The Histone Methyltransferase SETDB1 Controls T Helper Cell Lineage Integrity by Repressing Endogenous Retroviruses

Véronique Adoue, Bénédicte Binet, Agathe Malbec, Joanna Fourquet, Paola Romagnoli, Joost Pm van Meerwijk, Sebastian Amigorena, Olivier Joffre

The histone methyltransferase SETDB1 controls T helper cell lineage integrity by repressing endogenous retroviruses

Véronique Adoue[1,4,5], Bénédicte Binet[1,4], Agathe Malbec[1,4], Joanna Fourquet[1], Paola Romagnoli[1], Joost P M van Meerwijk[1], Sebastian Amigorena[2,3], Olivier P Joffre[1,5,6].

[1] CPTP, Université de Toulouse, CNRS, Inserm, UPS, Toulouse, France
[2] Centre de Recherche, Institut Curie, 75005 Paris, France
[3] INSERM, U932, F-75005 Paris, France

[4] These authors contributed equally to this work.
[5] Correspondence: olivier.joffre@inserm.fr (O.J.), veronique.adoue@inserm.fr (V.A.)
[6] Lead contact

Running title: SETDB1 critically controls the Th1 gene expression program

SUMMARY

Upon activation, naïve $CD4^+$ T cells differentiate into distinct T cell subsets via processes reliant on epigenetically regulated, lineage-specific developmental programs. Here we examined the function of the histone methyltransferase SETDB1 in T helper (Th) cell differentiation. *Setdb1*$^{-/-}$ naïve $CD4^+$ T cells exhibited exacerbated Th1 priming, and when exposed to a Th1-instructive signal, *Setdb1*$^{-/-}$ Th2 cells crossed lineage boundaries and acquired a Th1 phenotype. SETDB1 did not directly control Th1 gene promoter activity, but relied instead on deposition of repressive H3K9me3 mark at a restricted and cell type-specific set of endogenous retroviruses (ERVs) located in the vicinity of genes involved in immune processes. Refined bioinformatic analyses suggest that these retrotransposons regulate Th1 gene *cis*-regulatory elements or act as Th1 gene enhancers. Thus, H3K9me3 deposition by SETDB1 ensures Th cell lineage integrity by repressing a repertoire of ERVs that have been exapted into *cis*-regulatory modules to shape and control the Th1 gene network.

INTRODUCTION

T lymphocytes protect vertebrates against a wide variety of endogenous and exogenous dangers. Their efficacy comes at least in part from their ability to adapt their phenotype and function to the threat detected by the cells of the innate immune system. Depending on the nature and strength of the signals delivered by these cells and the surrounding tissues, T lymphocytes mobilize different networks of transcription factors to induce distinct developmental programs that coordinate the acquisition of lineage-specific and danger-adapted phenotypes and functions (O'Shea and Paul, 2010; Wilson et al., 2009). This plasticity is best illustrated by naïve $CD4^+$ T cells, which are able to differentiate into multiple distinct effector populations.

The transcription factors mobilized in response to environmental signals orchestrate a massive remodeling of the epigenetic landscape of T cells (Kanno et al., 2012; Wilson et al., 2009). These dynamic changes in chromatin composition and compaction are necessary to set up and stabilize gene expression programs and to allow their faithful transmission to the progeny. Indeed, interfering with the post-translational modifications of histones or with DNA methylation critically affects the differentiation and stability of effector and memory T cells (Allan et al., 2012; Wilson et al., 2009; Xiao et al., 2016; Young et al., 1994). In $CD4^+$ T lymphocytes, epigenetic remodeling is largely coordinated by STAT proteins and by the master regulators specific to each lineage, such as T-bet and GATA-3 for, respectively, Th1 and Th2 cells (Kanno et al., 2012; O'Shea et al., 2011). These transcriptional regulators fine-tune the balance between T helper cell determination and plasticity by directing the deposition of permissive epigenetic marks at lineage-specific *cis*-regulatory elements, and by targeting repressive epigenetic pathways to the loci associated with alternative fates (Kanno et al., 2012; O'Shea et al., 2011; Vahedi et al., 2012; Wilson et al., 2009).

Trimethylation of Histone H3 on Lysine 9 (H3K9me3) has varied roles in the control of genome functions (Mozzetta et al., 2015). This epigenetic mark was first implicated in the scaffolding and function of constitutive heterochromatin (Lachner et al., 2001; Peters et al., 2001). H3K9me3 deposition at promoters of genes encoding developmental regulators is necessary to repress these loci and maintain embryonic stem cell pluripotency (Bilodeau et al., 2009). In adult cells, H3K9me3-dependent repression of gene expression in euchromatin and facultative heterochromatin is also important to define and maintain cell identity (Allan et al., 2012; 2016; Liu et al., 2015). However, the repertoires of loci and genomic elements that are targeted, as well as the molecular mechanisms at work remain poorly characterized. H3K9me3 also accumulates on the body of active genes where it may affect transcription elongation and alternative splicing (Saint-André et al., 2011; Vakoc et al., 2005). H3K9me3 is thus a versatile chromatin mark, with multiple and at times potentially opposing functions.

Several lysine methyltransferases trimethylate H3K9. These include SUV39H1, SUV39H2 and SETDB1, which all belong to the SUV39H family (Mozzetta et al., 2015). SUV39H1 and SUV39H2 were first identified as key components of constitutive heterochromatin (Peters et al., 2001; 2002), whereas SETDB1 was initially found to be involved in the dynamic repression of gene transcription at euchromatin and facultative heterochromatin (Schultz et al., 2002). SUV39H1 can also repress euchromatic gene expression through H3K9me3 deposition at promoters (Allan et al., 2012; Liu et al., 2015), whereas the maintenance of H3K9me3 at pericentromeric heterochromatin during DNA replication might depend on a stepwise process involving H3K9 mono- and tri-methylation by SETDB1 and SUV39H1, respectively (Loyola et al., 2009). In embryonic stem cells, these two enzymes also collaborate to repress ERVs (Bulut-Karslioglu et al., 2014). Since various cell types use these repeat elements as *cis*-regulatory modules to shape and control gene networks (Chuong et al., 2017), SETDB1 and SUV39H1 may therefore also control cell integrity through deposition of H3K9me3 at ERVs.

In T cells, while SETDB1 is implicated in OX40-dependent repression of the *Il17a* locus in Th17 cells (Xiao et al., 2016), SUV39H1 controls Th2 cell stability by depositing H3K9me3 at the *Ifng* promoter (Allan et al., 2012). However, the deregulation of the *Ifng* locus observed in *Suv39h1*$^{-/-}$ cells cannot by itself explain a loss of Th2 cell integrity. Other critical Th1 cell lineage-specific loci might therefore be controlled by H3K9me3-dependent repressive mechanisms. In addition, while a clear H3K9me3 signal is detected at the gene encoding T-bet in Th2 cells, SUV39H1 has no impact in the deposition of the repressive mark at this locus (Allan et al., 2012). Together with the fact that H3K9me3 disappearance at euchromatin and facultative heterochromatin is limited in SUV39H1-deficient cells (Peters et al., 2002), these observations suggest that other H3K9me3-dependent epigenetic pathways critically control Th2 cell stability.

Here we examined the effects of SETDB1-dependent H3K9me3 deposition on CD4^{+} T cell differentiation. We found that SETDB1 restricts Th1 cell priming and ensures Th2 cell integrity. Unlike their wild-type counterparts, *Setdb1*$^{-/-}$ Th2 cells readily expressed Th1-associated genes when exposed to the Th1-instructing cytokine interleukin (IL)-12. SETDB1 repressed Th1-related loci by depositing H3K9me3 at a subset of ERVs that flank and repress Th1 enhancers or behave themselves as *cis*-regulatory elements of Th1 genes. Our findings reveal a repertoire of ERVs that have been co-opted to behave as Th1-specific *cis*-regulatory modules, and outline a model wherein H3K9me3 deposition by SETDB1 locks the Th1 gene network and ensures Th cell lineage integrity by repressing these repeat elements.

RESULTS

Th1 priming is enhanced in the absence of SETDB1

To analyze the role of SETDB1 in $CD4^+$ T cell differentiation and plasticity, we generated mice homozygous for a LoxP-flanked *Setdb1* allele and expressing (*Setdb1*$^{-/-}$), or not (*Setdb1*$^{+/+}$), the CRE recombinase under the control of the *Cd4* promoter. This strategy resulted in the almost complete absence of SETDB1 from CD4 single-positive thymocytes (**Figure S1A**). As SETDB1-deficiency was not compensated by other methyltransferases targeting H3K9 (**Figure S1B**), we also observed a marked loss of H3K9me3 from naïve *Setdb1*$^{-/-}$ $CD4^+$ T cells (**Figure S1C)**. The use of the *Cd4-Cre* transgene, which induces SETDB1 deletion relatively late in ontogeny, allowed for normal intrathymic T cell development. Indeed, the total number of cells in the thymus, the relative proportions of the four main populations of thymocytes, and the proportion of mature $CD4^+$ T cells were similar in *Setdb1*$^{-/-}$ mice and control littermates (**Figures S1D–S1H**).

In peripheral lymphoid tissues, we detected no consequences of T cell-specific SETDB1-deficiency on other populations of immune cells (**Figures S2A–S2C**). SETDB1 was previously implicated in survival in various cell types. For example, conditional deletion of the enzyme in mice expressing the CRE recombinase under the control of the *Mb1* promoter abolishes the B cell lineage (Collins et al., 2015). The impact of *Setdb1* deletion on T cell survival was less pronounced: despite substantially increased activity of caspase-3/7, we only observed a partial loss of the T cell pool (**Figures S2D–S2G**).

To obtain a global view of the changes in gene expression induced by *Setdb1* deletion, we performed RNA sequencing (RNA–seq) on naïve *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ $CD4^+$ T cells. Most of the differentially expressed genes were more expressed in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ cells (**Figure S3A**), consistent with a globally repressive effect of H3K9me3 on gene transcription. Among the overexpressed genes, those involved in cell division were particularly enriched (**Figures S3B** and **S3C**). The proportion of $CD4^+$ T cells expressing the nuclear antigen Ki67 being higher in *Setdb1*$^{-/-}$ mice than in control littermates (**Figure S3D**), this increased expression of cell division-related genes most likely resulted from the observed lymphopenia rather than from a direct impact of *Setdb1* deletion on the regulation of these genes. Moreover, we found no particular enrichment for H3K9me3 domains at cell cycle-related genes (**Figure S3E**).

To assess if *Setdb1* deletion could affect T cell function, we next analyzed the differential transcription of a gene set related to Th cell differentiation. We found no major differences in the expression of these genes between *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ cells (**Figure 1A** and **Table S1**), despite the presence of H3K9me3 domains close to loci involved in lymphocyte-mediated immunity (**Figure S3E**). The vast majority of the genes were equally expressed in *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ cells, most of the differentially expressed loci were transcribed at very

low levels, and no lineage-specific transcriptomic signature appeared when focusing on deregulated genes. This lack of effect of SETDB1 deficiency on naïve $CD4^+$ T cell programming was further confirmed when we analyzed the production of lineage-specific mediators following acute *ex vivo* stimulation (**Figure 1B**). Together, these observations show that *Setdb1*$^{-/-}$ cells are not *a priori* biased towards a specific Th lineage.

To test whether SETDB1 regulated Th cell lineage commitment in response to environmental signals, we analyzed *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ $CD4^+$ T cell fate in an IL-12-mediated Th1 differentiation assay. As expected from our experiments measuring caspase 3/7 activity *ex vivo* (**Figures S2F** and **S2G**), SETDB1 deficiency impaired T cell survival at early time points (**Figure S3F**). However, a significant proportion of cells remained viable and showed normal activation upon TCR triggering (**Figures S3G** and **S3H**). As T cell differentiation depends on cell cycle progression, we next analyzed the proliferative response of activated $CD4^+$ T cells. There were no differences between control and mutant cells (**Figure 1C**), which displayed similar proliferation indexes and percentages of divided cells (**Figure 1D**). Upon exposure to IL-12, *Setdb1*$^{-/-}$ cells also expressed T-bet at a similar level as their *Setdb1*$^{+/+}$ counterparts (**Figures 1E** and **1F**). However, SETDB1-deficiency led to greater acquisition of lineage-specific functions. Indeed, both percentages of cells producing Th1-related cytokines and amounts of cytokine synthesized per cell were higher in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ cells (**Figures 1G–1J**). This exacerbated production of cytokines was not the result of a global transcriptional derepression, since we did not evidence any aberrant secretion of soluble mediators related to alternative lineages (**Figure 1J**). It was also not the result of a greater sensitivity of the *Setdb1*$^{-/-}$ cells to IL-12 (**Figures S3I–S3K**). Together, these results highlight a key role for SETDB1 in regulating the magnitude of Th1 responses.

Impaired acquisition of the Th2 phenotype by SETDB1-deficient cells

Most of the genes encoding lineage-specific cytokines in naïve $CD4^+$ T cells have both permissive and repressive epigenetic marks on their promoters and enhancers. They are thus poised for transcription to guarantee the plasticity, while also preserving the identity of the cells. The enhanced Th1 response observed in *Setdb1*$^{-/-}$ cells might, therefore, result from a loss of H3K9me3 at these *cis*-regulatory regions, and this may potentially affect other lineages. To test this hypothesis, we cultured *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ naïve $CD4^+$ T cells in Th2 polarizing conditions. The proliferative response and viability of *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ cells were comparable at day six (**Figures 2A-2C**). Moreover, *Setdb1*$^{-/-}$ cells seemed to commit to the Th2 lineage similar to their control counterparts, with almost all cells expressing GATA-3 (**Figures 2D** and **2E**) and no aberrant expression of T-bet (**Figure 2F**). Production of IL-13 and IL-4 was also similar between *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ cells (**Figures 2G**-**2I**). Thus, in contrast to what we observed in Th1-polarizing conditions, there was no

enhanced production of Th2 lineage-specific mediators by *Setdb1*$^{-/-}$ cells grown in the presence of IL-4. In fact, global transcriptional profiling revealed that the level of expression of the Th2 signature genes was significantly lower in *Setdb1*$^{-/-}$ Th2 cells than in their wild-type counterparts, although both wild-type and mutant naïve CD4^{+} T cells efficiently switched on the Th2 program upon exposure to IL-4 (**Figures 2J** and **2L**). This impaired induction of the Th2 gene network correlated with lower expression of GATA-3 (**Figures 2D** and **2E**) and with decreased chromatin accessibility at Th2 gene enhancers (**Figure 2K**). Moreover, unlike their wild-type counterparts, *Setdb1*$^{-/-}$ cells grown in Th2-polarizing conditions also produced small amounts of IFN-γ (**Figures 2G**-**2I**). This IFN-γ 'leak' may result from defective repression of Th1-related loci in Th2 cells, which could potentially antagonize the Th2 gene expression program and lead to functional and phenotypic instability.

SETDB1 is required for stable Th2 cell commitment

To assess if SETDB1 controlled Th2 cell plasticity, we cultured *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ cells in Th2-polarizing conditions and then switched to culture in Th1-polarizing medium. In agreement with the Th1/Th2 paradigm, the control *Setdb1*$^{+/+}$ Th2 cells remained stable (**Figures 3A–3C**). By contrast, a large fraction of the *Setdb1*$^{-/-}$ cell population secreted IFN-γ, this phenomenon being even more pronounced after four days of culture (**Figures 3A** and **3B**). IFN-γ secretion was accompanied by decreased expression of GATA-3 and increased expression of T-bet (**Figures 3D** and **3E**). In fact, SETDB1-deficiency allowed the virtually complete reprogramming of Th2 cells upon exposure to Th1-instructing signals, with extinction of Th2 gene expression and induction of a large part of the Th1 gene set (**Figures 3F** and **3G** and **Table S2**). This plasticity was not the result of a bias in *Setdb1*$^{-/-}$ cells programming due to lymphopenia (**Figures S4A–S4C**).

SETDB1 plays a key role in silencing ERVs (Bulut-Karslioglu et al., 2014; Matsui et al., 2010). Ectopic expression of these retrotransposons can lead to activation of the nucleic acid-sensing machinery and, eventually, to production of type I IFNs (Chiappinelli et al., 2015). Together with IL-12 and IFN-γ, type I IFNs can reprogram Th2 cells into stable cells producing IFN-γ and expressing both GATA-3 and T-bet (Hegazy et al., 2010). Activation of ERVs in *Setdb1*$^{-/-}$ Th2 cells might thus account for the increased plasticity that we observed: ERV-induced secretion of IFN-α and IFN-β might reprogram the Th2 cells in combination with exogenous IL-12 and the observed aberrant production of IFN-γ. However, we found no abnormal levels of type I IFN mRNA in *Setdb1*$^{-/-}$ cells (**Figures S4D** and **S4E**), and neutralization of IFN-γ did not prevent *Setdb1*$^{-/-}$ Th2 cells from switching to a Th1 phenotype (**Figures S4F** and **S4G**). To assess directly if the ectopic expression of Th1-instructive mediators by *Setdb1*$^{-/-}$ Th2 cells might account for their phenotypic instability, we co-cultured

Setdb1$^{-/-}$ and *Setdb1*$^{+/+}$ Th2 cells in Th1 polarizing conditions. In this setting, *Setdb1*$^{-/-}$ cells still showed substantial plasticity, while their control counterparts did not (**Figures S4H–S4J**). Therefore, SETDB1 critically controls Th2 cell commitment through a cell-intrinsic mechanism.

To test the role of SETDB1 in CD4^{+} T cell programming *in vivo*, we next immunized *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ mice with the 1W1K variant of the I-E alpha chain immunodominant peptide formulated in RIBI adjuvant. Using 1W1K/I-A^{b} tetramers, we first showed that antigen-specific *Setdb1*$^{-/-}$ CD4^{+} T cells expanded and accumulated in the lymph nodes draining the site of immunization (**Figures 3H** and **3I**). The frequency of tetramer-positive cells was lower than in control mice but this quantitative defect did not prevent *Setdb1*$^{-/-}$ cells from differentiating into Th1 and Th2 effector cells. However, relative to the size of the antigen-specific T cell compartment, the frequency of CD4^{+} T cells producing IFN-γ was strongly increased in *Setdb1*$^{-/-}$ as compared to *Setdb1*$^{+/+}$ mice (**Figure 3J**). This exacerbated Th1 response shifted the Th1/Th2 balance towards Th1 dominant immunity (**Figure 3K**). Therefore, as observed *in vitro*, SETDB1 controls the Th1 gene expression program *in vivo*.

SETDB1-dependent H3K9 trimethylation at a subset of ERVs

To determine how SETDB1 controlled Th2 cell commitment and stability, we performed high-throughput chromatin immunoprecipitation (ChIP) of H3K9me3 in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells. We first analyzed H3K9me3 signal at gene promoters. As expected from the literature, our genome-wide analysis revealed an inverse correlation between the deposition of the mark and gene expression (**Figure S5A**). However, SETDB1-deficiency did not significantly impact on H3K9me3 deposition at Th1-related genes (**Figure S5B**). We then analyzed the genome-wide distribution of H3K9me3 domains in *Setdb1*$^{+/+}$ Th2 cells. We observed no more peaks at gene bodies or promoters than would be expected by a random distribution (**Figures 4A** and **4B**). By contrast, we found statistically significant enrichment of H3K9me3 domains at enhancers (defined as non-promoter H3K4me1^{+} genomic regions) and at ERVs (**Figures 4A** and **4B** and **Table S3**). Since 74% of the peaks at enhancers also overlapped with ERVs (**Figure 4C**), these observations suggest that the ERVs rather than the enhancers themselves are the main targets for H3K9 trimethylation. To test this hypothesis, we analyzed the distribution of H3K9me3 across the length of individual H3K9me3^{+} ERV and enhancer sequences. The H3K9me3 signal clearly peaked at and aligned with the center of the ERVs (**Figures 4D** and **4E**). By contrast, the signal appeared randomly distributed across the enhancer sequences, with only a modest accumulation on their flanking regions. In agreement with the repressive role of the histone mark and with the status of the *cis*-regulatory elements, H3K9me3 signal was higher on the flanks of poised than of active

enhancers, and the mark only overlapped inactive *cis*-regulatory elements (**Figure 4F**). H3K9me3 accumulation over the enhancer sequence might thus result from a spreading from neighboring ERVs (**Figure 4G**), or might be due to physical overlap between the enhancers and the retroelements (**Figure 4H**). Whatever the model, we observed that H3K9me3 domains were closer to the center of the ERVs than to the center of the enhancers (**Figure 4I**). Therefore, these results support the hypothesis that H3K9me3 is deposited at a subset of ERVs that overlap or flank enhancers, which might lead to their repression.

Finally, to determine which lysine methyltransferase is necessary for H3K9me3 deposition at ERVs, we determined if the retroelements marked by H3K9me3 in wild-type cells were still covered by the repressive mark in *Setdb1*$^{-/-}$ or *Suv39h1*$^{-/-}$ Th2 cells. While SUV39H1-deficiency had no major impact on H3K9me3 deposition at these genomic locations, most of the peaks disappeared from the ERVs in the SETDB1-deficient cells (**Figure 4J**), even if some residual signal persisted at certain locations (**Figures S6A–S6C**). Together, these data indicate that SETDB1 targets H3K9me3 at a subset of ERVs in Th2 cells, and that some of these retrotransposons overlap or flank enhancers.

Increased expression of ERVs and neighboring genes in *Setdb1*$^{-/-}$ cells

Recent studies have provided evidence for the hypothesis that transposable elements are co-opted for the regulation of host gene networks (Chuong et al., 2017; 2016). The impact of SETDB1 deletion on CD4^{+} T cell fate might, therefore, result from a loss of H3K9me3 at ERVs that behave as *cis*-regulatory modules of transcription, and/or regulate the activity of enhancers. To test this hypothesis, we analyzed the consequences of SETDB1 deletion on ERV accessibility and activity, on the status of their nearest enhancers, and on the expression of associated genes. We first compared ERV expression levels in *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ Th2 cells. In *Setdb1*$^{-/-}$ Th2 cells, the expression of 22% of the ERVs that lost H3K9me3 was deregulated, 77% of which being overexpressed (**Figures 5A** and **5B**). As the expression levels were very low, we used other parameters to confirm that H3K9me3 disappearance from ERVs led to local chromatin remodeling. Loss of H3K9me3 at ERVs in *Setdb1*$^{-/-}$ Th2 cells also correlated with the accumulation and spreading of permissive histone marks and with increased chromatin accessibility (**Figures 5C and S6A–S6C**). Since we observed a decompaction of the chromatin on both sides of the retrotransposons overexpressed in mutant cells (**Figure 5C**), we hypothesized that the *cis*-regulatory elements that flank ERVs might be de-repressed in the absence of SETDB1. To test this hypothesis, we analyzed the activation status of enhancers located in the vicinity of ERVs marked by H3K9me3 in *Setdb1*$^{+/+}$ cells and whose expression was increased in *Setdb1*$^{-/-}$ cells. The enhancers associated with ERVs that were activated following H3K9me3 disappearance were themselves more expressed in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ Th2 cells (**Figure 5D**) As

expected from their increased transcription, they also accumulated permissive histone marks in mutant cells (**Figure 5E**). We then tested if this cascade of events resulted in deregulation of gene expression. Even if the Th2 gene expression program was impaired in *Setdb1*$^{-/-}$ Th2 cells (**Figures 2J** and **2L**), our analysis showed that the genes associated with enhancers flanking or overlapping ERVs whose expression was increased in *Setdb1*$^{-/-}$ cells were also significantly more expressed (**Figure 5F**). Finally, we observed a positive correlation between the degree of gene expression change per gene between *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ cells and the number of activated retroelements associated with a gene (**Figure 5G**). Therefore, we propose that SETDB1-dependent H3K9me3 deposition at ERVs inactivates neighboring enhancers and thus participates to the silencing of their target genes.

H3K9me3^{+} ERVs mark the Th1 enhancer landscape in Th2 cells

In our functional assays, SETDB1 deletion led to enhanced Th1 priming and to Th2 cell instability. As discussed above, loss of regulation of the Th1 gene network might underlie these observations. Based on our epigenetic and transcriptomic studies, we postulated that SETDB1 controls Th1 gene expression by either repressing ERVs operating as *cis*-regulatory elements of these genes, and/or regulating the activity of their enhancers. To test this hypothesis, we first assigned biological significance to the ERVs marked by H3K9me3 in a SETDB1-dependent manner by analyzing the annotations of their nearby genes. We observed a strong association of the retrotransposons with genes involved in immune processes, including leukocyte activation and cytokine production (**Figure 6A**). This distribution was cell type-specific as there was very little overlap between the ERVs marked by H3K9me3 in Th2 cells and those marked in adipocytes (**Figure S7A**). Moreover, the ERVs marked by H3K9me3 in white adipose cells were associated with genes that have no direct link with immunity (**Figures S7B** and **S7C**). As IFN-γ plays a critical role in Th1 cell programming, we next investigated whether these ERVs more specifically targeted IFN-γ-stimulated genes (ISGs). In contrast to H3K9me3^{-} ERVs or to the repertoire of H3K9me3^{+} ERVs found in adipocytes, the ERVs marked by H3K9me3 in Th2 cells were located in the vicinity of ISGs (**Figure 6B**). Motif enrichment analysis of H3K9me3^{+} ERV sequences strengthened this observation; it revealed a strong enrichment for the binding sites of STAT1, the main transcription factor responsible for the diverse cellular effects induced by IFN-γ (**Figure 6C**), as well as for other critical Th1-related transcription factors. The “Upstream Regulator Analysis” of our RNA-seq data also identified these five transcription factors as very likely to be responsible for the differences in gene expression observed in *Setdb1*$^{-/-}$ vs *Setdb1*$^{+/+}$ Th2 cells upon culture in Th1-polarizing conditions (**Figures 6D** and **S7D**). To strengthen these *in silico*-based predictions, we next tested whether ERV sequences marked by H3K9me3 in Th2 cells were associated with the genomic localizations of STAT1 and

STAT4 in Th1 cells. In contrast to H3K9me3^{-} ERVs or to the repertoire of H3K9me3^{+} ERVs found in adipocytes, we observed that the ERVs marked by the repressive histone mark in Th2 cells were strongly enriched near Th1-specific STAT1 and STAT4 binding sites (**Figures S7E** and **S7F**). In fact, almost 2,000 H3K9me3^{+} retroelements were associated with Th1-specific STAT1/4 genomic locations in Th2 cells (**Figures 6E–6H**). In addition, whereas a substantial number of these ERVs overlapped STAT ChIP-seq peaks, and were thus very likely to behave as Th1 gene enhancers, most of them only flanked the transcription factors binding sites (**Figure S7G**). To confirm that SETDB1 deposits H3K9me3 at a subset of ERVs associated with Th1 enhancers in Th2 cells, we analyzed the location of the retroelements relative to Th1 enhancers. We found 4,411 putative Th1 enhancers associated with H3K9me3^{+} ERVs in Th2 cells (**Figure 6I**). In contrast, we did not detect any enrichment of the Th1-specific *cis*-regulatory elements at H3K9me3^{+} ERVs in adipocytes (**Figure 6J**). Together these data suggest that the differences in stability observed between *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ Th2 cells are very likely explained by SETDB1 causing H3K9me3 deposition at, and thus suppression of, ERVs that overlap or flank a large network of Th1-specific enhancers. Since 70% of the Th1 *cis*-regulatory elements that were covered by the repressive mark overlapped an ERV onto which H3K9me3 signal was centered (**Figures 6K** and **6L**), this epigenetic silencing pathway is probably central for the H3K9me3-dependent suppression of the Th1 gene network in Th2 cells.

SETDB1-dependent H3K9 trimethylation at ERVs represses Th1-specific enhancers

H3K9me3^{+} ERVs marked a large repertoire of Th1 enhancers in Th2 cells. To test whether the deposition of H3K9me3 at these genomic elements was associated with their repression, we first analyzed the activation status of the Th1 enhancers associated with H3K9me3^{+} ERVs in wild-type Th2 cells. As expected, most of these *cis*-regulatory elements were repressed (**Figure 7A**), and we detected an accumulation of the histone mark over “poised” or “ghost” enhancers but not on active *cis*-regulatory elements (**Figure 7B**). To test whether the histone mark had a causal role in the repression of Th1 enhancers, we next compared their status in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells. H3K9me3 disappearance led to a strong increase in chromatin accessibility at Th1 enhancers associated with ERVs whose expression was increased in a mutant context (**Figure 7C**). This derepression correlated with an accumulation of permissive histone marks on the enhancer sequences (**Figures 7D–7G**). Among the large set of Th1 enhancers associated with at least one ERV marked by H3K9me3 in Th2 cells, we identified the conserved non-coding sequence (CNS) located 17–20kb downstream of the *Ifng* gene (Shnyreva et al., 2004). While this region was poised in naïve T cells, with a strong H3K4me1 signal flanked by a large domain of H3K9me3, it lost competence upon Th2 cell commitment, with an accumulation of H3K9me3 and a complete

loss of H3K4me1 (**Figure 7F**). In contrast, this enhancer exhibited a diminished H3K9me3 signal in *Setdb1*$^{-/-}$ Th2 cells, which may make this region, as well as the hundreds of other Th1 enhancers associated with an ERV marked by H3K9me3 in wild-type cells, accessible to Th1-specific transcription factors. Consistent with this hypothesis, the H3K4me1 and ATAC-seq signals at *Ifng* $CNS_{17\text{-}20}$ were substantially higher in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ Th2 cells (**Figures 7G and 7H**). We identified many other critical Th1-related genes whose expression might be regulated by an ERV acting as a *cis*-regulatory module. They included those encoding T-bet, the "master regulator" of the lineage, and other critical transcriptional regulators such as STAT4, IRF1 and RUNX3 (**Figure 7I**). These genes had at least one enhancer associated with an ERV marked by H3K9me3 in Th2 cells and were, in addition, more expressed in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ Th2 cells upon culture in Th1-inducing conditions. In conclusion, our data reveal that Th2 cell stability is controlled at the level of chromatin by the SETDB1-dependent deposition of H3K9me3 at a restricted set of ERVs flanking or behaving as Th1-gene enhancers.

DISCUSSION

Up to 10% of the mouse genome is comprised of ERVs, which had long been considered to be junk DNA sequences. Recently, however, regulatory functions over gene expression have been assigned to transposable elements. In mouse $CD4^+$ T cells, we documented that a set of ERVs enriched in binding sites for pro-Th1 transcription factors overlapped or flanked the enhancers of genes from the Th1 cell transcriptomic signature. We further showed that the accessibility of these repeat elements was regulated at the epigenetic level by SETDB1. Indeed, in *Setdb1*$^{-/-}$ cells, the lack of deposition of H3K9me3 at this subset of ERVs correlated with their activation and with the increased expression of their closest genes. At the cellular level, this deregulation of gene expression translated into increased Th2 cell plasticity and enhanced Th1 cell priming. Together, these data suggest that SETDB1 controls Th2 cell integrity by repressing a restricted and cell-type specific repertoire of ERVs.

While the use of *Setdb1*$^{-/-}$ cells allowed us to establish cause and effect links between SETDB1 depletion, H3K9me3 disappearance and ERV-derepression, we have not strictly demonstrated that the ERVs marked by H3K9me3 in Th2 cells acted as Th1 gene enhancers. The most direct way to prove that SETDB1 controlled the Th1 gene expression program through the regulation of ERVs that behave as Th1 gene enhancers would have been to selectively inactivate the transposable elements marked by H3K9me3 in Th2 cells using CRISPR/Cas9. Unfortunately, due to the large number of ERVs and to the absence of a consensus sequence to target, we have not been able to perform this experiment.

ERVs potentially control Th1 genes expression through two non-mutually exclusive mechanisms: they may either behave as *cis*-regulatory elements, or they may regulate chromatin accessibility at nearby enhancers. Our evidence that the binding motifs for critical Th1-associated transcription factors were enriched in H3K9me3$^+$ ERV sequences suggests that the ERVs directly act as *cis*-regulatory elements. The existence of such a subset of regulatory ERVs, which might have shaped the Th1 transcriptional network over time, is supported by a recent study showing that ERVs containing binding sites for IFN-induced transcription factors are necessary for AIM2 inflammasome activation (Chuong et al., 2016). However, when we fractionated the repertoire of H3K9me3$^+$ ERVs associated with Th1 enhancers in Th2 cells, we observed that most of them only flanked STAT-1/4 binding sites. This result suggests that ERVs regulate the Th1 gene network mainly by modulating the activity of the Th1 enhancers located in their vicinity. Interestingly, the H3K9me3$^+$ ERVs that flanked Th1 enhancers accumulated at a distance of 3 to 5 kb from the STAT peaks. This distribution of the retrotransposons overlaps with the distribution of the H3K9me3 signal observed on the flanking regions of enhancers whose activity is regulated by this histone mark in DC and fibroblasts (Zhu et al., 2012). Although the authors of that study do

not implicate SETDB1 in H3K9me3 deposition and do not identify ERVs as the targeted genomic elements, they correlate the accumulation of H3K9me3 at this location with the repression of adjacent enhancer activity. This observation reinforces, and extends to other cell types, our model supporting that retrotransposons are the genetic elements that are targeted by the H3K9me3-dependent silencing machinery to regulate enhancer activity in a cell type-specific manner. The underlying molecular mechanism probably relies on local heterochromatin spreading from ERVs to nearby regulatory elements as suggested by our ChIP-seq data and by studies from the literature (Rebollo et al., 2011).

SETDB1 is ubiquitously expressed in mouse tissues and more than 900,000 ERVs are dispersed through the mouse genome. Nevertheless, the repertoire of ERVs under the control of SETDB1 was limited and highly cell type-specific. Only 17,349 ERVs were associated with H3K9me3 domains in Th2 cells. The vast majority of these ERVs were not associated with H3K9me3 domains in white adipose cells. Moreover, the biological functions of genes associated with the ERVs marked by H3K9me3 in Th2 and white adipose cells were fully different since they were associated with genes involved in immune processes only in differentiated lymphocytes. The enrichment of H3K9me3 at a specific set of ERVs might be explained by the fact that SETDB1 is recruited to the chromatin by Krüppel-associated box zinc-finger proteins (KZFPs) using the scaffold protein TRIM28 as a molecular intermediate. The mouse genome encodes hundreds of KZFPs, whose expression depends on the cell type and its physiological state (Imbeault et al., 2017). These transcriptional regulators have different DNA binding sites and ERVs are one of their main genomic targets. Indeed, KZFPs have co-evolved with transposable elements and use evolutionarily conserved regions located mainly within their regulatory sequences to control gene expression (Chuong et al., 2017; Imbeault et al., 2017). The cell-type specific and SETDB1-dependent deposition of H3K9me3 at ERVs that we observed in Th2 cells was therefore probably orchestrated by a specific set of KZFPs that await identification.

Of the three lysine methyltransferases from the SUV39H family, only SETDB1 was necessary for silencing ERVs in differentiated lymphocytes. This finding is consistent with those obtained in neural progenitor cells and immortalized mouse embryonic fibroblasts, in which SUV39H1-deficiency does not severely affect ERV silencing (Bulut-Karslioglu et al., 2014). Although they apparently argue against a direct collaboration of SUV39H1 and SETDB1 in H3K9 trimethylation at ERVs, our data do not fully exclude cooperation between these two enzymes in regulating Th2 cell commitment. In fact, the H3K9me3-dependent epigenetic regulation of $CD4^+$ T cell differentiation involves both SUV39H1 and SETDB1. In differentiated Th2 cells, SUV39H1 controls H3K9me3 deposition at the *Ifng* promoter(Allan et al., 2012) and our data demonstrate that SETDB1 regulates the entire Th1 gene network through repression of ERVs overlapping or flanking Th1-specific enhancers. To guarantee

Th2 cell stability in a changing environment, two non-redundant epigenetic silencing pathways therefore converge to lock the Th1 transcriptional program at different genomic locations.

In conclusion, our data support that SETDB1 controls CD4 T cell identity by repressing ERVs that flank or overlap Th1-specific enhancers. This enzyme is thus a potential target for drugs that might be useful, for example, to promote Th1 cell differentiation in various infectious diseases, or to prevent harmful Th2 responses in allergic disorders.

FIGURES LEGENDS

Figure 1. Th1 priming is enhanced in the absence of SETDB1.

(A) Expression of T helper-related genes by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ naïve CD4^{+} T cells. The names of the genes that were differentially expressed in the two genotypes (adjusted p-value<0.1 and fold difference >2 or <0.5) are highlighted in red. (B) Cytokine production by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ naïve CD4^{+} T cells, as measured flow cytometry. (C) Proliferation profiles of *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells cultured in Th1 medium. (D) Percentage of divided cells (left) and proliferation index (right) as calculated from the data in (C). (E) Expression of T-bet by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells after 6 days of culture in Th1 medium, as measured by flow cytometry. (F) Average expression of T-bet by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells after 6 days of culture in Th1 medium. (G) Production of GM-CSF and IFN-γ by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells after 6 days of culture in Th1-inducing conditions. (H, I) Percentage of CD4^{+} T cells-producing cytokine (left) and average cytokine production per cell (right) after 6 days of culture in Th1 medium. (J) Production of cytokines by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells following 6 days of culture in Th1-inducing conditions. Data are representative of two (C) or three (B, E and G) independent experiments. Data are represented as mean ± SD of two (D) or three (J) independent experiments or of three biological replicates from one representative experiment out of three performed (F, H, I). **p*<0.05, ***p*<0.01, ****p*<0.001 (unpaired Student's t-test). See also Figure S1-S3 and Table S1.

Figure 2. Impaired acquisition of the Th2 phenotype by SETDB1-deficient cells.

Setdb1$^{+/+}$ and *Setdb1*$^{-/-}$ naïve CD4^{+} T cells were cultured for 3 (A, B) or 6 (C–H) days in Th2-polarizing conditions. (A) Proliferation profiles, as determined by CTV dilution. (B) Percentages of divided cells and proliferation index, as calculated from the data in (A). (C) Percentage of live (acridin orange-positive and propidium iodide-negative) cells. (D) Expression of GATA-3, as determined by flow cytometry. (E) Percentage of GATA-3-expressing cells and average expression of GATA-3 per cell. (F) Average expression of T-bet, as measured by flow cytometry. (G) Production of IL-13 and IFN-γ by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells after culture in Th2-inducing conditions, as determined by flow cytometry. (H) Percentage of cytokine-producing CD4^{+} T cells. (I) Production of Th1- (IFN-γ) and Th2-related (IL-13, IL-4) cytokines by *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ CD4^{+} T cells following culture in Th2-polarizing conditions. (J) Box and whisker (min to max) diagrams representing the mRNA levels of 79 Th2-related genes in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ naïve and Th2 cells. (K) Average ATAC-seq signal at Th2 enhancers in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells. (L) Th2-related genes expression in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ naïve and Th2 cells. Data

are representative of three (A) or eight (D, G) independent experiments. Data are represented as mean ± SEM of eight independent experiments (E, F and H) or are means ± SD from three (I), five (B) or six (C) independent biological replicates from three independent experiments. **p<0.01, ***p<0.001 (paired Student's t-test). The data in (J, L) and (K) are from three and two independent biological replicates, respectively.

Figure 3. SETDB1-deficiency leads to impaired Th2 cell commitment *in vitro* and to deregulation of the Th1/Th2 balance *in vivo*.

(A-G) Unless stated otherwise, *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells were analyzed after being cultured for two days in Th1 polarizing conditions. (A) Production of IL-13 and IFN-γ, as determined by flow cytometry after two days (top) and four days (bottom) of culture in Th1 medium. (B) Percentages of cells producing IL-13 and/or IFN-γ after two days of culture, as calculated from the data in (A). (C) Percentages of cells producing IL-4, as determined by flow cytometry. (D) Expression of GATA-3 and T-bet, as determined by flow cytometry. (E) Average expression of GATA-3 and T-bet, as calculated from the data in (D). (F) Th2- and Th1-related gene expression. Differentially expressed genes (adjusted p-value <0.1 and fold difference >2 or <0.5) are indicated by the gene names in red. (G) GSEA of *Setdb1*$^{-/-}$ Th2 cells, with genes ranked on the basis of expression in *Setdb1*$^{-/-}$ Th2 cells relative to that in their *Setdb1*$^{+/+}$ counterparts. (H-K) *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ mice were immunized *s.c.* with 1W1K peptide in RIBI adjuvant. Eight days after, the draining lymph nodes were collected and the antigen-specific CD4^{+} T cell response was analyzed. The frequency of antigen-specific cells among CD4^{+} T cells was first determined by flow cytometry using the 1W1K/I-A^{b} tetramer. Representative dot-plots (H) and frequency of 1W1K/I-A^{b+}CD44^{+} cells among CD4^{+} T cells for each mouse (I) are shown. (J) Percentages of IFN-γ producing cells among 1W1K/I-A^{b+} CD4^{+} T cells, as determined by flow cytometry. The percentages are expressed relative to the size of the antigen-specific CD4^{+} T cell compartment, as determined in (I). (K) IFN-γ and IL-13 production by CD4^{+} T cells was determined by ELISA following antigen-specific restimulation. The IFN-γ to IL-13 ratio is shown for each mouse. Data are representative of eight independent experiments (A, D) or are represented as mean ± SEM (B, E) of eight independent experiments or as mean ± SD of three independent biological replicates I. The data in (F) and (G) are from two (*Setdb1*$^{+/+}$) or three (*Setdb1*$^{-/-}$) independent biological replicates. *p<0.05, ***p<0.001 (unpaired Student's t-test). See also Figure S4 and Table S2.

Figure 4. SETDB1-dependent H3K9 trimethylation at a subset of ERVs.

(A) Random and observed genomic distribution of H3K9me3 domains in *Setdb1*$^{+/+}$ Th2 cells, as measured by ChIP–seq. (B) Relative enrichment of H3K9me3 domains at indicated genomic elements. ***p<0.001 Pearson's chi-squared test. (C) Numbers of H3K9me3 domains found at ERVs, enhancers, or both. (D) H3K9me3 signal distribution across each H3K9me3^{+} enhancer or ERV sequence. The “0” on the X-axis corresponds to the center of the indicated genomic element. (E) Average H3K9me3 signal profiles at H3K9me3^{+} ERVs and enhancers. (F) Average H3K9me3 signal profiles at Th2 enhancers associated with H3K9me3^{-} (“H3K9me3^{-} ERVs”) or H3K9me3^{+} ERVs showing (“Active Enh.”) or not (“Poised Enh.”) enrichment for H3K27ac in Th2 cells. (G) H3K9me3 signal (median) at enhancers that did not overlap with H3K9me3 domains and that were located at a distance of 0 to 1kb (red), 1 to 5kb (blue) or more than 5kb (black) from an ERV marked by H3K9me3. (H) Number of H3K9me3 domains overlapping ERVs and enhancers which themselves overlapped (o) or were mutually exclusive (non-overlapping, n.o.). (I) Box plots representing the distance between H3K9me3 domains and ERVs or enhancers in the two situations described in (H). ***p<0.001 (Wilcoxon test). (J) Percentage of H3K9me3 domains overlapping ERVs in *Setdb1*$^{-/-}$ and *Suv39h1*$^{-/-}$ Th2 cells (relative to *Setdb1*$^{+/+}$ Th2 cells). All data are from two independent biological replicates for each genotype. See also Figure S5 and Table S3.

Figure 5. Increased expression of ERVs and neighboring genes in *Setdb1*$^{-/-}$ cells

Analyses were focused on ERVs overlapping a SETDB1-dependent H3K9me3 domain in Th2 cells. ERVs overlapping gene bodies or promoters were excluded from the analysis. (A) Proportions and numbers of ERVs that were differentially (“Changed”) or similarly (“Stable”) expressed between *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells are shown on the left. The percentages of differentially expressed ERVs that were up-regulated (“Up”) or down-regulated (“Down”) are shown on the right. (B) Expression levels of all ERVs that were differentially expressed in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells. (C) Average ATAC-Seq signal profiles at ERVs that overlapped (H3K9me3^{+}), or not (H3K9me3^{-}), with H3K9me3 domains in Th2 cells and that were more expressed (“H3K9me3^{+} up”, fold change >10, expression in *Setdb1*$^{-/-}$≥1), or not (H3K9me3^{+}), in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ cells. The data represent the signal measured in *Setdb1*$^{-/-}$ Th2 cells relative to that measured in their wild-type counterparts. (D) Expression levels of the 238 enhancers located in the vicinity of ERVs that overlapped with H3K9me3 domains in *Setdb1*$^{+/+}$ cells and were more expressed in *Setdb1*$^{-/-}$ cells. ***p<0.001 Pearson's chi-squared test. (E) Average H3K4me1 or H3K27ac signal profiles at Th2 enhancers associated with H3K9me3^{+} ERVs that were more expressed in *Setdb1*$^{-/-}$ Th2 cells. (F) GSEA of *Setdb1*$^{-/-}$ Th2 cells, with genes ranked on the basis of their expression in *Setdb1*$^{-/-}$ vs *Setdb1*$^{+/+}$ Th2 cells. The gene set comprises all the

genes associated with enhancers located in the vicinity of ERVs that overlapped with H3K9me3 domains and were more expressed in *Setdb1*$^{-/-}$ cells (left panel). A random selection of 500 genes associated with enhancers that remained silent in *Setdb1*$^{-/-}$ cells was used as control (right panel). (G) Gene expression change between *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ Th2 cells for genes that were expressed in at least one population and that were located in the vicinity of H3K9me3^{+} ERVs that remained silent in *Setdb1*$^{-/-}$ cells or were more expressed in *Setdb1*$^{-/-}$ than in *Setdb1*$^{+/+}$ cells. *p<0.05, ***p<0.001 (Wilcoxon test). All transcriptomic data are from three independent biological replicates. ChIP- and ATAC-seq data are from two independent biological replicates. See also Figure S6.

Figure 6. H3K9me3^{+} ERVs mark the Th1 enhancer landscape in Th2 cells.

Only ERVs that did not overlap with Th2 enhancers were analyzed. (A) Biological functions assigned to SETDB1-dependent H3K9me3^{+} ERVs by GREAT. (B) Frequency histograms of absolute distances from each H3K9me3^{-} (Th2) or H3K9me3^{+} (Th2 or adipocytes) ERV to the nearest ISG. Statistical significance of the observed enrichment within the first 10 kb of the nearest ISG was assessed by chi-squared test. (C) Th1-related transcription factor binding sites showing enrichment within the 13,303 ERVs sequences. (D) The transcriptional regulators that might account for the differences in gene expression observed between *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells following culture in Th1-polarizing conditions were identified by using the Upstream Regulator Analysis module of IPA. The activation z-score infers the activation state of the transcriptional regulators in *Setdb1*$^{-/-}$ cells by comparing the differences in gene expression observed in *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ cells to the predicted effect (activating or inhibiting) of each transcriptional regulator on these genes. Hashtags highlight transcriptional regulators whose binding sites are enriched at ERVs. Only transcriptional regulators with an overlap p-value of less than 0.01 and a z-score lower than -2 or greater than 2 are shown. (E, G) Absolute numbers of H3K9me3^{+} ERVs associated with Th1-specific STAT4 (E) or STAT1 (G) binding sites in Th2 cells (black bars) compared to counts obtained with a random distribution (shuffle, white bar). (F, H) Relative enrichment of H3K9me3^{+} or H3K9me3^{-} ERVs at Th1-specific STAT4 (F) or STAT1 (H) peaks in adipocytes and Th2 cells. (I) Numbers of putative Th1 enhancers associated with SETDB1-dependent H3K9me3^{+}H3K4me1^{-} ERVs in Th2 cells. (J) Relative enrichment of Th1 enhancers at H3K9me3^{+} ERVs in adipocytes and Th2 cells. (K) Relative proportions of Th1 enhancers intersecting with H3K9me3 domains that overlapped, or not, with ERVs in Th2 cells. (L) Box plots representing the distance between H3K9me3 domains and ERVs or Th1 enhancers. ***p<0.001 (chi-squared test (B and E-J) or Wilcoxon test (L)). See also Figure S7.

Figure 7. SETDB1-dependent H3K9 trimethylation at ERVs represses Th1-specific enhancers.

(A) Proportions of Th1 enhancers associated with H3K9me3^{+} ERVs in Th2 cells that were either active (H3K4me1^{+}H3K27ac^{+}) or repressed (H3K4me1^{+}H3K27ac^{-} or H3K4me1^{-} H3K27ac^{-}). The percentages of repressed enhancers that were poised (H3K4me1^{+}H3K27ac^{-}) or "ghost" (H3K4me1^{-}H3K27ac^{-}) are shown on the right. (B) Average H3K9me3 signal profiles in *Setdb1*$^{+/+}$ Th2 cells at Th1 enhancers that were associated with H3K9me3^{+} ERVs and that were either active, poised or repressed in *Setdb1*$^{+/+}$ Th2 cells. (C) Average ATAC-seq signal profiles in *Setdb1*$^{-/-}$ Th2 cells at Th1 enhancers that were repressed and associated with H3K9me3^{+} ERVs in *Setdb1*$^{+/+}$ Th2 cells and that were more expressed, or not ("Silent"), in the absence of SETDB1. (D, E) Average H3K4me1 (D) and H3K27ac (E) signal profiles in *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ Th2 cells at Th1 enhancers that were repressed and associated with H3K9me3^{+} ERVs in *Setdb1*$^{+/+}$ Th2 cells and that were more expressed, or not ("Silent"), in the absence of SETDB1. (F,G) IGV snapshots of the *Ifng* locus. Colored boxes indicate peaks for H3K4me1 (green) and H3K9me3 (red). Black boxes correspond to ERV coordinates. (H) Enrichment of H3K4me1 at $CNS_{17\text{-}20}$ of the *Ifng* gene in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells as measured by ChIP–qPCR. Data are means ± SEM of three independent biological replicates. (I) Th1 gene network generated with the IPA software showing the main Th1-related genes that had at least one enhancer in the vicinity (± 5kb) of a SETDB1-dependent H3K9me3^{+}H3K4me1^{-} ERV in Th2 cells.

STAR METHODS

KEY RESOURCES TABLE

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by the corresponding authors Véronique Adoue (veronique.adoue@inserm.fr) and Olivier P Joffr.fr).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

Suv39h1-deficient mice were kindly provided by T. Jenuwein (Peters et al., 2001). The *Setdb1* mutant mouse strain (common strain name EPD0028_1_B07; international strain designation B6Dnk;B6N-*Setdb1*$^{tm1a(EUCOMM)Wtsi}$) was established as part of the International Mouse Phenotyping Consortium (EMMA ID: EM:04052) at the German Research Center for Environmental Health (Helmholtz Zentrum, Muenchen). The targeting vector was composed of the promoterless L1L2_gt1 cassette inserted in the L3L4_pZero_kan plasmid backbone. The construct was microinjected into C57BL/6 ES cells (JM8.N4 parental cell line) and the L1L2_gt1 cassette was inserted at position 95350414 of chromosome 3, upstream of *Setdb1* exon 4. The cassette was composed of a lacZ–neomycin sequence flanked by Flp Recombinase Target (FRT) sites and followed by a loxP sequence. An additional loxP site was inserted downstream of *Setdb1* exon 4 at position 95349598. Additional information on the *Setdb1* mutant mouse strain can be found at https://www.infrafrontier.eu/search?keyword=EM:04052. Mice with a conditional ready *Setdb1* allele (*Setdb1*^{fl}) were generated by intercrossing *Setdb1* mutant mice with mice expressing the Flipper recombinase under the control of the ubiquitous *Rosa26* promoter. Conditional *Setdb1*-deficient mice (*Setdb1*$^{-/-}$) were obtained by intercrossing *Setdb1*$^{fl/fl}$ and *CD4*-CRE mutant mice. All the mice were bred and housed at the Regional Centre of Functional Exploration and Experimental Resources (CREFRE, UMS006/INSERM). Sex-matched 6- to 12-week-old mice were used and compared in all experiments. All experiments involving animals were conducted according to animal study protocols approved by the local ethics committee (# 16-U1043-JVM-496 and 16-U1043-JVM-20).

METHOD DETAILS

Naïve CD4^{+} T cell isolation

Spleen and lymph nodes (mesenteric, inguinal, axillary, brachial and cervical) were collected and digested with Liberase TM and DNAse I (Sigma). Single-cell suspensions were then pooled and depleted of erythrocytes by osmotic shock (Red Blood Cell Lysis

buffer, Sigma). $CD4^+$ T cells were enriched by negative selection by using antibodies specific for CD16/32 (2.4G2), I-A/I-E (M5/114.15.2), CD8α (H59) and B220 (RA3-6B2), and Dynabeads sheep anti-rat IgG (Thermo Fisher Scientific). Naïve $CD4^+$ T cells, defined as $CD4^+CD25^-CD62L^{high}CD44^{low}$, were labeled with fluorochrome-conjugated monoclonal antibodies specific for CD4 (GK1.5, BD Biosciences), CD25 (PC61, BD Biosciences), CD62L (MEL14, Thermo Fisher Scientific) and CD44 (IM7, BD Biosciences), and purified from the enriched fraction of $CD4^+$ T cells by fluorescence-activated cell sorting (FACS Aria, BD Biosciences).

Th cell cultures

Naïve $CD4^+$ T cells were cultured for three days in 96-well flat bottom plates coated with 10 μg/mL anti-CD3ε antibody (145-2C11, InVivoMab™, BioXcell) in RPMI 1640 Glutamax™ supplemented with 1 mM sodium pyruvate, non-essential amino acids, 10 mM HEPES, 100 units/mL penicillin, 100 μg/mL streptomycin, 50 mM 2β-mercaptoethanol, 10% fetal calf serum (all from Thermo Fisher Scientific) and 1 μg/mL anti-CD28 antibody (37.51, InVivoMab™, BioXcell). Unless stated otherwise, Th1 medium also contained 10 ng/mL recombinant mouse IL-12 (R&D Systems) and 10 μg/mL anti-IL-4 neutralizing antibody (11B11, InVivoMab™, BioXcell). Th2 medium contained 50 ng/mL recombinant mouse IL-4 (R&D systems) and 10 μg/mL anti-IFN-γ neutralizing antibody (XMG1.2, InVivoMab™, BioXcell). At day 3, the cells were re-plated in the same conditioning medium but without the anti-CD3ε and anti-CD28 antibodies and with 30 IU/mL recombinant IL-2 (Proleukin). To test for Th2 cell lineage commitment, cells were harvested at day 6, extensively washed in complete medium, and re-plated in Th1-polarizing conditions as indicated above. To assess the role of the IFN-γ pathway in Th2 cell plasticity, Th1 medium was supplemented with 10 μg/mL anti-IFN-γ. In co-culture experiments, *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells were differentiated separately, mixed at a 1:3 ratio, and then plated in Th1 culture conditions.

T cell proliferation and differentiation analysis by flow cytometry

To analyze intracellular transcription factor expression upon Th cell differentiation, cells were collected at the requires time points, stained with the fixable viability dye eFluor 506 (Thermo Fisher Scientific), and labeled with fluorochrome-conjugated antibodies specific for T-bet (ebio4B10, Thermo Fisher Scientific) and GATA-3 (TWAJ, Thermo Fisher Scientific) by means of the Transcription Factor Staining Buffer Set (Thermo Fisher Scientific). For intracellular cytokine staining, cells were first stimulated at 37°C with 20 ng/mL phorbol 12-myristate 13-acetate (Millipore) and 1 μg/mL ionomycin (Millipore) for 5 hours in the presence of GolgiStop™ (BD Biosciences). Cells were then labelled with the fixable viability

dye eFluor 506 and stained with fluorochrome-coupled antibodies specific for IL-13 (ebio13A, Thermo Fisher Scientific), IFN-γ (XMG1.2, Thermo Fisher Scientific), GM-CSF (MP1-22E9, BD Biosciences), TNF (MP6-XT22, Thermo Fisher Scientific) or IL-4 (11B11, BD Biosciences) by using the Cytofix/Cytoperm™ Fixation/Permeabilization Kit (BD Biosciences). When indicated, naïve $CD4^+$ T cells were labeled prior to culture with 0.5 μM CellTrace Violet (Thermo Fisher Scientific). Flow cytometry was performed by using a LSRII Fortessa cytometer (BD Biosciences) or MACSQuant analyzer 10 (Myltenyi) and the data were analyzed by using FlowJo software (Tree Star).

Mouse phenotyping

To determine the frequency and phenotype of immune cells in primary and secondary lymphoid organs, thymus, spleen and lymph nodes were collected from *Setdb1*$^{-/-}$ and *Setdb1*$^{+/+}$ mice and single-cell suspensions were obtained by mechanical disruption. Cells were then incubated on ice in FACS buffer (PBS, 3 mM EDTA, 3% fetal calf serum) containing 10 μg/mL anti-CD16/32 antibody for 20 minutes. Fluorochrome-conjugated antibodies were added at saturating concentrations, and cell suspensions were incubated on ice and protected from light for a further 20 minutes. For intracellular staining, cells were fixed and permeabilized by using the Transcription Factor Staining Buffer Set (Thermo Fisher Scientific) according to the manufacturer's instructions. The following antibodies were used for phenotyping: anti-TCR-β (H57-597), anti-CD4 (GK1.5), anti-NKp46 (29A1.4), anti-CD11b (M1/70), anti-CD19 (1D3), anti-CD25 (PC61), anti-CD69 (H1.2F3), anti-Siglec-F (E50-2440), anti-H-$2K^b$ (AF6-88.5) and anti-Ki67 (B56) all from BD Biosciences; anti-CD8β (ebioH35-17.2), anti-CD8α (53-6.7), anti-PDCA1 (ebio927), anti-I-A/I-E (M5/114), anti-CD11c (N418), anti-Gr1 (RB6-8C5), anti-B220 (RA3-6B2), anti-CD44 (IM7) and anti-CD62L (MEL-14), all from Thermo Fisher Scientific. Dendritic cells (DC) were gated based on $I\text{-}A^b$ and CD11c expression ($CD19^-TCR\text{-}\beta^-CD11c^+I\text{-}Ab^+$) and CD8α, CD11b and PDCA-1 were used as markers to identify the conventional type 1 (cDC1, $CD8\alpha^+CD11b^-$), conventional type 2 (cDC2, $CD8\alpha^-CD11b^+$) and plasmacytoid (pDC, $PDCA\text{-}1^+$) sub-populations. Monocytes/macrophages (Macro) and B cells (B cell) were defined as $lin^-CD11c^-CD11b^+SSC\text{-}A^{low}Gr\text{-}1^{low/-}$ and $TCR\text{-}\beta^-CD19^+B220^+$, respectively. Neutrophils (Neutro), natural killer cells (NK) and eosinophils (Eosino) were identified as $lin^-CD11c^-CD11b^+SSC\text{-}A^{high}Gr\text{-}1^+$, $TCR\text{-}\beta^-NKP46^+$ and $lin^-CD11c^-CD11b^+SSC\text{-}A^{high}Gr\text{-}1^-$, respectively. Flow cytometry was performed by using a LSRII Fortessa cytometer (BD Biosciences) and the data were analyzed by using FlowJo software (Tree Star).

***Ex vivo* measurement of apoptosis**

Single-cell suspensions of spleen, mesenteric lymph nodes and thymus were obtained as described above. Spleen and lymph node cells were labeled with antibodies specific for TCR-β (H57-597, Thermo Fisher Scientific) and CD4 (GK1.5, BD Biosciences), whereas thymocytes were stained with antibodies specific for CD4 and CD8β (ebioH35-17.2, ThermoFisher Scientific). Apoptotic cells were then labeled by using the Cell Event™ Caspase-3/7 Green Detection Reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. Flow cytometry was performed by using a LSRII Fortessa cytometer (BD Biosciences) and the data were analyzed by using FlowJo software (Tree Star).

Measurement of cytokines in cell culture supernatants

Following 6 days of culture in Th1 or Th2 polarizing conditions, T cells were collected and extensively washed in complete medium. The differentiated cells (7.5×10^4 per well) were cultured overnight in 96-well flat bottom plates coated with anti-CD3ε antibody in complete culture medium containing anti-CD28 antibody. The concentrations of cytokines in the cell culture supernatants were then measured by flow cytometry using the FlowCytomix Kit (a bead-based multiple cytokines detection system) according to the manufacturer's instructions (FlowCytomix, eBiosciences). Flow cytometry was performed by using a MACSquant Q10 flow cytometer (Miltenyi).

Phospho-STAT4 (Tyr693) staining

Naïve $CD4^+$ T cells were cultured for three days in 96-well flat bottom plates in complete medium supplemented with 30 UI/mL of recombinant IL-2 (Peprotech) at a ratio of 1 to 1 with Dynabeads Mouse T-Activator CD3/CD28 (Thermo Fisher Scientific). After 3 washes in complete medium, 10^5 cells were stimulated for 30 minutes with recombinant mouse IL-12 (10 ng/mL). Cells were then fixed with BD Phosflow™ Lyse/Fix Buffer, permeabilized with BD Phosflow™ Perm Buffer III in the presence of 10mM Sodium Fluoride and 1mM Sodium Orthovanadate (Sigma), and stained with fluorochrome-conjugated monoclonal antibodies specific for phosphorylated STAT4 (pY693, clone 38/pSTAT4, BD). Flow cytometry was performed by using a LSRII Fortessa cytometer and the data were analyzed by using FlowJo software.

Mixed hematopoietic bone marrow chimeras

Bone marrow from femurs and tibias was collected from $CD45.2^+CD90.1^+$ C57BL/6 mice and from *Setdb1*$^{+/+}$ or *Setdb1*$^{-/-}$ $CD45.2^+CD90.2^+$ littermates. Single-cell suspensions were washed in complete medium. $CD90.1^+$ and $CD90.2^+$ cells were then eliminated using HO-22 and AT83 hybridoma supernatants, respectively, and rabbit complement (TCS

Biosciences). Following treatment with DNase I (Sigma), single-cell suspensions were washed three times in PBS and filtered through a 70 μm cell strainer. $CD45.2^+CD90.1^+$ cells were mixed at a ratio of 3 to 7 with either *Setdb1*$^{+/+}$ or *Setdb1*$^{-/-}$ $CD45.2^+CD90.2^+$ cells. 10^7 cells were then injected intravenously into γ-irradiated $CD45.1^+$ C57BL/6 hosts (11 Gy; ^{137}Cs source) that were kept on antibiotic-containing water (0.28% pediatric suspension of Bactrim; Roche) for the next 4 weeks. Spleen and lymph nodes were isolated from these mice at least 6 weeks after injection of the bone marrow cells.

Immunization

Mice were immunized subcutaneously on each side of the base of the tail with 20 μg of 1W1K peptide (EAWGALANKAVDKA, Genecust) in 100 μL of RIBI adjuvant (Sigma). 8 days after immunization, the lymph nodes draining the site of immunization were collected and homogenized. To determine the frequency of Ag-specific $CD4^+$ T cells, 4.10^6 cells were incubated with the 1W1K-I-A^b tetramer (NIH tetramers core facility) for 2h at room temperature. Cells were then labelled with fluorochrome-conjugated monoclonal antibodies specific for CD4 (RM4.5), CD44 (IM7), CD8α (53-6.7) and B220 (RA3-6B2) and with the fixable viability dye eFluor 506 (Thermo Fisher Scientific) according to the manufacturer's instructions. To determine the frequency of antigen-specific $CD4^+$ T cells producing IFN-γ, 4.10^6 lymph node cells were stimulated overnight with 7μM of the 1W1K peptide in complete medium and in the presence of Brefeldin A (5μg/mL, Sigma) and GolgiStop™ (BD Biosciences) during the last 5h. Cells were then harvested and stained with fluorochrome-conjugated monoclonal antibodies specific for CD4 and CD44 and with the fixable viability dye eFluor 506. Following fixation and permeabilization (Cytofix/Cytoperm™ Fixation/Permeabilization Kit, Thermo Fisher Scientific), cells were finally stained with fluorochrome-coupled antibodies specific for IFN-γ. To determine the Th1/Th2 balance of the antigen-specific T cell response, 2.10^6 cells were stimulated for 72h with 7μM of 1W1K peptide in HL-1 medium (Lonza) supplemented with 2 mM of glutamine, 100 units/mL penicillin and 100 μg/mL streptomycin. IL-13 and IFN-γ levels were then measured by ELISA in the cell culture supernatants.

Western blotting

The different subpopulations of thymocytes were sorted on a FACS Aria (BD Biosciences) based on their expression of CD4 and CD8. Naïve $CD4^+$ T cells were purified from the spleen and lymph nodes as described above. Cells were lysed in 1X NuPAGE LDS sample buffer and 1X NuPAGE sample reducing agent (Thermo Fisher Scientific). Whole cell lysates were then sonicated briefly and proteins were separated by SDS-PAGE on 4–12%

Bis-Tris gels (Thermo Fisher Scientific), transferred onto nitrocellulose membranes (BA-S 83 Optitran, GE Healthcare Life Sciences), and probed with antibodies specific for SETDB1 (ab107225, Abcam), total H3 (ab1791, Abcam), H3K9me3 (D4W1U, Cell Signaling Technology) or beta ACTIN (ab8227, Abcam). The bands were detected by using Amersham ECL western blotting detection reagent (GE Healthcare Life Sciences) and the ChemiDoc XRS+ imaging system (Bio-Rad) after staining with secondary antibodies coupled to horseradish peroxidase. Images were analyzed by using Image Lab software (Bio-Rad).

RNA–seq sample preparation and analysis

Total RNA was extracted by using the RNeasy Micro Kit (Qiagen) and its quality was assessed on a 2100 Bioanalyzer (Agilent Technologies). Only high-quality RNA (i.e. RNA of integrity number > 7) was subsequently used to prepare the libraries according to the ScriptSeq RNA–seq protocol (Illumina). Quality controls of the libraries were performed by using standard methods, including quantification by Qubit (Thermo Fisher Scientific) and assessment of size distribution by using the 2100 Bioanalyzer. Samples were indexed and sequenced on an Illumina HiSeq 2500 or 3000 (paired-end reads of 100 or 150 bp, respectively). After trimming of adaptor sequences (Cutadapt 1.3) and removal of low-quality bases (−q value, < 15), high-quality reads were aligned to the mouse reference genome mm10 (Genome Reference Consortium) by using TopHat version 2.0.5 (Trapnell et al., 2009). Count of the reads mapping to each gene was performed using Htseq-count. Differential expression analysis was performed by using the DESeq package (Bioconductor software) (Anders and Huber, 2010). An adjusted P value of < 0.1 (P value adjusted for multiple testing with the Benjamini–Hochberg procedure) was used as cutoff to select the genes differentially expressed.

ChIP, semi-quantitative PCR and library preparation and sequencing

ChIP was performed as previously described (Lee et al., 2006). Briefly, following cell lysis, the chromatin was sonicated with a Bioruptor Pico (Diagenode) to obtain fragments of 100-300 bp. In each assay, we used 5-50 million cells and 2-10 µg of antibody specific for H3K9me3 (ab8898, Abcam), H3K27ac (Ab4729, Abcam) or H3K4me1 (ab8895, Abcam). Immunoprecipitation was performed by using Dynabead® Protein G (Thermo Fisher Scientific). Library preparation was carried out by using the TruSeq ChIP Sample Preparation Kit (Illumina). Library quality was assessed by using the 2100 Bioanalyzer and sequencing was performed on an Illumina HiSeq 3000 (paired-end reads of 150 bp). When indicated, semi-quantitative PCR was performed on a Light Cycler® 480 (Roche) using LightCycler 480 SYBR Green I Master (Roche). Primers specific for *Ifng* $CNS_{17\text{-}20}$ (forward:

tccctagactctgccactct; and reverse: gctcaccatcaataggcgtg) and for the glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) promoter (forward: gctccttgcccttccagatt and reverse: cccttcccaccctgttcatc) were used. The results were expressed as the percentage of input DNA normalized to the signal from the *Gapdh* promoter. More details on STAT1, STAT4 and H3K4me1 ChIP-seq data in Th1 cells can be found in (Vahedi et al., 2012).

ChIP–seq data processing

Reads were filtered as described for RNA-seq and aligned to the mm10 reference genome by using BWA v.0.7.10 (Li and Durbin, 2009). H3K4me1 and H3K27ac peaks were detected by using the 'broad' option of MACS2 v.2.1.0 (Zhang et al., 2008). To detect H3K9me3 domains, we used the R Bioconductor package CSAW v.1.4.1 (Lun and Smyth, 2015). The minimum mapping quality score and the FDR threshold were set to 50 and 0.05, respectively. We tested 11 window sizes ranging from 200 to 600 bp. In the end, we selected 300 bp as it allowed for the most accurate detection of H3K9me3-enriched domains, as determined using the visualization tool IGV. Differential binding windows were clustered in regions with the 'mergeWindows' function and the Benjamini-Hochberg method was applied to control the False Discovery Rate (FDR) across all detected clusters ('combineTests' function). For vizualizing H3K4me1 and H3K27ac signals in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells, reads were extended by 200 bp because of low reads coverage.

ATAC-seq sample preparation and analysis

ATAC-seq was performed as previously described (Buenrostro et al., 2015) with some modifications. Briefly, 50 000 cells were lysed in ice-cold lysis buffer and the transposition reaction was performed using the Tn5 transposase at 37 °C for 30 min. DNA was purified using the Qiagen MinElute kit (Qiagen). The libraries were prepared using the Nextera DNA Library Prep Kit (Illumina) and purified using AMPure XP beads (Beckman) following a double-sided protocol to remove primer dimers and large fragments. Samples were performed in duplicates, multiplexed and sequenced on NextSeq-500 (75 bp paired-end reads) at the Transcriptomic & Genomic Platform Marseille Luminy (TGML, Marseille, France). Reads were aligned with BWA mem (version 0.7.12-r1039). Samples were normalized by scaling reads numbers according to the signal found at promoter of genes which were highly (expression level > 5000) and equally (0.9 < adjusted p-value) expressed in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ samples, as determined by RNA-seq.

ERV reconstruction

Annotations of ERV elements were downloaded from the UCSC Genome Browser (assembly GRCm38, release of RepeatMasker: 2012-02-07). We used the four major

subfamilies (*ERV1*, *ERVK*, *ERVL* and *ERVL-MaLR*) of LTRs and excluded elements for which the curator was unsure of the classification. We merged ERV fragments from the same family (identical 'repName') into a single ERV when located within 20 bp, as previously described (Göke et al., 2015). Count of the reads mapping to each ERV was performed using Htseq-count(Anders et al., 2015) and normalization was performed with DESeq. ERVs with an expression score ≥1 were considered as expressed.

Bioinformatics analyses

R (https://www.R-project.org), SAMtools (Li et al., 2009) and the BEDtools suite v2.22.1 (Quinlan and Hall, 2010) were used to analyze high-throughput sequencing files. To determine the genome-wide distribution of H3K9me3 domains, we defined the different genomic regions as follows: gene body coordinates were extracted from assembly GRCm38; promoters were defined as transcription start sites +1kb/-2kb; enhancers were identified as H3K4me1 peaks with no overlap with promoters; ERV coordinates were rebuilt from the RepeatMasker database as described above. As a control, we randomly distributed H3K9me3 domains through the genome using the shuffle sub-command of the BEDtools suite. R package 'Genomation' was used to visualize genomic intervals (Akalin et al., 2015). To represent to what extend the retroelements marked by H3K9me3 in wild-type cells were still covered by the repressive mark in *Setdb1*$^{-/-}$ or *Suv39h1*$^{-/-}$ Th2 cells, we expressed the number of peaks that remained in mutant contexts as percentage of wild-type. Biological functions analysis of H3K9me3 ChIP–seq peak coordinates was performed by using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010) with default settings and using the 'single nearest gene' option (each gene is assigned a regulatory domain that extends in both directions (within 100kb) to the midpoint between the gene's TSS and the nearest gene's TSS but no more than the maximum extension in one direction). For Gene Ontology analysis, the "enrichment analysis" tool from the Gene Ontology Consortium was used (http://geneontology.org). For analysis of motif enrichment, we used AME software of the MEME Suite version 4.11.3 with defaults options (0.05 ≥ adjusted p-value) (McLeay and Bailey, 2010); the HOCOMOCOv11 MOUSE was used as the input transcription factor motif database. We also used the gene set enrichment analysis (GSEA) software (http://software.broadinstitute.org/gsea/index.jsp) with default settings except for the 'Collapse dataset to gene symbols' and 'the permutation type' which were set as 'false' and 'gene set', respectively. Heat maps were generated by using matrix2png version 1.2.1 (http://www.chibi.ubc.ca/matrix2png). To measure H3K9me3, H3K4me1 or H3K27ac signal, we used bamCoverage from the deepTools suite (v2.3.4) to generate normalized bedgraph files (we normalized by sequencing depth and ignored chrX, chrY and chrM) and Bedtools map to calculate the average signal. Correlation analysis

between H3K9me3^{+} or H3K9me3^{-} ERVs and IFN-γ-stimulated genes (ISG) was performed as described elsewhere (Chuong et al., 2016). The absolute distance between ERVs and the nearest ISG (n=4,276) was first determined for all 13,303 H3K9me3^{+}H3K4me1^{-} retroelements. The distances were then grouped by 10 kb bin sizes. As control, a similar analysis was performed using an equal number of randomly sampled H3K9me3^{-} ERVs. Sampling was repeated 10 times and the mean number of elements was used. Statistical significance was determined for the first 10 kb bin by chi-squared test.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical parameters including the exact value and significance of n and precision measures (Mean +/- SEM or SD) as well as statistical significance are reported in the figures and figure legends. Unless stated otherwise, asterisks denote statistical significance as calculated by Student's t test in GraphPad PRISM 6 (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). When large sets of unpaired data were compared, Pearson's chi-squared test was calculated in R to determine whether the observed difference between the sets of data arose by chance.

DATA AVAILABILITY

Data Resources

Raw and processed data files from ChIP–seq and RNA–seq experiments have been deposited in the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE101546.

SUPPLEMENTAL TABLES

Table S1. Related to Figure 1. mRNA levels of genes related to Th cell differentiation in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ naïve CD4^{+} T cells.

Table S2. Related to Figure 3. mRNA levels of genes related to the Th1- or Th2-lineages in *Setdb1*$^{+/+}$ and *Setdb1*$^{-/-}$ Th2 cells following culture in Th1-polarizing conditions.

Table S3. Related to Figure 4. Family, name and coordinates of H3K9me3^{+} ERVs in *Setdb1*$^{+/+}$ Th2 cells.

ACKNOWLEDGMENTS

We greatly acknowledge T. Jenuwein (Max Planck Institute, Freiburg, Germany) for providing the *Suv39h1*-deficient mice. We also thank F.-E. L'Faqihi-Olive, V. Duplan-Eche and A.-L. Iscache for technical assistance at the flow cytometry facility of INSERM U1043, the personnel of the US006 ANEXPLO/CREFRE animal facility for expert animal care, D. Rozet for administrative assistance, the GeT and Bioinformatics platforms from the Genotoul (Toulouse, Région Occitanie, France), the Transcriptomic & Genomic Platform Marseille Luminy (TGML, Marseille, France), and the Genomic and Transcriptomic platform from the CPTP (Toulouse, Région Occitanie, France). We are also grateful to members of the "T cell-mediated Immune Tolerance" laboratory for advice and discussions. We would like to thank D. Dunia for his careful reading of the manuscript. This work was supported by Agence Nationale de la Recherche (ANR JCJC "EpiTreg" to O.J.), Région Occitanie (NVEQ 2014 to O.J.) and Fondation pour la Recherche Médicale (FRM, AJE201212 to O.J.). B.B. was funded by a fellowship from the Association pour la Recherche sur le Cancer (ARC). A.M. was supported by fellowships from the Région Midi-Pyrénées and FRM (FDT20170437348).

AUTHOR CONTRIBUTIONS

V.A., B.B. and A.M. performed the main experiments, analyzed data and contribute to writing the manuscript. J.F. assisted V.A. with ChIP-Seq and RNA-Seq data analyses. V.A., P.R., J.P.M.v.M. and S.A supervised the study. O.J conceived and supervised the study and wrote the manuscript.

CORRESPONDING AUTHORS

Correspondence should be addressed to Olivier P Joffre (olivier.joffre@inserm.fr) or Véronique Adoue (veronique.adoue@inserm.fr).

DECLARATION OF INTERESTS

O.P.J., V.A., B.B. and A.M. have previously filed a patent application based on the use of SETDB1 as a target to treat Th2-mediated and -related diseases. The remaining authors declare no conflict of interest.

REFERENCES

Akalin, A., Franke, V., Vlahoviček, K., Mason, C.E., Schübeler, D., 2015. Genomation: a toolkit to summarize, annotate and visualize genomic intervals. Bioinformatics 31, 1127–1129.

Allan, R.S., Zueva, E., Cammas, F., Schreiber, H.A., Masson, V., Belz, G.T., Roche, D., Maison, C., Quivy, J.-P., Almouzni, G., Amigorena, S., 2012. An epigenetic silencing pathway controlling T helper 2 cell lineage commitment. Nature 487, 249–253.

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169.

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

Buenrostro, J.D., Wu, B., Chang, H.Y., Greenleaf, W.J., 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr. Protoc. Mol. Biol. 109, 21.29. 1-21.29.9.

Bulut-Karslioglu, A., La Rosa-Velázquez, De, I.A., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galán, C., Winter, G.E., Engist, B., Gerle, B., O'Sullivan, R.J., Martens, J.H.A., Walter, J., Manke, T., Lachner, M., Jenuwein, T., 2014. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. Mol. Cell 55, 277–290.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., Makarov, V., Budhu, S., Buhu, S., Slamon, D.J., Wolchok, J.D., Pardoll, D.M., Beckmann, M.W., Zahnow, C.A., Merghoub, T., Mergoub, T., Chan, T.A., Baylin, S.B., Strick, R., 2015. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell 162, 974–986.

Chuong, E.B., Elde, N.C., Feschotte, C., 2017. Regulatory activities of transposable elements: from conflicts to benefits. Nat. Rev. Genet. 18, 71–86.

Chuong, E.B., Elde, N.C., Feschotte, C., 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science 351, 1083–1087.

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

Göke, J., Lu, X., Chan, Y.-S., Ng, H.-H., Ly, L.-H., Sachs, F., Szczerbinska, I., 2015. Dynamic Transcription of Distinct Classes of Endogenous Retroviral Elements Marks Specific Populations of Early Human Embryonic Cells. Cell Stem Cell 16, 135–141.

Hegazy, A.N., Peine, M., Helmstetter, C., Panse, I., Fröhlich, A., Bergthaler, A., Flatz, L., Pinschewer, D.D., Radbruch, A., Löhning, M., 2010. Interferons Direct Th2 Cell Reprogramming to Generate a Stable GATA-3+T-bet+ Cell Subset with Combined Th2 and Th1 Cell Functions. Immunity 32, 116–128.

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

Kanno, Y., Vahedi, G., Hirahara, K., Singleton, K., O'Shea, J.J., 2012. Transcriptional and Epigenetic Control of T Helper Cell Specification: Molecular Mechanisms Underlying Commitment and Plasticity. Annu Rev. Immunol. 30, 707–731.

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

Lee, T.I., Johnstone, S.E., Young, R.A., 2006. Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. 1, 729–748.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Liu, J., Magri, L., Zhang, F., Marsh, N.O., Albrecht, S., Huynh, J.L., Kaur, J., Kuhlmann, T., Zhang, W., Slesinger, P.A., Casaccia, P., 2015. Chromatin Landscape Defined by Repressive Histone Methylation during Oligodendrocyte Differentiation. J Neurosci. 35, 352–365.

Loyola, A., Tagami, H., Bonaldi, T., Roche, D., Quivy, J.-P., Imhof, A., Nakatani, Y., Dent, S.Y.R., Almouzni, G., 2009. The HP1α–CAF1–SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. EMBO Rep. 10, 769–775.

Lun, A.T.L., Smyth, G.K., 2015. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res. 44, e45–e45.

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

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G., 2010. GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28, 495–501.

McLeay, R.C., Bailey, T.L., 2010. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. BMC Bioinformatics 11, 165.

Mozzetta, C., Boyarchuk, E., Pontis, J., Ait-Si-Ali, S., 2015. Sound of silence: the properties and functions of repressive Lys methyltransferases. Nat. Rev. Mol. Cell. Biol. 16, 499–513.

O'Shea, J.J., Lahesmaa, R., Vahedi, G., Laurence, A., Kanno, Y., 2011. Genomic views of STAT function in CD4+ T helper cell differentiation. Nat. Rev. Immunol. 11, 239–250.

O'Shea, J.J., Paul, W.E., 2010. Mechanisms Underlying Lineage Commitment and Plasticity of Helper CD4+ T Cells. Science 327, 1098–1102.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M., Jenuwein, T., 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107, 323–337.

Peters, A.H.F.M., Mermoud, J.E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., Jenuwein, T., 2002. Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. Nat. Genet. 30, 77–80.

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

Rebollo, R., Karimi, M.M., Bilenky, M., Gagnier, L., Miceli-Royer, K., Zhang, Y., Goyal, P., Keane, T.M., Jones, S., Hirst, M., Lorincz, M.C., Mager, D.L., 2011. Retrotransposon-Induced Heterochromatin Spreading in the Mouse Revealed by Insertional Polymorphisms. PLoS Genet. 7, e1002301.

Saint-André, V., Batsché, E., Rachez, C., Muchardt, C., 2011. Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. Nat. Struct. Mol. Biol. 18, 337–344.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., Rauscher, F.J., III, 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. 16, 919–932.

Shnyreva, M., Weaver, W.M., Blanchette, M., Taylor, S.L., Tompa, M., Fitzpatrick, D.R., Wilson C.B., 2004. Evolutionarily conserved sequence elements that positively regulate IFN-γ expression in T cells. Proc. Natl. Acad. Sci. 101, 12622–12627.

Trapnell, C., Pachter, L., Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111.

Vahedi, G., Takahashi, H., Nakayamada, S., Sun, H.-W., Sartorelli, V., Kanno, Y., O'Shea, J.J., 2012. STATs shape the active enhancer landscape of T cell populations. Cell 151, 981–993.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., Blobel, G.A., 2005. Histone H3 Lysine 9 Methylation and HP1γ Are Associated with Transcription Elongation through Mammalian Chromatin. Mol. Cell 19, 381–391.

Wilson, C.B., Rowell, E., Sekimata, M., 2009. Epigenetic control of T-helper-cell differentiation. Nat. Rev. Immunol. 9, 91–105.

Xiao, X., Shi, X., Fan, Y., Wu, C., Zhang, X., Minze, L., Liu, W., Ghobrial, R.M., Lan, P., Li, X.C., 2016. The Costimulatory Receptor OX40 Inhibits Interleukin-17 Expression through Activation of Repressive Chromatin Remodeling Pathways. Immunity 44, 1271–1283.

Young, H.A., Ghosh, P., Ye, J., Lederer, J., Lichtman, A., Gerard, J.R., Penix, L., Wilson, C.B., Melvin, A.J., McGurn, M.E., 1994. Differentiation of the T helper phenotypes by analysis of the methylation state of the IFN-gamma gene. J Immunol. 153, 3603–3610.

Zhang, Y., Tao, L., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based Analysis of ChIP-Seq (MACS). Genome Biol. 9, R137.

Zhu, Y., van Essen, D., Saccani, S., 2012. Cell-type-specific control of enhancer activity by H3K9 trimethylation. Mol. Cell 46, 408–423.

Figure 1

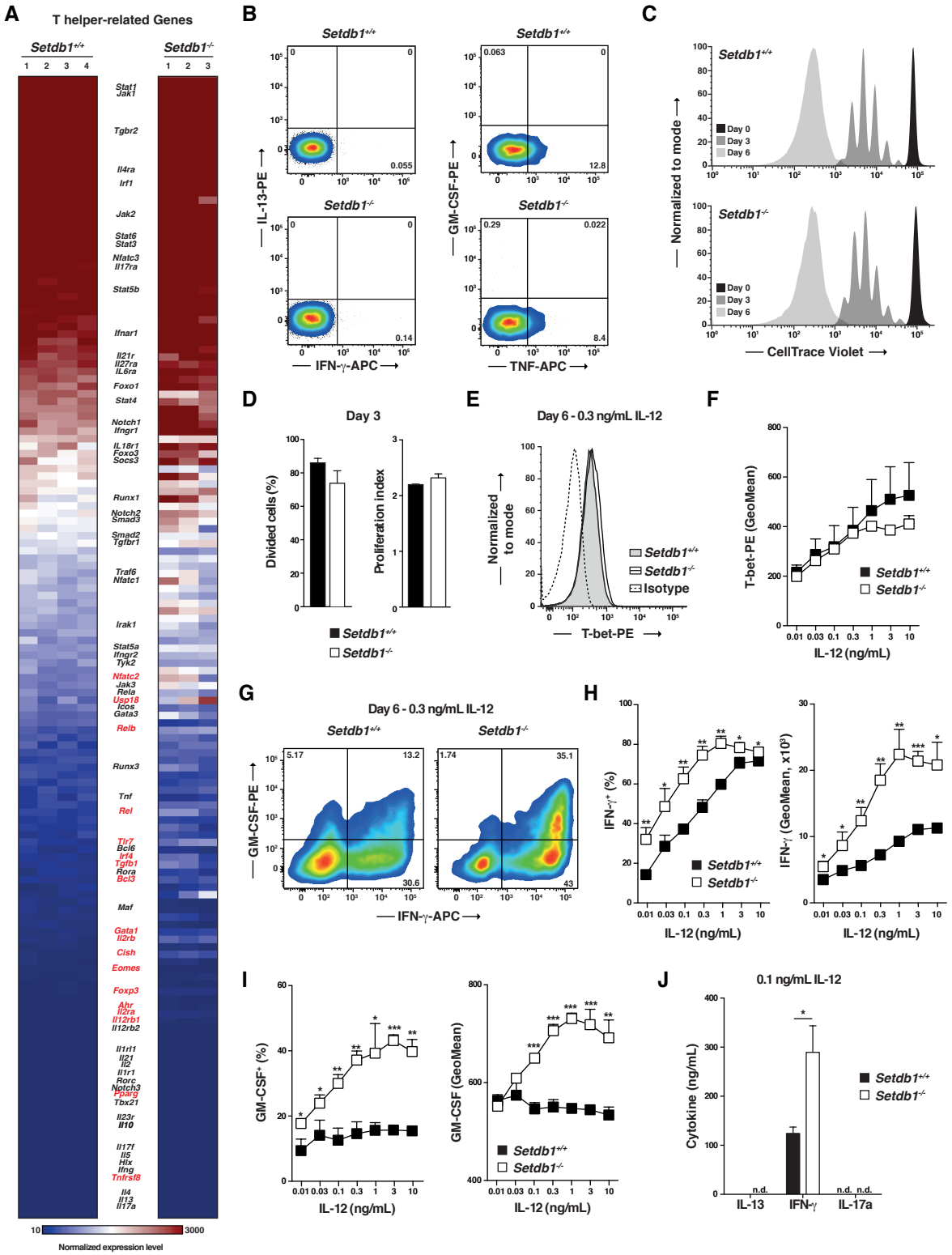

A
T helper-related Genes
Setdb1+/+
Setdb1-/-
Normalized expression level
B
IL-13-PE
IFN-γ-APC
GM-CSF-PE
TNF-APC
C
Normalized to mode
Day 0
Day 3
Day 6
CellTrace Violet
D
Day 3
Divided cells (%)
Proliferation index
E
Day 6 - 0.3 ng/mL IL-12
Isotype
T-bet-PE
F
T-bet-PE (GeoMean)
IL-12 (ng/mL)
G
Day 6 - 0.3 ng/mL IL-12
H
IFN-γ+ (%)
IFN-γ (GeoMean, x10³)
I
GM-CSF+ (%)
GM-CSF (GeoMean)
J
0.1 ng/mL IL-12
Cytokine (ng/mL)
IL-13
IFN-γ
IL-17a

Figure 2

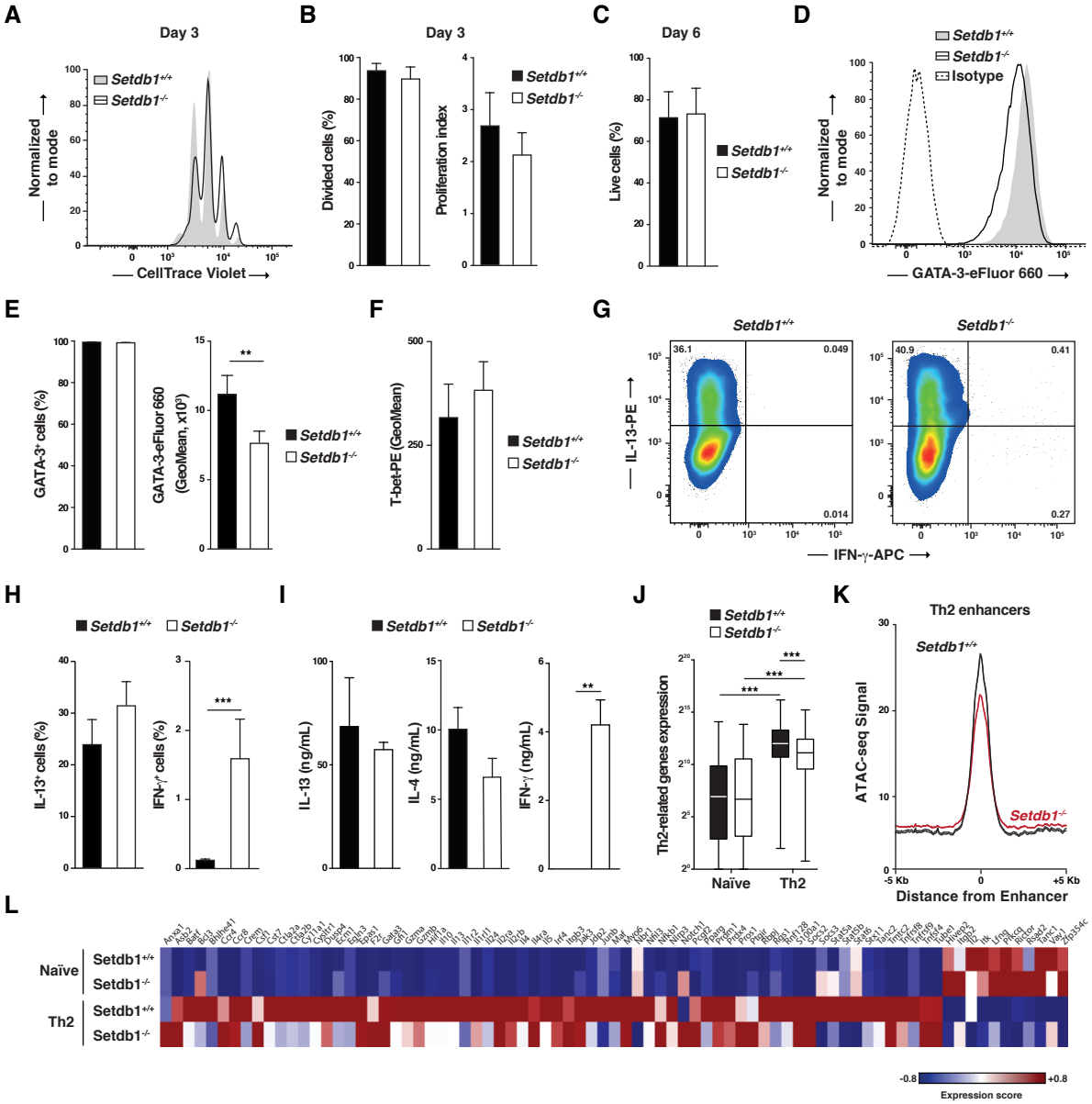

A
Day 3
Setdb1+/+
Setdb1-/-
Normalized to mode
CellTrace Violet
B
Day 3
Divided cells (%)
Proliferation index
Setdb1+/+
Setdb1-/-
C
Day 6
Live cells (%)
Setdb1+/+
Setdb1-/-
D
Setdb1+/+
Setdb1-/-
Isotype
Normalized to mode
GATA-3-eFluor 660
E
GATA-3+ cells (%)
GATA-3-eFluor 660 (GeoMean, x10³)
**
Setdb1+/+
Setdb1-/-
F
T-bet-PE (GeoMean)
Setdb1+/+
Setdb1-/-
G
Setdb1+/+
Setdb1-/-
36.1
0.049
0.014
40.9
0.41
0.27
IL-13-PE
IFN-γ-APC
H
Setdb1+/+
Setdb1-/-
IL-13+ cells (%)
IFN-γ+ cells (%)
***
I
Setdb1+/+
Setdb1-/-
IL-13 (ng/mL)
IL-4 (ng/mL)
IFN-γ (ng/mL)
**
J
Setdb1+/+
Setdb1-/-
Th2-related genes expression
***
***
***
Naïve
Th2
K
Th2 enhancers
Setdb1+/+
Setdb1-/-
ATAC-seq Signal
-5 Kb
0
+5 Kb
Distance from Enhancer
L
Naïve
Th2
Setdb1+/+
Setdb1-/-
-0.8
+0.8
Expression score

Figure 3

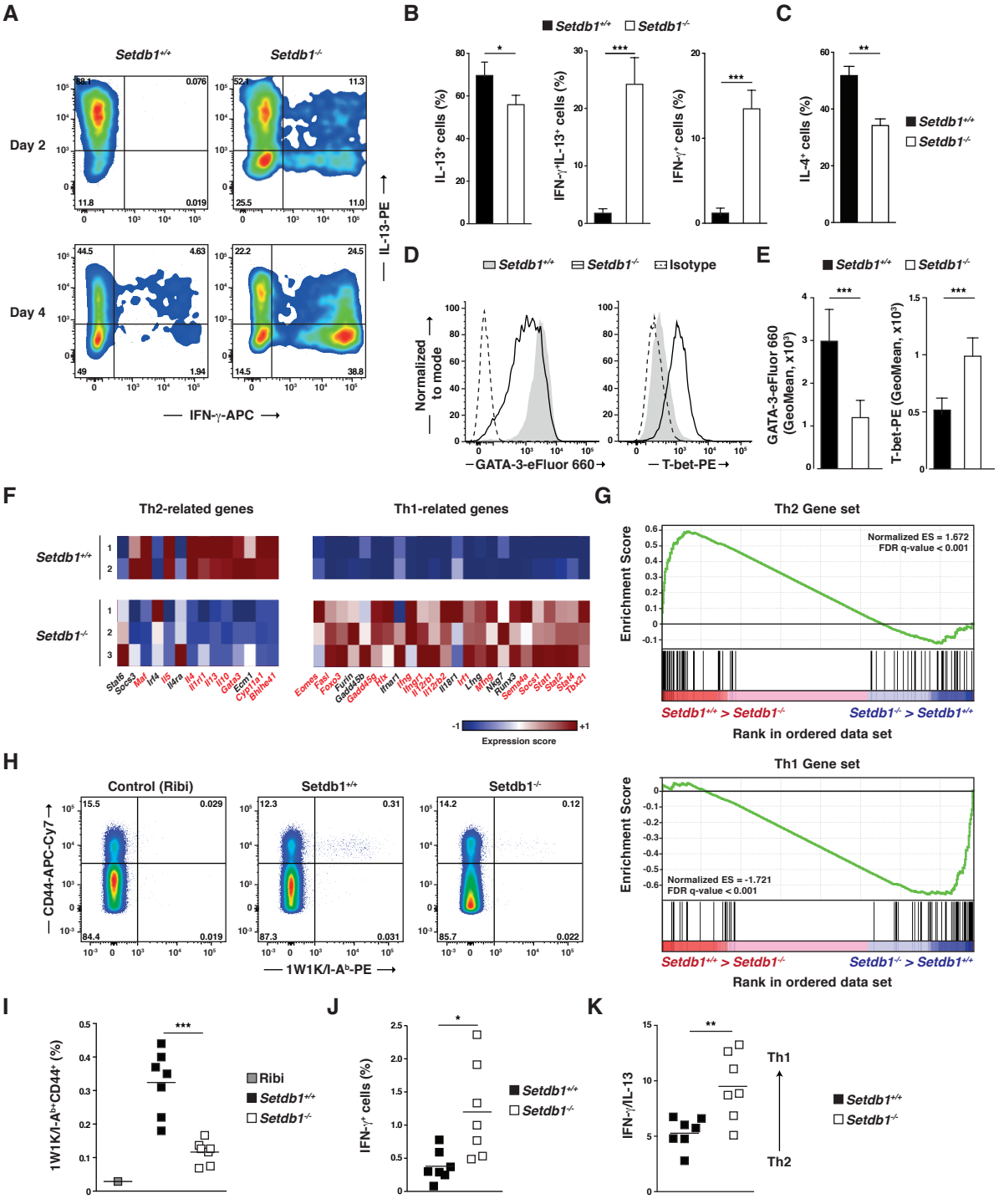

A
Setdb1+/+
Setdb1-/-
Day 2
Day 4
IL-13-PE
IFN-γ-APC
B
Setdb1+/+ Setdb1-/-
IL-13+ cells (%)
IFN-γ+IL-13+ cells (%)
IFN-γ+ cells (%)
C
IL-4+ cells (%)
D
Setdb1+/+ Setdb1-/- Isotype
Normalized to mode
GATA-3-eFluor 660
T-bet-PE
E
GATA-3-eFluor 660 (GeoMean, x10³)
T-bet-PE (GeoMean, x10³)
F
Th2-related genes
Th1-related genes
Stat6 Socs3 Maf Irf4 Il5 Il4ra Il4 Il1rl1 Il13 Il10 Gata3 Ecm1 Cyp11a1 Bhlhe41
Eomes Fasl Foxo3 Furin Gadd45g Gadd45b Tbx Ifnar1 Ifng Ifngr1 Il12rb1 Il12rb2 Il18r1 Irf1 Lfng Mthfd Nkg7 Runx3 Sema4a Socs1 Stat1 Stat2 Stat4 Tbx21
Expression score
G
Th2 Gene set
Enrichment Score
Normalized ES = 1.672
FDR q-value < 0.001
Setdb1+/+ > Setdb1-/-
Setdb1-/- > Setdb1+/+
Rank in ordered data set
H
Control (Ribi)
Setdb1+/+
Setdb1-/-
CD44-APC-Cy7
1W1K/I-Aᵇ-PE
Th1 Gene set
Normalized ES = -1.721
FDR q-value < 0.001
I
1W1K/I-Aᵇ+CD44+ (%)
Ribi
Setdb1+/+
Setdb1-/-
J
IFN-γ+ cells (%)
K
IFN-γ/IL-13
Th1
Th2

Figure 4

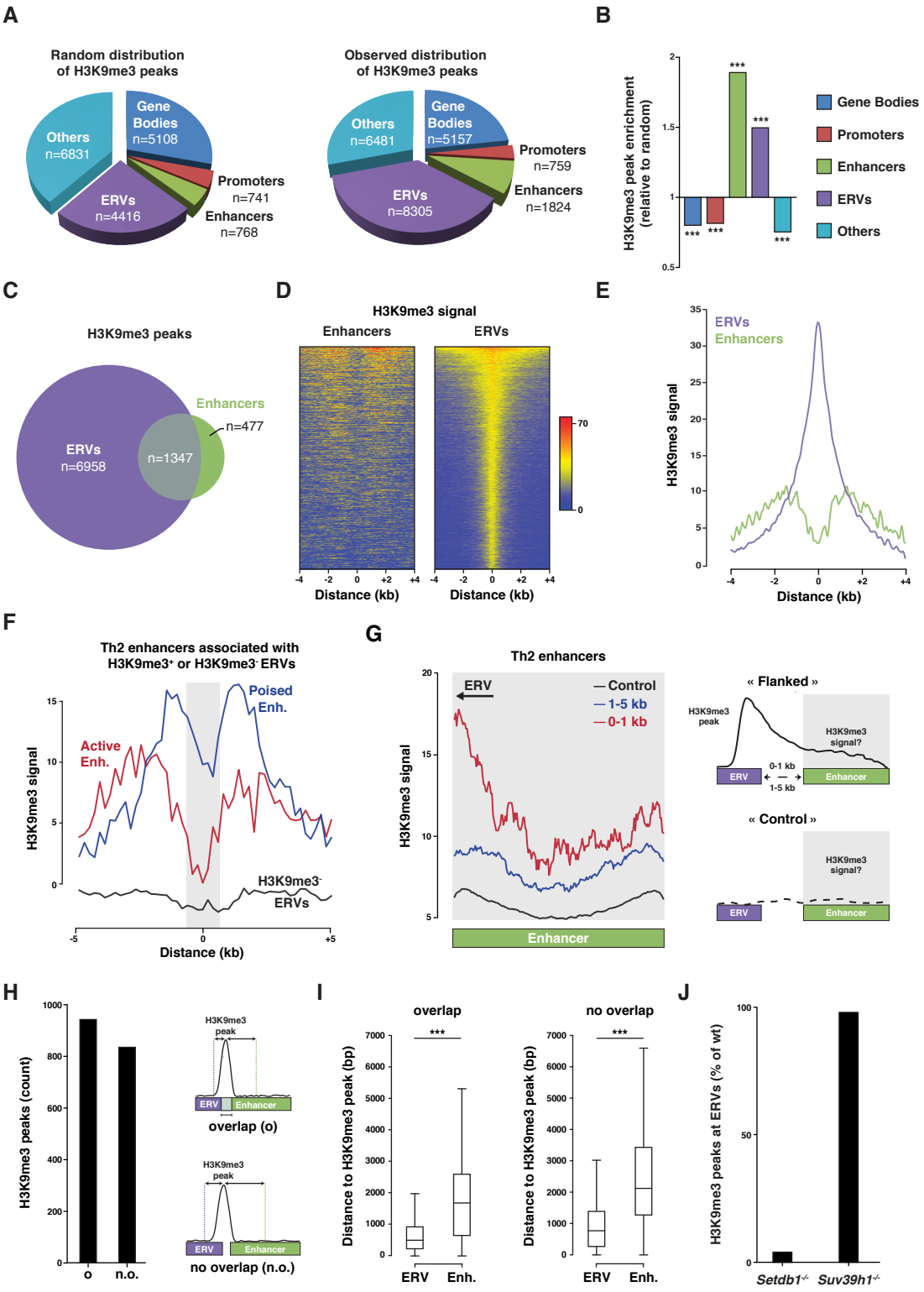

A
Random distribution of H3K9me3 peaks
Gene Bodies n=5108
Promoters n=741
Enhancers n=768
ERVs n=4416
Others n=6831
Observed distribution of H3K9me3 peaks
Gene Bodies n=5157
Promoters n=759
Enhancers n=1824
ERVs n=8305
Others n=6481
B
H3K9me3 peak enrichment (relative to random)
Gene Bodies
Promoters
Enhancers
ERVs
Others
C
H3K9me3 peaks
ERVs n=6958
n=1347
Enhancers n=477
D
H3K9me3 signal
Enhancers
ERVs
Distance (kb)
E
ERVs
Enhancers
H3K9me3 signal
Distance (kb)
F
Th2 enhancers associated with H3K9me3+ or H3K9me3- ERVs
Poised Enh.
Active Enh.
H3K9me3- ERVs
H3K9me3 signal
Distance (kb)
G
Th2 enhancers
ERV
Control
1-5 kb
0-1 kb
H3K9me3 signal
Enhancer
« Flanked »
H3K9me3 peak
H3K9me3 signal?
0-1 kb
1-5 kb
« Control »
H
H3K9me3 peaks (count)
o
n.o.
H3K9me3 peak
ERV
Enhancer
overlap (o)
no overlap (n.o.)
I
overlap
no overlap
Distance to H3K9me3 peak (bp)
ERV
Enh.
J
H3K9me3 peaks at ERVs (% of wt)
Setdb1-/-
Suv39h1-/-

Figure 5

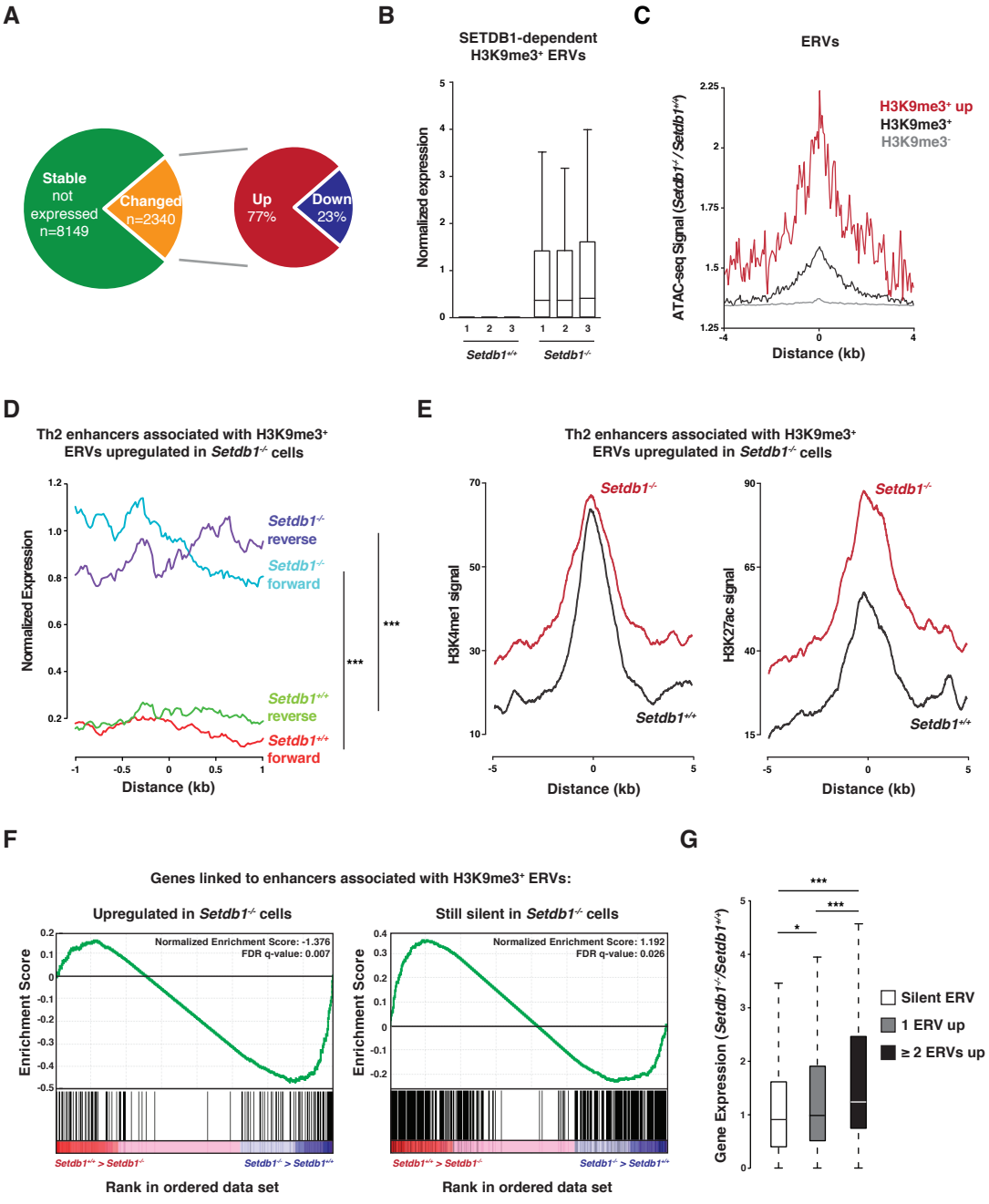

A
Stable
not
expressed
n=8149
Changed
n=2340
Up
77%
Down
23%
B
SETDB1-dependent
H3K9me3+ ERVs
Normalized expression
Setdb1+/+
Setdb1-/-
C
ERVs
H3K9me3+ up
H3K9me3+
H3K9me3-
ATAC-seq Signal (Setdb1-/-/Setdb1+/+)
Distance (kb)
D
Th2 enhancers associated with H3K9me3+
ERVs upregulated in Setdb1-/- cells
Normalized Expression
Setdb1-/-
reverse
Setdb1-/-
forward
Setdb1+/+
reverse
Setdb1+/+
forward
***
***
Distance (kb)
E
Th2 enhancers associated with H3K9me3+
ERVs upregulated in Setdb1-/- cells
H3K4me1 signal
H3K27ac signal
Setdb1-/-
Setdb1+/+
Distance (kb)
F
Genes linked to enhancers associated with H3K9me3+ ERVs:
Upregulated in Setdb1-/- cells
Still silent in Setdb1-/- cells
Normalized Enrichment Score: -1.376
FDR q-value: 0.007
Normalized Enrichment Score: 1.192
FDR q-value: 0.026
Enrichment Score
Setdb1+/+ > Setdb1-/-
Setdb1-/- > Setdb1+/+
Rank in ordered data set
G
Gene Expression (Setdb1-/-/Setdb1+/+)
***
***
*
Silent ERV
1 ERV up
≥ 2 ERVs up

Figure 6

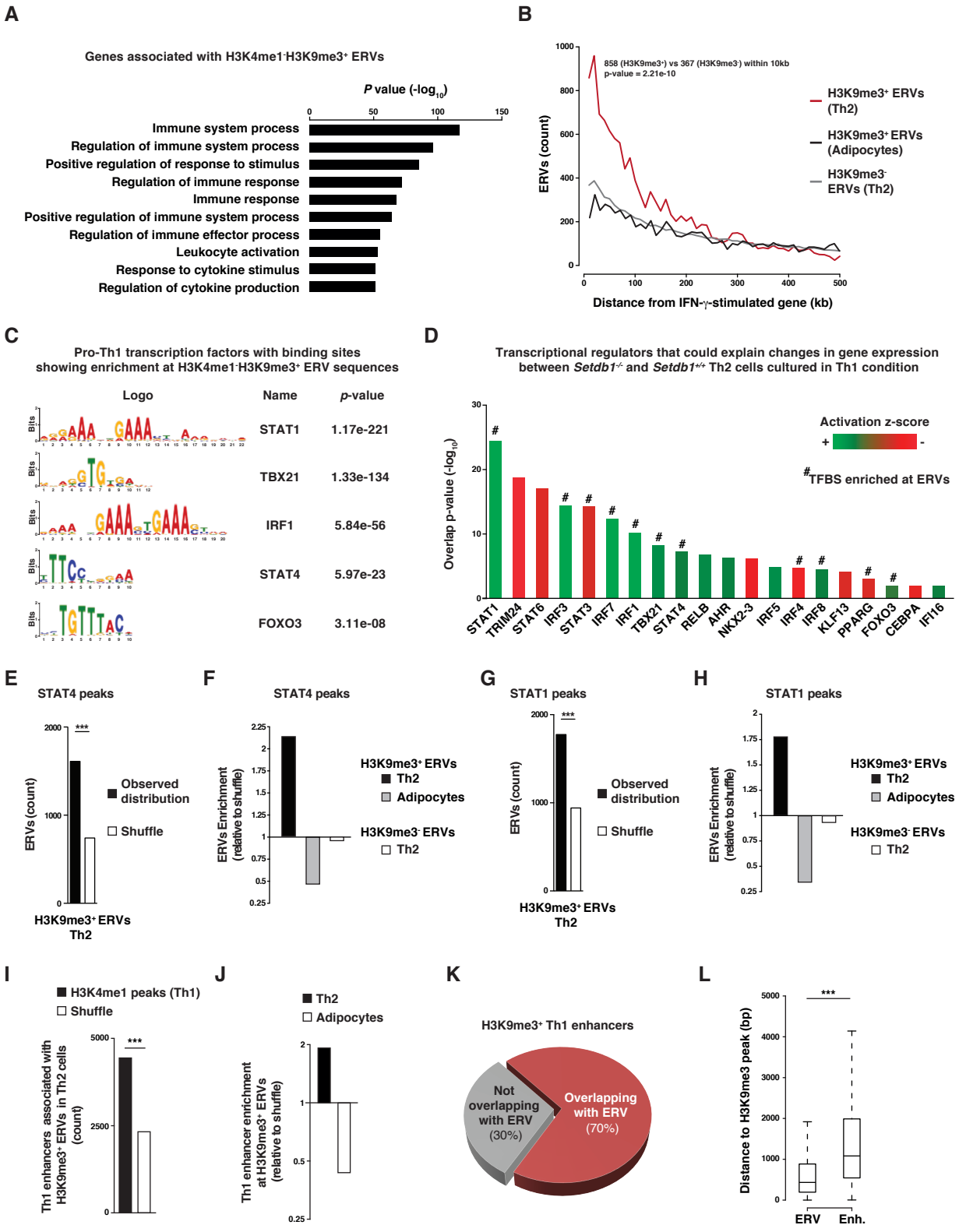
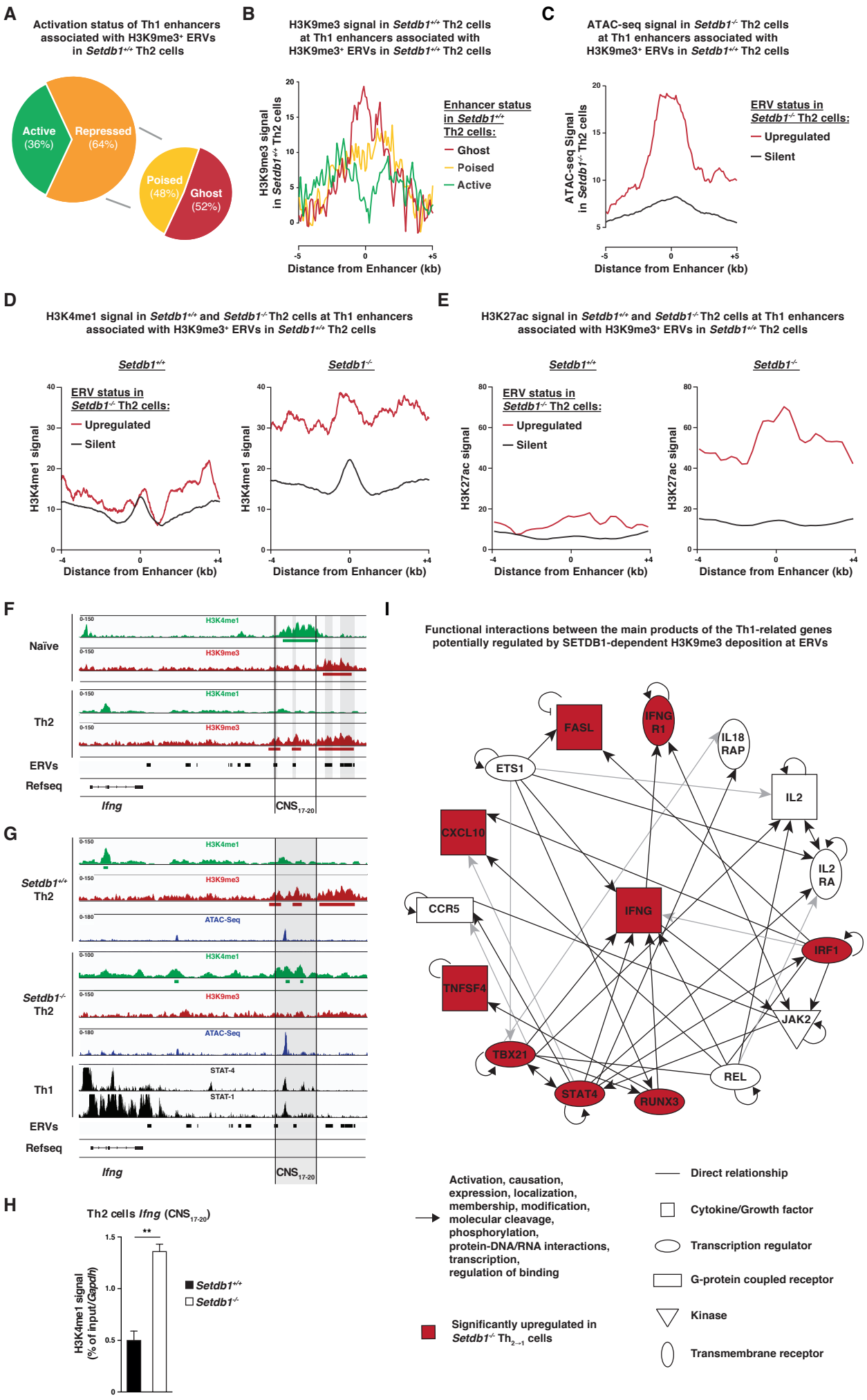
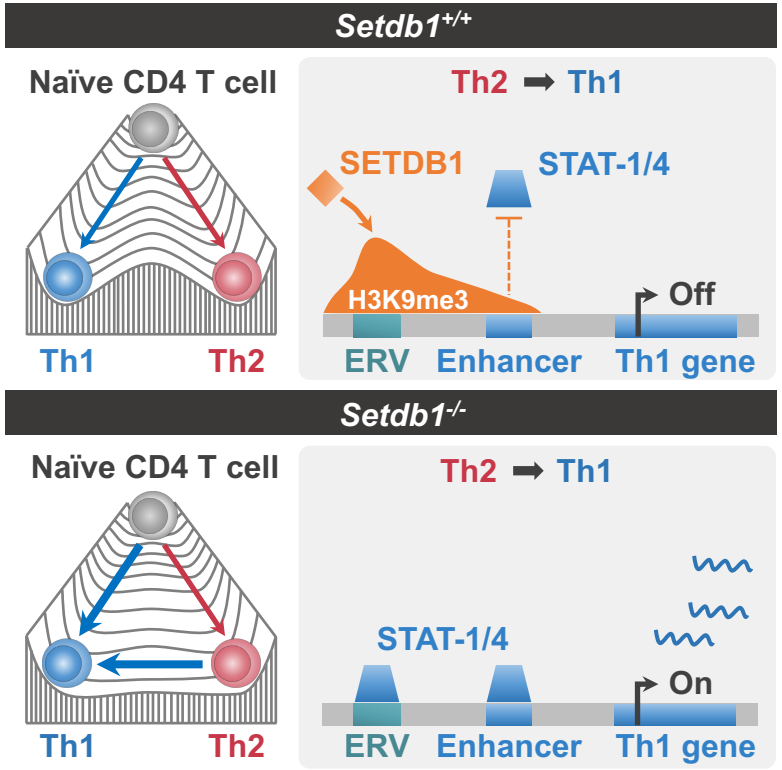

A
Genes associated with H3K4me1-H3K9me3+ ERVs
P value (-log10)
0
50
100
150
Immune system process
Regulation of immune system process
Positive regulation of response to stimulus
Regulation of immune response
Immune response
Positive regulation of immune system process
Regulation of immune effector process
Leukocyte activation
Response to cytokine stimulus
Regulation of cytokine production
B
858 (H3K9me3+) vs 367 (H3K9me3-) within 10kb
p-value = 2.21e-10
H3K9me3+ ERVs (Th2)
H3K9me3+ ERVs (Adipocytes)
H3K9me3- ERVs (Th2)
ERVs (count)
Distance from IFN-γ-stimulated gene (kb)
C
Pro-Th1 transcription factors with binding sites showing enrichment at H3K4me1-H3K9me3+ ERV sequences
Logo
Name
p-value
STAT1
1.17e-221
TBX21
1.33e-134
IRF1
5.84e-56
STAT4
5.97e-23
FOXO3
3.11e-08
D
Transcriptional regulators that could explain changes in gene expression between Setdb1-/- and Setdb1+/+ Th2 cells cultured in Th1 condition
Overlap p-value (-log10)
Activation z-score
+
-
#TFBS enriched at ERVs
STAT1
TRIM24
STAT6
IRF3
STAT3
IRF7
IRF1
TBX21
STAT4
RELB
AHR
NKX2-3
IRF5
IRF4
IRF8
KLF13
PPARG
FOXO3
CEBPA
IFI16
E
STAT4 peaks
ERVs (count)
Observed distribution
Shuffle
H3K9me3+ ERVs Th2
F
STAT4 peaks
ERVs Enrichment (relative to shuffle)
H3K9me3+ ERVs
Th2
Adipocytes
H3K9me3- ERVs
Th2
G
STAT1 peaks
ERVs (count)
Observed distribution
Shuffle
H3K9me3+ ERVs Th2
H
STAT1 peaks
ERVs Enrichment (relative to shuffle)
H3K9me3+ ERVs
Th2
Adipocytes
H3K9me3- ERVs
Th2
I
H3K4me1 peaks (Th1)
Shuffle
Th1 enhancers associated with H3K9me3+ ERVs in Th2 cells (count)
J
Th2
Adipocytes
Th1 enhancer enrichment at H3K9me3+ ERVs (relative to shuffle)
K
H3K9me3+ Th1 enhancers
Not overlapping with ERV (30%)
Overlapping with ERV (70%)
L
Distance to H3K9me3 peak (bp)
ERV
Enh.

Figure 7


A
Activation status of Th1 enhancers associated with H3K9me3+ ERVs in Setdb1+/+ Th2 cells
Active (36%)
Repressed (64%)
Poised (48%)
Ghost (52%)
B
H3K9me3 signal in Setdb1+/+ Th2 cells at Th1 enhancers associated with H3K9me3+ ERVs in Setdb1+/+ Th2 cells
H3K9me3 signal in Setdb1+/+ Th2 cells
Distance from Enhancer (kb)
Enhancer status in Setdb1+/+ Th2 cells:
Ghost
Poised
Active
C
ATAC-seq signal in Setdb1-/- Th2 cells at Th1 enhancers associated with H3K9me3+ ERVs in Setdb1+/+ Th2 cells
ATAC-seq Signal in Setdb1-/- Th2 cells
Distance from Enhancer (kb)
ERV status in Setdb1-/- Th2 cells:
Upregulated
Silent
D
H3K4me1 signal in Setdb1+/+ and Setdb1-/- Th2 cells at Th1 enhancers associated with H3K9me3+ ERVs in Setdb1+/+ Th2 cells
Setdb1+/+
Setdb1-/-
H3K4me1 signal
Distance from Enhancer (kb)
ERV status in Setdb1-/- Th2 cells:
Upregulated
Silent
E
H3K27ac signal in Setdb1+/+ and Setdb1-/- Th2 cells at Th1 enhancers associated with H3K9me3+ ERVs in Setdb1+/+ Th2 cells
Setdb1+/+
Setdb1-/-
H3K27ac signal
Distance from Enhancer (kb)
ERV status in Setdb1-/- Th2 cells:
Upregulated
Silent
F
Naïve
Th2
ERVs
Refseq
Ifng
CNS17-20
H3K4me1
H3K9me3
G
Setdb1+/+ Th2
Setdb1-/- Th2
Th1
ERVs
Refseq
Ifng
CNS17-20
H3K4me1
H3K9me3
ATAC-Seq
STAT-4
STAT-1
H
Th2 cells Ifng (CNS17-20)
H3K4me1 signal (% of input/Gapdh)
**
Setdb1+/+
Setdb1-/-
I
Functional interactions between the main products of the Th1-related genes potentially regulated by SETDB1-dependent H3K9me3 deposition at ERVs
FASL
IFNG R1
IL18 RAP
ETS1
IL2
CXCL10
IL2 RA
CCR5
IFNG
IRF1
TNFSF4
JAK2
TBX21
REL
STAT4
RUNX3
Activation, causation, expression, localization, membership, modification, molecular cleavage, phosphorylation, protein-DNA/RNA interactions, transcription, regulation of binding
Significantly upregulated in Setdb1-/- Th2→1 cells
Direct relationship
Cytokine/Growth factor
Transcription regulator
G-protein coupled receptor
Kinase
Transmembrane receptor


Setdb1+/+
Naïve CD4 T cell
Th1
Th2
Th2 ➡ Th1
SETDB1
STAT-1/4
H3K9me3
Off
ERV
Enhancer
Th1 gene
Setdb1-/-
Naïve CD4 T cell
Th1
Th2
Th2 ➡ Th1
STAT-1/4
On
ERV
Enhancer
Th1 gene