1* KO mice, in which the regulator is deleted not during early embryogenesis but some 3 weeks after birth, there was increased expression of *Bglap3* but not of the IAP, suggesting an uncoupling of the regulation of the two LTRs of the retroelement (Figure 6F). Also in this setting, we observed that the IAP Bglap remained highly methylated (Figure S6G).

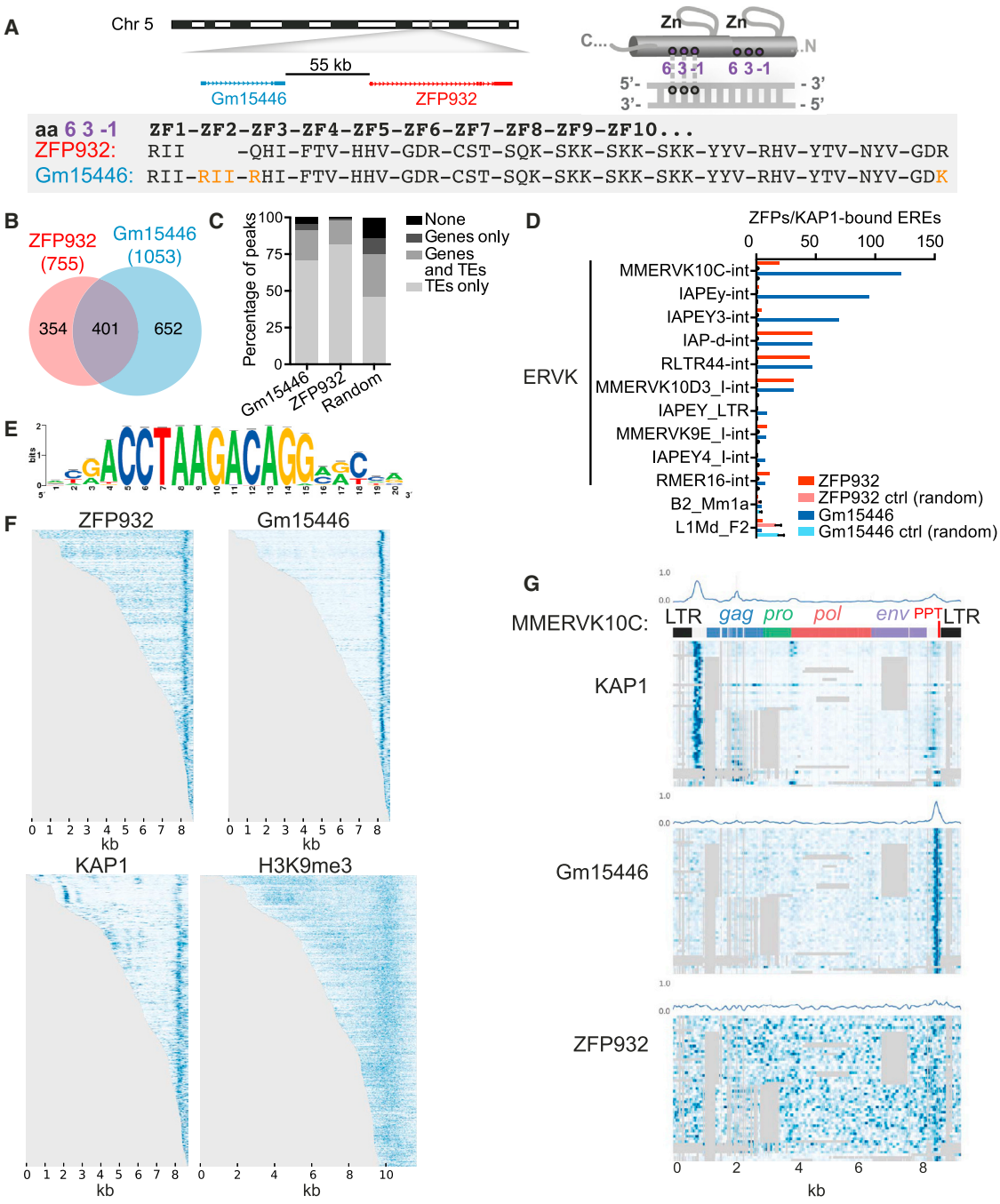


Figure 3. ZFP932 and Its Paralog Gm15446 Bind with KAP1 to the 3′ End of Distinct Subsets of ERVK

(A) Representation of genomic location of *Zfp932* and *Gm15446* (left), and scheme of the three amino acids important for DNA recognition on a zinc finger domain (right). Lower panel depicts the comparison of the three amino acids predicted to contact DNA of each zinc finger of ZFP932 and Gm15446, highlighting differences in orange.

(B) Venn diagram of ZFP932 and Gm15446 binding sites determined by ChIP-seq in murine ESCs.

(C) Percentage of ZFP932, Gm15446, or random control peaks on genic and repeated regions. Random control is based on the mean of 100 random shuffling of the peaks of Gm15446 (KRAB-ZFP ChIP with more peaks).

(D) Top ten ERE families bound by ZFP932 or Gm15446 together with KAP1 in murine ESCs. SINEs and LINEs (bottom) served as negative controls. A control with 100 random shuffling of the binding sites is also shown.

(E) Predicted DNA-binding motif for ZFP932 and Gm15446.

(F) ChIP-seq coverage plots in murine ESCs on all EREs bound by ZFP932 or Gm15446, ranked by size. Each row is independently normalized, with enrichment proportional to the darkness of the blue color. For H3K9me3, representation extends 1 kb up and downstream of the ERE boundaries.

(legend continued on next page)

Finally, we asked if the *Bglap3* gene is expressed in physiological conditions and how does it correlate with the ZFPs transcription. By analyzing CAGE expression data we observed that, in tissues like placenta, mammary gland, and pancreas, *Bglap3* transcripts are expressed and correlate with lower *Zfp932*/*Gm15446* levels (the long transcript driven by the IAP 5′ LTR, and the short transcript by the 3′ LTR) (Figure 6G). Collectively, these data establish that the KRAB/KAP1 system uses TE-residing platforms to regulate gene expression in adult cells.

DISCUSSION

This work establishes that TEs and their KRAB-ZFP controllers can regulate gene expression in adult tissues. Our results thus invalidate a generally accepted model that assumes that the KRAB/KAP1 system irreversibly silences TEs during early embryonic development (Maksakova et al., 2008; Mikkelsen et al., 2007; Rowe and Trono, 2011; Walsh et al., 1998; Wolf et al., 2015). Furthermore, our study strongly suggests that *Krab-Zfp* genes are not selected simply to inactivate TEs but also to domesticate their transcriptional potential for the benefit of the host.

Here, through a target-centric functional screen, we identified several KRAB-ZFPs responsible for the sequence-specific repression of TEs in murine ESCs, including two KRAB-ZFP paralogs that recognized partly overlapping members of the ERVK family. Not only could we verify that these two KRAB-ZFPs controlled these ERVs and neighboring genes in ESCs but we further discovered that they were expressed in a broad range of somatic murine cells. We could also demonstrate that, in several differentiated cell types, including in vivo in the liver, KAP1 and KRAB-ZFPs were still bound to and controlled TEs and modulated the expression of nearby genes via promoter and enhancer effects, even if with a smaller global impact than observed in ESCs. Remarkably, this process occurred without apparent alteration of the DNA methylation status of these loci but was primarily involved histone-based changes.

Although DNA methylation is traditionally considered as a silencing mechanism, emerging evidence indicates that its impact on gene expression is far more diversified (Jones, 2012; Schubeler, 2012). We had previously observed that docking of the KRAB/KAP1 complex at genomic loci leads to their DNA methylation when it occurs in early embryonic cells but not in their differentiated counterparts, where it induces repression solely by histone-based, hence fully reversible, changes (Groner et al., 2012; Quenneville et al., 2012; Wiznerowicz et al., 2007). The present data further indicate that protein-restricted chromatin modifications can reignite transcription at these sites in adult cells, in spite of persistent DNA methylation. This is consistent with recent observations that some ERVs are activated without changes in their DNA methylation status in B cells depleted for the histone methyltransferase SETDB1 (Collins et al., 2015), and that, in human ESCs, transcription can occur from some highly methylated KAP1-controlled promoters (Turelli et al., 2014).

Whether in the context of murine or human ESCs, we had previously noted that KAP1 depletion does not systematically trigger the activation of all KAP1-bound TEs (Rowe et al., 2010; Turelli et al., 2014). Here, we observed that, in adult cells as well, the range of ERVs activated upon KAP1 or ZFP932/Gm15446 depletion differed depending on the cellular environment. In C2C12, for example, many IAPEz elements bound by KAP1 were not significantly de-repressed upon its removal (not illustrated). At the *Gm13251* locus, KAP1 depletion activated two ERVs in liver and in ESCs, but not in MEFs or C2C12 cells. Correspondingly, in *Setdb1* KO B lymphocytes, the upregulation of selected ERVs correlated with the presence within their promoters of binding sites for B cell-specific transcription factors (Collins et al., 2015). It is likely that this type of restriction broadly applies to KRAB-ZFP-controlled loci, the expression of which is conditioned not only by removal or biochemical inactivation of their epigenetic repressors but also by the presence of a proper set of activators, likely tissue specific. In addition, specific loci may be subjected to dominant influences imposed by the local chromatin configuration, and other control mechanisms, such as small RNAs based, may be at play in some environments (Bierhoff et al., 2014; Heras et al., 2013; Marchetto et al., 2013; Pezic et al., 2014). Finally, KAP1 can undergo locus-specific post-translational modifications switching its function from corepressor to coactivator (Singh et al., 2015), and KRAB-ZFPs could also be subjected to this type of regulation, leading to the recruitment of different sets of effector complexes.

Our demonstration that KRAB-ZFPs are involved in controlling TEs not only in embryonic but also differentiated cells is consistent with the widespread expression of these proteins in adult organisms, whether mouse or human (Lizio et al., 2015), and with the finding that a great majority of these transcription factors have TEs as their preferred genomic targets (Najafabadi et al., 2015) (and our unpublished results). It also contributes to explain the broad range of phenotypes induced by the conditional KO of *Kap1* in the mouse, even though it should be kept in mind that the master regulator also carries out some KRAB-ZFP- and TE-independent functions (Iyengar and Farnham, 2011; Iyengar et al., 2011; McNamara et al., 2016; Singh et al., 2015).

Our study unveils several other unsuspected aspects of the KRAB/KAP1-mediated control of TEs. First, we found that a given KRAB-ZFP can recognize different subgroups of TEs, e.g., Gm6871 repressing members of both the LINE and SINE families. Second, we determined that the two ERVK-targeting KRAB-ZFPs identified here tether KAP1 near the 3′ end of their retroviral targets, not at the PBS region or close to their promoter as previously documented for MLV, several other IAPs, and most LINEs (Castro-Diaz et al., 2014; Jacobs et al., 2014; Rowe et al., 2010; Wolf and Goff, 2009; Wolf et al., 2015). Interestingly, this 3′ end peak overlaps with the 3′ PPT, a sequence important for reverse transcription. While it readily explained how ZFP932 and Gm15446 can control the impact of the IAP Bglap 3′ LTR on the *Bglap3* gene, it was more surprising to observe that this distal location also served to repress the 5′ LTR of this element,

(G) Coverage plot on multiple alignment of different ChIP-seqs in ESCs on "full-length" (>5 kb) MMERVK10C-int bound by ZFP932 or Gm15446. Repbase MMERVK10C-int consensus is represented on top. Mean of binding coverage is illustrated in each plot. Each row is independently normalized, with enrichment proportional to the darkness of the blue color. Gray areas correspond to gaps in multiple alignments. Error bars represent SD.
See also Figure S3.

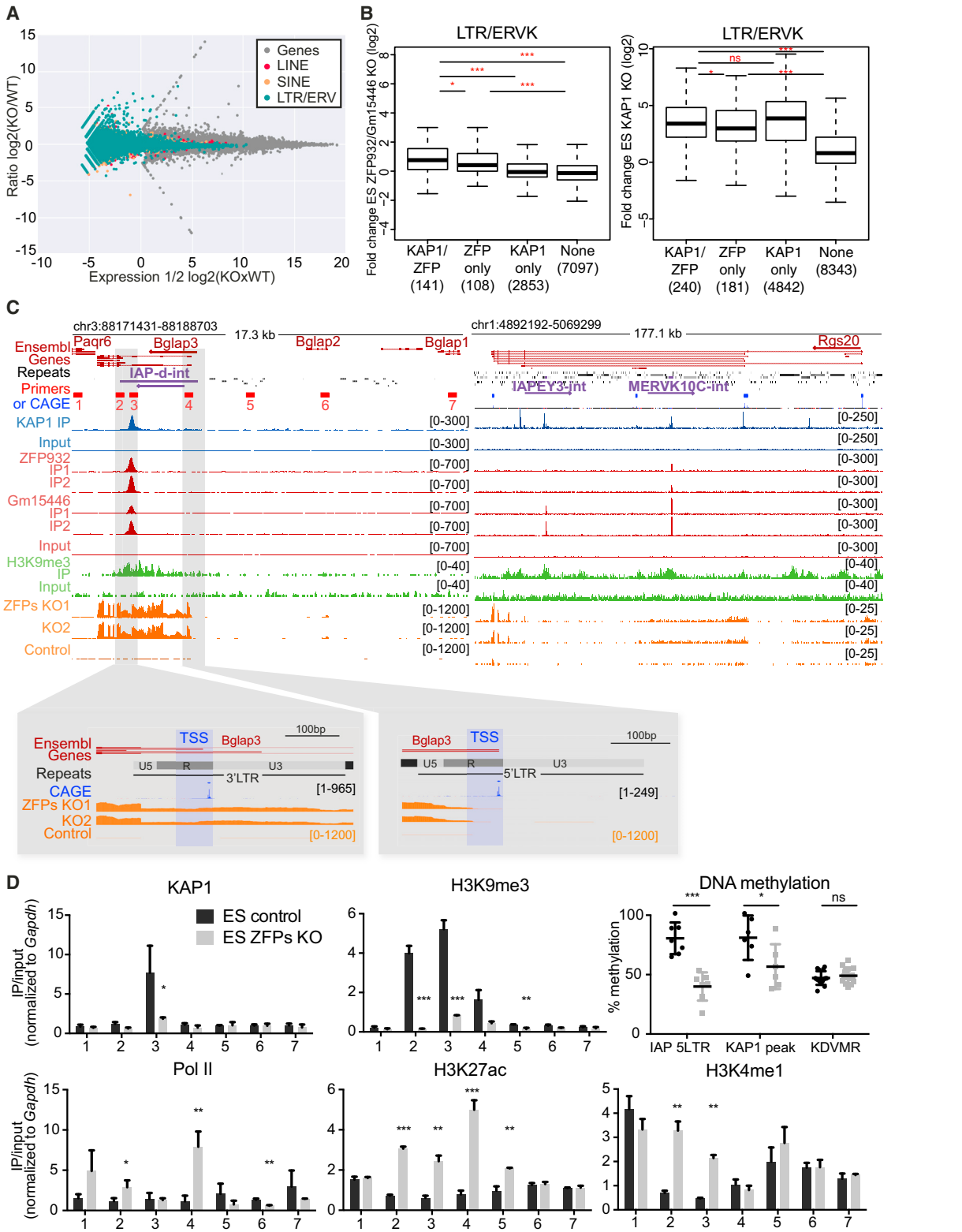


Figure 4. ZFP932/Gm15446/KAP1 Binding Regulates Expression of ERVs and Nearby Genes in Murine ESCs

(A) MA-plot of RNA-seq from *Zfp932/Gm15446* KO versus control ESCs. Expression of genes and TEs is shown.

(B) Boxplots of ERVK fold expression in *Zfp932/Gm15446* KO (left) or *Kap1* KO (right) ESCs.

(legend continued on next page)

as demonstrated by the spreading of chromatin marks and its upregulation in ESCs or mice where these two KRAB-ZFPs were depleted. Third, we observed the differential regulation of the two LTRs of the same ERV, and consequently of the gene placed under their influence, via one KRAB/KAP1-binding site. Depletion of ZFP932/Gm15446 from the earliest times of embryonic development activated the expression of both *Bglap3* and its controlling IAP in the liver of adult mice, whereas the deletion of *Kap1* in this organ after birth only de-repressed the *Bglap3* gene, but not its IAP. It thus seems that the regulation of the 5′ and 3′ LTR of this endogenous retrovirus can be uncoupled. In the absence of obvious differences in the DNA methylation status of these two transcription units, the molecular basis of their distinctive behavior remains to be identified. We could even speculate that, as previously reported (Wolf et al., 2015), the depletion of KRAB/KAP1 regulation at different moments of development could lead to differential TE control due to deposition of other epigenetic marks. Finally, we found that KAP1 can be recruited to several regions of the same ERV, likely via distinct KRAB-ZFPs. It strongly suggests that a TE can be regulated in temporally and functionally differential fashions, hence that the complexity of KRAB/KAP1-mediated regulation of TEs and their gene targets is much greater than envisioned so far.

Our results support a model in which the evolutionary selection of KRAB-ZFP genes is not only an arms race aimed at silencing TEs but also the instrument of their domestication. It is likely that TEs and their KRAB-ZFPs regulators modulate multiple aspects of the biology of tetrapods, superimposing a species-specific layer of control over canonical, conserved regulatory pathways. Indeed, a large fraction of recognizable mobile elements in the genome are unique to the corresponding species and its close relatives, both in sequence and location; accordingly, levels of orthology between KRAB-ZFPs of different organisms are limited. Therefore, KRAB-ZFPs and their targets must play major roles in the speciation of higher vertebrates, including humans.

EXPERIMENTAL PROCEDURES

Cell Culture and Mouse Work

Murine ESCs and MEFs wild-type and KO for *Kap1* were cultured and generated as previously described (Rowe et al., 2013a, 2013b). Mouse C2C12 myoblasts were cultured as described (Singh et al., 2015). Two clonal *Krab-zfps* KO cell lines were generated using an integrase-defective lentiviral vector containing the pLentiCRISPR with an sgRNA for *Zfp932* and *Gm15446* or a control luciferase sgRNA. KD experiments were performed with specific shRNA in pLKO lentiviral vectors. Hepatocyte-specific *Kap1* KO mice were generated and genotyped according to Bojkowska et al. (2012). ZFP932/Gm15446 KD mice were generated by lentiviral transgenesis as described (Rowe et al., 2013a). All shRNAs and sgRNAs used in this study are listed in Supplemental Experimental Procedures. All animal experiments were approved by the local veterinary office and carried out in accordance with the EU Directive (2010/63/EU) for the care and use of laboratory animals.

Functional Screen

DNA target sequences to be tested were chosen from the overlap of publicly available murine ESC ChIP-seq data for KAP1 (GEO: GSE41903), SETDB1 (GEO: GSE18371), H3K9me3 (GEO: GSE41903), and absence of ZFP57 (GEO: GSE31183). The $PBS^{Lys1,2}$ sequence plus 19 sequences corresponding to KAP1 peaks were selected based on the presence of TEs or of interesting KAP1 targets (such as the 3′ end of ZFP genes) (see Table S1). These sequences were cloned upstream of a PGK-GFP cassette, tested for repression in murine ESCs, and only repressed sequences were tested in the screen. Lentiviral vectors with these sequences were used to transduce 293T cell lines, which were sorted for GFP presence, generating stable cell lines. The screen was performed by reverse transfection of a library of 211 murine KRAB-ZFPs (Table S2), in 96-well plates, in an automated fashion, in triplicates. Cells were harvested on day 6 after transfection, lysed, and GFP fluorescence was measured by plate reader (excitation at 485 nm and emission at 520 nm). Total protein content was quantified using BCA (BCA Protein Assay Reagent, Thermoscientific), and used to calculate normalized GFP fluorescence. Candidate hits were identified by selecting the ten KRAB-ZFPs with the lowest normalized fluorescence values per 96-well plate, and only the ones that were present in all three replicates were considered. Hits were identified by transfection of the 293T cell line of interest with the candidate KRAB-ZFPs, in 24-well plates, with FACS readout after 6 days. The detailed protocol is given in Supplemental Experimental Procedures.

RT-PCR and RNA-Seq

Total RNA was extracted and DNase-I treated using a spin column-based RNA purification kit (Macherey-Nagel). cDNA was prepared with SuperScript II reverse transcriptase (Invitrogen). Primers (listed in Supplemental Experimental Procedures) were used for SYBR green qPCR (Applied Biosystems), and specificity was confirmed with dissociation curves. For mRNA sequencing, 100-bp single-end RNA-seq libraries were prepared using the Illumina TruSeq Stranded mRNA reagents (Illumina). Cluster generation was performed with the resulting libraries using the Illumina TruSeq SR Cluster Kit v4 reagents. Sequencing was performed with Illumina HiSeq 2500 in 100-bp reads run. The RNA-seq reads were mapped to the mm9 genome using TopHat (Kim et al., 2013), allowing multimapped reads to be randomly assigned once among the mapped loci. Gene counts were generated with the HTseq-count program using default parameters, and TE counts were computed using BEDTools (multicov). Sequencing depth normalization and differential expression analyses were performed using the voom function of the Bioconductor package LIMMA (Law et al., 2014).

ChIP-PCR and ChIP-Seq

ChIP and library preparation were done according to Rowe et al. (2013b), with modifications as described in Supplemental Experimental Procedures. Sequencing was performed with Illumina HiSeq 2500 in 100-bp reads run. Reads were mapped to the mouse genome assembly mm9 using Bowtie 2 (Langmead and Salzberg, 2012), using the sensitive-local mode. Peaks were called using MACS (Zhang et al., 2008) or SICER software (for histone modification marks) (Zang et al., 2009), with total input as control.

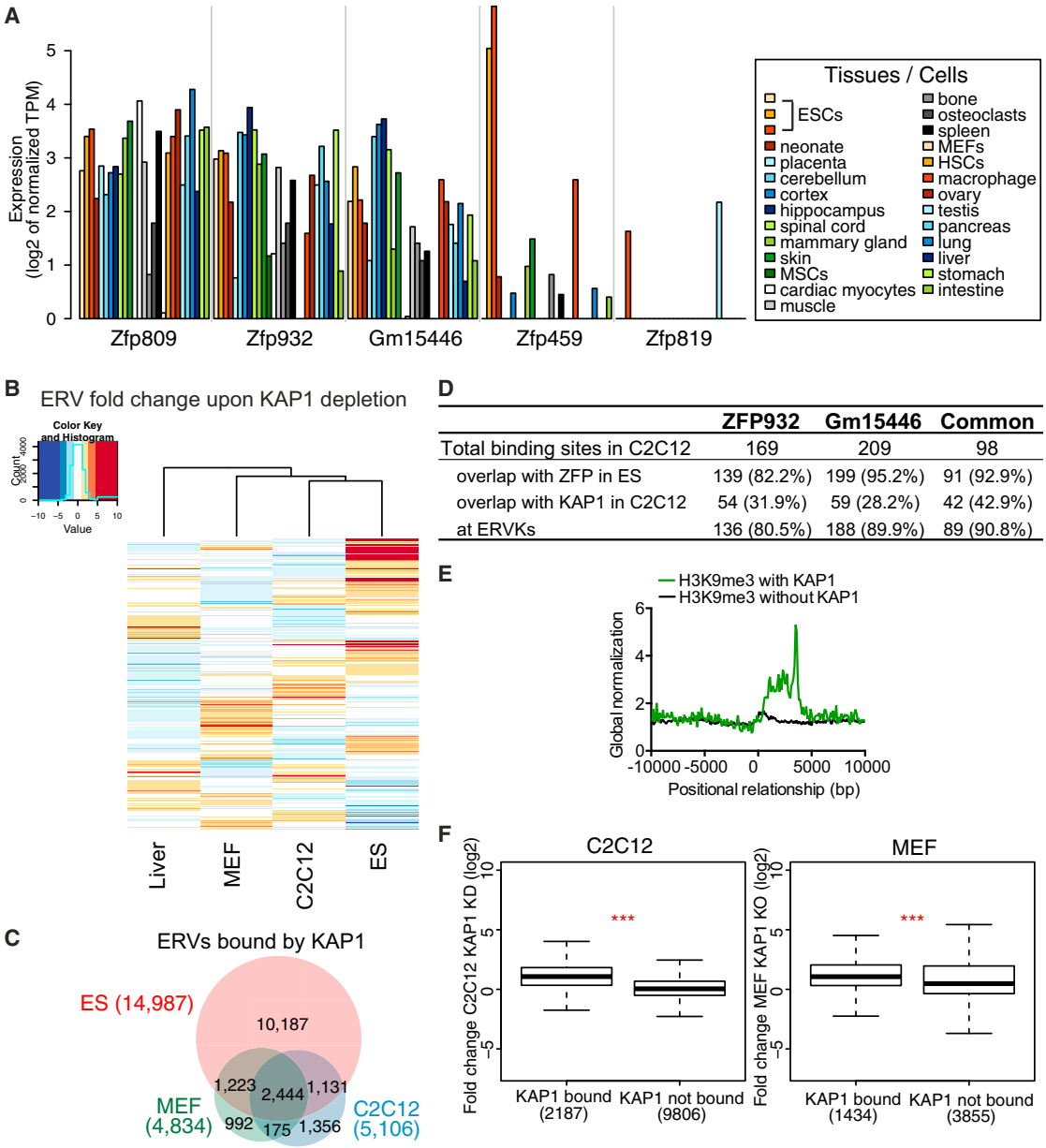
DNA Methylation Analysis

Genomic DNA was extracted, converted using an Epitect Bisulfite kit (Qiagen), and used in two rounds of PCR followed by PCR product purification. Pyrosequencing and bisulfite sequencing were performed as previously described (Fasching et al., 2015; Rowe et al., 2013a).

(C) Scheme of *Bglap3* (left) and *Rgs20* (right) genomic loci. ChIP-seq and RNA-seq signal are depicted, along with the primers used in (D) (numbered 1–7) and FANTOM CAGE data corresponding to TSSs. Orientation of the genes and TEs is indicated with an arrow. Repeats of interest are highlighted in purple. Lower panels correspond to two regions of zoom in the *Bglap3* locus, indicating the TSSs of the gene. CAGE data correspond to mapped TSS peaks and to the max signal of CAGE reads obtained in one of the tissues analyzed by the FANTOM5 consortium.

(D) ChIP-PCR analysis of different epigenetic marks and pyrosequencing determination of DNA methylation at the *Bglap3* locus in *Zfp932*/*Gm15446* KO versus control ESCs. Location of primers is depicted in (C).

Error bars represent SD, $^{*}p < 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, ns, not significant, Student's t test was used in (D) and Wilcoxon test in (B). See also Figure S4.



	ZFP932	Gm15446	Common
Total binding sites in C2C12	169	209	98
overlap with ZFP in ES	139 (82.2%)	199 (95.2%)	91 (92.9%)
overlap with KAP1 in C2C12	54 (31.9%)	59 (28.2%)	42 (42.9%)
at ERVKs	136 (80.5%)	188 (89.9%)	89 (90.8%)

Figure 5. The KRAB/KAP1 System Controls TE Expression in Differentiated Tissues

(A) *Krab-zfp* genes mRNA expression according to FANTOM 5 CAGE data. MSCs, mesenchymal stem cells; HSCs, hematopoietic stem cells.

(B) Heatmap of ERVs average fold change expression (linear scale) upon KAP1 depletion in different tissues. Liver = *Kap1* KO versus control, C2C12 = KAP1 KD versus control, MEF = *Kap1* KO versus control, ES = *Kap1* KO versus control.

(C) Venn diagram of ERVs bound by KAP1 in ESCs, C2C12 cells, or MEF cells.

(D) Comparative table of ZFP932, Gm15446, or common ZFP932 and Gm15446 binding sites in C2C12, determined by ChIP-seq. Absolute values and percentage relative to total binding sites are given.

(E) Distribution of H3K9me3 ChIP-seq with or without the presence of KAP1, in C2C12, relative to the 5′ end of ERVK elements (5′ end corresponds to 0 bp).

(F) Boxplot of ERVK fold change expression in C2C12 KAP1 KD or MEF *Kap1* KO versus control. ***p $\leq$ 0.001, Wilcoxon test.

See also Figure S5.

Bioinformatics Analyses and Statistics

All genome-wide TE analyses were performed using a merged repeats track generated in house (by using RepeatMasker 3.2.8 and merging homonymous ERV-int integrants or attributed LTRs within 400 bp or less). Genomic region analyses were done with BEDTools (Quinlan and Hall, 2010). Motif search was performed using with RSAT (Medina-Rivera et al., 2015), using unbound ERVK elements as control. For coverage plots, ChIP-seq signals on each feature of interest were extracted from the bigwig file and scaled between 0 and 1; in some cases, multiple alignment was performed with MAFFT. Statistical difference was assessed by Student's t test, except for Figures 4B, 5F, and 6A for which the Wilcoxon test was applied. Error bars represent ±1SD. R version 3.1.2 (http://www.R-project.org) or GraphPad

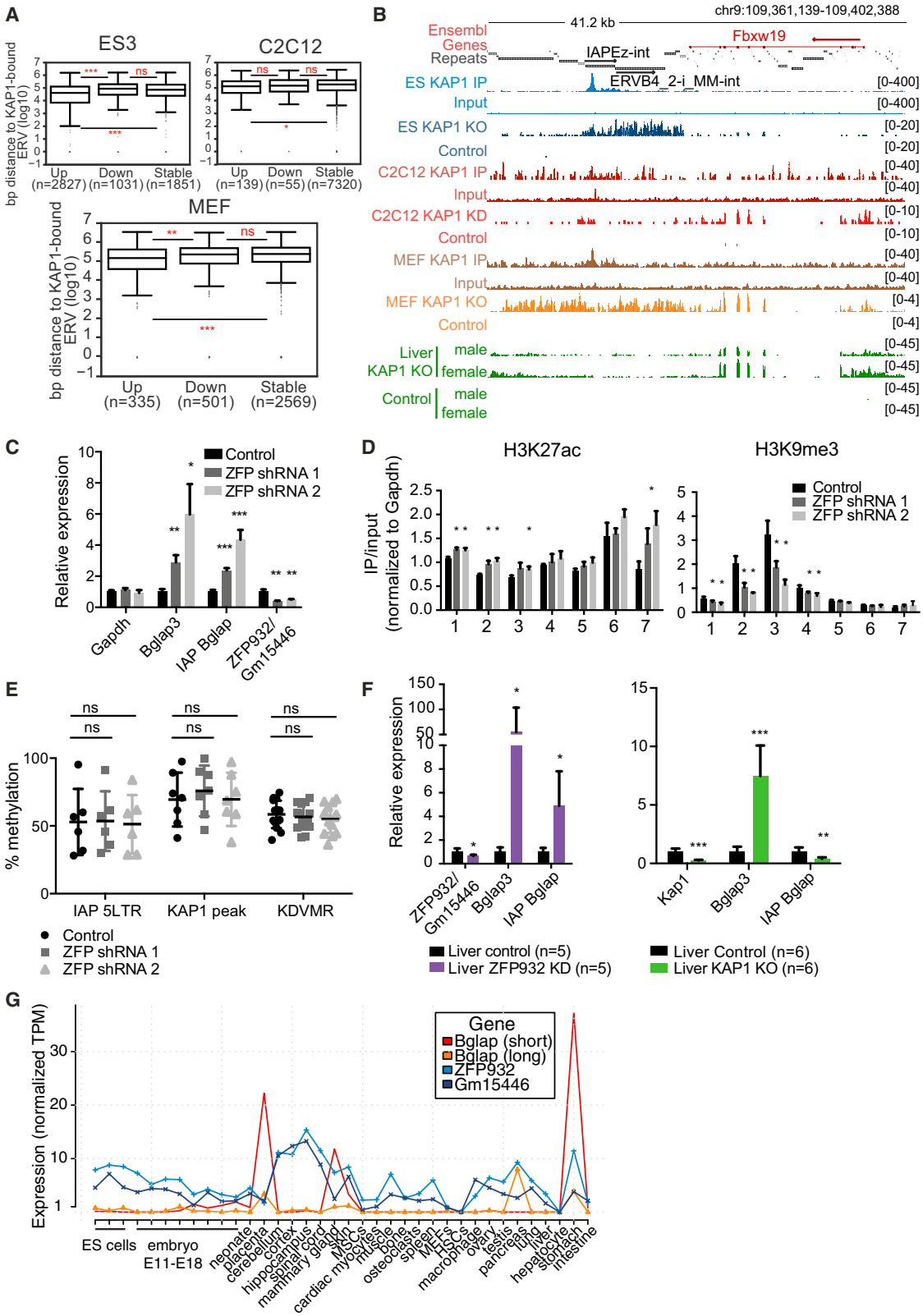


Figure 6. TEs and Their KRAB/KAP1 Controllers Influence Gene Expression in Adult Tissues

(A) Distance to nearest KAP1-bound ERV of upregulated (fold change ≥2, adjusted p ≤ 0.05), stable, or downregulated (fold change ≤ 2, adjusted p ≤ 0.05) genes in ESC *Kap1* KO, C2C12 KAP1 KD, or MEF *Kap1* KO.

(legend continued on next page)

Prism version 4.0 (http://www.graphpad.com) were used for statistical analyses. Detailed bioinformatics analyses are provided in Supplemental Experimental Procedures.

ACCESSION NUMBERS

All next-generation sequencing data have been submitted to the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) database under the accession number GEO: GSE74278.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.devcel.2016.02.024.

AUTHOR CONTRIBUTIONS

G.E. conceived the study, designed and performed experiments, analyzed and interpreted data, and wrote the manuscript. M.C. and A.K. designed, performed, and analyzed experiments. J.D. analyzed data. A.C. performed experiments and made intellectual contributions. S.O. performed experiments. M.I. analyzed data and made intellectual contributions. H.M.R. designed experiments and helped conceive the study. P.T. made intellectual contributions and helped supervise the project. D.T. conceived the study, designed experiments, interpreted data, and wrote the manuscript. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

We thank B. Deplancke and C. Gubelmann for part of the murine KRAB-ZFP library; E. Planet and A. Kapopoulou for help with data analysis; C. Raclot and S. Verp for technical assistance; I. Barde and M. Friedli for advice; the University of Lausanne Genomics Core Facility for sequencing; the EPFL Biomolecular Screening facility for robotics and advice; A. Reymond (CIG, UNIL, Lausanne) for use of the pyrosequencer; Vital-IT for computing; and the members of the Trono lab for discussions. This work was financed through grants from the Swiss National Foundation, the European Union (FP7/2007-2013/REA no. 290123), and the European Research Council to D.T.

Received: November 8, 2015
Revised: February 10, 2016
Accepted: February 24, 2016
Published: March 21, 2016

REFERENCES

Barde, I., Rauwel, B., Marin-Florez, R.M., Corsinotti, A., Laurenti, E., Verp, S., Offner, S., Marquis, J., Kapopoulou, A., Vanicek, J., et al. (2013). A KRAB/KAP1-miRNA cascade regulates erythropoiesis through stage-specific control of mitophagy. Science *340*, 350–353.

Bierhoff, H., Dammert, M.A., Brocks, D., Dambacher, S., Schotta, G., and Grummt, I. (2014). Quiescence-induced LncRNAs trigger H4K20 trimethylation and transcriptional silencing. Mol. Cell *54*, 675–682.

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

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

Bourque, G., Leong, B., Vega, V.B., Chen, X., Lee, Y.L., Srinivasan, K.G., Chew, J.L., Ruan, Y., Wei, C.L., Ng, H.H., et al. (2008). Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res. *18*, 1752–1762.

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

Chuong, E.B. (2013). Retroviruses facilitate the rapid evolution of the mammalian placenta. Bioessays *35*, 853–861.

Collins, P.L., Kyle, K.E., Egawa, T., Shinkai, Y., and Oltz, E.M. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. Proc. Natl. Acad. Sci. USA *112*, 8367–8372.

Corsinotti, A., Kapopoulou, A., Gubelmann, C., Imbeault, M., Santoni de Sio, F.R., Rowe, H.M., Mouscaz, Y., Deplancke, B., and Trono, D. (2013). Global and stage specific patterns of Kruppel-associated-box zinc finger protein gene expression in murine early embryonic cells. PLoS One *8*, e56721.

Dupressoir, A., Lavialle, C., and Heidmann, T. (2012). From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. Placenta *33*, 663–671.

Emerson, R.O., and Thomas, J.H. (2009). Adaptive evolution in zinc finger transcription factors. PLoS Genet. *5*, e1000325.

Fasching, L., Kapopoulou, A., Sachdeva, R., Petri, R., Jonsson, M.E., Manne, C., Turelli, P., Jern, P., Cammas, F., Trono, D., et al. (2015). TRIM28 represses transcription of endogenous retroviruses in neural progenitor cells. Cell Rep. *10*, 20–28.

Fort, A., Hashimoto, K., Yamada, D., Salimullah, M., Keya, C.A., Saxena, A., Bonetti, A., Voineagu, I., Bertin, N., Kratz, A., et al. (2014). Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. Nat. Genet. *46*, 558–566.

Friedli, M., and Trono, D. (2015). The developmental control of transposable elements and the evolution of higher species. Annu. Rev. Cell Dev. Biol. *31*, 13.11–13.23.

Groner, A.C., Tschopp, P., Challet, L., Dietrich, J.E., Verp, S., Offner, S., Barde, I., Rodriguez, I., Hiiragi, T., and Trono, D. (2012). The Kruppel-associated box repressor domain can induce reversible heterochromatization of a mouse locus in vivo. J. Biol. Chem. *287*, 25361–25369.

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

He, Z., Cai, J., Lim, J.W., Kroll, K., and Ma, L. (2011). A novel KRAB domain-containing zinc finger transcription factor ZNF431 directly represses Patched1 transcription. J. Biol. Chem. *286*, 7279–7289.

Heras, S.R., Macias, S., Plass, M., Fernandez, N., Cano, D., Eyras, E., Garcia-Perez, J.L., and Caceres, J.F. (2013). The Microprocessor controls the activity of mammalian retrotransposons. Nat. Struct. Mol. Biol. *20*, 1173–1181.

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

(B) Representative genomic region with ERV and *Fbxw19* gene regulated by KAP1 in differentiated cells. KAP1-depleted and control RNA-seq densities of different tissues/cells are depicted along with KAP1 ChIP-seq data.
(C) mRNA expression of *Bglap3* gene and IAP Bglap in ZFP932/Gm15446 KD in C2C12 cells versus control (normalized to *ActinG* and *Tbp*).
(D) ChIP-PCR of chromatin marks on *Bglap3* locus in ZFP932/Gm15446 KD versus control C2C12 cells. Numbers correspond to primers in Figure 4C.
(E) DNA methylation analysis by pyrosequencing on IAP Bglap in same cells.
(F) mRNA expression of *Bglap3* gene and IAP Bglap in ZFP932/Gm15446 KD or *Kap1* KO mice liver compared with control (normalized to *Gapdh* and *ActinB*).
(G) *Zfp932*, *Gm15446*, and *Bglap3* gene expression in different murine tissues/cells according to FANTOM 5 CAGE data. MSCs, mesenchymal stem cells; HSCs, hematopoietic stem cells.
Error bars represent SD, *$p < 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ns, not significant, Wilcoxon test was used in (A) and Student's t test in (C–F). See also Figure S6.

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

Iyengar, S., Ivanov, A.V., Jin, V.X., Rauscher, F.J., 3rd, and Farnham, P.J. (2011). Functional analysis of KAP1 genomic recruitment. Mol. Cell Biol. *31*, 1833–1847.

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

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

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

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29.

Liu, H., Chang, L.H., Sun, Y., Lu, X., and Stubbs, L. (2014). Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies. Genome Biol. Evol. *6*, 510–525.

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., Abugessaisa, I., Fukuda, S., Hori, F., Ishikawa-Kato, S., et al. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. *16*, 22.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63.

Maksakova, I.A., Mager, D.L., and Reiss, D. (2008). Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. Cell Mol. Life Sci. *65*, 3329–3347.

Marchetto, M.C., Narvaiza, I., Denli, A.M., Benner, C., Lazzarini, T.A., Nathanson, J.L., Paquola, A.C., Desai, K.N., Herai, R.H., Weitzman, M.D., et al. (2013). Differential L1 regulation in pluripotent stem cells of humans and apes. Nature *503*, 525–529.

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.

McNamara, R.P., Reeder, J.E., McMillan, E.A., Bacon, C.W., McCann, J.L., and D'Orso, I. (2016). Kap1 recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II. Mol. Cell *61*, 39–53.

Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J.A., Delerce, J., Jaeger, S., Blanchet, C., Vincens, P., Caron, C., et al. (2015). RSAT 2015: regulatory sequence analysis tools. Nucleic Acids Res. *43*, W50–W56.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Najafabadi, H.S., Mnaimneh, S., Schmitges, F.W., Garton, M., Lam, K.N., Yang, A., Albu, M., Weirauch, M.T., Radovani, E., Kim, P.M., et al. (2015). C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nat. Biotechnol. *33*, 555–562.

Pezic, D., Manakov, S.A., Sachidanandam, R., and Aravin, A.A. (2014). piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes Dev. *28*, 1410–1428.

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

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rebollo, R., Miceli-Royer, K., Zhang, Y., Farivar, S., Gagnier, L., and Mager, D.L. (2012). Epigenetic interplay between mouse endogenous retroviruses and host genes. Genome Biol. *13*, R89.

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

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

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

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

Schubeler, D. (2012). Molecular biology. Epigenetic islands in a genetic ocean. Science *338*, 756–757.

Singh, K., Cassano, M., Planet, E., Sebastian, S., Jang, S.M., Sohi, G., Faralli, H., Choi, J., Youn, H.D., Dilworth, F.J., et al. (2015). A KAP1 phosphorylation switch controls MyoD function during skeletal muscle differentiation. Genes Dev. *29*, 513–525.

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

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

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

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–562.

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

Wolf, D., and Goff, S.P. (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.

Wolf, D., Hug, K., and Goff, S.P. (2008). TRIM28 mediates primer binding site-targeted silencing of Lys1,2 tRNA-utilizing retroviruses in embryonic cells. Proc. Natl. Acad. Sci. USA *105*, 12521–12526.

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

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952–1958.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.