nature cell biology

RESOURCE

https://doi.org/10.1038/s41556-019-0403-5

Inducible histone K-to-M mutations are dynamic tools to probe the physiological role of site-specific histone methylation in vitro and in vivo

Justin Brumbaugh[1,2,3,4,5,6,11,12], Ik Soo Kim[3,7,10,12], Fei Ji[1,6,8], Aaron J. Huebner[1,2,3,4,5,6], Bruno Di Stefano[1,2,3,4,5,6], Benjamin A. Schwarz[1,2,3,4,5,6], Jocelyn Charlton[4,9], Amy Coffey[1,2,3,4,5,6], Jiho Choi[1,2,3,4,5,6], Ryan M. Walsh[1,2,3,4,5,6], Jeffrey W. Schindler[2,3,5,6], Anthony Anselmo[1,6,8], Alexander Meissner[4,9,10], Ruslan I. Sadreyev[1,6,7], Bradley E. Bernstein[3,7,10], Hanno Hock[2,3,5,6]* and Konrad Hochedlinger[1,2,3,4,5,6,8,10]*

Development and differentiation are associated with profound changes to histone modifications, yet their in vivo function remains incompletely understood. Here, we generated mouse models expressing inducible histone H3 lysine-to-methionine (K-to-M) mutants, which globally inhibit methylation at specific sites. Mice expressing H3K36M developed severe anaemia with arrested erythropoiesis, a marked haematopoietic stem cell defect, and rapid lethality. By contrast, mice expressing H3K9M survived up to a year and showed expansion of multipotent progenitors, aberrant lymphopoiesis and thrombocytosis. Additionally, some H3K9M mice succumbed to aggressive T cell leukaemia/lymphoma, while H3K36M mice exhibited differentiation defects in testis and intestine. Mechanistically, induction of either mutant reduced corresponding histone trimethylation patterns genome-wide and altered chromatin accessibility as well as gene expression landscapes. Strikingly, discontinuation of transgene expression largely restored differentiation programmes. Our work shows that individual chromatin modifications are required at several specific stages of differentiation and introduces powerful tools to interrogate their roles in vivo.

During development and tissue homeostasis, immature progenitor cells give rise to differentiated cells without altering genomic content. Post-translational histone modifications are thought to guide these processes by altering gene expression[1,2]. For example, trimethylation (me3) of H3K36 (H3K36me3) and H3K4 (H3K4me3) are associated with actively transcribed regions, while H3K9me3 and H3K27me3 correlate with repressed regions. However, the biological function of specific modifications is difficult to determine through mutational analyses[3] as canonical histones are encoded by many genes in the mammalian genome[4]. As a result, the function of individual histone marks has been inferred from the mutational analysis of histone-modifying enzymes. Lysine methylation of histone H3 is regulated by multiple, redundant methyltransferases (KMTs) and demethylases (KDMs). Since these enzymes also have non-histone substrates and act as scaffolds to recruit other factors, some KMT and KDM knockout phenotypes may be unrelated to the histone marks[5–8]. Hence, the direct biological impact of histone tail modifications remains incompletely understood.

Lysine to methionine (K-to-M) or lysine-to-isoleucine (K-to-I) point mutations on the variant histone H3.3 suppress methylation levels at their respective sites[9–14]. Such mutations reportedly interfere with the active site of their cognate SET-domain-containing KMTs[9,10,12,15,16], resulting in a global decrease of individual histone marks. K-to-M mutations effectively function as hypomorphs, allowing the study of histone marks in cases where genetic disruption of the corresponding KMTs is lethal[17]. Of note, H3.3K36M mutations have been identified in bone and soft-tissue sarcomas[10–12], and in head and neck cancer[14,18]. Moreover, H3.3K9M mutations in *Drosophila* lead to defects in position effect variegation[15]. However, the impact of site-specific H3 methylation in untransformed primary cells or in mice remains incompletely explored. Here, we utilize K-to-M histone mutations in vivo and in vitro to modulate levels of histone methylation during complex biological processes, including tissue homeostasis, cellular differentiation and cancer.

Results

Transgenic system to suppress H3K9 and H3K36 methylation. We targeted complementary DNAs encoding histone H3.3K9M, H3.3K36M and wild-type H3.3 (hereafter referred to as H3K9M, H3K36M and H3, respectively) into mouse embryonic stem (ES) cells using a site-specific, single-copy integration system[19,20] (Fig. 1a). Our inducible system facilitated rapid and specific expression of the histone constructs following doxycycline administration (Fig. 1b; Extended Data Fig. 1a). Mutant histones partitioned with the nuclear fraction, which suggested that they were properly incorporated into chromatin (Fig. 1b). Consistent with previous reports, expression

[1]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA. [2]Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA, USA. [3]Center for Cancer Research, Massachusetts General Hospital, Boston, MA, USA. [4]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. [5]Harvard Stem Cell Institute, Harvard University, Cambridge, MA, USA. [6]Harvard Medical School, Boston, MA, USA. [7]Department of Pathology, Harvard Medical School, Boston, MA, USA. [8]Department of Genetics, Harvard Medical School, Boston, MA, USA. [9]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany. [10]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [11]Present address: Department of Molecular, Cellular, and Developmental Biology, University of Colorado-Boulder, Boulder, CO, USA. [12]These authors contributed equally: Justin Brumbaugh, Ik Soo Kim. *e-mail: Hock.Hanno@mgh.harvard.edu; khochedlinger@mgh.harvard.edu

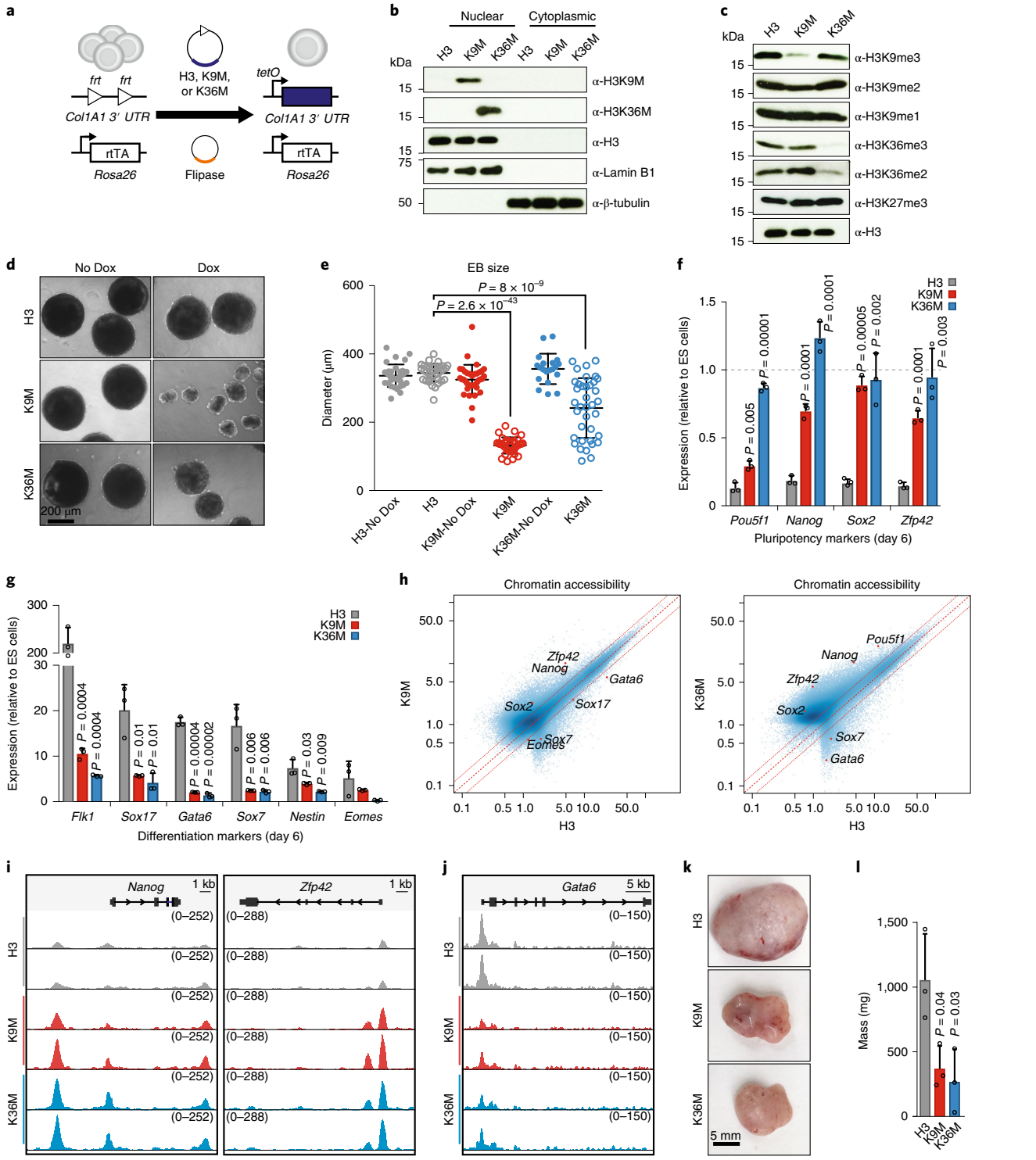


Fig. 1 | Doxycycline-inducible K-to-M histone mutants globally suppress site-specific histone methylation and impair differentiation of ES cells. **a**, Schematic of the strategy used to generate cells harbouring inducible histone constructs. *Col1A1* (3′ UTR) is the locus targeted for mutant histone insertion; *Rosa26* is the locus targeted for rtTA insertion; *tetO* is a doxycycline-inducible promoter. **b**, Western blot analysis of nuclear and cytoplasmic fractions from ES cells expressing mutant histones. The H3 loading control is the same image as in **c**. **c**, Western blot analysis for the indicated histone modifications in ES cells expressing mutant histones. The H3 loading control is the same image as in **b**. **d**, Images of EBs at day 9 of induction with and without expression of H3K9M and H3K36M. **e**, Quantification of EB diameters for each condition in technical replicate (H3, $n=32$; H3 + doxycycline (Dox), $n=34$; H3K9M, $n=37$; H3K9M + Dox, $n=33$; H3K36M, $n=20$; H3K36M + Dox, $n=37$). The centre bar represents the mean and the whiskers represent the standard deviation of the mean. **f**, qRT-PCR results for pluripotency markers at day 6 of induction. **g**, qRT-PCR results for differentiation markers at day 6 of induction. **h**, Scatter plots comparing ATAC-seq peak coverage for each mutant histone sample compared with H3 control. **i**, Gene tracks showing ATAC-seq data for pluripotency genes. **j**, Gene tracks showing ATAC-seq data for differentiation-associated genes. **k**, Images of teratomas expressing H3K9M or H3K36M. **l**, Quantification of teratoma mass for each condition in biological triplicate ($n=3$ teratomas). For **f**, **g** and **l**, columns represent the mean and error bars represent the standard deviation of the mean for $n=3$ independent experiments. For **e**–**g** and **l**, statistical significance was determined using two-tailed unpaired Student's *t*-test. See Source data for full membrane western blot images. Data in **b**, **c**, **d** and **k** are representative of three independent experiments.

of H3K9M and H3K36M dramatically reduced the global levels of H3K9me3 and H3K36me3, respectively[9,10,14–16] (Fig. 1c). Dimethyl marks at both sites were also suppressed, albeit less appreciably, and H3K27me3 levels were slightly elevated with H3K36M expression. Importantly, we did not observe crosstalk between mutant histones (that is, expression of H3K9M did not alter H3K36 methylation and vice versa) and expression of H3 had no effect on the methylation status of either residue (Fig. 1c; Extended Data Fig. 1a).

H3K9M and H3K36M expression impairs ES cell differentiation. To study the impact of our mutants on pluripotent stem cell differentiation, we generated embryoid bodies (EBs). Expression of both H3K9M and H3K36M yielded significantly smaller EBs compared with the control (Fig. 1d,e), which suggested that there was a defect in differentiation. Consistent with this observation, EBs expressing H3K9M and H3K36M retained expression of the pluripotency genes *Pou5f1* (also known as *Oct4*), *Nanog*, *Sox2* and *Zfp42* (also known as *Rex1*) compared with control EBs (Fig. 1f). Moreover, both types of mutant EBs expressed markedly lower levels of the differentiation markers *Flk1*, *Eomes*, *Gata6*, *Sox7*, *Sox17* and *Nestin* relative to control EBs (Fig. 1g). Assay for transposase-accessible chromatin using sequencing (ATAC-seq) analysis of these samples suggested that chromatin was more accessible overall in the mutant EBs, which was consistent with a less differentiated state (36,341 accessible regions in control, 44,540 in H3K9M and 61,500 in H3K36M; Extended Data Fig. 1b). For example, chromatin was more open in regions associated with crucial pluripotency genes (for example, *Pou5f1, Nanog* and *Zfp42*) that remained active in EBs expressing H3K9M or H3K36M (Fig. 1h,i). Conversely, chromatin associated with differentiation markers (for example, *Gata6* and *Sox7*) was closed in mutant EBs compared with controls (Fig. 1h,j). In line with the observed defect in EB differentiation, cells expressing either H3K9M or H3K36M generated significantly smaller teratomas following injection into the flanks of immunocompromised mice (Fig. 1k,l). Histological analyses revealed that control cells formed well-differentiated teratomas containing structures characteristic of all three germ layers (Extended Data Fig. 1c). By contrast, H3K36M, and to a lesser extent H3K9M, produced poorly differentiated teratomas, with large, undifferentiated areas and fewer differentiated structures (Extended Data Fig. 1c). Together, these data suggest that H3K9 and H3K36 methylation are crucial for both exiting pluripotency and proper differentiation.

To test whether individual suppression of H3K9 and H3K36 KMTs phenocopy K-to-M mutations, we transduced Rex1-dGFP ES cells[21] with short hairpin RNAs (shRNAs) targeting H3K9-specific KMTs (*G9a* or *Setdb1*) or H3K36-specific KMTs (*Nsd1* or *Setd2*) (Extended Data Fig. 1d), allowing for a sensitive, flow-cytometry-based readout of pluripotency exit. Indeed, suppression of the H3K36 KMTs *Nsd1* or *Setd2* partially recapitulated the differentiation block observed with H3K36M expression at day 2 of EB formation, which coincided with a modest reduction in H3K36me3 levels (Extended Data Fig. 1e,f). These results are consistent with previous findings in *Setd2* knockout ES cell lines[22]. By contrast, neither *G9a* nor *Setdb1* suppression impaired differentiation in our exit system. However, western blot analyses revealed that knockdown of these enzymes was insufficient to reduce H3K9me3 levels (Extended Data Fig. 1g). Previous work has shown that disruption of H3K9 KMTs is compensated by alternative KMTs[23], which probably explains the lack of an observable reduction in H3K9me3 following H3K9 KMT knockdown. Of note, shRNAs that more efficiently suppressed *G9a* and *Setdb1* led to widespread differentiation in ES cells under self-renewal conditions, thus precluding the use of more potent hairpins, which is in line with previous results[24]. These difficulties in interpreting KMT shRNA knockdown experiments therefore underscore the power of our K-to-M mutations to study histone methylation.

Expression of H3K9M and H3K36M perturbs tissue homeostasis. We next produced transgenic mice from our targeted ES cells and used doxycycline to induce expression in experimental and control animals at 6–8 weeks of age (Fig. 2a). Strikingly, H3K36M mice developed poor body condition and became moribund within 4–7 weeks of induction (Fig. 2b). A postmortem examination showed that male H3K36M mice exhibited testicular atrophy, with deterioration of the seminiferous tubule and concomitant loss of mature testis cells (Extended Data Fig. 2a). RNA sequencing (RNA-seq) analysis of testes after 4 weeks of doxycycline treatment suggested that the transcriptional programme associated with undifferentiated spermatogonia was unaffected in H3K36M mice based on gene set enrichment analysis (GSEA) and spermatogonia marker genes (for example, *Stra8*, *Kit* and *Prdm9*) (Extended Data Fig. 2b,c, left panels). However, we found a marked depletion of the more mature spermatocyte-specific programme (for example, *Piwil2, Meioc* and *Hormad1*) (Extended Data Fig. 2b,c, right panels). These results point to a specific role for H3K36 methylation between the spermatogonia and spermatocyte stages of differentiation, and resemble the phenotype of *Setd2* knockout mice[25].

In the intestine, PAS$^+$ goblet and lysozyme$^+$ Paneth cells were absent in H3K36M mice, while the typical crypt-villus structure of the epithelium was preserved, thus suggesting that there were defects in secretory lineage commitment (Extended Data Fig. 2d,e). Indeed, RNA-seq analysis demonstrated that H3K36M intestinal cells were depleted for secretory-lineage-associated transcriptional programmes (Extended Data Fig. 2f, left panel). Key regulators of secretory differentiation such as *Atoh1* and *Gfi1* were strongly reduced (Extended Data Fig. 2g, left panels). Likewise, the mRNA expression signature characteristic of goblet cells was disrupted in the intestine of mice expressing H3K36M, and transcript levels for the goblet cell markers *Spdef* and *Muc2* were severely reduced (Extended Data Fig. 2f,g, right panels). Thus, H3K36M disrupts distinct differentiation programmes in the intestine.

H3K36M mice also exhibited splenomegaly, thymic atrophy and lack of red marrow in leg bones (Fig. 2c,d; Extended Data Fig. 2h). Histological sections and cell counts revealed markedly decreased bone marrow cellularity in H3K36M mice, with a complete absence of megakaryocytes and few terminally differentiated erythroid cells compared with both H3K9M mice and rtTA mice, which served as littermate controls (Fig. 2e,f). Complete blood counts (CBCs) pointed to a significant increase in white blood cell numbers (predominantly from circulating early erythroid cells), reduced platelet numbers and severely reduced haemoglobin levels in H3K36M-expressing mice (Fig. 2g). By contrast, H3K9M mice displayed partially reciprocal abnormalities, including a fourfold decrease in white blood cells and a 1.5-fold increase in platelets (Fig. 2g). We confirmed strong reduction of H3K9me3 and H3K36me3 levels in H3K9M and H3K36M progenitor cells (lineage$^-$, c-Kit$^+$ and Sca1$^-$ (LKS)), respectively, using western blot analysis (Fig. 2h). However, levels of total H3 were not appreciably elevated with doxycycline administration, which suggested that the mutant transgenes are expressed at physiological levels. Together, these results point to widespread and distinct functions of H3K9 and H3K36 methylation in tissue maintenance and blood cell homeostasis.

K-to-M mutations influence cell fate at discrete points during haematopoiesis. To refine our haematopoietic analysis, we first examined the myeloid lineages after 4 weeks of mutant histone induction, before mice were overtly sick. Consistent with the severe anaemia observed in H3K36M mice, we observed a dramatic increase in early erythroid cells, such as colony forming unit-erythroid progenitors (CFUEs) and pro-erythroblasts, with a concomitant decrease in more mature stages, specifically Ter119$^+$ erythroblasts and enucleated erythrocytes (Fig. 3a,b, see gate I). Notably, the immature erythroid cells expressed abnormally low

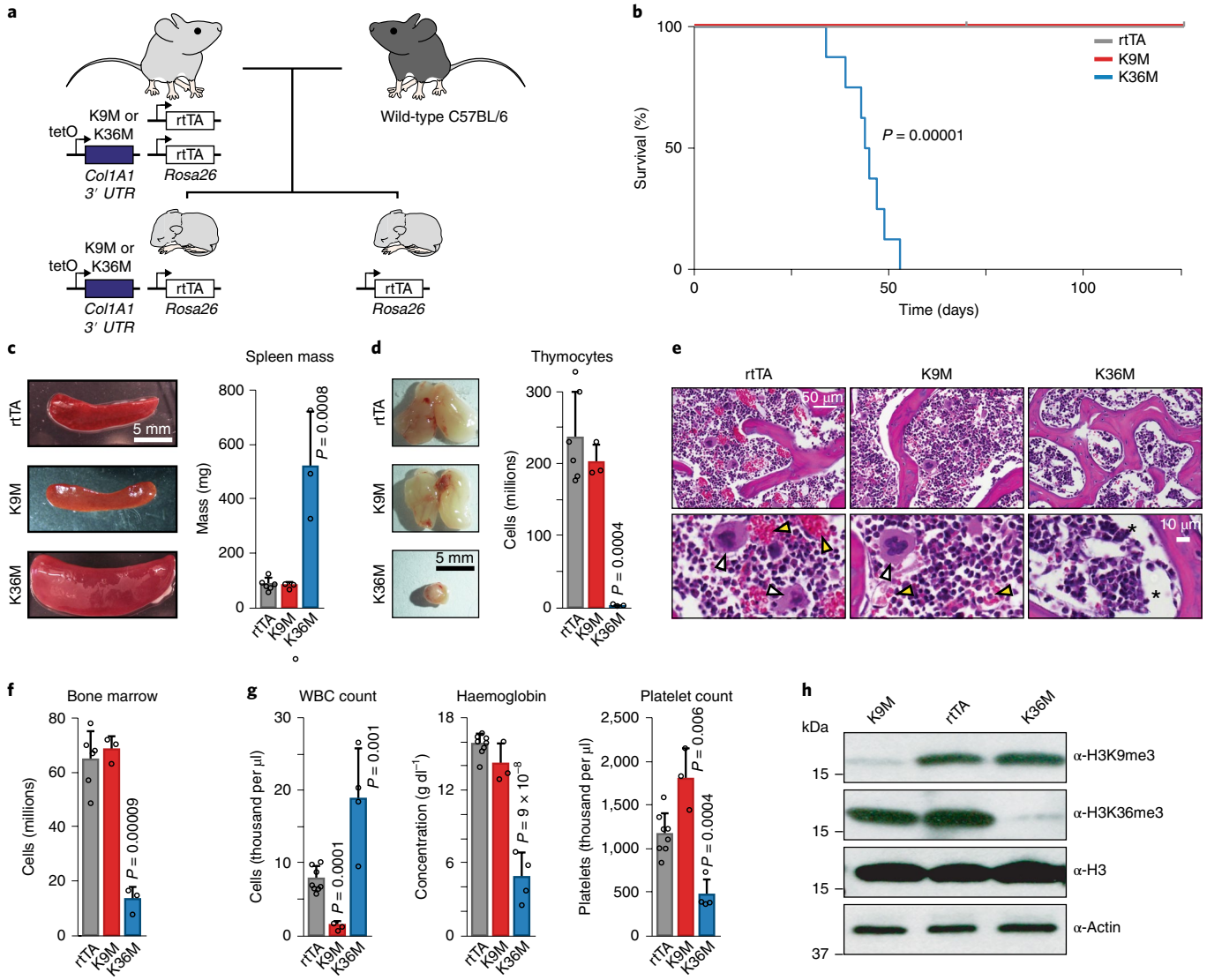


Fig. 2 | In vivo expression of H3K36M leads to rapid death with severe anaemia and organ abnormalities. a, Schematic of the strategy used to generate transgenic mice with littermate controls. **b**, Survival curve for transgenic mice following doxycycline administration (rtTA, $n=10$ mice; H3K9M, $n=3$; H3K36M, $n=8$). Statistical significance was determined using two-sided log-rank (Mantel–Cox) test. **c**, Representative images (left) of spleens taken from mice induced for 4 weeks and quantified by mass (right; rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). **d**, Representative thymus images taken from mice induced for 4 weeks with total thymocytes quantified by cell count (rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). **e**, Histological analysis of bone from mice induced for 4 weeks. Scale bar, 50 μm (upper panels) and 10 μm (lower panels). White arrowheads indicate megakaryocytes, yellow arrowheads indicate mature erythroid cells; asterisks indicate areas with low cellularity. **f**, Bone marrow cellularity quantified by cell count (rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). **g**, CBC data for white blood cells, haemoglobin and platelets sampled at 6 weeks (as H3K36M mice became moribund). rtTA, $n=8$; H3K9M, $n=3$; H3K36M, $n=4$. **h**, Western blot analysis for H3K9me3 and H3K36me3 in LKS cells sorted from transgenic mice following 1 week of doxycycline administration. See Source data for full membrane western blot images. Data in **e** and **h** are representative of three independent experiments. For **c**, **d**, **f** and **g**, columns represent the mean and error bars represent the standard deviation of the mean for biological replicates, and statistical significance was determined using two-tailed unpaired Student's *t*-test.

levels of the progenitor marker c-Kit (Extended Data Fig. 3a,b), which suggested that they were not simply expanded progenitors. Indeed, May–Grünwald–Giemsa (MGG) staining revealed that the abnormal cells exhibit morphologies characteristic of both immature progenitors (pro-erythroblasts) and more mature precursors (erythroblasts), despite the absence of the erythroblast marker Ter119 (Fig. 3a,c). Critically, haemoglobin levels were significantly lower by 4 weeks of H3K36M induction and dropped precipitously by death at 5–6 weeks of induction, thus confirming a severe differentiation defect (Fig. 2g; Extended Data Fig. 3c). H3K9M mice did not present overt defects in erythropoiesis (Fig. 3a–c; Extended Data Fig. 3a–c). Collectively, these data demonstrate that global loss of H3K36 methylation, but not H3K9 methylation, results in a robust block in erythroid differentiation, yet preserves early, albeit abnormal and ineffective, erythropoiesis.

Considering the abnormal platelet counts observed with both mutants (Fig. 2g), we next quantified megakaryocyte progenitors (MKPs) by flow cytometry following 4 weeks of doxycycline induction. MKPs in H3K9M mice increased nearly threefold while H3K36M expression caused a subtle but significant decrease in MKP numbers (Fig. 3d,e). Serial CBCs revealed that H3K36M and H3K9M expression had reciprocal effects on platelet numbers as

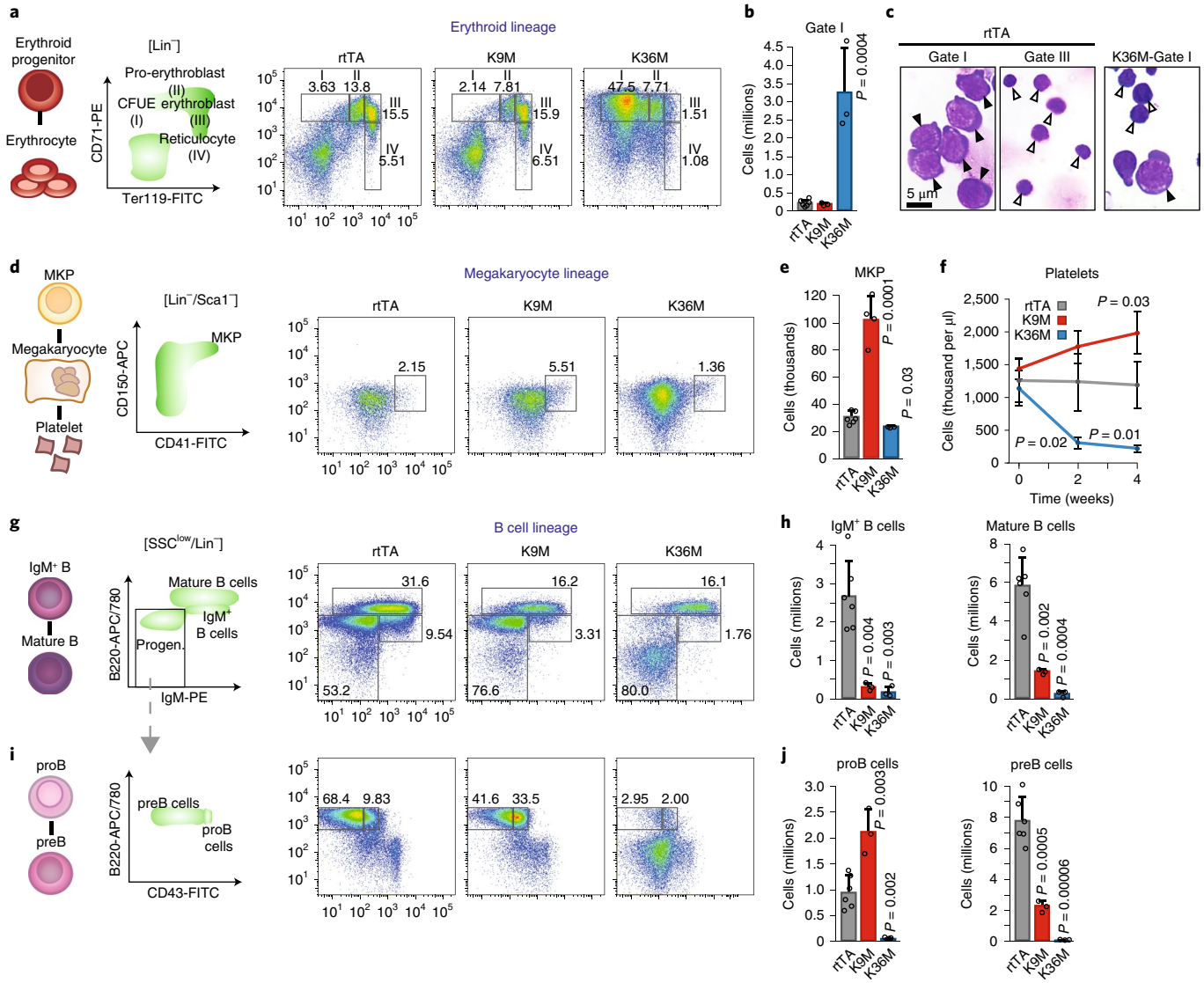


Fig. 3 | In vivo expression of H3K36M dramatically alters myeloid and lymphoid differentiation. a, Flow cytometry analyses of erythroid cells. Frequencies for each cell type are indicated as a percentage of the parent gate. **b**, Quantification of erythroid progenitors (rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). Statistical significance was determined using two-tailed unpaired Student's *t*-test. **c**, Cytospin images of erythroid progenitors sorted from H3K36M and rtTA mice. Black arrowheads indicate pro-erythroblasts, white arrowheads indicate erythroblasts. **d**, Flow cytometry analysis of MKPs. Frequencies for each cell type are indicated as a percentage of the parent gate. **e**, Quantification of MKPs. Columns represent the mean and error bars represent standard deviation of the mean for biological replicates (rtTA, $n=6$; H3K9M, $n=4$; H3K36M, $n=3$). **f**, Time course analysis for platelets assessed by CBC. Columns represent the mean and error bars represent the standard deviation of the mean for biological replicates (rtTA, $n=5$; H3K9M, $n=4$; H3K36M, $n=3$). **g**, Flow cytometry analyses for B cells. Frequencies for each cell type are indicated as a percentage of the parent gate. **h**, Quantification of mature B cells and IgM⁺ B cells (rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). **i**, Flow cytometry analyses for B cell progenitors. Frequencies for each cell type are indicated as a percentage of the parent gate. **j**, Quantification of proB cells and preB cells (rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). Schematics in **a**,**d**,**g**, and **i** show differentiation pathways for the indicated cells. For **b**, **e**, **f**, **h** and **j**, columns represent the mean and error bars represent the standard deviation of the mean for biological replicates, and statistical significance was determined using two-tailed unpaired Student's *t*-test (**b**, **e**, **h** and **j**) or the Holm–Sidak method (**f**). Data in **a**, **c**, **d**, **g** and **i** are representative of three independent experiments.

early as 2 weeks after induction (Fig. 3f). Moreover, two out of three examined H3K9M mice suffered spleen infarction (Extended Data Fig. 3d), which may be due to increased platelet numbers and/or altered platelet function. Thus, reduced H3K36 and H3K9 methylation has opposite effects on the megakaryocyte lineage.

During flow cytometry analysis of H3K36M mouse samples, we noted a decrease in cells exhibiting high side scatter (SSC), a measure of granularity that identifies mature granulocytes. Accordingly, bone marrow cells positive for the pan-granulocyte marker Gr1 exhibited reduced SSC compared with both control and H3K9M cells (Extended Data Fig. 3e,f). Moreover, sorted granulocytes from H3K36M animals appeared underdeveloped after MGG staining, with thicker, less segmented nuclei and more immature (blue) cytoplasm compared with control cells (Extended Data Fig. 3g). We conclude that H3K36M expression impedes granulocyte maturation without affecting early granulocyte specification.

Strikingly, mice expressing either H3K9M or H3K36M exhibited reduced B cell numbers (Fig. 3g,h). Analysis of B cell progenitors revealed that H3K9M induction led to an accumulation of proB cells at the expense of preB cells, which indicated that there

was a specific block during the proB-to-preB stage of differentiation (Fig. 3i,j). By contrast, mice expressing H3K36M were nearly devoid of early B cells, which suggested that there was an earlier block in lymphoid differentiation. Supporting this notion, H3K36M mice exhibited thymic atrophy (Fig. 2d), with a severe reduction in T cells positive for both CD4 and CD8 compared with H3K9M mice (Extended Data Fig. 3h,i). These data show that H3K36 methylation levels affect early lymphoid specification, while H3K9 methylation is required for the transition of more committed progenitor states.

H3K9M and H3K36M have distinct effects on haematopoietic stem and progenitor cell maintenance. Within the bone marrow, the LKS population contains all long-term haematopoietic stem cells (LT-HSCs) as well as short-term HSCs (ST-HSCs) and multipotent progenitors (MPPs). H3K9M induction led to a robust increase in the number of LKS cells (Fig. 4a,b), which suggested that H3K9 methylation limits their expansion. A more detailed analysis using Flt3 and SLAM markers indicated that the increase in LKS cells was due to increased abundances of ST-HSCs and MPPs (Fig. 4c–f), while LT-HSCs were not significantly altered (Fig. 4f, right panel). Consistent with this notion, a large proportion of the LKS population in H3K9M mice expressed high levels of Flt3, which marks lymphoid-primed multipotent progenitors (LMPPs)[26,27] (Fig. 4c,d). In stark contrast, expression of H3K36M reduced the number of LKS cells, LT-HSCs and LMPPs (Fig. 4a–f), thus suggesting a potent role of H3K36 methylation during very early lymphoid differentiation as well as stem cell maintenance, which is consistent with a recent report that implicated *Setd2* in haematopoietic stem and progenitor cell (HSPC) self-renewal[28]. Together, these data show that H3K36 and H3K9 methylation have profound, partially opposing effects on early HSPC populations.

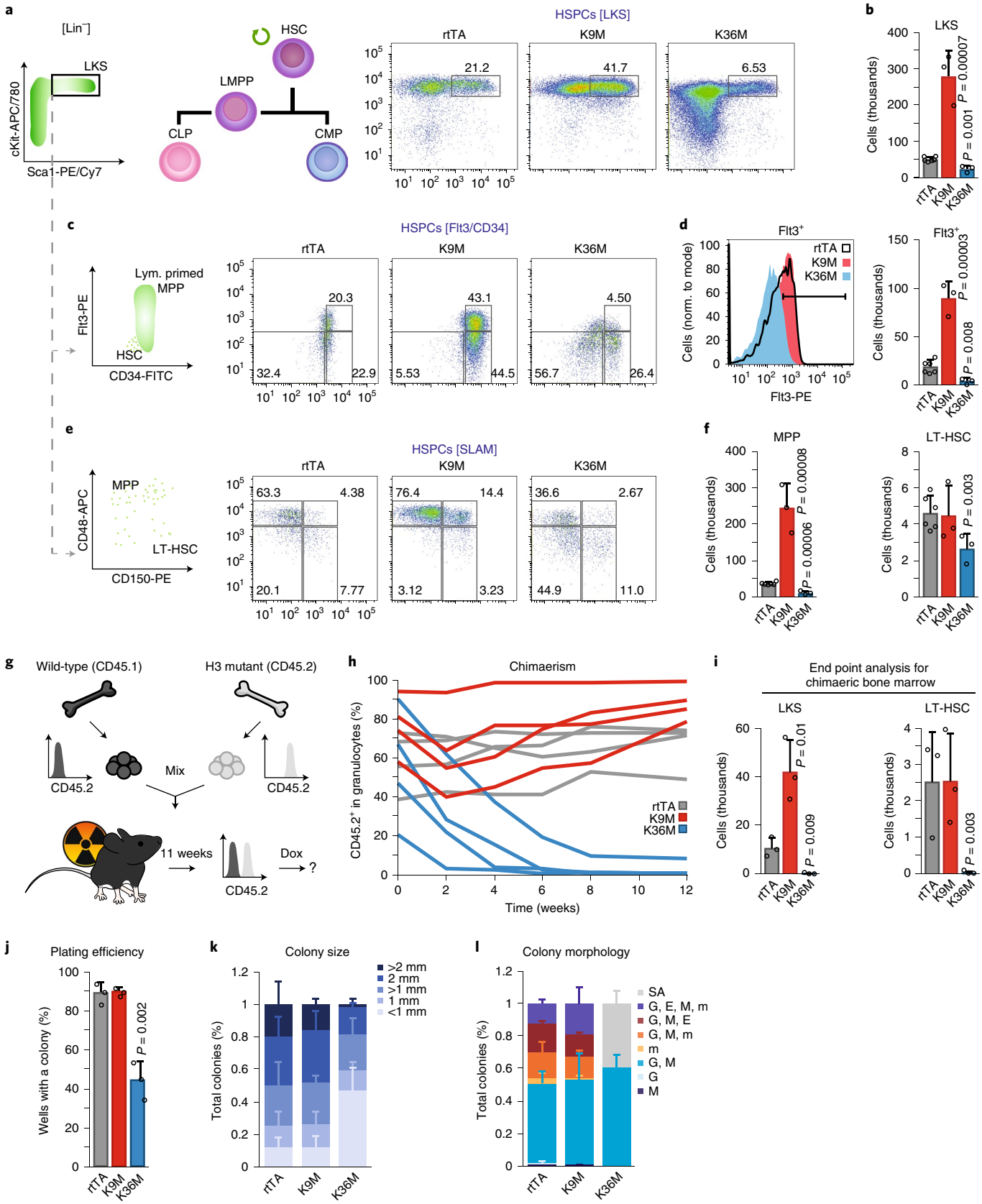
To study the intrinsic role of these mutants in HSCs, we analysed haematopoiesis after competitive bone marrow transplantation into mice with an unperturbed microenvironment (Fig. 4g). Recipients were induced with doxycycline 11 weeks after transplantation to ensure complete engraftment (Fig. 4h). H3K36M-expressing granulocytes (CD45.2$^+$, mutant-donor-derived) progressively diminished compared with rtTA control granulocytes and became undetectable by 12 weeks (Fig. 4h; Extended Data Fig. 4a). By contrast, the contribution from H3K9M-expressing granulocytes remained robust over this time period (Fig. 4h; Extended Data Fig. 4a). To confirm that these effects reflected impaired HSCs, we examined the levels of donor-derived HSPCs via flow cytometry at the conclusion of the time course. Indeed, H3K36M donor-derived HSCs were absent after 12 weeks of induction (Fig. 4i; Extended Data Fig. 4b). Moreover, granulocytes exhibited an abnormal SSC profile akin to H3K36M transgenic mice, which indicated that the observed granulocyte defect is intrinsic to blood cells (Extended Data Fig. 4c). Similar to transgenic mice, the pool of H3K9M-expressing LKS cells and LMPPs was expanded while the number of LT-HSCs was not significantly altered in H3K9M bone marrow recipients, thus confirming that these phenotypes are also cell-intrinsic (Fig. 4i; Extended Data Fig. 4b). These results demonstrate that both H3K36M and H3K9M induction causes cell-autonomous early haematopoietic defects.

To study the HSC defect at a single-cell level in vitro, we induced mutant and control mice for 4 weeks and sorted single LT-HSCs into 96-well plates in the presence of a replete cytokine cocktail and doxycycline. H3K36M-expressing LT-HSCs showed markedly diminished plating efficiency and generated smaller colonies compared with both control and H3K9M cells (Fig. 4j,k). Scoring for morphology, we observed severe lineage restriction in colonies derived from H3K36M-expressing LT-HSCs (Fig. 4l). Specifically, no erythroid differentiation was appreciable by light microscopy after H3K36M induction, although cells with morphologies characteristic of arrested, early erythroid lineage differentiation were detected after MGG staining (Extended Data Fig. 4d,e). These findings recapitulate the abnormalities observed in the bone marrow and demonstrate that the erythroid differentiation defect is lineage-intrinsic. Moreover, no megakaryocytes were detected in methylcellulose assays after H3K36M induction, as observed in mice (Fig. 4l; Extended Data Fig. 4d,e). By contrast, colonies from HSCs expressing H3K9M comprised all differentiated cell types at expected frequencies (Fig. 4l; Extended Data Fig. 4d,e). Collectively, these in vitro results corroborate our conclusion that H3K36M, but not H3K9M, induction profoundly perturbs HSC proliferation and differentiation.

H3K9M and H3K36M expression elicits aberrant transcriptional programmes. To investigate the molecular mechanisms by which our mutants affect cell fate, we analysed gene expression patterns in lineage$^-$, c-Kit$^+$ HSPCs using RNA-seq following 4 weeks of doxycycline induction. In general, gene expression changes following H3K9M induction were subtle, while induction of H3K36M led to robust expression differences. Compared with control, 441 and 2,306 genes were differentially expressed (twofold or greater change in expression, false discovery rate (FDR) of <0.05) in H3K9M-induced and H3K36M-induced HSPCs, respectively. Multidimensional scaling analysis indicated that biological replicates were highly consistent, and expression profiles for H3K9M, H3K36M and control were distinct (Fig. 5a). Importantly, genes associated with specific haematopoietic lineages were dysregulated between mutant histone samples and control, and correlated with the haematopoietic abnormalities observed in vivo (Fig. 5b). For example, we detected a striking increase in the expression of erythroid lineage regulators such as *Gata1*, *Zfpm1*, *Klf1* and *Gfi1b* in HSPCs following H3K36M induction, which is consistent with increased early erythropoiesis in these mice. Similarly, downregulation of important HSC and granulocyte regulators such as *Gfi1* (refs. [29,30]) and *Cebpa*[31,32] correlated with the HSC defect and impaired granulocyte maturation in H3K36M mice. By contrast, regulators of the B cell lineage, including *Pax5* and *Ebf1*, were notably upregulated in HSPCs following H3K9M induction in

Fig. 4 | Expression of H3K36M and H3K9M have opposing, cell autonomous effects on HSPC function. a, Flow cytometry analysis of LKS cells. Frequencies are shown as a percentage of the parent gate. **b**, Quantification of LKS cells (biological replicates, rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$). **c**, Analysis of HSPC populations gated on LKS cells. Frequencies are shown as a percentage of the parent gate. **d**, Flt3$^+$ distribution within LKS cells (left) with quantification of LMPPs (right), defined as Flt3^{high} cells (top 25% of Flt3$^+$ cells set on rtTA control). Biological replicates, rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$. **e**, Analysis of SLAM markers, gated on LKS cells. Frequencies are indicated as a percentage of the parent gate. **f**, Quantification of MPPs (left) and LT-HSCs (right). Biological replicates, rtTA, $n=6$; H3K9M, $n=3$; H3K36M, $n=3$. **g**, Scheme of competitive bone marrow transplant assay. **h**, Time course of the contribution of mutant or rtTA control bone marrow in a competitive bone marrow transplant assay. **i**, Quantification of donor-derived CD45.2$^+$ LKS cells and LT-HSCs at the conclusion of the competitive bone marrow transplant assays ($n=3$, biological replicates). **j**, Quantification of plating efficiency for methylcellulose assays (rtTA, $n=3$; H3K9M, $n=3$; H3K36M, $n=3$, biological replicates). **k**, Quantification of colony size from methycellulose assays (rtTA, $n=3$; H3K9M, $n=3$; H3K36M, $n=3$, biological replicates). **l**, Quantification of colony morphology from methycellulose assays (rtTA, $n=3$; H3K9M, $n=3$; H3K36M, $n=3$, biological replicates). E, erythrocyte; G, granulocyte; m, megakaryocyte; M, monocyte; SA, sparse, atypical. For **b**, **d**, **f**, **i** and **j**–**l**, columns represent the mean and error bars the standard deviations, and statistical significance was determined using two-tailed unpaired Student's *t*-test (**b**, **d**, **f**, **i** and **j**). Data in **a**, **c** and **e** are representative of three independent experiments.

accordance with the increase in LMPPs and proB cells. GSEA of haematopoietic gene expression signatures[33,34] revealed that HSPCs from H3K9M mice were enriched for transcriptional programmes characteristic of early haematopoietic progenitors (for example, MPPs) and lymphoid progenitors (for example, common lymphoid progenitors (CLPs)), thus corroborating the differentiation defects in these mice (Fig. 5c; Extended Data Fig. 5a). Conversely, H3K36M HSPCs exhibited a marked loss of gene expression patterns



associated with HSCs, but showed a strong enrichment for expression patterns related to early erythroid development (for example, pre-CFUE) (Fig. 5c; Extended Data Fig. 5a). Finally, H3K36M, H3K9M and control HSPCs showed similar transcriptional differences and lineage bias after as little as 1 week of doxycycline induction, which suggested that these programmes are directly affected by the loss of H3K36 and H3K9 methylation (Extended Data Fig. 5b,c). Thus, induction of H3K9M and H3K36M in early progenitors leads to expression changes of important haematopoietic genes, which probably contribute to the observed abnormalities in vivo.

H3K9M expression increases chromatin accessibility via loss of H3K9me3. We next performed chromatin immunoprecipitation with sequencing (ChIP-seq) for H3K9me3 on HSPCs purified from induced H3K9M mice. H3K9me3 deposition was predominantly found in extended heterochromatic areas in control cells, including intergenic regions, long interspersed nuclear elements and long terminal repeat elements (Extended Data Fig. 6a). Comparing H3K9M to control samples, we observed a genome-wide reduction of H3K9 methylation (Extended Data Fig. 6b). In total, 323 out of 576 domains (56%) with broad H3K9me3 enrichment in control cells had reduced methylation in H3K9M cells. Closer inspection revealed that domains with the highest H3K9me3 signal (40 out of 576; 7%) appeared to resist H3K9M-dependent loss of H3K9 methylation (Extended Data Fig. 6b). By contrast, regions that showed intermediate enrichment levels in control cells demonstrated a substantial loss of the H3K9me3 signal (Fig. 5d,e). Thus, H3K9M induction leads to a global reduction of H3K9 trimethylation across heterochromatic regions while sparing certain domains that normally exhibit high levels of H3K9me3.

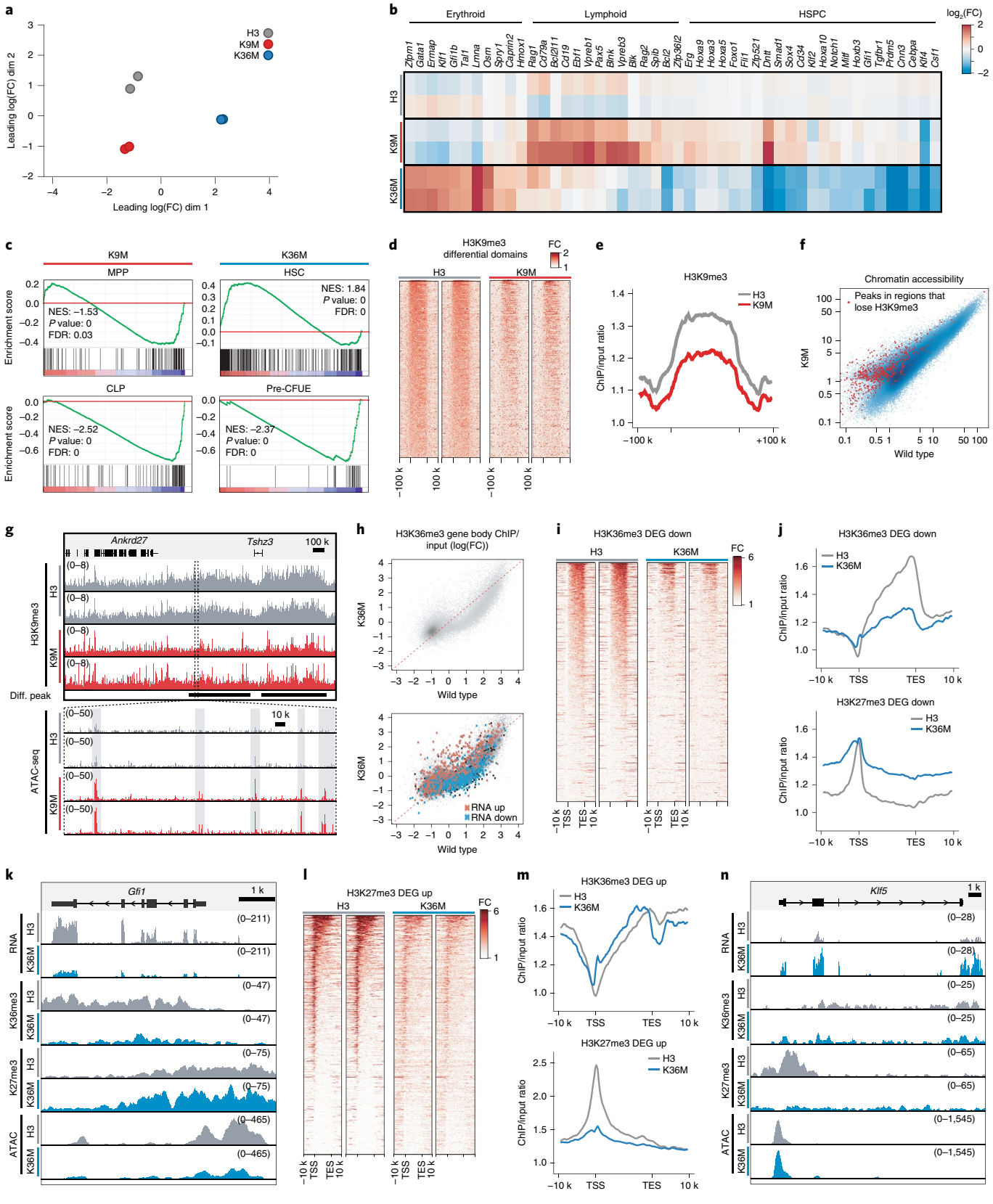
We failed to detect a direct correlation between altered H3K9me3 levels and gene expression in H3K9M HSPCs compared with control. This is probably due to the broad distribution of H3K9me3 over repetitive and gene-poor areas. We note that although H3K9me3 differences were smaller in magnitude, they tended to span larger areas (Fig. 5e). We therefore hypothesized that H3K9M may exert its regulatory effect indirectly by changing chromatin organization rather than by influencing specific genes or loci. In support of this notion, we discovered a strong association between loss of the H3K9me3 signal and increased chromatin accessibility when comparing our ChIP-seq data with ATAC-seq data from the same cells (Fig. 5f,g; Extended Data Fig. 6c). These results imply that H3K9 methylation predominantly affects chromatin accessibility in HSPCs, which secondarily influences gene expression of critical haematopoietic genes. The increased expression of repetitive elements, including long and short interspersed nuclear elements and long terminal repeats, we observed after H3K9M induction (Extended Data Fig. 6d) are probably due to more subtle changes in H3K9me3 levels that escaped detection.

H3K36M expression globally alters the chromatin landscape. We performed ChIP-seq for H3K36me3 and its antagonizing mark H3K27me3 in HSPCs expressing H3K36M and observed a genome-wide reduction of H3K36me3 (Extended Data Fig. 6e). In total, 49.2% of H3K36me3 domains were reduced 1.5-fold or greater in H3K36M cells. Consistent with its connection to active transcription, H3K36me3 in control cells was primarily associated with gene bodies (that is exon and intron regions; Extended Data Fig. 6f). Overall, H3K36me3 was depleted within genes that were downregulated following induction of H3K36M (Fig. 5h-j and Extended Data Fig. 6g), therefore suggesting a connection between loss of the histone mark and transcriptional silencing. Notably, a subset of loci that lost H3K36me3 showed an increased H3K27me3 signal at the promoter and in regions adjacent to the gene body (Fig. 5j; Extended Data Fig. 6h, i), which is in agreement with the antagonistic relationship between these two histone marks[35]. Increased H3K27me3 signal was observed over large domains (±100 kb from the transcription start site (TSS)), which suggests that there is a widespread effect of H3K36M on this mark (Extended Data Fig. 6j, upper panel). However, this H3K27me3 signal was not simply increased genome-wide, because H3K27me3 levels were unaltered at control loci (Extended Data Fig. 6j, lower panel). Several key haematopoietic regulators such as *Gfi1* and its co-factors *Prdm5*, *Klf4*, *Cebpa* and *Fli1* demonstrated this pattern, suggesting that they may be directly silenced by H3K36M-induced loss of H3K36 methylation and gain of H3K27 methylation (Fig. 5k). Transcriptional silencing of these genes was accompanied by reduced chromatin accessibility of associated regulatory elements, as judged by ATAC-seq profiles (Fig. 5k; Extended Data Fig. 6k–m). Moreover, we observed a global correlation between loss of H3K36 methylation and decreased chromatin accessibility when considering all genomic regions (Extended Data Fig. 6k). Previous work has suggested a link between genic H3K36 methylation and DNA methylation[36]. We therefore performed reduced representation bisulfite sequencing to compare DNA methylation between H3K36M-expressing and control HSPCs. We observed only modest decreases in DNA methylation over gene bodies (Extended Data Fig. 6n). Nevertheless, several haematopoietic genes were affected (Extended Data Fig. 6n), including *Gfi1*, which exhibited a decrease in DNA methylation at the gene body (Extended Data Fig. 6o), thus paralleling the observed loss of H3K36me3 (Fig. 5k).

A minority of H3K36me3 domains (3.3%) showed an increased signal within gene bodies after H3K36M induction. While expression of the corresponding transcripts increased (Extended Data Fig. 6g),

Fig. 5 | H3K36M and H3K9M induction in HSPCs leads to widespread changes in the chromatin landscape and altered gene expression. **a**, Multidimensional scaling analysis based on the union of differentially expressed genes (DEGs) between H3 control and H3K9M or H3K36M samples. Each circle represents a biological replicate. **b**, Expression values ($\log_2$(fold-change (FC)) relative to control) for select haematopoietic genes in biological duplicate from HSPCs sorted at 4 weeks of induction. **c**, GSEA analysis based on RNA-seq data for HSPCs sorted from H3K9M and H3K36M mice induced for 4 weeks (H3, $n=2$; H3K9M, $n=2$; H3K36M, $n=2$). Enrichment is shown for transcriptional signatures related to MPPs[50], CLPs[34], HSCs[33] and early erythroid progenitors (pre-CFUE[34]). Statistics were generated in accordance with the published GSEA algorithm[51]. NES, normalized enrichment score. **d**, Replicate heatmaps showing H3K9me3 levels at loci that lose methylation following H3K9M expression in HSPCs. **e**, A meta-analysis profile for H3K9me3 at loci that lose methylation following H3K9M induction in HSPCs. **f**, A correlation plot comparing chromatin accessibility between H3- and H3K9M-expressing HPSCs. Loci that lose H3K9me3 signal with H3K9M expression are highlighted. **g**, Representative tracks for a locus with a decreased H3K9me3 signal following H3K9M expression in HSPCs. **h**, A correlation plot showing the ratio of gene body H3K36me3 levels (upper). The lower panel displays the same data, but indicates differentially expressed transcripts between H3K36M- and H3-expressing cells. **i**, Replicate heatmaps showing H3K36me3 methylation at gene bodies for genes that are downregulated (DEG down) following H3K36M induction in HSPCs. TES, transcription end site. **j**, Meta-analysis profiles for H3K36me3 and H3K27me3 for genes that are downregulated following H3K36M induction in HSPCs. **k**, Representative tracks for *Gfi1*, a gene that showed decreased expression in HSPCs following H3K36M induction. **l**, Replicate heatmaps showing H3K27me3 methylation at gene bodies for genes that are upregulated (DEG up) following H3K36M induction in HSPCs. **m**, Meta-analysis profiles for H3K36me3 and H3K27me3 for genes that are upregulated following H3K36M induction in HSPCs. **n**, Representative tracks for *Klf5*, a gene that showed increased expression following H3K36M induction. DEG, differentially expressed genes.

they represented only a fraction of the upregulated genes following H3K36M induction (Extended Data Fig. 6p), suggesting that there may be a secondary, indirect mechanism of regulation. Indeed, when we examined chromatin patterns for upregulated genes, we discovered that H3K36me3 levels were largely unaffected. However, the TSSs of these genes exhibited a decreased H3K27me3 signal in H3K36M cells compared with controls (Fig. 5l,m). In line with the phenotypes we observed in H3K36M-expressing mice, this subset



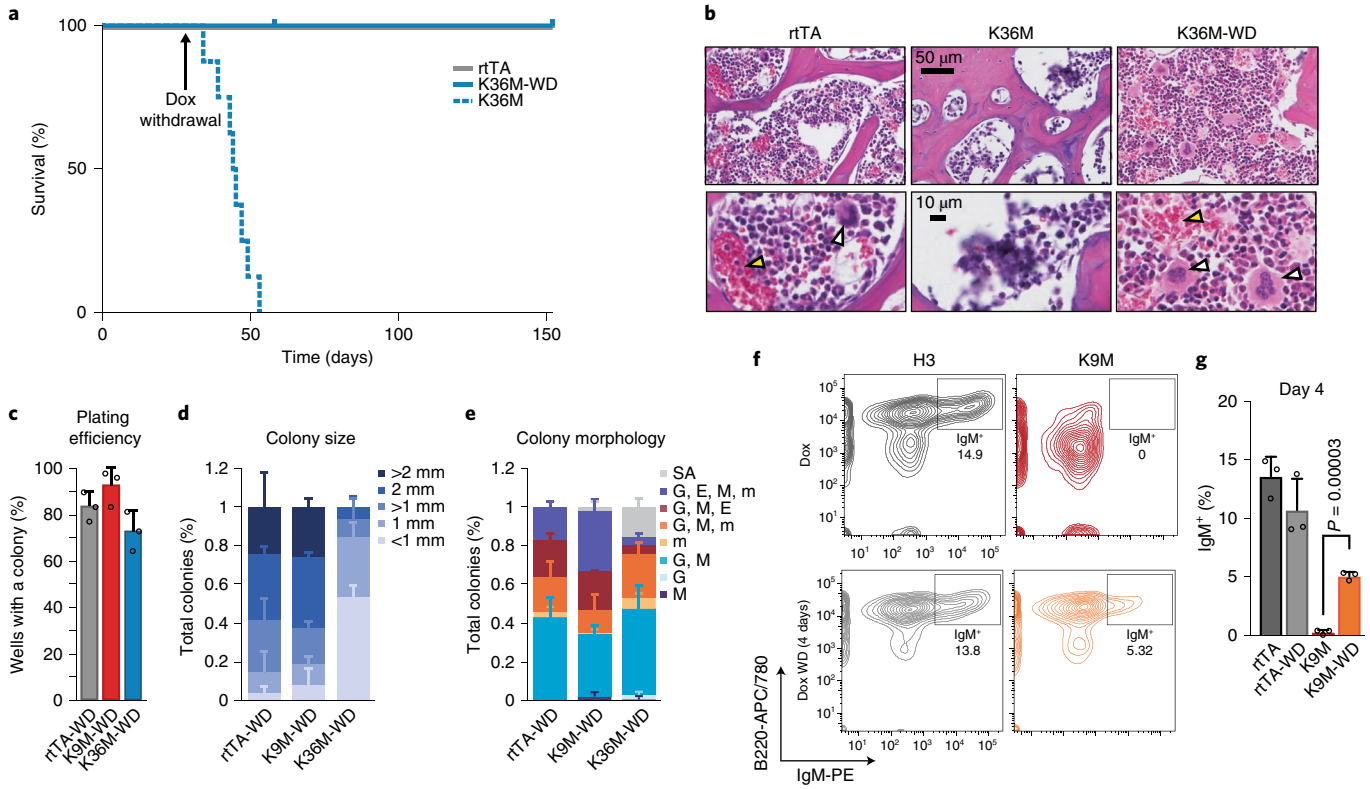


Fig. 6 | The physiological consequences of suppressing H3K9 and H3K36 methylation are reversible both in vivo and in vitro. a, A survival curve for transgenic mice following doxycycline administration and removal at 4 weeks (rtTA, $n = 6$ mice; H3K36M, $n = 8$ mice; H3K36M-WD, $n = 8$ mice). **b**, Histological analyses of bone from mice after 4 weeks of doxycycline induction and 4 weeks of doxycycline withdrawal. Scale bar, 50 μm (upper panels) and 10 μm (lower panels). White arrowheads indicate megakaryocytes, yellow arrowheads indicate mature erythroid cells. **c**, Quantification of plating efficiency for methylcellulose assays with doxycycline removal (rtTA, $n = 3$; H3K9M, $n = 3$; H3K36M, $n = 3$). **d**, Quantification of colony size from methylcellulose assays with doxycycline removal (rtTA, $n = 3$; H3K9M, $n = 3$; H3K36M, $n = 3$). **e**, Quantification of colony morphology from methylcellulose assays with doxycycline removal (rtTA, $n = 3$; H3K9M, $n = 3$; H3K36M, $n = 3$). **f**, Flow cytometry analysis for IgM$^+$B cells. Frequencies for each cell type are indicated as a percentage of the parent gate. **g**, Quantification of IgM$^+$B cells (H3, $n = 3$; H3K9M, $n = 3$; H3K36M, $n = 3$). Statistical significance was determined using a two-tailed unpaired Student's *t*-test (****$P < 0.0001$). For **c**–**e** and **g**, columns represent the mean and error bars represent the standard deviation of the mean for biological replicates. Data in **b** and **f** are representative of three biological replicates.

of genes included ectopically expressed haematopoietic regulators such as *Klf5*, *Zfpm1* and *Lmna* (Fig. 5n; Extended Data Fig. 6q,r). This finding is consistent with previous observations that H3K36M induction leads to a genomic change in the H3K27me3 signal, and correlates with the transcriptional reactivation of genes that are normally silenced by the PRC complex[10]. Collectively, our integrative genomic analyses suggest that H3K36M induction affects cell fate through both direct and indirect mechanisms.

H3K9M- and H3K36M-dependent differentiation defects are largely reversible. To test whether the HSC defects associated with H3K36M expression were reversible, we induced H3K36M mice and then withdrew doxycycline (referred to hereafter as H3K36M-WD). Indeed, H3K36M-WD mice survived long-term (Fig. 6a). Histological analysis of mice 4 weeks after doxycycline removal demonstrated robust recovery of bone marrow cellularity (Fig. 6b) and large numbers of megakaryocytes (Fig. 6b). Moreover, histological examination of the testes revealed robust expansion of spermatocytes, indicating recovery of testicular atrophy (Extended Data Fig. 7a). Likewise, Paneth and goblet cells were again detectable in H3K36M-WD mice, which suggests that the secretory lineage defect was rescued (Extended Data Fig. 7b).

To confirm these results in a controlled in vitro system, we performed single-cell methylcellulose assays using LT-HSCs from H3K9M and H3K36M mice that had been induced for 4 weeks, followed by culture in doxycycline-free media. Notably, the plating efficiency for H3K36M HSCs recovered to levels commensurate with control and H3K9M mice following doxycycline withdrawal (Fig. 6c), although colony sizes remained smaller for H3K36M cells following the 12-day differentiation time course (Fig. 6d). This observation suggested that the self-renewal capacity of H3K36M cells recovered, but the proliferative potential was not fully restored during the examined time period following transgene removal. Importantly, however, these colonies contained all the expected differentiated cell types, including megakaryocytes and mature erythroid cells, consistent with reversal of the differentiation defects (Fig. 6e; Extended Data Fig. 7c).

To determine whether the consequences of H3K9M expression were reversible, we examined B cell maturation[37] using a well-defined in vitro system. We again induced mutant and control mice with doxycycline for 4 weeks and then sorted proB cells (Hardy fraction C[37]) onto ST2 stromal cells with or without doxycycline. H3K9M cells maintained in doxycycline were unable to differentiate into IgM$^+$ B cells, which was consistent with the in vivo phenotype and confirmed the lineage-intrinsic nature of the B cell differentiation defect (Fig. 6f,g). However, H3K9M-WD samples demonstrated a significant recovery of differentiation capacity after only 4 days of culture without doxycycline (Fig. 6f,g). Critically, western blot analysis revealed that histone methylation levels recovered following doxycycline withdrawal both in vitro (EB assays)

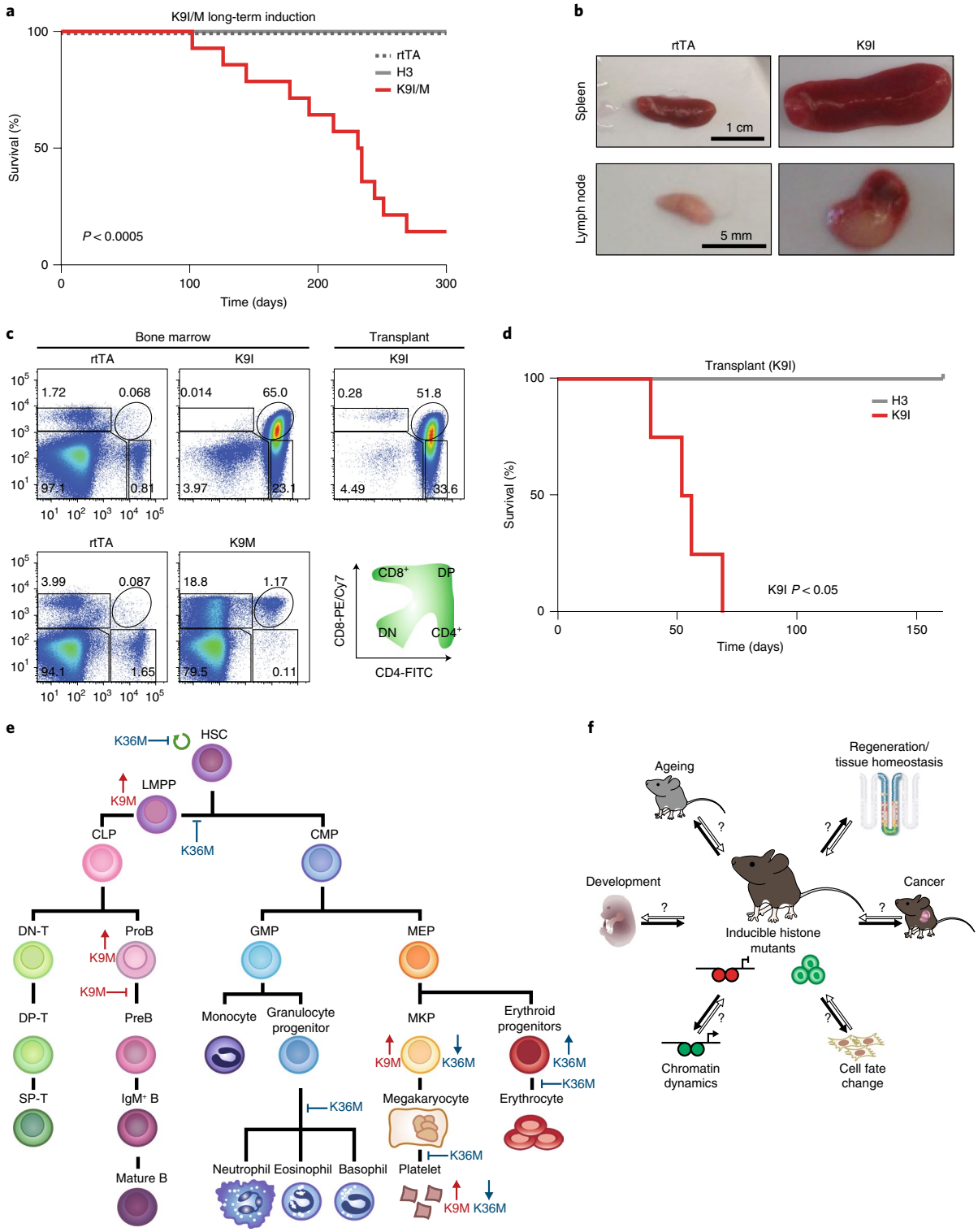


Fig. 7 | Prolonged expression of H3K9M/I is associated with T cell leukaemia/lymphoma. a, A survival curve for H3K9M/I transgenic mice with long-term doxycycline administration (rtTA, $n=4$ mice; H3, $n=8$; H3K9M/I, $n=14$). Statistical significance was determined using a two-sided log-rank (Mantel–Cox) test. **b**, Representative spleens and lymph nodes from control (rtTA) and K9I mice. Scale bar, 1 cm (spleens) or 5 mm (lymph nodes). H3, $n=3$ mice; K9I, $n=4$ mice; one representative example of two independent transplant experiments from independent primary leukaemia-bearing mice). **c**, Flow cytometry analyses for T cells in bone marrow of mice with long-term doxycycline induction (left) and a representative transplant recipient (right). Frequencies for each cell type are indicated as a percentage of the parent gate (H3, $n=3$ mice; K9I, $n=4$ mice; one representative example of two independent transplant experiments from independent primary leukaemia-bearing mice). **d**, A survival curve for transplant recipients of bone marrow from a mouse that developed acute T cell leukaemia/lymphoma or control bone marrow after long-term doxycycline induction (H3, $n=3$; K9I, $n=4$; one representative example of two independent transplant experiments from independent primary leukaemia-bearing mice). Statistical significance was determined using a two-sided log-rank (Mantel–Cox) test. **e**, A summary of the effects of H3K9M and H3K36M induction on haematopoiesis. **f**, Potential uses of the K-to-M system to study diverse biological processes. DP-T, double positive T cells; DN-T, double negative T cells; SP-T, single positive T cells; CMP, common myeloid progenitor; GMP, granulocyte macrophage progenitor; MEP, megakaryocyte erythroid progenitor.

and in vivo (HSPCs) (Extended Data Fig. 7d,e). Thus, the methylation changes and differentiation blocks imposed by H3K36M and H3K9M induction are largely reversible.

Long-term induction of H3K9M or H3K9I leads to aggressive T cell leukaemia. Previous studies have shown that replacement of histone lysine residues with isoleucine (for example, H3K9I) blocks methylation in a manner similar to methionine substitution[9,38]. Analogous to our K-to-M mutants, we generated doxycycline-inducible H3K9I ES cells and mice. We then examined haematopoiesis after 4 weeks of doxycycline treatment. Similar to H3K9M mice, H3K9I mice exhibited expanded MKPs, increased platelet counts (Extended Data Fig. 8a–c) and a block during the proB cell-to-preB cell transition (Extended Data Fig. 8d–g), as well as increased MPPs and LMPPs but normal numbers of LT-HSCs (Extended Data Fig. 8h–m). These results demonstrate that H3K9I mice closely phenocopy H3K9M mice.

Chronic loss of H3K9 methylation has been implicated in impaired viability and increased tumorigenesis[39]. To explore the long-term impact of H3K9M or H3K9I induction in vivo, we followed cohorts of H3K9M, H3K9I, rtTA and H3 wild-type mice continuously treated with doxycycline. Remarkably, most H3K9M and H3K9I mice but none of the control mice died by 12 months of doxycycline treatment (Fig. 7a). Analysis of four moribund animals (one H3K9M mouse and three H3K9I mice) showed enlarged spleens and lymph nodes, and histological examination revealed infiltration by blast cells, suggestive of leukaemia/lymphoma (Fig. 7b). Flow cytometry analysis of bone marrow from these mice confirmed increased numbers of cells positive for CD4 or CD8 or aberrant cells positive for both CD4 and CD8 (Fig. 7c), which are indicative of T cell leukaemia/lymphoma. We unequivocally confirmed leukaemia by demonstrating engraftment of cells and lethality (1–2 months) after transplantation into syngeneic recipients (Fig. 7d). Thus, chronic H3K9I/M expression compromises long-term survival and predisposes mice to aggressive T cell leukaemia/lymphoma.

Discussion

Using histone mutants as tools, we probed the impact of two crucial histone modifications on cell identity and differentiation in mammals. Our study identified unexpectedly specific and context-dependent roles for H3K36 and H3K9 methylation in cell fate regulation, including pluripotent stem cell differentiation, testis and gut regeneration as well as blood cell homeostasis (Fig. 7e). We anticipate that our mouse models will be equally useful in studying the role of histone methylation in other biological contexts. Indeed, during the preparation of our manuscript, a study was published that characterized the effects of ectopic H3K36M expression on adipogenesis in mice[40]. In agreement with our observations in blood, those authors found that expression of H3K36M impairs differentiation of progenitors by altering the distribution of H3K36me3 and H3K27me3.

Our system holds key advantages over knockout approaches for histone-modifying enzymes. First, our K-to-M and K-to-I mutations elicit global suppression of methylation at their respective sites, without the need for genetic manipulation of multiple factors. Second, expression of these mutations results in decreased levels, but not complete absence of histone marks, which is advantageous in cases where a complete loss of histone methylation is toxic. Finally, our inducible systems can probe the reversibility of observed cellular phenotypes and underlying chromatin changes. Indeed, we demonstrate that the effect of histone K-to-M mutations is largely reversible, as differentiation defects recovered just days after doxycycline withdrawal. We recognize that transgenic and heterogeneous expression of histone mutants in certain cellular contexts[19,20] is a potential caveat of our approach compared to KMT knockouts. Moreover, the mechanisms by which histone mutants inhibit methylation may be complex and involve sequestering of KMTs or KDMs on chromatin[16,41–44], aberrant histone acetylation and exclusion of KMTs[17,45] and/or global redistribution of histone marks[46]. Regardless of the precise mechanisms, our inducible mutant mice offer a fresh perspective at understanding the physiological role of histone methylation, thus complementing conventional knockout strategies for KMTs.

Previous studies have suggested that K-to-M mutations drive tumorigenesis by locking cells into a more primitive developmental state[10,14,47]. Our observations that H3K36M, H3K9M and H3K9I mice exhibit differentiation blocks in multiple lineages and that a subset of H3K9M and H3K9I mice succumb to T cell leukaemia/lymphoma support this concept[10,12,14]. We speculate that an increase in the fraction of progenitor cell populations, such as LMPPs in the case of T cell leukaemia in H3K9M and H3K9I mice, provides an expanded pool of cells that select for secondary oncogenic mutations. Further studies are needed to test this hypothesis and to establish whether the oncogenic effects of K-to-M mutants are reversible. Our tools may also prove useful to study histone modifications and physiological processes that are associated with changes in chromatin such as tissue regeneration following injury[48] and ageing[49] (Fig. 7f). Finally, our system will enable the study of tissue-specific effects of K-to-M mutants in combination with models that express rtTA under cell type-specific promoters. In summary, our study has uncovered unexpected, direct effects for H3K36 and H3K9 methylation in cell fate and provides a powerful resource to study the functional role of chromatin signalling in both normal and pathological contexts.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41556-019-0403-5.

Received: 5 June 2019; Accepted: 12 September 2019;
Published online: 28 October 2019

References

1. Soshnev, A. A., Josefowicz, S. Z. & Allis, C. D. Greater than the sum of parts: complexity of the dynamic epigenome. *Mol. Cell* **62**, 681–694 (2016).
2. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).
3. Pengelly, A. R., Copur, O., Jackle, H., Herzig, A. & Muller, J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* **339**, 698–699 (2013).
4. Maze, I., Noh, K. M., Soshnev, A. A. & Allis, C. D. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat. Rev. Genet.* **15**, 259–271 (2014).
5. Miller, S. A., Mohn, S. E. & Weinmann, A. S. Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol. Cell* **40**, 594–605 (2010).
6. Shpargel, K. B., Sengoku, T., Yokoyama, S. & Magnuson, T. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. *PLoS Genet.* **8**, e1002964 (2012).
7. Kim, E. et al. Phosphorylation of EZH2 activates STAT3 signaling via STAT3 methylation and promotes tumorigenicity of glioblastoma stem-like cells. *Cancer Cell* **23**, 839–852 (2013).
8. Xu, K. et al. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science* **338**, 1465–1469 (2012).
9. Lewis, P. W. et al. Inhibition of PRC2 activity by a gain-of-function *H3* mutation found in pediatric glioblastoma. *Science* **340**, 857–861 (2013).
10. Lu, C. et al. Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* **352**, 844–849 (2016).
11. Behjati, S. et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat. Genet.* **45**, 1479–1482 (2013).
12. Fang, D. et al. The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science* **352**, 1344–1348 (2016).
13. Schwartzentruber, J. et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).

14. Papillon-Cavanagh, S. et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat. Genet.* **49**, 180–185 (2017).
15. Herz, H. M. et al. Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science* **345**, 1065–1070 (2014).
16. Jayaram, H. et al. *S*-adenosyl methionine is necessary for inhibition of the methyltransferase G9a by the lysine 9 to methionine mutation on histone H3. *Proc. Natl Acad. Sci. USA* **113**, 6182–6187 (2016).
17. Mohammad, F. et al. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nat. Med.* **23**, 483–492 (2017).
18. Mohammad, F. & Helin, K. Oncohistones: drivers of pediatric cancers. *Genes Dev.* **31**, 2313–2324 (2017).
19. Beard, C., Hochedlinger, K., Plath, K., Wutz, A. & Jaenisch, R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* **44**, 23–28 (2006).
20. Hochedlinger, K., Yamada, Y., Beard, C. & Jaenisch, R. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* **121**, 465–477 (2005).
21. Wray, J. et al. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol.* **13**, 838–845 (2011).
22. Zhang, Y. et al. H3K36 histone methyltransferase Setd2 is required for murine embryonic stem cell differentiation toward endoderm. *Cell Rep.* **8**, 1989–2002 (2014).
23. Inagawa, M. et al. Histone H3 lysine 9 methyltransferases, G9a and GLP are essential for cardiac morphogenesis. *Mech. Dev.* **130**, 519–531 (2013).
24. Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* **23**, 2484–2489 (2009).
25. Zuo, X. et al. The histone methyltransferase Setd2 is required for expression of acrosin-binding protein 1and protamines and essential for spermiogenesis in mice. *J. Biol. Chem.* **293**, 9188–9197 (2018).
26. Sitnicka, E. et al. Key role of flt3 ligand in regulation of the common lymphoid progenitor but not in maintenance of the hematopoietic stem cell pool. *Immunity* **17**, 463–472 (2002).
27. Adolfsson, J. et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* **121**, 295–306 (2005).
28. Zhang, Y. L. et al. Setd2 deficiency impairs hematopoietic stem cell self-renewal and causes malignant transformation. *Cell Res.* **28**, 476–490 (2018).
29. Hock, H. et al. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* **431**, 1002–1007 (2004).
30. Hock, H. et al. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity* **18**, 109–120 (2003).
31. Ye, M. et al. C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. *Nat. Cell Biol.* **15**, 385–394 (2013).
32. Zhang, D. E. et al. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proc. Natl Acad. Sci. USA* **94**, 569–574 (1997).
33. Ng, S. Y., Yoshida, T., Zhang, J. & Georgopoulos, K. Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. *Immunity* **30**, 493–507 (2009).
34. Pronk, C. J. et al. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell* **1**, 428–442 (2007).
35. Yuan, W. et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286**, 7983–7989 (2011).
36. Baubec, T. et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**, 243–247 (2015).
37. Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D. & Hayakawa, K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* **173**, 1213–1225 (1991).
38. Lehnertz, B. et al. H3(K27M/I) mutations promote context-dependent transformation in acute myeloid leukemia with RUNX1 alterations. *Blood* **130**, 2204–2214 (2017).
39. Peters, A. H. et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
40. Zhuang, L. et al. Depletion of Nsd2-mediated histone H3K36 methylation impairs adipose tissue development and function. *Nat. Commun.* **9**, 1796 (2018).
41. Shan, C. M. et al. A histone H3K9M mutation traps histone methyltransferase Clr4 to prevent heterochromatin spreading. *eLife* **5**, e17903 (2016).
42. Yang, S. et al. Molecular basis for oncohistone H3 recognition by SETD2 methyltransferase. *Genes Dev.* **30**, 1611–1616 (2016).
43. Justin, N. et al. Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nat. Commun.* **7**, 11316 (2016).
44. Jiao, L. & Liu, X. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* **350**, aac4383 (2015).
45. Piunti, A. et al. Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. *Nat. Med.* **23**, 493–500 (2017).
46. Fang, D. et al. H3.3K27M mutant proteins reprogram epigenome by sequestering the PRC2 complex to poised enhancers. *eLife* **7**, e36696 (2018).
47. Funato, K., Major, T., Lewis, P. W., Allis, C. D. & Tabar, V. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* **346**, 1529–1533 (2014).
48. Blanpain, C. & Fuchs, E. Stem cell plasticity. Plasticity of epithelial stem cells in tissue regeneration. *Science* **344**, 1242281 (2014).
49. Booth, L. N. & Brunet, A. The aging epigenome. *Mol. Cell* **62**, 728–744 (2016).
50. de Graaf, C. A. et al. Haemopedia: an expression atlas of murine hematopoietic cells. *Stem Cell Rep.* **7**, 571–582 (2016).
51. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

Acknowledgements

The authors thank M. Handley, A. Galvin, M. Gesner and E. Surette of the MGH/HSCI Rodent Histopathology Core for technical assistance. They also thank members of the Hochedlinger Lab for helpful discussions. A.J.H. is supported by an American Cancer Society—New England Division—Ellison Foundation Postdoctoral Fellowship (PF-15-130-01-DDC). B.D.S. was supported by an EMBO long-term fellowship (no. ALTF 1143-2015) and a MGH Tosteson and FMD postdoctoral fellowship. K.H. was supported by funds from the MGH, NIH (R01 HD058013-06) and the Gerald and Darlene Jordan Chair in Regenerative Medicine. H.H. was supported by a Hyundai Hope on Wheels Scholar Grant. J.B. is grateful for support from the NIH (1F32HD078029-01A1). R.I.S. was supported by funds from the NIH (P30-DK40561).

Author contributions

J.B., H.H. and K.H. conceived the study and wrote the manuscript. J.B., B.A.S., A.J.H., J. Choi, J. Charlton, J.W.S., A.C., I.S.K., B.D.S., A.M., B.B. and R.M.W. designed and performed the experiments and analysed the data. F.J., A.A. and R.I.S. the performed bioinformatics analysis.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41556-019-0403-5.

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-019-0403-5.

Correspondence and requests for materials should be addressed to H.H. or K.H.

Reprints and permissions information is available at www.nature.com/reprints.

Methods

Plasmid generation and cloning. The coding sequence for *H3f3a* was amplified using Phusion polymerase (Thermo Fisher Scientific) from mouse ES cell cDNA using the following primers: ccgaattcgccaccATGGCTCGTACAAAG; ccgaattcTTAAGCACGTTCTCCG. Following purification of the PCR product using a gel extraction kit (Qiagen), A-tailing was performed using Taq polymerase (New England Biolabs) according to the manufacturer's recommendations. The resulting DNA fragment was inserted into pCR2.1 (Thermo Fisher Scientific) via TOPO-TA cloning. The plasmid was sequence verified and the coding sequence for *H3f3a* was subcloned into pBS31 using EcoRI to generate pBS31-H3.3 (New England Biolabs). Targeting plasmids for mutant histone H3.3 were generated by performing site-directed mutagenesis on pBS31-H3.3 using a QuikChange II Site-Directed Mutagenesis kit (Agilent Technologies) according to the manufacturer's recommendations. Primers used are listed below:

H3K9M: actgcccgcaTGtccaccggtggtaaagca; ttaccaccggtggaCAtgcgggcagtctgctt; *H3K36M*: ggaggggtgaTgaaacctcatcgttacagg; taacgatgaggtttcAtcacccctccagta. All constructs were sequence verified.

ES cell culture and targeting. Mouse ES cells were cultured on irradiated mouse embryonic fibroblasts in growth media (500 ml KO-DMEM (Life Technologies) 15% fetal bovine serum (FBS; Hyclone), 1× non-essential amino acids (Life Technologies), 1× Glutamax (Life Technologies), 1000 U ml^{-1} leukaemia inhibitory factor, 55 μM β-mercaptoethanol (Sigma)).

For targeting, cells were expanded to 80% confluence, trypsinized and collected. The cells were then washed once in PBS and suspended in 500 μl of PBS. Next, 40 μg of pCAGGS-flpE-puro and 40 μg of pBS31-H3.3 or pBS31-H3.3K9M or pBS31-H3.3K36M were added and the mixture was transferred to a 4-m electroporation cuvette (Bio-Rad). The cells were electroporated in a Gene Pulser Xcell system (Bio-Rad) on time constant mode with the voltage at 800 V and time constant of 0.2 ms. The cells were then seeded at clonal density onto DR4 irradiated mouse embryonic fibroblasts (GlobalStem) in a 10-cm^2 plate and cultured in growth media for 24 h. The medium was then replaced with growth medium containing 200 μg ml^{-1} hygromycin to select for targeted clones. Individual colonies were selected, expanded and tested via PCR to ensure proper integration.

EB formation. EBs were generated using the hanging drop method. Briefly, ES cells were trypsinized and pre-plated directly onto plastic tissue culture dishes for 1 h in ES cell growth media to remove fibroblasts. Following centrifugation, the cells were resuspended at a density of 13,000 cells per ml in EB medium (15% FBS (Hyclone), 1× non-essential amino acids (Life Technologies), 1× Glutamax (Life Technologies), 0.4% 1-tioglycerol (Sigma-Aldrich), 1 mM sodium pyruvate (Life Technologies), 10 μg ml^{-1} iron-saturated transferrin (Sigma-Aldrich) and 50 μg ml^{-1} ascorbic acid (Sigma-Aldrich)) supplemented with 2 μg ml^{-1} doxycycline (Sigma-Aldrich). The cells were then cultured in 30 μl hanging drops for 3 days. At that time, EBs were collected and seeded into non-adherent plates and incubated with constant agitation for an additional 3–6 days. Medium was changed every 48 h.

Blastocyst injection and generation of transgenic mice. To obtain transgenic mice, ES cells targeted with the mutant histone constructs were injected into blastocysts at embryonic day 3.5 as previously described[52,53]. KH2 ES cells have a B6/129 background. High-grade chimaeras were crossed to C57/B6 wild-type mice and the offspring were genotyped to determine germline transmission. These mice were then crossed to mice carrying constitutive M2-rtTA in a C57/B6 background. Both male and females were used in this study at 6–8 weeks of age and were induced by adding 2 mg ml^{-1} doxycycline to drinking water. Mice used in this study were housed and bred in specific-pathogen-free rooms located in the AAALAC-accredited Center for Comparative Medicine vivarium at Massachusetts General Hospital. Mice were housed in ventilated cages on a standard 12–12 h light cycle. All procedures involving mice adhered to the guidelines of the approved Massachusetts General Hospital Institutional Animal Care and Use Committee protocol no. 2006N000104. The main transgenic strains will be deposited to The Jackson Laboratory under Stock No. 034364 (Col1a1-tetOP-H3.3 WT), Stock No. 034365 (Col1a1-tetOP-H3.3K36M) and Stock No. 034366 (Col1a1-tetOP-H3.3K9M).

Tissue section staining. Periodic acid–Schiff (PAS) staining of sectioned intestine was performed using a PAS kit (Sigma-Aldrich) according to the manufacturer's instructions. Immunohistochemistry for lysozyme was carried out on intestinal sections using a Vectastain elite ABC HRP kit (Vector Laboratories), lysozyme antibody (PA5-16668, Thermo Fisher) and a DAB peroxidase Substrate kit (Vector Laboratories) with haematoxylin counterstain (Vector Laboratories).

CBC with differential. Whole blood was drawn via retro-orbital bleeding using heparinized microhaematocrit capillary tubes (Thermo Fisher Scientific) into EDTA-coated microvette tubes (Sarstedt). CBCs with differential were analysed on an Element HT5 machine (Heska).

Flow cytometry. Whole blood was prepared for flow cytometry by lysing blood cells in ACK buffer (150 mM NH_4Cl, 10 mM $KHCO_3$, 0.2 mM EDTA, all from Sigma-Aldrich) for 2 min at room temperature. The cells were then centrifuged at 350×*g* for 8 min, resuspended in FACS buffer (3% FBS (Hyclone) in PBS) and filtered through 0.2-μm mesh. The cells were again centrifuged at 350×*g* for 8 min and stained on ice in residual volume for 1 h with the following antibodies: CD3e (APC; BD Biosciences, clone 145-2C11), CD45.2 (Pacific Blue; BioLegend, clone 104), Mac1 (FITC; eBiosciences, Clone M1/70), B220/CD45R (PE/APC; BD Biosciences, clone RA3-6B2), c-Kit (eFluor 780; eBiosciences, clone 2B8), Gr1 (PE; eBiosciences, clone RB6-8C5) and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer, and propidium iodide was added at a concentration of 1 μg ml^{-1} as a viability dye.

Bone marrow from euthanized mice was collected by flushing both femurs and tibiae using a 27.5 gauge needle with 10 ml of culture medium (500 ml DMEM (Life Technologies) 10% FBS (Hyclone), 1× non-essential amino acids (Life Technologies), 1× Glutamax (Life Technologies), 55 μM β-mercaptoethanol (Sigma)). The bone marrow was then dissociated via trituration with a 16 gauge needle. The resulting cells were centrifuged at 350×*g* for 8 min and resuspended in 5 ml of cold ACK lysis buffer and incubated for 7.5 min on ice. The lysis was then quenched immediately with 20 ml of culture medium and again centrifuged at 350×*g* for 8 min. The resulting cells were resuspended in 10 ml of FACS buffer, filtered through 0.2-μm mesh, and counted. The cells were aliquoted and stained as follows:

Pre/proB stain (1 million cells): lineage markers (Ter119 (PE/Cy5; eBiosciences, clone TER-119), CD3e (PE/Cy5; eBiosciences, clone 145-2C11), Mac1 (PE/Cy5; eBiosciences, clone M1/70), Gr1 (PE/Cy5; eBiosciences, clone RB6-8C5)); CD43 (FITC; BD Biosciences, clone S7); B220/CD45R (eFluor 780; eBiosciences, clone RA3-6B2); IgM (PE; eBiosciences, clone II/41); CD45.2 (Pacific Blue; BioLegend, clone 104); and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer and 4,6-diamidino-2-phenylindole (DAPI) was added at a concentration of 1 μg ml^{-1} as a viability dye.

Hardy fraction proB stain (5 million cells): lineage markers (Ter119 (PE/Cy5; eBiosciences, clone TER-119), CD3e (PE/Cy5; eBiosciences, clone 145-2C11), Mac1 (PE/Cy5; eBiosciences, clone M1/70), Gr1 (PE/Cy5; eBiosciences, clone RB6-8C5)); CD43 (FITC; BD Biosciences, clone S7); B220/CD45R (eFluor 780; eBiosciences, clone RA3-6B2); IgM (PE/Cy5; eBiosciences, clone II/41), CD45.2 (Pacific Blue; BioLegend, clone 104); Fc block (BD Biosciences); CD249 (PE; BD Biosciences, clone BP-1); and CD24 (APC; BioLegend, clone M1/69). Following staining, cells were washed twice with FACS buffer and DAPI was added at a concentration of 1 μg ml^{-1} as a viability dye.

Main stain (1 million cells): Gr1 (APC; BD Biosciences, clone RB6-8C5); Mac1 (PE; BD Biosciences, clone M1/70); B220/CD45R (eFluor 780; eBiosciences, clone RA3-6B2); Ter119 (FITC; eBiosciences, clone TER-119); CD45.2 (Pacific Blue; BioLegend, clone 104); and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer and DAPI was added at a concentration of 1 μg ml^{-1} as a viability dye.

SLAM (5 million cells): lineage markers (B220/CD45R (PE/Cy5; eBiosciences, clone RA3-6B2), Ter119 (PE/Cy5; eBiosciences, clone TER-119), CD3e (PE/Cy5; eBiosciences, clone 145-2C11), Mac1 (PE/Cy5; eBiosciences, clone M1/70), Gr1 (PE/Cy5; eBiosciences, clone RB6-8C5)); Sca1 (PE/Cy7; eBiosciences, clone D7); CD150 (PE; BioLegend, clone TC15-12F12.2); CD48 (APC; BioLegend, clone HM48-1); CD45.2 (Pacific Blue; BioLegend, clone 104); c-Kit (eFluor 780; eBiosciences, clone 2B8); and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer and DAPI was added at a concentration of 1 μg ml^{-1} as a viability dye.

LKS (5 million cells): lineage markers (B220/CD45R (PE/Cy5; eBiosciences, clone RA3-6B2), Ter119 (PE/Cy5; eBiosciences, clone TER-119), CD3e (PE/Cy5; eBiosciences, clone 145-2C11), Mac1 (PE/Cy5; eBiosciences, clone M1/70), Gr1 (PE/Cy5; eBiosciences, clone RB6-8C5)); Sca1 (PE/Cy7; eBiosciences, clone D7), c-Kit (APC; BioLegend, clone 2B8); CD34 (FITC; eBiosciences, clone RAM34); Flt3 (PE; eBiosciences, clone A2F10); CD45.2 (Pacific Blue; BioLegend, clone 104); and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer and DAPI was added at a concentration of 1 μg ml^{-1} as a viability dye.

Erythroid stain (1 million cells): lineage markers (B220/CD45R (PE/Cy5; eBiosciences, clone RA3-6B2), Ter119 (PE/Cy5; eBiosciences, clone TER-119), CD3e (PE/Cy5; eBiosciences, clone 145-2C11), Mac1 (PE/Cy5; eBiosciences, clone M1/70), Gr1 (PE/Cy5; eBiosciences, clone RB6-8C5), c-Kit (APC; BioLegend, clone 2B8), Ter119 (FITC; eBiosciences, clone TER-119), CD71 (PE; eBiosciences, clone R17217 (RI7 217.1.4)); CD45.2 (Pacific Blue; BioLegend, clone 104); and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer and DAPI was added at a concentration of 1 μg ml^{-1} as a viability dye.

Megakaryocyte precursor stain (5 million cells): lineage markers (B220/CD45R (PE/Cy5; eBiosciences, clone RA3-6B2), Ter119 (PE/Cy5; eBiosciences, clone TER-119), CD3e (PE/Cy5; eBiosciences, clone 145-2C11), Mac1 (PE/Cy5; eBiosciences, clone M1/70), Gr1 (PE/Cy5; eBiosciences, clone RB6-8C5)); Sca1 (PE/Cy7; eBiosciences, clone D7); c-Kit (eFluor 780; BioLegend, clone 2B8); CD41 (FITC; BD Biosciences, clone MWReg30); CD16/CD32 (PE; BD Biosciences, clone 2.4G2); CD150 (APC; BioLegend, clone TC15-12F12.2); and CD105 (Pacific Blue; BioLegend, clone MJ7/18). Following staining, cells were washed twice with FACS buffer and propidium iodide was added at a concentration of 1 μg ml^{-1} as a viability dye.

Thymocytes were collected by mechanically dissociating the thymus. The cells were filtered through 0.2-µm mesh and counted. Thymocytes (5 million) were stained as follows: CD4 (FITC; eBiosciences, clone RM4-5); CD8a (PE/Cy7; eBiosciences, clone 53-6.7); c-Kit (APC; BioLegend, clone 2B8); CD25 (PE; eBiosciences, clone PC61.5); CD45.2 (Pacific Blue; BioLegend, clone 104); and Fc block (BD Biosciences). Following staining, cells were washed twice with FACS buffer and DAPI was added at a concentration of 1 µg ml^{-1} as a viability dye.

Thereafter, cells were analysed on an LSRII flow cytometer (BD Biosciences) or collected on an Aria II flow cytometer (BD Biosciences) using Diva v.6.1.2 (BD Biosciences), and data analyses were performed using the software FlowJo v.10.2 (FlowJo).

Cytospin. Cytospin slides were performed using a Cytospin 3 centrifuge (Thermo Shandon) and single cytology funnels (Thermo Fisher Scientific). Funnels were prepared by applying 100 µl of PBS to the spin funnel and centrifuging at 1,000 r.p.m. for 1 min. Then, 200,000 cells were resuspended in 100 µl of PBS, applied to the funnel and centrifuged for 1 min at 1,000 r.p.m. Slides dried overnight at room temperature and were stained for 2 min in May–Grünwald stain (neat; Sigma-Aldrich), 13 min in Giemsa stain (diluted 1:40 with distilled water; Sigma-Aldrich) and then washed twice in clean water for 1 min. Slides were dried overnight and mounted with Permount (Thermo Fisher Scientific) and coverslipped.

Clonal methylcellulose assays. Methylcellulose assays were carried out in Methocult GF 3434 media (Stem Cell Technologies) according to the manufacturer's recommendations, except that granulocyte–macrophage colony-stimulating factor and thrombopoietin (both from Peprotech) were added to final concentrations of 5 ng ml^{-1} and 50 ng ml^{-1}, respectively. Where indicated, doxycycline was added to a final concentration of 2 µg ml^{-1}. Single, LT-HSCs were sorted into 96-well plates containing the methylcellulose, using the SLAM stain described above. After 12 days of incubation at 37 °C, the colonies were scored for size. At day 14, the colonies were scored for morphology. Select colonies were extracted for cytospin by adding 200 µl of PBS directly to the colony and incubating at 37 °C for 10 min. The cells were then transferred to a fresh tube and washed twice with PBS before performing the cytospin protocol as described above.

Bone marrow transplantation. For competitive bone marrow transplantation, 6-week-old B6.SJL (CD45.1) recipient mice were irradiated with a total dose of 900 rad. Bone marrow was collected from uninduced donor mice (rtTA control, H3K9M or H3K36M) and a competitor mouse (B6.SJL (CD45.1)) as described above and then counted. Each recipient received 500,000 bone marrow cells from the donor mouse and 500,000 cells from the competitor mouse. The cells were suspended in 200 µl of PBS and injected retro-orbital into the mice. The mice were then housed for 11 weeks to ensure complete engraftment before induction with doxycycline (2 mg ml^{-1} in drinking water).

For leukaemia transplants, we used 6-week-old female B6.SJL (CD45.1) mice as recipients, irradiated at a dose 900 rad. Each recipient was injected with 1 million bone marrow cells. The bone marrow cells were previously frozen in 10% dimethylsulfoxide and thawed by reconstituting in 20% FBS in PBS. The bone marrow cells were then washed twice with PBS before injection into recipient mice. Each recipient received 500,000 bone marrow cells from the donor mouse and 500,000 cells from the competitor mouse (B6.SJL (CD45.1)). The cells (in 200 µl of PBS) were injected retro-orbital into the mice.

Quantitative RT–PCR. For quantitative RT–PCR (qRT-PCR), ES cells and EBs were first pelleted and lysed using buffer RLT (Qiagen). Crude lysate was passed through a QIAshredder column (Qiagen) and RNA was purified using a RNeasy plus mini kit (Qiagen) using the provided protocol. cDNA was generated using a High-Capacity RNA-to-cDNA kit using the provided protocol (Applied Biosystems). Quantitative PCR was performed on a lightcycler 480 system (Roche) using Brilliant III Sybr-Green (Agilent) and cDNA from each sample in triplicate. The efficiencies of each primer set used was found to be >90%, and fold-change was calculated using the $\Delta\Delta C_t$ method. The following primers were used:

Pou5f1 F: GGGGCTGTATCCTTTCCTCT;
Pou5f1 R: GCTGGTGCCTCAGTTTGAAT;
Nanog F: TGATTCAGAAGGGCTCAGCA;
Nanog R: CTTCCAGATGCGTTCACCAG;
Sox2 F: AGTGGTACGTTAGGCGCTTC;
Sox2 R: TGGACATTTGATTGCCATGT;
Flk1 F: GCTCCTGACTACACTACCCC;
Flk1 R: CCCAAATGCTCCACCAACTC;
Nestin F: GCCACTGAAAAGTTCCAGCT;
Nestin R: AGGGACATCTTGAGGTGTGC;
Gata6 F: TTCTACACAAGCGACCACCT;
Gata6 R: TTGAGGTCACTGTTCTCGGG;
Gapdh F: AGGTCGGTGTGAACGGATTTG;
Gapdh R: TGTAGACCATGTAGTTGAGGTCA;
Rex1 F: GTTCGTCCATCTAAAAAGGG;
Rex1 R: TAGTCCATTTCTCTAATGCCC;
Sox17 F: CTTGGAAGGCGTTGACCTTG;
Sox17 R: ACTTGTAGTTGGGGTGGTCC;
Sox7 F: ATGAGAGGAAACGTCTGGCA;
Sox7 R: CCTTCCATGACTTTCCCAGC;
Eomes F: TCGTGGAAGTGACAGAGGAC;
Eomes R: AGCTGGGTGATATCCGTGTT.

Western blotting. Cells grown in culture were prepared for western blotting by nuclear isolation. First, 400 µl of cold nuclear isolation buffer (50 mM Tris-HCl pH 8, 15 mM NaCl, 60 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 250 mM sucrose, 1 mM dithiothreitol, 0.6% IGEPAL (all from Sigma-Aldrich)) supplemented with complete mini protease inhibitors (Roche) was gently added to 1×10^7 cells. The suspension was incubated for 5 min on ice. Nuclei were then centrifuged at 960×*g* for 5 min. The resulting pellet was gently washed twice in nuclear isolation buffer and lysed in RIPA buffer (50 mM Tris-HCL (pH 8), 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100 and 1 mM EDTA (all from Sigma-Aldrich) supplemented with complete mini protease inhibitors (Roche) and 0.01 U µl^{-1} benzonase (Novagen). Sorted LKS cells were centrifuged for 5 min at 350×*g* and lysed directly in RIPA buffer.

The lysates were then incubated on ice for 10 min and sonicated 3 times for 30 s using a Branson tip sonicator with a 30-s pause between pulses. The lysates were then cleared for 5 min at maximum speed in a bench-top centrifuge (Eppendorf) to remove cell debris. The supernatant was boiled together with Laemmli sample buffer (100 mg ml^{-1} SDS, 250 mM Tris pH 6.8, 1 mg ml^{-1} bromophenol blue and 50% glycerol (all from Sigma-Aldrich) and loaded into 4–20% mini-Protean TGX precast protein gels (Bio-Rad). Protein was transferred to nitrocellulose membranes (Bio-Rad) and blocked for 1 h in 5% powdered milk in Tris-buffered saline and Tween 20. The following primary antibodies were used: H3 (Abcam, 1791; 1:10,000 dilution); H3K9me3 (Abcam, 8898; 1:1,000); H3K36me3 (Abcam 9050; 1:1,000); H3K36M (RevMAb Bioscience, 31-1085-00; 1:10,000); β-actin (Cell Signaling Technologies, 5125S; 1:10,000); H3K9M (Millipore, ABE579; 1:1,000); Lamin-B1 (Cell Signaling Technologies, 15068; 1:2,000), β-tubulin (Cell Signaling Technologies, 5346; 1:2,000); H3K9me1 (Abcam, 8896; 1:1,000); H3K9me2 (Abcam, 1220; 1:1,000); H3K36me2 (Millipore, 07-274; 1:1,000); H3K27me3 (Millipore, 07-449; 1:1,000). Goat, anti-rabbit-HRP-conjugated (Thermo Fisher Scientific, 31460; 1:2,000 dilution) was used as the secondary antibody. Proteins were detected using Immobilon western chemiluminescent HRP substrate (Millipore).

Pluripotency exit assay. We cultured mouse ES cells harbouring a Rex1-dGFP reporter (Rex1GFPd2)[25] in naive culture medium (DMEM/F12 medium and neurobasal medium (1:1 ratio), supplemented with 1× non-essential amino acids, 1 mM sodium pyruvate, 2 mM L-glutamine, 50 µM β-mercaptoethanol, N2 and B27 supplements (all from Life Technologies), 1 µM PD0325901 (Selleck Chemicals), 3 µM CHIR99021 (Selleck Chemicals) and leukaemia inhibitory factor (LIF) (1,000 U ml^{-1}, produced in-house). Exit from pluripotency was induced via EB formation as described above, and Rex1-dGFP expression was evaluated by flow cytometry. The shRNAs used to knockdown histone methyltransferases were obtained from the Molecular Profiling Laboratory of the MGH Cancer Center. Details of each construct are provided below:

G9a: TRCN0000301235;
Nsd1: TRCN0000123381;
Setd2: TRCN0000238536;
Setdb1: TRCN0000092975.

RNA-seq. RNA-seq libraries were constructed using a SMARTer v3 kit (Clontech) and sequenced on an Illumina HiSeq2500 instrument, resulting in approximately 30 million reads per sample on average. Transcriptome mapping was performed using STAR[54] with the Ensembl annotation of mm9 or mm10 (for day 7 analysis of HSPCs) reference genome. Read counts for individual genes were produced using HTSeq[55]. Differential expression analysis was performed using the package EdgeR[56] after normalizing read counts and including only those genes with count per million reads (CPM) of >1 for one or more samples. Differentially expressed genes were defined based on the criteria of >twofold change in expression value and FDR of <0.05. Enrichment of functional gene sets from a previously published study[57] was analysed on the ranked lists of whole-transcriptome expression values using GSEA[51] with default parameters. RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) repository with the accession numbers GSE103521 (day 7 timepoint for HSPCs) and GSE126829 (4 week timepoint for HSCPs, testes and intestine).

Teratoma assay. Approximately 2×10^6 cells were resuspended in 300 µl of growth medium and injected subcutaneously into the flanks of an NOD/SCID mouse (Jackson Laboratories). Host mice were maintained on doxycycline (2 mg ml^{-1}). Teratomas were monitored and removed once the tumour size reached 1 cm or tumours became ulcerated. Tumours were then weighed and processed for haematoxylin and eosin staining analysis.

In vitro maturation of proB cells. Hardy fraction C was sorted directly onto ST2 stromal cells in RPMI medium (Sigma) supplemented with 5% FBS (Hyclone), 1×

non-essential amino acids (Life Technologies), 1× Glutamax (Life Technologies) and 55 µM β-mercaptoethanol (Sigma). The cells were incubated for 4 days at 37 °C before analysis by flow cytometry as described above.

Reduced representation bisulfite sequencing. DNA was extracted from 100,000 cells using a Zymo Quick-DNA Universal kit following the manufacturer's instructions. Isolated genomic DNA was then quantified using a Qubit fluorometer and analysed for quality on a 1% agarose gel. A total of 10–60 ng of each sample was made into a reduced representation bisulfite sequencing (RRBS) library with a NuGen Ovation RRBS Methyl-seq kit, following the manufacturer's instructions, but using only half the recommended reaction volumes. The bisulfite conversion was carried out using a Qiagen Epitect Fast kit. The samples underwent 12 cycles of amplification post-conversion, and a 1× SPRI cleanup. Subsequent libraries were sequenced on an Illumina Hiseq2500.

Raw demultiplexed fastq reads were trimmed to remove Illumina adapters using cutadapt (–quality-cutoff 20 –overlap 3 –minimum-length 25), then using the Nugen python script trimRRBSdiversityAdaptCustomers.py. Trimmed reads were then aligned to the mm9 genome using bsmap (-v 0.1 -s 12 -q 20 -w 100 -S 1 -u -R). Methylation calling was performed using mcall with standard parameters. Bedtools was used to intersect the RRBS data with gene body locations (RefSeq, UCSC), and then R was used to compute the average methylation in control and mutant samples and generate scatter plots. The sequencing data for RRBS have been deposited in the GEO with the accession number GSE123676.

ATAC-seq. Approximately 60,000 cells were washed once with 100 µl PBS and suspended in 50 µl lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.2% IGEPAL CA-630). The suspension of nuclei was then centrifuged for 10 min at 500 × *g* at 4 °C, followed by the addition of 50 µl transposition reaction mix (25 µl TD buffer, 2.5 µl Tn5 transposase and 22.5 µl nuclease-free H_2O) and incubated at 37 °C for 30 min. DNA was isolated using a MinElute kit (Qiagen). Libraries were amplified by PCR (13 cycles). After the PCR reaction, the library was selected for fragments between 100 and 1,000 bp with AmpureXP beads (Beckman Coulter). Libraries were purified using a Qiaquick PCR kit (Qiagen) and their integrity was confirmed on a Bioanalyzer (Agilent) before sequencing. ATAC-seq libraries were sequenced on an Illumina HiSeq 2500 instrument in the paired-end 50 bp setting, resulting in approximately 40 million read pairs per sample on average. The reads were mapped to the mm9 reference genome using Burrows-Wheeler Alignment (BWA). Alignments were filtered for uniquely mapped non-mitochondrial reads and duplicates were removed. Peak calling was carried out using HOTSPOT, and peaks with less than 1 CPM in both replicates were removed. As a result, we identified 40,000–80,000 peaks per sample, which showed high consistency between biological duplicates. The union of these peak sets was used to calculate the ATAC-seq tag density over each peak region across all samples. Differentially accessible peaks were identified using edgeR, with the cut-offs of at least 1.5-fold difference and FDR of <0.01. The data have been deposited in the GEO repository with the accession number GSE126829.

ChIP-seq. Cultured cells were fixed in 1% formaldehyde and stored at −80 °C. Fixed cells were lysed in 1% SDS buffer (50 mM Tris-HCl, 1% SDS, 0.25% sodium deoxycholate (DOC), protease inhibitors (Roche)) and sheared (Branson 450) for 4.5 min with 45% amplitude on cold block (4 °C). The sheared chromatin was incubated overnight at 4 °C with the following antibodies: H3K36me3 (Abcam, ab9050), H3K27me3 (Millipore, 07-449) and H3K9me3 (Abcam, ab8898). Antibody-bound DNA was captured using protein G magnetic Dynabeads (Life Technologies) and washed with RIPA buffer (0.1% SDS, 0.1% DOC, 1% Triton X-100, 1 mM EDTA, 10 mM Tris-HCl pH 8.1, 150 mM NaCl), high-salt RIPA buffer (0.1% SDS, 0.1% DOC, 1% Triton X-100, 1 mM EDTA, 10 mM Tris-HCl pH 8.1, 500 mM NaCl), LiCl buffer (10 mM Tris-HCl pH 8.1, 250 mM LiCl, 0.5% Triton X-100, 0.5% DOC) and PBS. Eluted DNA was incubated overnight at 65 °C to reverse crosslinking and treated with RNase A followed by proteinase K. Bead-purified ChIP DNA was used to generate sequencing libraries by end repair, A base addition and ligation of barcoded sequencing adapters (KAPA HyperPrep kit, Roche). Barcoded fragments were amplified by PCR. Libraries were sequenced (paired-end) using a HiSeq2500 instrument.

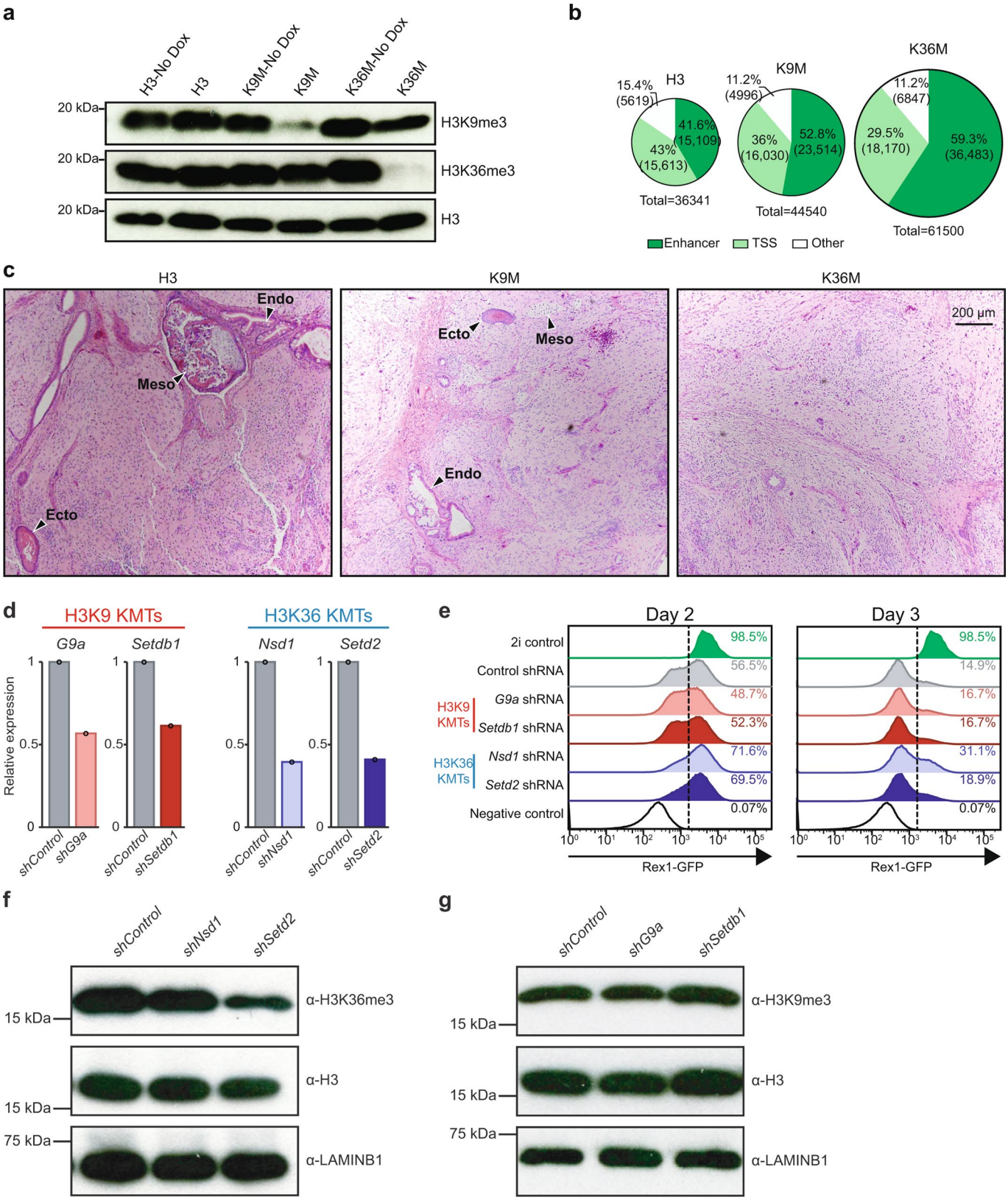
For sequence analysis, reads were aligned to mm9 using BWA 0.7.4, and multi-mapped reads and PCR duplicates were removed (Picard Tools). Peak calling was performed using the package Homer (v.4.8) with matched inputs. The following peak calling parameters were used: H3K4me3 and H3K27me3 and H3K9me3, –style histone, –q 0.1 –f 1.5; broad range of H3K9me3, -region –size 50000 –minDist 1000. For merged replicate peaks, we used mergePeaks with –d given. ChIP tracks were converted to a tiled data file (TDF) by igvtools count, and then visualized using Integrative Genomics Viewer (http://www.broadinstitute.org/igv/). To detect differential peaks in mutant samples (K36M and K9M) versus wild type, we merged replicate peaks and calculated enrichment values for each peak, then scored as loss or gained if the values pass 1.5-fold enrichment and 10% FDR cut-off (DESeq2). The following differential peak calling parameters were used: getDifferentialPeaksReplicates.pl –t K36M (K9M), –b WT –DESeq2 –q 0.1 –f 1.5 (Homer). Nearby genes from the peaks were annotated based on the distance to nearest TSS of RefSeq from peak files. Read density heatmaps were calculated with normalized signal by total number of signal, across distance from the centre of the peak using the data matrix generated by annotatePeaks.pl –ghist (Homer), and then visualized by cluster 3.0 and TreeView. Average read density was quantified for the histogram with centred peaks, and for meta-gene profiles with upstream, gene body (0–100%, RefSeq) and downstream of gene clusters using makeMetaGeneProfile.pl rna mm9 (Homer). To determine broad regions of H3K9me3 ChIP-seq tag enrichment, we analysed tag counts in a 100-kb window sliding with a 50-kbp step across the chromosome length, and calculated the ratio of ChIP tag count to input count in each window. Regions of H3K9me3 enrichment were generated by merging adjacent enriched windows with ChIP over input ratio of >1.5 that were separated by 100-kbp or less. ChIP-seq data are available in the GEO repository with the accession number GSE125703.

Statistics and reproducibility. Statistics tests are described in the corresponding figure legends. Unless otherwise indicated, statistical analyses were carried out using GraphPad Prism. For Fig. 4a,c,e, data are representative of three independent experiments. For Fig. 4h, data are representative of two biological replicates. For Fig. 5g,k,n, data are representative of two biological replicates. For Extended Data Fig. 6c, data are representative of two biological replicates. For Extended Data Fig. 6l,m,q,r, data are representative of two biological replicates. For Extended Data Fig. 1a,c,e–g, data are representative of three independent experiments. For Extended Data Fig. 2a,d,e,h, data are representative of three independent experiments. For Extended Data Fig. 4b–e, data are representative of three independent experiments. For Extended Data Fig. 7a–e, data are representative of three independent experiments. For Extended Data Fig. 8a,d,f,h,j,l, data are representative of three independent experiments. All other representative data are described in the figure legends.

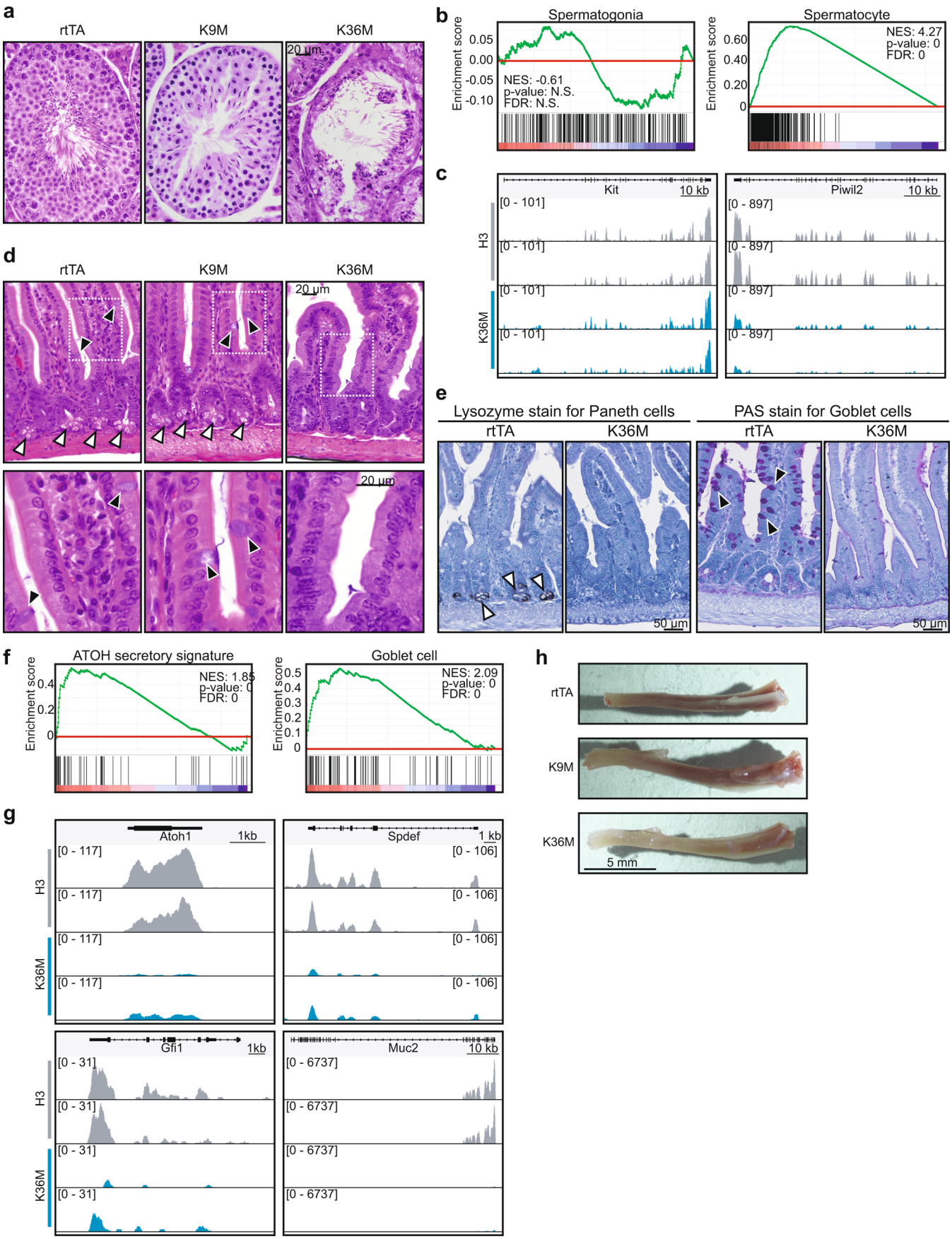
Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

References

52. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl Acad. Sci. USA* **90**, 8424–8428 (1993).
53. Eggan, K. et al. Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. *Proc. Natl Acad. Sci. USA* **98**, 6209–6214 (2001).
54. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
55. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
56. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
57. Lo, Y. H. et al. Transcriptional regulation by ATOH1 and its target SPDEF in the intestine. *Cell. Mol. Gastroenterol. Hepatol.* **3**, 51–71 (2017).
58. Green, C. D. et al. A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Dev. Cell* **46**, 651–667.e10 (2018).
59. Haber, A. L. et al. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).

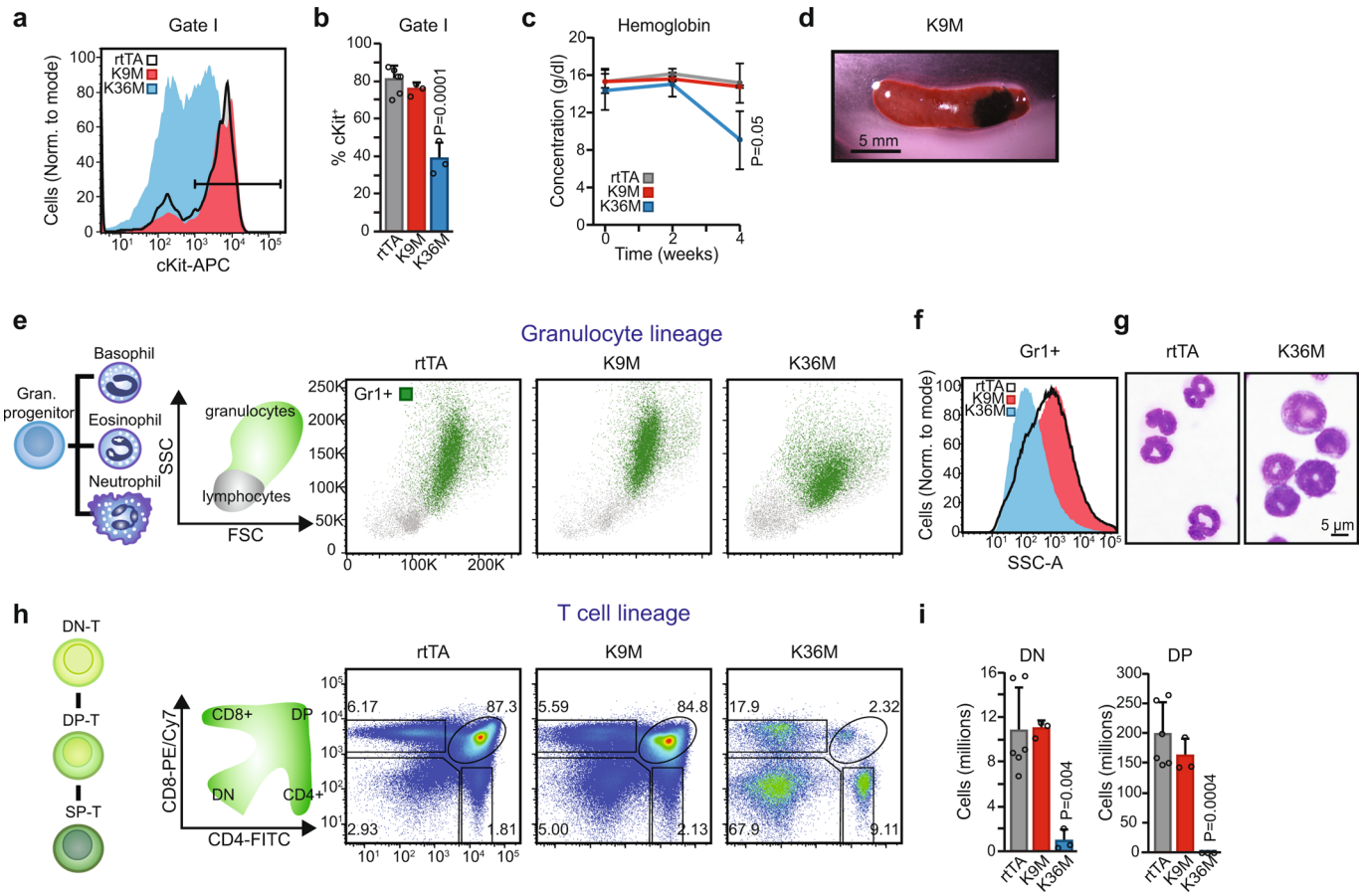


Extended Data Fig. 1 | Expression of H3K9M and H3K36M suppresses methylation and blocks differentiation of pluripotent stem cells. (**a**) Western blot analysis for tri-methyl H3K9 and H3K36 with and without mutant histone expression in mouse ES cells. (**b**) Pie charts showing the distribution of ATAC-seq peaks in EBs expressing H3, H3K9M, and H3K36M. (**c**) Images of teratomas derived from ES cells expressing H3, H3K9M, or H3K36M. Scale bar = 200 μm; endo = endoderm, meso = mesoderm, ecto = ectoderm. (**d**) qRT-PCR analysis for knockdown of H3K9 and H3K36 KMTs. One sample is shown. (**e**) Flow cytometry for Rex1-GFP at days 2 and 3 of embryoid body formation. (**f**) Western blot analysis showing H3K36me3 levels following knockdown of H3K36 KMTs. (**g**) Western blot analysis showing H3K9me3 levels following knockdown of H3K9 KMTs. See source data for full membrane Western blot images. Data in a,c,e,f,g are representative of 3 independent experiments.

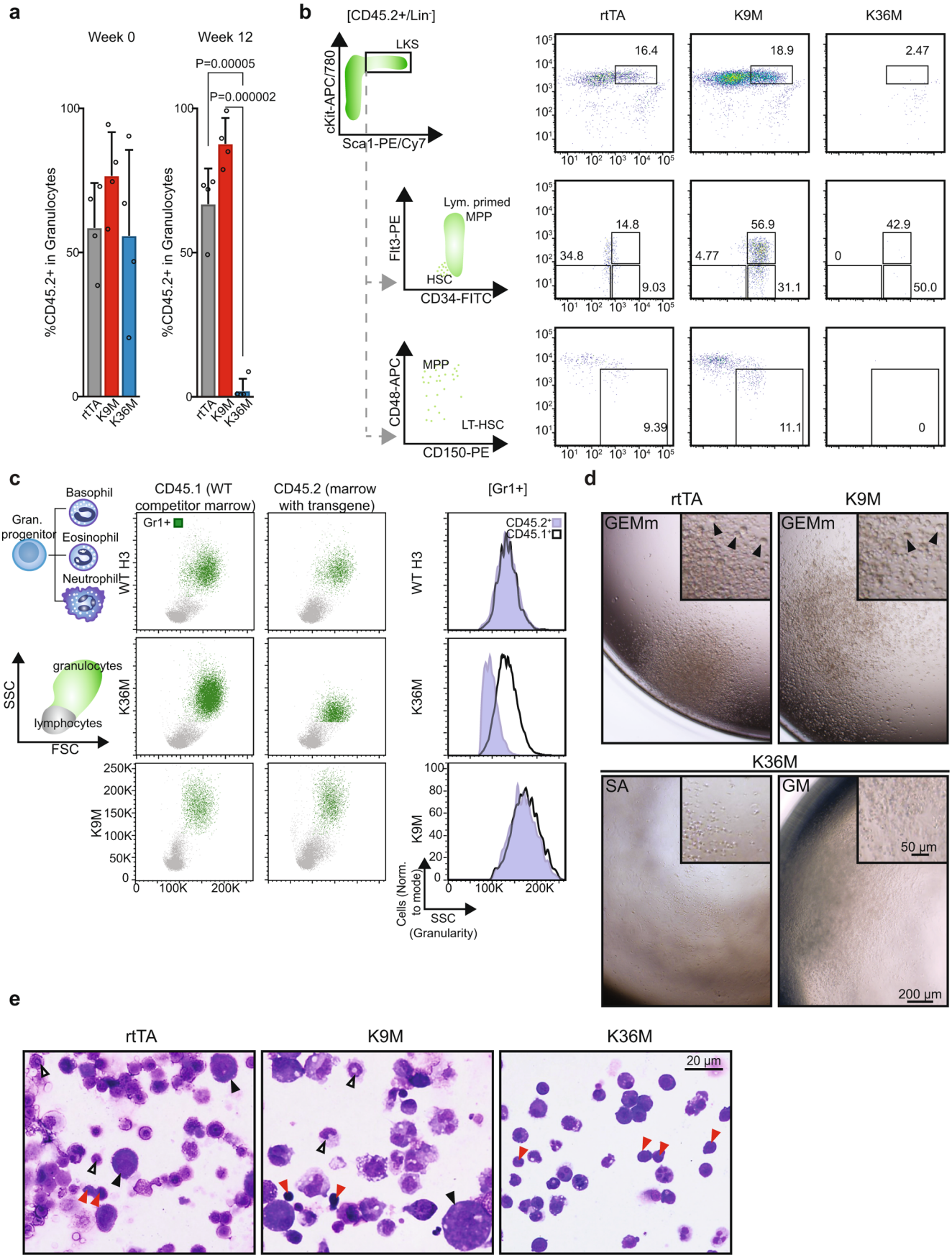


Extended Data Fig. 2 | see figure caption on next page.

Extended Data Fig. 2 | *In vivo* expression of H3K9M and H3K36M disrupts tissue homeostasis. (**a**) Histological analysis of testis in mice induced for 4 weeks. Scale bar = 20 µm. (**b**) GSEA based on RNA-seq data from testes isolated from H3K9M and H3K36M mice induced for 4 weeks (H3, n = 2; H3K9M, n = 2; H3K36M, n = 2). Enrichment is shown for transcriptional signatures[58] related to spermatogenesis. Statistics were generated in accordance with the published GSEA algorithm[51] (**c**) Gene tracks showing expression for genes characteristic of spermatogonia (left panel) or spermatocytes (right panel). (**d**) Histological analysis of intestine in mice induced for 4 weeks. Black and white arrowheads indicate goblet cells and paneth cells, respectively. Scale bar = 20 µm. (**e**) Staining for markers of Paneth cells (Lysozyme; left panel) and goblet cells (periodic acid Schiff; right panel) (**f**) GSEA based on RNA-seq data from intestine isolated from H3K9M and H3K36M mice induced for 4 weeks (H3, n = 2; H3K9M, n = 2; H3K36M, n = 2). Enrichment is shown for transcriptional signatures related to secretory lineages[57] and Goblet cells[59]. Statistics were generated in accordance with the published GSEA algorithm[51]. (**g**) Gene tracks showing expression for genes characteristic of secretory lineages (left panel) or goblet cells (right panel). (**h**) Tibia from mice expressing mutant histones for 4 weeks. Scale bar = 5 mm. Data in a,d,e,h are representative of 3 independent experiments.

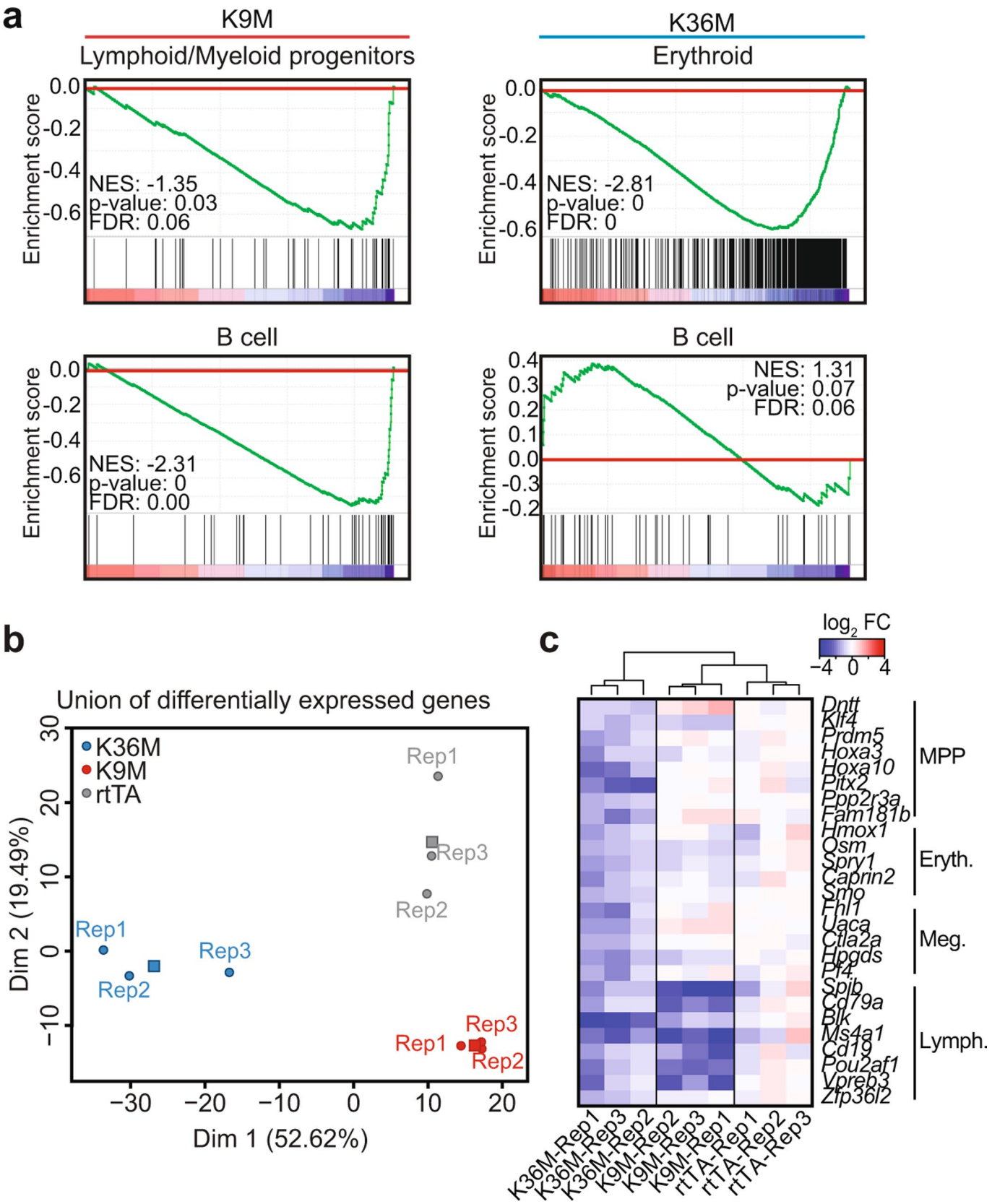


Extended Data Fig. 3 | H3K9M and H3K36M induction cause abnormalities in the erythroid, megakaryocyte, granulocyte and T cell lineages. (**a**) cKit expression in early erythroid progenitors. Data are representative of three independent experiments. (**b**) Quantification of cKit positive erythroid progenitors (right panel). Columns represent the mean and error bars represent standard deviation of the mean for biological replicates (rtTA, n = 6; H3K9M, n = 3; H3K36M, n = 3). Statistical significance was determined using a two-tailed unpaired Student's t-test. (**c**) Time course analysis for hemoglobin assessed by CBC. Columns represent the mean and error bars represent standard deviation of the mean for biological replicates (rtTA, n = 5; H3K9M, n = 4; H3K36M, n = 3). Statistical significance was determined using the Holm-Sidak method. (**d**) An infarcted spleen from a H3K9M mouse induced for 4 weeks. Scale bar = 5 mm. Infarctions were present in 2 out of 3 mice inspected. (**e**) Flow cytometry analysis for granulocyte cells. Gr1$^+$ cells are shown in green. (**f**) Granularity of transgenic cells assessed by side scatter. (**g**) Cytospin images of Gr1$^+$ granulocytes sorted from H3K36M and rtTA mice. Scale bar = 5 µm. (**h**) Flow cytometry analysis for thymocytes. Frequencies for each cell type are indicated as a percentage of the parent gate. (**i**) Quantification of CD4/CD8 double negative thymocytes (left panel) and double positive CD4/CD8 thymocytes (right panel). Columns represent the mean and error bars represent standard deviation of the mean for biological replicates (rtTA, n = 6; H3K9M, n = 3; H3K36M, n = 3). Statistical significance was determined using a two-tailed unpaired Student's t-test. Data in e,f,g,h are representative of 3 independent experiments.

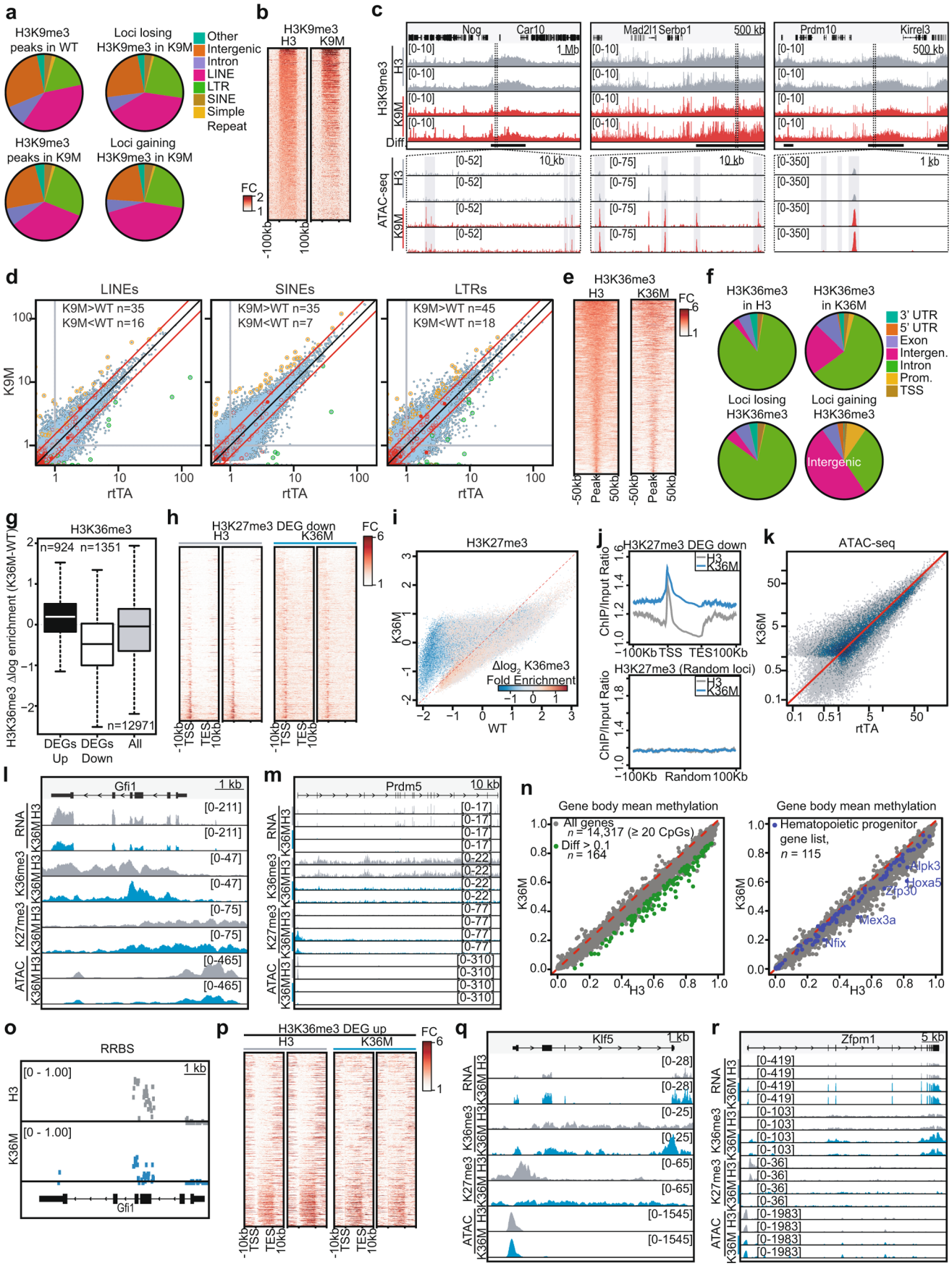


Extended Data Fig. 4 | see figure caption on next page.

Extended Data Fig. 4 | Mutant histone expression impacts HSC function in competitive bone marrow transplantation assays. (**a**) Frequency of granulocytes derived from donor bone marrow ($CD45.2^+$) harboring control or mutant histones before and after 12 weeks of induction with doxycyline. Columns represent the mean and error bars represent standard deviation of the mean for biological replicates (n = 4). Statistical significance was determined using a two-tailed unpaired Student's t-test. (**b**) Flow cytometry analysis for donor-derived, $CD45.2^+$ LKS cells, hematopoietic progenitor cells, and stem cells at the conclusion of competitive bone marrow transplant assays. Frequencies for each cell type are indicated as a percentage of the parent gate. (**c**) Flow cytometry showing granularity of both wild-type competitor marrow (CD45.1) and marrow expressing histone transgenes (CD45.2) as assessed by side scatter. (**d**) Brightfield images of colonies in methylcellulose. Black arrowheads indicate megakaryocytes. Scale bar = 200 μm or 50 μm. (**e**) May Grünwald Giemsa stains of cells grown in methylcellulose. Representative images are shown for rtTA, H3K9M, and H3K36M. Scale bar = 20 μm. Black arrowheads indicate megakaryocytes, white arrowheads indicate granulocytes, and red arrowheads indicate erythroid cells. Data in b,c,d,e are representative of 3 independent experiments.

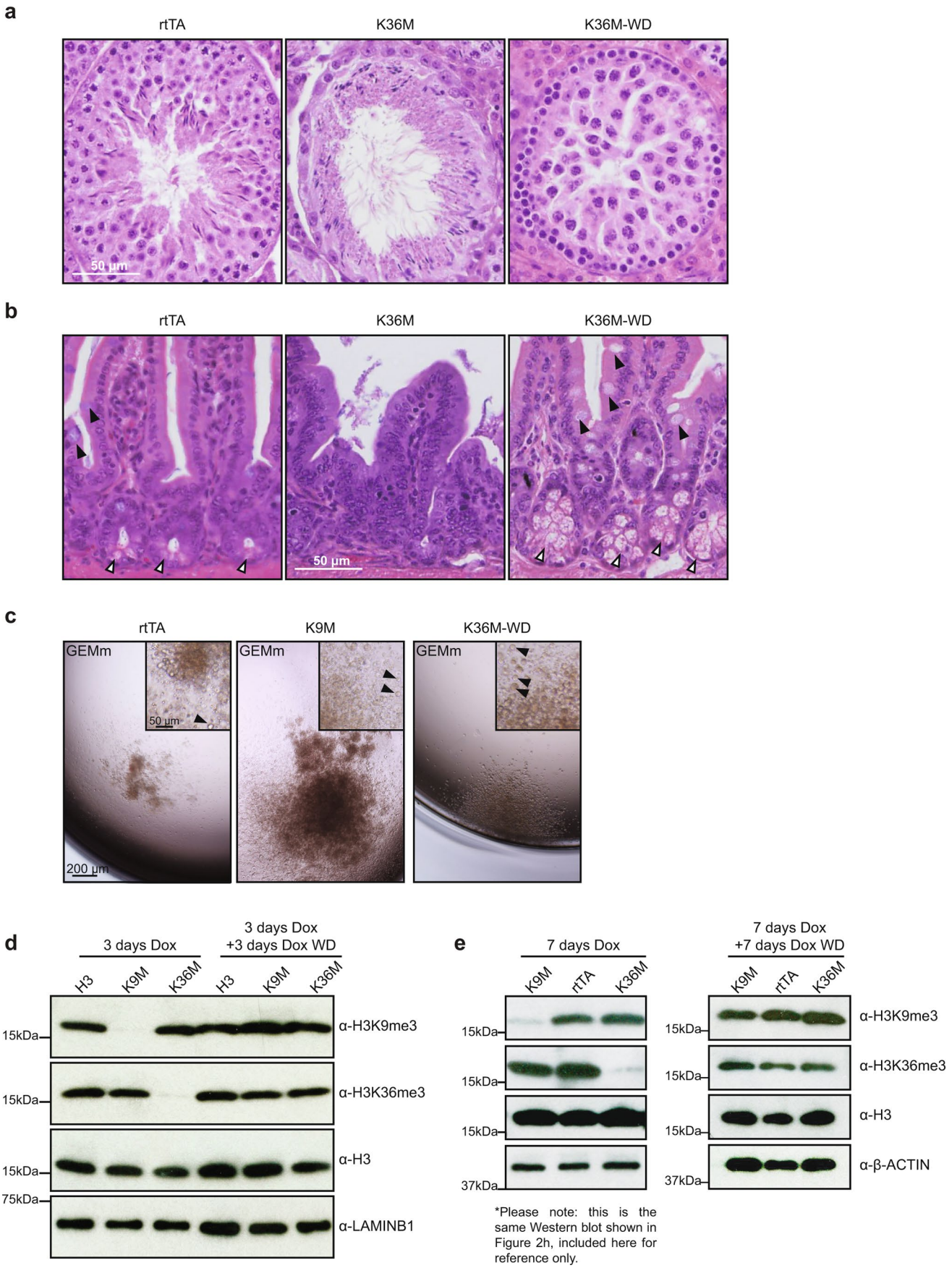


Extended Data Fig. 5 | Mutant histone expression impacts HSC gene expression. (**a**) GSEA analysis based on RNA-seq data for hematopoietic stem and progenitor cells sorted from H3K9M and H3K36M mice induced for 4 weeks (H3, n = 2; H3K9M, n = 2; H3K36M, n = 2). Enrichment is shown for transcriptional signatures related to Lymphoid/Myeloid progenitors (CMP, LMPP, and proB)[33] and B cells[50] (K9M) as well as erythroid cells[33] and B cells[50] (K36M). Statistics were generated in accordance with the published GSEA algorithm[51]. (**b**) Principal component analysis based on the union of differentially expressed genes between rtTA control and H3K9M or H3K36M samples. Each circle represents a biological replicate (rtTA, n = 3; H3K9M, n = 3; H3K36M, n = 3) while squares represent the mean. (**c**) Expression (RPKM) values for select genes relative to average expression in rtTA replicates (shown in $\log_2$ scale) for biological triplicates from LKS cells sorted at day 7 of induction.



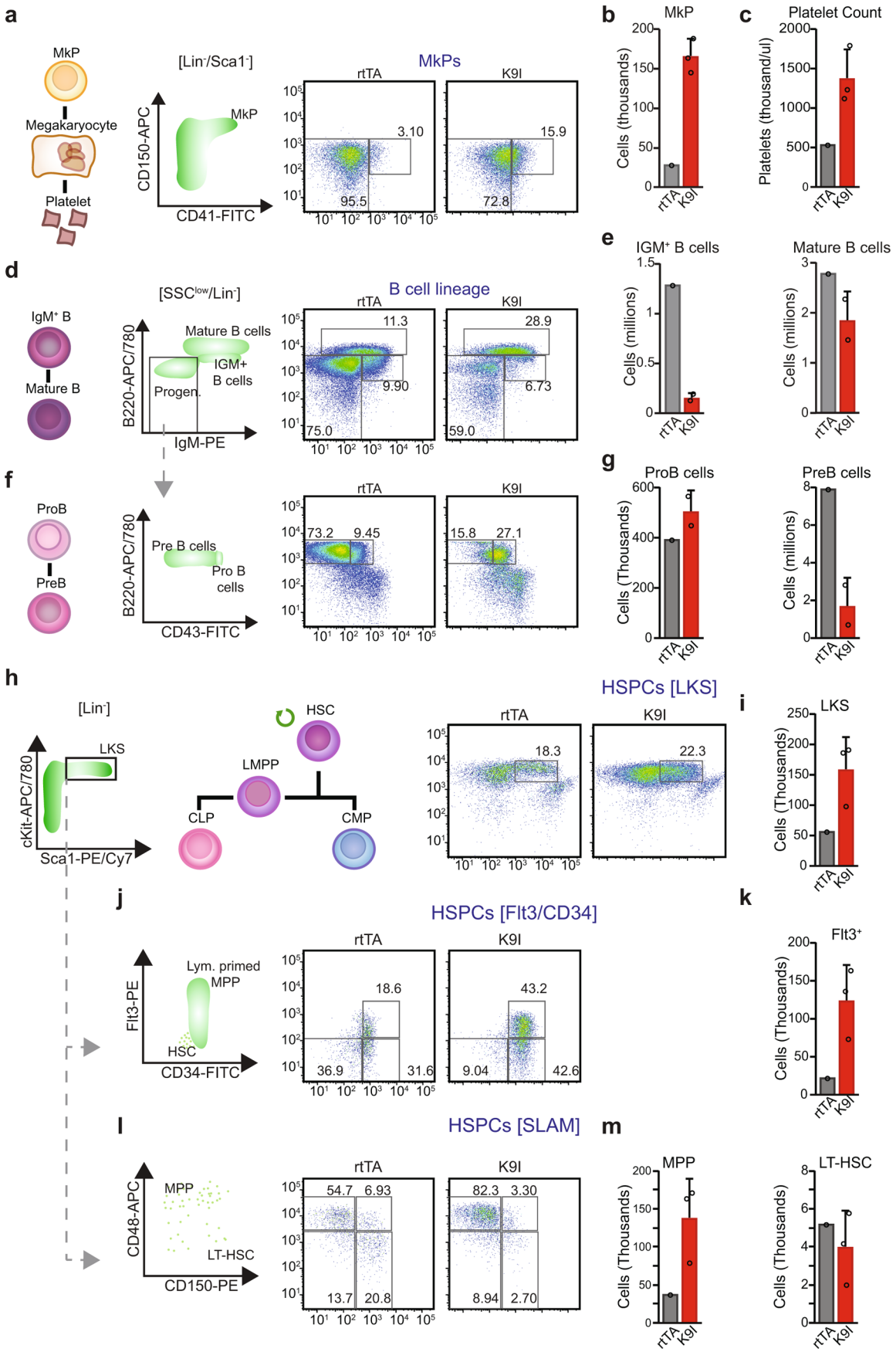
Extended Data Fig. 6 | see figure caption on next page.

Extended Data Fig. 6 | H3K36M/H3K9M induction change the chromatin landscape in HSPCs. (**a**) Distribution of H3K9me3 across genomic elements. (**b**) Heat-maps showing H3K9me3 methylation levels genome-wide following four weeks of H3 or H3K9M induction (averages, two replicates). (**c**) Representative tracks for loci with decreased H3K9me3 signal following H3K9M expression. (**d**) RNA-seq scatter plot based on RPKM. Red lines marks 1.5-fold difference between WT and K9M, gray line marks RPKM = 1 (cutoff for expressed repeats). Differentially expressed repeats (1.5-fold, FDR < 0.1; RNA-Seq data) are marked in orange and green dots. Red circles mark loci where H3K9me3 is differentially bound. Repeats overlapping protein-coding exons were excluded. (**e**) Heat-maps showing H3K36me3 methylation levels genome-wide following four weeks of H3 or H3K36M induction (averages, two replicates). (**f**) Distribution of H3K36me3 across genomic elements. (**g**) Levels of H3K36me3 in genes with increased expression, decreased expression, and all genes in H3K36M vs control HSPCs. The center bar represents the mean and the whiskers represent the standard deviation of the mean. (**h**) Replicate heat-maps showing H3K27me3 methylation for genes that are downregulated following H3K36M induction. (**i**) Correlation plot comparing H3K27me3 levels in H3K36M vs. H3 HSPCs. Levels of H3K36me3 overlaid on data points. (**j**) Meta-analysis showing H3K27me3 levels for a 200 kb region surrounding genes that are downregulated with H3K36M expression (top) and at random regions (bottom). (**k**) ATAC-seq scatter plot. Loci with reduced H3K36me3 following H3K36M expression are in blue. (**l**) Replicate tracks for *Gfi1* in Fig. 5g. (**m**) Representative tracks for *Prdm5*, which is downregulated following H3K36M expression. (**n**) Correlation plots comparing gene body methylation between H3 or H3K36M HSPCs. Significant differences in green (left), hematopoietic genes in blue (right). (**o**) Gene tracks for *Gfi1* showing DNA methylation in H3 or H3K36M HSPCs. (**p**) Replicate heat-maps showing H3K36me3 methylation at gene bodies for genes upregulated in H3K36M HSPCs. (**q**) Replicate tracks for *Klf5* in Fig. 5h. (**r**) Representative tracks for *Zfpm1*, which is upregulated following H3K36M expression. N numbers in d represent repetitive elements. N numbers in g,n represent genes. All n numbers are stated in the figure.



Extended Data Fig. 7 | see figure caption on next page.

Extended Data Fig. 7 | Reversibility of K-to-M dependent phenotypes. (**a**) Histological analysis of testis in mice induced for 4 weeks with an additional 4 weeks of dox withdrawal. Scale bar = 50 μm. (**b**) Histological analysis of intestine in mice induced for 4 weeks with an additional 4 weeks of dox withdrawal. Black and white arrowheads indicate goblet cells and paneth cells, respectively. Scale bar = 50 μm. (**c**) Brightfield images of colonies in methylcellulose with dox removal. Black arrowheads indicate megakaryocytes. Scale bar = 200 μm or 50 μm. (**d**) Western blot analysis for tri-methyl H3K9 and H3K36 from EBs induced for 3 days with dox and EBs induced for 3 days with dox followed by 3 days of dox withdrawal. (**e**) Western blot analysis for H3K9me3 and H3K36me3 from HSPCs in mice induced for 7 days with dox and mice induced for 7 days with dox followed by 7 days of dox withdrawal. See source data for full membrane Western blot images. Data in a-e are representative of 3 independent experiments.



Extended Data Fig. 8 | see figure caption on next page.

Extended Data Fig. 8 | K9I mice phenocopy K9M mice. (**a**) Flow cytometry analysis of megakaryocyte progenitors (MkPs) cells, 4 weeks of dox. Frequencies as percentages of the parent gate. (**b**) Quantification of MkPs in bone marrow. Columns represent the mean +/− standard deviation (rtTA, n = 1; H3K9I, n = 3, biological replicates). (**c**) Platelet counts after 4 weeks of dox. Columns represent the mean +/− standard deviation (rtTA, n = 1; H3K9I, n = 3, biological replicates). (**d**) Flow cytometry of B cells in the bone marrow, 4 weeks of dox. Frequencies as percentages of the parent gates. (**e**) Quantification of mature B cells and IGM^+ B cells. Columns represent the mean +/− standard deviation (rtTA, n = 1; H3K9I, n = 2, biological replicates). (**f**) Flow cytometry analysis of B cell progenitors in bone marrow, 4 weeks of dox. Frequencies as percentages of parent gates. (**g**) Quantification of ProB cells and PreB cells. Columns represent the mean +/− standard deviation (rtTA, n = 1; H3K9I, n = 2, biological replicates). (**h**) Flow cytometry analysis of LKS cells. Frequencies as percentages of parent gates. (**i**) Quantification of LKS cells in bone marrow. Columns represent the mean +/− standard deviation (rtTA, n = 1; H3K9I, n = 3, biological replicates). (**j**) Flow cytometry of stem cell populations gated on LKS cells. Frequencies indicated as percentages of parent gates. (**k**) Quantification of LMPPs, defined as $Flt3^{high}$ cells (top 25% of $Flt3^+$ cells set on rtTA control). Columns represent means and error bars standard deviation (rtTA, n = 1; H3K9I, n = 3, biological replicates). (**l**) Flow cytometry of SLAM markers gated on LKS cells. Frequencies as percentages of parent gates. (**m**) Quantification of MPPs (left panel) and long-term HSCs (right panel) in bone marrow. Columns represent the mean +/− standard deviation (rtTA, n = 1; H3K9I, n = 3 biological replicates). Data in a,d,f,h,j,l are representative of 3 independent experiments.

nature research

Corresponding author(s): Hanno Hock and Konrad Hochedlinger

Last updated by author(s): Aug 29, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For flow cytometry, cells were analyzed on an LSRII flow cytometer (BD Biosciences) or collected on an Aria II flow cytometer (BD Biosciences) using Diva v6.1.2 (BD Biosciences). Quantitative PCR was performed on a lightcycler 480 system (Roche) using Brilliant III Sybr-Green (Agilent) and cDNA from each sample in triplicate. The efficiencies of each primer set used was found to be >90%, and fold change was calculated using the ΔΔCt method. RNA-Seq libraries were constructed using the SMARTer v3 kit (Clontech) and sequenced on Illumina HiSeq2500 instrument, resulting in approximately 30 million reads per sample on average. For Reduced Representation Bisulfite Sequencing (RRBS), libraries were sequenced on an Illumina Hiseq2500. ATAC-seq libraries were sequenced on Illumina HiSeq 2500 instrument in the paired-end 50 bp setting, resulting in approximately 40 million read pairs per sample on average. For ChIP-seq, libraries were sequenced (paired-end) with HiSeq2500 instrument.
Data analysis	For flow cytometry, data analysis was performed using FlowJo software v10.2 (FlowJo, LLC). For RNA-seq, Transcriptome mapping was performed with STAR59 using the Ensembl annotation of mm9 or mm10 (for day 7 analysis of HSPCs) reference genome. Read counts for individual genes were produced using HTSeq60. Differential expression analysis was performed using the EdgeR package61 after normalizing read counts and including only those genes with CPM > 1 for one or more samples. Differentially expressed genes were defined based on the criteria of >2-fold change in expression value and false discovery rate (FDR) < 0.05. Enrichment of functional gene sets from a previously published study53 was analyzed on the ranked lists of whole-transcriptome expression values using GSEA62 with default parameters. RNA-seq data is deposited in the Gene Expression Omnibus repository with the accession numbers GSE103521 (day 7 time point for HSPCs) and GSE126829 (4 week time point for HSCPs, testes, and intestine). For Reduced Representation Bisulfite Sequencing (RRBS), raw demultiplexed fastq reads were trimmed to remove Illumina adapters using cutadapt (--quality-cutoff 20 --overlap 3 --minimum-length 25), then using the Nugen python script trimRRBSdiversityAdaptCustomers.py. Trimmed reads were then aligned to the mm9 genome using bsmap (-v 0.1 -s 12 -q 20 -w 100 -S 1 -u -R). Methylation calling was performed using mcall with standard parameters. Bedtools was used to intersect the RRBS data with gene body locations (RefSeq, UCSC), and then R was used to compute the average methylation in control and mutant samples and generate scatter plots. The sequencing data for RRBS was deposited in the Gene Expression Omnibus with the accession number GSE123676. ATAC-seq reads were mapped to the mm9 reference genome using BWA. Alignments were filtered for uniquely mapped none-mitochondrial reads and duplicates were removed. Peak calling were carried out with HOTSPOT and peaks with less than 1 count per million reads (CPM) in both replicates were removed. As a result we

identified 40,000 – 80,000 peaks per sample, which showed high consistency between biological duplicates. The union of these peak sets was used to calculate the ATAC-seq tag density over each peak region across all samples. Differentially accessible peaks were identified using edgeR with the cutoffs of at least 1.5-fold difference and false discovery rate (FDR) < 0.01. The data is deposited in the Gene Expression Omnibus repository with the accession number GSE126829. For ChIP-seq, analysis, reads were aligned to mm9 using BWA 0.7.4, and multi-mapped reads and PCR duplicates were removed (Picard Tools). Peak calling was performed using the Homer package (HOMER v4.8) with matched inputs. Peak calling parameters: H3K4me3 and H3K27me3 and H3K9me3, –style histone, –q 0.1 –f 1.5; broad range of H3K9me3, -region –size 50000 –minDist 1000. For merged replicate peaks, we used mergePeaks with –d given. ChIP tracks were converted to a tiled data file (TDF) by igvtools count and then visualized using Integrative Genomics Viewer (IGV, http://www.broadinstitute.org/igv/). To detect differential peaks in mutant samples (K36M and K9M) versus WT, we merged replicate peaks and calculated enrichment values for each peak, then scored as loss or gained if the values pass 1.5 fold enrichment and 10% FDR cutoff (DESeq2). Differential peak calling parameters: getDifferentialPeaksReplicates.pl –t K36M (K9M), –b WT –DESeq2 –q 0.1 –f 1.5 (HOMER). Nearby genes from the peaks were annotated based on the distance to nearest TSS of RefSeq from peak files. Read density heat maps were calculated with normalized signal by total number of signal, across distance from the center of the peak using the data matrix generated by annotatePeaks.pl –ghist (HOMER), and then visualized by cluster 3.0 and TreeView. Average read density was quantified for the histogram with centered peaks, and for meta-gene profiles with upstream, gene body (0%-100%, RefSeq) and downstream of gene clusters using makeMetaGeneProfile.pl rna mm9 (HOMER). To determine broad regions of H3K9me3 ChIP-seq tag enrichment, we analyzed tag counts in a 100 Kb window sliding with a 50 Kbp step across the chromosome length, and calculated the ratio of ChIP tag count to input count in each window. Regions of H3K9me3 enrichment were generated by merging adjacent enriched windows with ChIP over input ratio > 1.5 that were separated by 100 Kbp or less. ChIP-seq data is available in the Gene Expression Omnibus repository with the accession number GSE125703. Unless otherwise indicated, statistical analysis was carried out using GraphPad Prism.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA-seq data are deposited in the Gene Expression Omnibus repository under the accession numbers GSE103521 (Supplementary Figure 5) and GSE126829 (Figure 5 and Supplementary Figures 2 and 5). RRBS data are deposited in the Gene Expression Omnibus repository under the accession number GSE123676 (Supplementary Figure 6). ATAC-seq data are deposited in the Gene Expression Omnibus repository under the accession number GSE126829 (Figures 1 and 5; Supplementary Figure 1 and 6). ChIP-seq data are deposited in the Gene Expression Omnibus repository under the accession number GSE125703 (Figure 5 and Supplementary Figure 6).

All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical method was used to determine sample size. Sample sizes were determined based on the authors' previous experience.
Data exclusions	No data were excluded from this manuscript.
Replication	All attempts at reproducing experiments reported here were successful. Reproducibility statements are included in the figure legends or methods section for each figure.
Randomization	Randomization was not applicable to this manuscript since no such comparisons (between sample groups) were made.
Blinding	Blinding was also not applicable since no comparisons between sample groups were made.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

AB info:
Primary antibodies used:
H3 (Abcam 1791, 1:10,000 dilution)
H3K9Me3 (Abcam, 8898, 1:1000 dilution)
H3K36Me3 (Abcam 9050, 1:1000 dilution)
H3K36M (RevMAb Bioscience 31-1085-00, 1:10,000 dilution)
β-ACTIN (Cell Signaling Technologies 5125S, 1:10,000)
H3K9M (Millipore ABE579, 1:1000)
LaminB1 (Cell Signaling Technologies 15068, 1:2000)
B-Tubulin (Cell Signaling Technologies 5346, 1:2000)
H3K9Me1 (Abcam 8896, 1:1000)
H3K9Me2 (Abcam 1220,1:1000)
H3K36Me2 (Millipore 07-274, 1:1000)
H3K27Me3 (Millipore 07-449, 1:1000)
Lysozyme antibody (Thermo Fisher PA5-16668, 1:100 dilution)

The secondary antibody used for Western analysis was:
goat, anti-rabbit-HRP-conjugated (Thermo Fisher Scientific 31460, 1:2000 dilution)

Anitbodies used for flow cytometry were:
CD4 (FITC; eBiosciences; MCD0401, clone RM4-5, dilution: 0.1 ul per 1 million cells in residual volume)
CD8a (PE/Cy7; clone 53-6.7; eBiosciences, 25-0081-82, dilution: 0.1 ul per 1 million cells in residual volume)
cKit (APC; Biolegend, 105812; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume)
CD25 (PE; eBiosciences, 12-0251-82, clone PC61.5, dilution: 0.1 ul per 1 million cells in residual volume)
CD45.2 (Pacific Blue; Biolegend, 109820; clone 104, dilution: 0.1 ul per 1 million cells in residual volume)
Fc block (BD Biosciences, 553141, clone 2.4G2, dilution: 0.1 ul per 1 million cells in residual volume)
B220/CD45R (PE/Cy5; eBiosciences, 15-0452-82; clone RA3-6B2, dilution: 0.1 ul per 1 million cells in residual volume)
Ter119 (PE/Cy5; eBiosciences, 15-5921-82; clone TER-119, dilution: 0.1 ul per 1 million cells in residual volume)
CD3e (PE/Cy5; eBiosceices, 15-0031-82; clone 145-2C11, dilution: 0.1 ul per 1 million cells in residual volume)
Mac1 (PE/Cy5; eBiosciences, 15-0112-82; clone M1/70, dilution: 0.1 ul per 1 million cells in residual volume)
Gr1 (PE/Cy5; eBiosciences, 15-5931-82; clone RB6-8C5, dilution: 0.1 ul per 1 million cells in residual volume)
Sca1 (PE/Cy7; eBiosciences, 25-5981-82; clone D7, dilution: 0.1 ul per 1 million cells in residual volume)
cKit (eFlour 780; Biolegend, 105802; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume)
CD41 (FITC; BD Biosciences, 553848; clone MWReg30, dilution: 0.1 ul per 1 million cells in residual volume)
CD16/CD32 (PE; BD Biosciences, 561727; clone 2.4G2, dilution: 0.1 ul per 1 million cells in residual volume)
CD150 (APC; Biolegend, 115910; clone TC15-12F12.2, dilution: 0.1 ul per 1 million cells in residual volume)
CD105 (Pacific Blue; Biolegend, 120412; clone MJ7/18, dilution: 0.1 ul per 1 million cells in residual volume)
cKit (APC; Biolegend, 105812; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume)
Ter119 (FITC; eBiosciences, 11-5921-82; clone TER-119, dilution: 0.1 ul per 1 million cells in residual volume)
CD71 (PE; EBiosciences, 12-0711-82; clone R17217 (RI7 217.1.4), dilution: 0.1 ul per 1 million cells in residual volume)
CD34 (FITC; eBiosciences, 14-0341-82; clone RAM34, dilution: 0.3 ul per 1 million cells in residual volume)
Flt3 (PE; eBioscience, 12-1351-82; clone A2F10, dilution: 0.1 ul per 1 million cells in residual volume) CD150 (PE; Biolegend, 115904; clone TC15-12F12.2, dilution: 0.1 ul per 1 million cells in residual volume)
CD48 (APC; Biolegend, 103412; clone HM48-1, dilution: 0.1 ul per 1 million cells in residual volume)
cKit (eFluor 780; eBiosciences, 47-1171-82; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume)
Gr1 (APC; BD Biosciences, 553129; clone RB6-8C5, dilution: 0.1 ul per 1 million cells in residual volume)
Mac1 (PE; BD Biosciences, 553311; clone M1/70, dilution: 0.1 ul per 1 million cells in residual volume)
IGM (PE/Cy5; eBiosciences, 15-5790-82; clone II/41, dilution: 0.1 ul per 1 million cells in residual volume)
CD249 (PE, BD Biosciences, 553735; clone BP-1, dilution: 0.1 ul per 1 million cells in residual volume)
CD24 (APC, Biolegend, 101814; clone M1/69, dilution: 0.1 ul per 1 million cells in residual volume)
CD43 (FITC; BD Biosciences, 553270; clone S7, dilution: 0.1 ul per 1 million cells in residual volume)
B220/CD45R (eFluor 780; eBiosciences, 47-0452-82; clone RA3-6B2, dilution: 0.1 ul per 1 million cells in residual volume)
IGM (PE; eBiosciences, 14-5790-82; clone II/41, dilution: 0.1 ul per 1 million cells in residual volume)

Validation

H3 (Abcam 1791, 1:10,000 dilution), validation/reference: https://www.abcam.com/histone-h3-antibody-nuclear-loading-control-and-chip-grade-ab1791.html
H3K9Me3 (Abcam, 8898, 1:1000 dilution), validation/references: https://www.abcam.com/histone-h3-tri-methyl-k9-antibody-

chip-grade-ab8898.html
H3K36Me3 (Abcam 9050, 1:1000 dilution), validation/references: https://www.abcam.com/histone-h3-tri-methyl-k36-antibody-chip-grade-ab9050.html
H3K36M (RevMAb Bioscience 31-1085-00, 1:10,000 dilution), validation/references: https://www.revmab.com/index.php/product/anti-histone-h3-k36m-rabbit-monoclonal-antibody-clone-rm193-histone-h3-k36m-mutant/
β-ACTIN (Cell Signaling Technologies 5125S, 1:10,000), validation/references: https://www.cellsignal.com/products/antibody-conjugates/b-actin-13e5-rabbit-mab-hrp-conjugate/5125
H3K9M (Millipore ABE579, 1:1000), validation/references: http://www.emdmillipore.com/US/en/product/Anti-Histone-H3.3-K9M,MM_NF-ABE579
LaminB1 (Cell Signaling Technologies 15068, 1:2000), validation/references: https://www.cellsignal.com/products/antibody-conjugates/lamin-b1-d9v6h-rabbit-mab-hrp-conjugate/15068
B-Tubulin (Cell Signaling Technologies 5346, 1:2000), validation/references: https://www.cellsignal.com/products/antibody-conjugates/b-tubulin-9f3-rabbit-mab-hrp-conjugate/5346
H3K9Me1 (Abcam 8896, 1:1000), validation/references: https://www.abcam.com/histone-h3-mono-methyl-k9-antibody-chip-grade-ab8896.html
H3K9Me2 (Abcam 1220,1:1000), validation/references: https://www.abcam.com/histone-h3-di-methyl-k9-antibody-mabcam-1220-chip-grade-ab1220.html
H3K36Me2 (Millipore 07-274, 1:1000), validation/references: http://www.emdmillipore.com/US/en/product/Anti-dimethyl-Histone-H3-Lys36-Antibody,MM_NF-07-274
H3K27Me3 (Millipore 07-449, 1:1000). validation/references:
Lysozyme antibody (Thermo Fisher PA5-16668, 1:100 dilution), validation/references: http://www.emdmillipore.com/US/en/product/Anti-trimethyl-Histone-H3-Lys27-Antibody,MM_NF-07-449
goat, anti-rabbit-HRP-conjugated (Thermo Fisher Scientific 31460, 1:2000 dilution), validation/references: https://www.thermofisher.com/order/genome-database/generatePdf?productName=Rabbit%20IgG%20(H+L)&assayType=PRANT&detailed=true&productId=31460
CD4 (FITC; eBiosciences; MCD0401, clone RM4-5, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD4-Antibody-clone-RM4-5-Monoclonal/MCD0401
CD8a (PE/Cy7; clone 53-6.7; eBiosciences, 25-0081-82, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD8a-Antibody-clone-53-6-7-Monoclonal/25-0081-82
cKit (APC; Biolegend, 105812; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/apc-anti-mouse-cd117-c-kit-antibody-72
CD25 (PE; eBiosciences, 12-0251-82, clone PC61.5, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD25-Antibody-clone-PC61-5-Monoclonal/12-0251-82
CD45.2 (Pacific Blue; Biolegend, 109820; clone 104, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/pacific-blue-anti-mouse-cd45-2-antibody-3108
Fc block (BD Biosciences, 553141, clone 2.4G2, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/applications/research/b-cell-research/surface-markers/mouse/purified-rat-anti-mouse-cd16cd32-mouse-bd-fc-block-24g2/p/553141
B220/CD45R (PE/Cy5; eBiosciences, 15-0452-82; clone RA3-6B2, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD45R-B220-Antibody-clone-RA3-6B2-Monoclonal/15-0452-82
Ter119 (PE/Cy5; eBiosciences, 15-5921-82; clone TER-119, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/TER-119-Antibody-clone-TER-119-Monoclonal/15-5921-82
CD3e (PE/Cy5; eBiosceices, 15-0031-82; clone 145-2C11, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD3e-Antibody-clone-145-2C11-Monoclonal/15-0031-82
Mac1 (PE/Cy5; eBiosciences, 15-0112-82; clone M1/70, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD11b-Antibody-clone-M1-70-Monoclonal/15-0112-82
Gr1 (PE/Cy5; eBiosciences, 15-5931-82; clone RB6-8C5, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/Ly-6G-Ly-6C-Antibody-clone-RB6-8C5-Monoclonal/15-5931-82
Sca1 (PE/Cy7; eBiosciences, 25-5981-82; clone D7, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/Ly-6A-E-Sca-1-Antibody-clone-D7-Monoclonal/25-5981-82
cKit (eFlour 780; Biolegend, 105802; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/purified-anti-mouse-cd117-c-kit-antibody-77
CD41 (FITC; BD Biosciences, 553848; clone MWReg30, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/applications/research/b-cell-research/surface-markers/mouse/fitc-rat-anti-mouse-cd41-mwreg30/p/553848
CD16/CD32 (PE; BD Biosciences, 561727; clone 2.4G2, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/applications/research/b-cell-research/surface-markers/mouse/pe-rat-anti-mouse-cd16cd32-24g2/p/561727
CD150 (APC; Biolegend, 115910; clone TC15-12F12.2, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/apc-anti-mouse-cd150-slam-antibody-2894
CD105 (Pacific Blue; Biolegend, 120412; clone MJ7/18, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/pacific-blue-anti-mouse-cd105-antibody-6080
cKit (APC; Biolegend, 105812; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/apc-anti-mouse-cd117-c-kit-antibody-72
Ter119 (FITC; eBiosciences, 11-5921-82; clone TER-119, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/TER-119-Antibody-clone-TER-119-Monoclonal/11-5921-82
CD71 (PE; EBiosciences, 12-0711-82; clone R17217 (RI7 217.1.4), dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD71-Transferrin-Receptor-Antibody-clone-R17217-RI7-217-1-4-Monoclonal/12-0711-82
CD34 (FITC; eBiosciences, 14-0341-82; clone RAM34, dilution: 0.3 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD34-Antibody-clone-RAM34-Monoclonal/14-0341-82
Flt3 (PE; eBioscience, 12-1351-82; clone A2F10, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD135-Flt3-Antibody-clone-A2F10-Monoclonal/12-1351-82
CD150 (PE; Biolegend, 115904; clone TC15-12F12.2, dilution: 0.1 ul per 1 million cells in residual volume), validation/references:

https://www.biolegend.com/en-us/products/pe-anti-mouse-cd150-slam-antibody-1369
CD48 (APC; Biolegend, 103412; clone HM48-1, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/apc-anti-mouse-cd48-antibody-3622
cKit (eFluor 780; eBiosciences, 47-1171-82; clone 2B8, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD117-c-Kit-Antibody-clone-2B8-Monoclonal/47-1171-82
Gr1 (APC; BD Biosciences, 553129; clone RB6-8C5, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/reagents/research/antibodies-buffers/immunology-reagents/anti-mouse-antibodies/cell-surface-antigens/apc-rat-anti-mouse-ly-6g-and-ly-6c-rb6-8c5/p/553129
Mac1 (PE; BD Biosciences, 553311; clone M1/70, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/applications/research/stem-cell-research/mesenchymal-stem-cell-markers-bone-marrow/mouse/negative-markers/pe-rat-anti-cd11b-m170/p/553311
IGM (PE/Cy5; eBiosciences, 15-5790-82; clone II/41, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/IgM-Antibody-clone-II-41-Monoclonal/15-5790-82
CD249 (PE, BD Biosciences, 553735; clone BP-1, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/applications/research/b-cell-research/surface-markers/mouse/pe-mouse-anti-mouse-cd249-l-51-bp-1/p/553735
CD24 (APC, Biolegend, 101814; clone M1/69, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.biolegend.com/en-us/products/apc-anti-mouse-cd24-antibody-2937
CD43 (FITC; BD Biosciences, 553270; clone S7, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: http://www.bdbiosciences.com/us/applications/research/b-cell-research/surface-markers/mouse/fitc-rat-anti-mouse-cd43-s7/p/553270
B220/CD45R (eFluor 780; eBiosciences, 47-0452-82; clone RA3-6B2, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/CD45R-B220-Antibody-clone-RA3-6B2-Monoclonal/47-0452-82
IGM (PE; eBiosciences, 14-5790-82; clone II/41, dilution: 0.1 ul per 1 million cells in residual volume), validation/references: https://www.thermofisher.com/antibody/product/IgM-Antibody-clone-II-41-Monoclonal/14-5790-82

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	KH2 ES cell lines were obtained from the Jaenisch lab and were previously reported and referenced in the text (Beard et al.).
Authentication	KH2 ES cells were authenticated in the referenced publication (Beard et al.).
Mycoplasma contamination	Cell lines were tested weekly for mycoplasma contamination using the LONZA Mycoplasma Detection Kit. All cell lines tested negative for Mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	KH2 ES cells have a B6/129 background. High-grade chimeras were crossed to B6/C57 wildtype mice. These mice were then crossed to mice carrying constitutive M2-rtTA in a B6/C57 background. Both male and females were used in this study at 6-8 weeks of age. For competitive bone marrow transplants, B6.SJL (CD45.1). DR4 and NOD/SCID mice were purchased from Jackson labs (https://www.jax.org/). DR4 males and females were mated and sacrificed at E13.5 to derive MEFs. NOD/SCID female mice were used for teratoma assays as described in the methods. Mice used in this study were housed and bred in Specific Pathogen Free (SPF) rooms located in the AAALAC-accredited Center for Comparative Medicine vivarium at Massachusetts General Hospital. Mice were housed in ventilated cages on a standard 12:12 light cycle. All procedures involving mice adhered to the guidelines of the approved Massachusetts General Hospital Institutional Animal Care and Use Committee (IACUC) protocol no. 2006N000104.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve animals collected from the field.
Ethics oversight	All procedures involving mice adhered to the guidelines of the approved Massachusetts General Hospital Institutional Animal Care and Use Committee (IACUC) protocol no. 2006N000104.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	To review GEO accession GSE125703: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125703 Enter token cpabgceklxitnav into the box
Files in database submission	GSM3580640 WT-rep1-H3K9me3 GSM3580641 WT-rep1-H3K27me3 GSM3580642 WT-rep1-H3K36me3 GSM3580643 WT-rep1-input GSM3580644 WT-rep2-H3K9me3 GSM3580645 WT-rep2-H3K27me3 GSM3580646 WT-rep2-H3K36me3 GSM3580647 WT-rep2-input GSM3580648 K9M-rep1-H3K9me3 GSM3580649 K9M-rep1-input GSM3580650 K9M-rep2-H3K9me3 GSM3580651 K9M-rep2-input GSM3580652 K36M-rep1-H3K27me3 GSM3580653 K36M-rep1-H3K36me3 GSM3580654 K36M-rep1-input GSM3580655 K36M-rep2-H3K27me3 GSM3580656 K36M-rep2-H3K36me3 GSM3580657 K36M-rep2-input
Genome browser session (e.g. UCSC)	We did not create a Genome browser session for this data.

Methodology

Replicates	ChIP experiments were performed in biological duplicate (each sample represents a separate mouse). All replicates were in good agreement and are represented in this manuscript.
Sequencing depth	Sequencing depth is listed for each sample below, with 95% of reads mapped to reference genome. All reads are 38bp pairend. Sample /Total Reads/Alignment rate K36M-rep1-H3K27me3/71,258,901/98.93% K36M-rep1-H3K36me3/31,747,104/97.35% K36M-rep1-WCE/11,853,907/98.91% K36M-rep2-H3K27me3/44,646,559/98.81% K36M-rep2-H3K36me3/25,281,375/97.63% K36M-rep2-WCE/9,514,515/98.57% K9M-rep1-H3K9me3/47,447,856/96.43% K9M-rep1-WCE/12,011,660/99.04% K9M-rep2-H3K9me3/45,531,997/96.66% K9M-rep2-WCE/24,385,036/99.06% WT-rep1-H3K27me3/60,250,956/98.49% WT-rep1-H3K36me3/37,535,527/98.58% WT-rep1-H3K9me3/47,491,613/97.05% WT-rep1-WCE/9,057,499/98.96% WT-rep2-H3K27me3/53,019,241/98.76% WT-rep2-H3K36me3/26,891,358/98.81% WT-rep2-H3K9me3/30,502,621/97.27% WT-rep2-WCE/19,350,119/99.03%
Antibodies	H3K36me3 (Abcam, ab9050), H3K27me3 (Millipore, 07-449) and H3K9me3 (Abcam, ab8898)
Peak calling parameters	For sequence analysis, reads were aligned to mm9 using BWA 0.7.4, and multi-mapped reads and PCR duplicates were removed (Picard Tools). Peak calling was performed using the Homer package (HOMER v4.8) with matched inputs. Peak calling parameters: H3K4me3 and H3K27me3 and H3K9me3, –style histone, –q 0.1 –f 1.5; broad range of H3K9me3, -region –size 50000 –minDist 1000. For merged replicate peaks, we used mergePeaks with –d given. ChIP tracks were converted to a tiled data file (TDF) by igvtools count and then visualized using Integrative Genomics Viewer (IGV, http://www.broadinstitute.org/igv/).
Data quality	Due to the nature of histone modification H3K36me3 and H3K9me3, we focused on H3K36me3 enrichment over gene body and K9me3 over 100Kb bin through the genome, instead of peak calling. 7636 genes are enriched with H3k36me3 above 2-fold over Input, 95% of which are active genes with RNA RPKM>1. 18.3% of genome are enriched with H3K9me3 enrichment above 1.5-fold over input.
Software	To detect differential peaks in mutant samples (K36M and K9M) versus WT, we merged replicate peaks and calculated enrichment values for each peak, then scored as loss or gained if the values pass 1.5 fold enrichment and 10% FDR cutoff (DESeq2). Differential peak calling parameters: getDifferentialPeaksReplicates.pl –t K36M (K9M), –b WT –DESeq2 –q 0.1 –f 1.5 (HOMER). Nearby genes from the peaks were annotated based on the distance to nearest TSS of RefSeq from peak files. Read density heat maps were calculated with normalized signal by total number of signal, across distance from the center of the peak using the data matrix generated by annotatePeaks.pl –ghist (HOMER), and then visualized by cluster 3.0 and TreeView. Average read density was quantified for the histogram with centered peaks, and for meta-gene profiles with

upstream, gene body (0%-100%, RefSeq) and downstream of gene clusters using makeMetaGeneProfile.pl rna mm9 (HOMER). To determine broad regions of H3K9me3 ChIP-seq tag enrichment, we analyzed tag counts in a 100 Kb window sliding with a 50 Kbp step across the chromosome length, and calculated the ratio of ChIP tag count to input count in each window. Regions of H3K9me3 enrichment were generated by merging adjacent enriched windows with ChIP over input ratio > 1.5 that were separated by 100 Kbp or less.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Whole blood was prepared for flow cytometry by lysing blood cells in ACK buffer (150 mM NH4Cl, 10 mM KHCO3, 0.2 mM EDTA (all Sigma Aldrich) for 2 minutes at room temperature. The cells were then centrifuged at 350 X g for 8 minutes, resuspended in FACS buffer (3% fetal bovine serum (Hyclone) in PBS), and filtered through 0.2 µm mesh. Bone marrow from euthanized mice was collected by flushing both femurs and tibiae using a 27½ gauge needle with 10 ml culture media (500 ml DMEM (Life Technologies) 10% fetal bovine serum (Hyclone), 1X non-essential amino acids (Life Technologies), 1X Glutamax (Life Technologies), 55 µM beta-mercaptoethanol (Sigma)). The bone marrow was then dissociated via trituration with a 16 gauge needle. The resulting cells were centrifuged at 350 X g for 8 minutes and resuspended in 5 ml of cold ACK lysis buffer and incubated for 7.5 minutes on ice. The lysis was then quenched immediately with 20 ml of culture media and again centrifuged at 350 X g for 8 minutes. The resulting cells were resuspended in 10 ml of FACS buffer, filtered through 0.2 µm mesh, and counted.
Instrument	Cells were analyzed on an LSRII flow cytometer (BD Biosciences) or collected on an Aria II flow cytometer (BD Biosciences).
Software	Data was collected using Diva v6.1.2 (BD Biosciences) and data analysis was performed using FlowJo software v10.2 (FlowJo, LLC).
Cell population abundance	For hematopoietic stem and progenitor cells, we reanalyzed purified population to ensure greater than 90% purity.
Gating strategy	Dead cells were identified with either DAPI or PI and gated out. Cell debris was gated out using a FSC vs SSC gate. Cell aggregates were excluded using a FSC-W vs FSC-A approach and then again using a SSC-W vs SSC-A approach. All other gating strategies are depicted as schematics within figures as schematics.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.