HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2017 March 01.

Published in final edited form as:

Nat Genet. 2016 September ; 48(9): 1014–1023. doi:10.1038/ng.3610.

DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells

Xiaotian Zhang[1,2,3,8], **Jianzhong Su**[4,8], **Mira Jeong**[1,2,8], **Myunggon Ko**[5,7], **Yun Huang**[6], **Hyun Jung Park**[4], **Anna Guzman**[1,2], **Yong Lei**[1,2], **Yung-Hsin Huang**[1,2], **Anjana Rao**[5], **Wei Li**[4], and **Margaret A. Goodell**[1,2,3]

[1]Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, Texas, United States

[2]Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, Texas, United States

[3]Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, Texas, United States

[4]Division of Biostatistics, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, United States

[5]La Jolla Institute for Allergy and Immunology, La Jolla, CA, United States

[6]Institute of Biosciences and Technology, Texas A&M Health Science Center, Houston, TX, United States

[7]School of Life Sciences, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea

Abstract

Mutations in the epigenetic modifiers *DNMT3A* and *TET2* non-randomly co-occur in lymphoma and leukemia despite their epistasis in the methylation-hydroxymethylation pathway. Using *Dnmt3a* and *Tet2* double knock-out (DKO) mice in which malignancy development is accelerated, we show that the DKO methylome reflects regions of independent, competitive and cooperative

Correspondence: wl1@bcm.edu, goodell@bcm.edu.

[8]These author contribute equally to the work

URLs

We used data and software from these sites: The Cancer Genome Atlas(TCGA) Data Portal, https://tcga-data.nci.nih.gov/tcga/
R language, http://www.r-project.org/.
Browser tracks for the data presented here can be accessed: http://dldcc-web.brc.bcm.edu/lilab/jianzhos/Tet2_3aDKO_Project/Tet2_3aDKO_hub.txt.

Accession Numbers

The RNA-seq, DNA methylation, histone ChIP-seq and CMS-IP data has been deposited in the GEO set: GSE72148. We also used data found at GSE49191 (associated with[40]).

Author Contributions

XZ and MAG conceived and discussed the project with YH, MGK, and AR. XZ analyzed phenotypes and performed shRNA knockdown with MJ. JS analyzed the WGBS and RNA-seq data with HP. MJ performed HSC sorting and WGBS and RNA-seq. MK and AR provided *Tet2*$^{-/-}$ mice. YH generated CMS libraries. XZ led the project and drafted the manuscript. All authors participated in discussions, data interpretation, and manuscript editing. MAG and WL provided funding and supervision.

activity. Expression of lineage-specific transcription factors, including the erythroid regulator *Klf1* is upregulated in DKO HSCs. DNMT3A and TET2 both repress *Klf1* suggesting a model of cooperative inhibition by the epigenetic modifiers. These data demonstrate a dual role for TET2 in promoting and inhibiting HSC differentiation, loss of which, along with DNMT3A, obstructs differentiation leading to transformation.

Introduction

DNA methylation at CpG dinucleotides plays a critical role in regulating gene expression in higher organisms. DNA methylation (mC) is introduced at cytosines by DNA methyltransferases -3A and -3B (DNMT3A and DNMT3B) and maintained by DNMT1. The Ten-Eleven-Translocation (TET) family of proteins (TET1,-2,-3) can oxidize mC to hydroxymethylcytosine (hmC)[1], a mark not effectively maintained by DNMT1, leading to loss of DNA methylation as cells divide[2]. Thus, action of these two protein families can lead to the opposite results: addition or removal of methylation. Importantly, they are also epistatic, because hydroxymethylation depends on the presence of mC introduced by DNMTs.

Mutations in *DNMT3A* and *TET2* are frequent in hematologic malignancies, occurring in a comparable range of lymphoid and myeloid disorders[3–10]. Furthermore, mutations in both are common in individuals with clonal hematopoiesis associated with aging[11–13], likely due to selective expansion of variant HSCs. *DNMT3A* and *TET2* mutations likely arise in the hematopoietic stem cell (HSC), conferring a competitive advantage[14], as mutations can be found in HSCs as well as the malignant clone in patients[15,16].

Loss of TET2 or DNMT3A also impacts HSCs in mice. Conditional ablation of *Dnmt3a* resulted in accumulation of HSCs, hampered differentiation, and multiple malignancies[17–19]. Similarly, *Tet2*$^{-/-}$ progenitors out-compete their normal counterparts and engender myeloid expansion reminiscent of chronic myelo-monocytic leukemia (CMML)[20–23]. Yet how loss-of-function in these epigenetic regulators leads to aberrant HSC function and malignancy is still poorly understood.

In some studies, *TET2* mutations correlated with decreased global hmC and increased mC in patients with acute myeloid leukemia (AML)[24–26]. In contrast, *DNMT3A* mutations lead to global hypomethylation with localized hypermethylation[27]. While the contrasting methylation changes are broadly consistent with our view of their biochemical actions on DNA (writing versus erasing mC), the resulting clinical outcomes are paradoxically similar, suggesting that these proteins also work in parallel to generate a similar cellular end-result. Reinforcing this view is the observation that, in addition to their independent mutation, *DNMT3A* and *TET2* mutations can be found together in some malignancies. In the TCGA AML cohort, 6/17 of patients with *TET2* mutations also have mutations in *DNMT3A* (chi-square test p value=0.0357). Likewise, in some T-cell lymphoma subtypes, over 80% of patients with *DNMT3A* mutations also carry *TET2* mutations[28–30]. Together these findings suggest cooperation between these DNA modifying enzymes. To explore these simultaneously parallel and antagonistic roles, we examined the roles of TET2 and

DNMT3A using single and double knockout mice, focusing on HSCs as the target of the initial changes that set the stage for malignancy development.

Results

Tet2 and *Dnmt3a* loss induces myeloid and lymphoid disease

We crossed *Dnmt3a* conditional knockout mice (*Dnmt3a*$^{f/f}$ Mx1-cre)[17] with *Tet2*$^{-/-}$ mice[20] to create *Dnmt3a-Tet2* double knockout (DKO) mice and controls (Supplementary Figure1a). Control mice include wildtype (WT), *Dnmt3a*$^{-/-}$ (as referred to after induced ablation of *Dnmt3a*) and *Tet2*$^{-/-}$. *Tet2* ablation induces expansion of mutant hematopoietic stem and progenitors cells (HSPCs) over WT counterparts[20–23], and increased serial replating capacity. We compared serial replating of all four genotypes in methylcellulose, finding that DKO cells exhibited the highest activity (Supplementary Figure1b) and appeared least differentiated (Supplementary Figure1c). To examine HSPC competitiveness *in vivo*, we performed whole bone marrow transplantation (BMT) of each genotype in competition with WT. The order of engraftment activity was DKO> *Tet2*$^{-/-}$ > *Dnmt3a*$^{-/-}$ ~ WT in total nucleated peripheral blood. In the myeloid compartment, both the DKO and *Tet2*$^{-/-}$ HSPCs showed higher engraftment than the *Dnmt3a*$^{-/-}$ (Figure 1a). These results indicate that *Dnmt3a* loss augments the competitive advantage of *Tet2*$^{-/-}$ cells.

DKO-transplanted mice developed hematologic disease earliest (Figure 1b) showing anemia and monocytosis as early as 4 months (Figure 1c). Peripheral blood of moribund mice revealed a marked myeloid bias in *Tet2*$^{-/-}$ and DKO transplanted mice (Figure 1d,e) and an erythropoiesis deficiency in DKO recipients (Figure 1f,g). Histology showed *Tet2*$^{-/-}$ recipients developed a myeloid bias 4–6 months after BMT while DKO recipients accumulated immature myeloid progenitors and developed dysmegakaryopoiesis (Supplementary Figure2a); around 10% of DKO animals developed bone marrow failure (Supplementary Figure2a). Only DKO animals showed disrupted spleen features and infiltration of non-hematopoietic organs (Figures S2b–d).

The impact of dual gene loss was prominent in HSPCs. Before transplantation (1 month after pIpC injection) and 4–6 months after transplantation, HSPCs were most abundant in DKO and *Tet2*$^{-/-}$ recipients (Figure 1h,i). Strikingly, the DKO showed marked increases in the Lin-Sca1+cKit+ (LSK) stem/progenitor cell pool (Figure 1i,k).

Interestingly, DKO recipients and some non-transplanted DKO animals developed lymphoid disease. Multiple B cell diseases occurred including mature B cell (B220+ CD19-) salivary gland infiltration, mature B cell lymphoma (largely B220+) that was transplantable and caused ALL in recipients, and B cell ALL (B220+ CD19−) (Supplementary Figure 3a,b,c). T cell thymic lymphoma was observed in 50% of DKO secondary recipients (Supplementary Figure 3d,e). These data support a tumor-suppressor role for DNMT3A and TET2 in lymphoid lineages, similar to that observed in mice with overexpression of *DNMT3A*^{R882H} in *Tet2*$^{-/-}$ mice[31].

Overall, DKO-transplanted mice display multiple hematologic abnormalities including CMML-like disease, bone marrow failure, and lymphoid diseases, establishing a synergistic impact of combined *Dnmt3a* and *Tet2* loss.

Lineage gene activation and HSC gene repression in DKO HSCs

To probe mechanisms of synergy, we examined gene expression differences among the four groups focusing on HSCs (LKS CD150+CD48–) to explore early changes that led to a "pre-leukemic" state. To maintain consistency with our previous *Dnmt3a* research[17], we used transplanted HSCs for all analysis. Many genes differentially expressed in the DKO were shared with the *Tet2*$^{-/-}$, with additional alterations compared to either single knockout alone (Figure 2a, Supplementary Table 1). Gene set enrichment analysis (GSEA) with lineage-specific fingerprint gene sets[32] revealed that genes associated with HSCs and nucleated red blood cells (RBC) was significantly altered (Figure 2b). While both *Tet2*$^{-/-}$ and DKO HSCs downregulated HSC genes (Figure 2b), expression of HSC-associated transcription factors (TFs) was similar (Supplementary Figure 4a–d). Gene expression differences could not be attributed to compensation from other DNA methylation regulators (Supplementary Figure 4e–g). The DKO compared with the *Tet2*$^{-/-}$ HSCs showed further upregulation of RBC signature genes, which were conversely down in *Dnmt3a*$^{-/-}$ HSCs. DKO HSCs also up-regulated a number of B cell genes (Figure 2b).

Further examination revealed up-regulation of genes specific to other lineages in *Tet2*$^{-/-}$ and DKO HSCs (Figure 2c, Supplementary Figure 4j). Strikingly, genes encoding regulators of RBC differentiation, *Klf1 (EKLF)* and the Epo receptor (*Epor*), were markedly up-regulated in DKO HSCs (Figure 2c), with expression levels across the genotypes in the order: DKO> *Tet2*$^{-/-}$>WT> *Dnmt3a*$^{-/-}$. This suggests TET2 represses these genes in WT HSCs, supporting the GSEA analysis. Master regulators of pan-lymphoid and B cell programs, *Ikzf1* and *Ebf1* respectively, followed a similar pattern, consistent with lymphoid disease development (Figure 2c and Supplementary Figure 4h). The pro-myeloid transcription factors *Cebpa and Cebpe* were up-regulated in *Tet2*$^{-/-}$ and DKO HSCs (Figure 2c, Supplementary Figure 4i), aligned with their myeloid differentiation bias. These results indicate *Dnmt3a* loss alone preserves the stem cell program, but in conjunction with *Tet2* loss, leads to de-repression of lineage-specific regulators, augmenting progenitor-like features observed with *Tet2* loss.

Klf1-regulated RBC gene expression promoted in DKO HSCs

We further examined the cluster of genes down in *Dnmt3a*$^{-/-}$ HSCs, up in *Tet2*$^{-/-}$, and further up-regulated in DKO HSCs (Figure 2a, area between dotted lines). Gene ontology analysis focused on distinctions between the DKO and *Tet2*$^{-/-}$ HSCs revealed enrichment of hematopoiesis- and immune response genes (Figure 2d). In particular, genes controlled by KLF1 were enriched (Figure 2d,e)[33,34]. Similarly, target genes of other lineage-specific transcription factors such as *Cebpa* and *Ebf1* were up in DKO HSCs (Supplementary Figure 4 k–l)[35,36]. This suggests that multiple lineage-specific and stem cell programs co-exist in the DKO HSCs, creating a self-renewal status which resembles the "seesaw" model in induce pluripotency[37]. Importantly, overexpression of *KLF1* and its downstream targets is also found in AML patients who harbor both *DNMT3A* and *TET2* mutations, but not in

patients carrying *DNMT3A* mutations alone or neither mutation (Supplementary Figure 5a,b). Thus, *Klf1* overexpression in DKO HSCs parallels that seen in patients and could be a characteristic, if not driving event, in leukemogenesis.

Klf1 and *Epor* erythroid genes promote DKO HSPC self-renewal

HSCs normally express differentiation-associated genes at a low level; such "lineage priming" may accelerate differentiation[38]. The up-regulation of RBC genes and other lineage-specific TFs in DKO HSCs suggests that they are more primed for differentiation than WT counterparts. Paradoxically, DKO results in mature erythroid deficiency and marked anemia (Figure 1f,g).

A similar scenario is observed in TCGA human AML patients who have increased expression of *KLF1* (Group 2 in TCGA AML analysis) and are enriched in *TET2* and *DNMT3A* co-mutations (4/6 *DNMT3A/TET2* double mutant patients fall into this cluster)[10]. Of Group 2 patients with increased *KLF1* expression, a few had erythroleukemia (n=2) and megakaryocytic leukemia (n=1), but the majority belonged to another non-erythro/megakaryocytic leukemia subtype biased towards myeloid development with a clinical manifestation of anemia. Overall, data from TCGA and our study indicate *KLF1* overexpression does not promote erythroid differentiation in the DKO, instead possibly impacting self-renewal.

To test this hypothesis, we performed shRNA knockdown of *Klf1* and *Epor* in DKO HSPCs. Upon *Klf1* and *Epor* knockdown, there was a clear reduction of colony forming units (CFU) from DKO HSPCs on semi-solid media (M3434 with EPO, IL3, IL6 and SCF) (Figure 3a). *Epor* knockdown resulted in a replating deficiency of DKO cells (Figure 3a). Furthermore, *Epor* and *Klf1* knockdown reduced expression of the self-renewal marker cKit on DKO HSPCs (Figure 3b), indicating *Epor* and *Klf1* promote self-renewal of DKO HSPCs *in vitro.* Knockdown of *Klf1* and *Epor* with additional shRNAs achieved similar effects (data not shown). These data suggested that pathways downstream of KLF1 and EPOR, such as JAK-STAT signaling, may contribute to DKO HSC self-renewal. Indeed, decreases in JAK2 and STAT5 expression and phosphorylation, as well as decreased expression of BCL-XL, were observed (Figure 3c). Treatment with the JAK2 inhibitor Ruxolitinib and BCL-XL-specific inhibitor WEHI-539[39] also reduced the CFU and replating capacity of DKO HSPCs (Figure 3d). Similar to the effect of *Klf1* and *Epor* knockdown, treatment with Ruxolitinib decreased the expression level of cKit in replated DKO cells (Figure 3e). Overall, these assays demonstrate that the RBC gene signature is contributing to abnormal self-renewal of DKO HSCs rather than promoting RBC differentiation.

DNMT3A and TET2 maintain DNA methylation in HSCs

In order to investigate DNA methylation changes we performed whole genome bisulfite sequencing (WGBS) in HSCs (Supplementary Figures 6a). Due to the complexity of comparing DNA methylation differences among the WT, single, and double mutants, we identified changes in a two-step process. First, we defined all differentially methylated regions (DMRs) in each mutant compared to WT (Figure 4a). Then, we examined the methylation level of those DMRs in the other genotypes, identifying dynamic methylation

alteration patterns (DMAPs). This strategy revealed 22,075 DMRs, which fell into 6 major DMAPs (Figure 4b; Supplementary Figures 6b–d). *Dnmt3a* loss was the major driver of DNA methylation alteration, with the majority of hypomethylated *Dnmt3a*$^{-/-}$ DMRs also hypomethylated in the DKO (Figures 4c,d). *Tet2*$^{-/-}$ HSCs displayed the least DMRs (Figure 4c).

By identifying DMAPs, we also discovered genotype-specific DMRs (e.g. Type I, II, VI in Figure 4b). In addition, some DMRs (Type IV) gained methylation in the *Tet2*$^{-/-}$ and lost methylation in the *Dnmt3a*$^{-/-}$, with intermediate methylation levels in the DKO, consistent with counteractive TET2 and DNMT3A action at these loci. Some of these correspond to the edges of large and small unmethylated regions (canyons and UMRs) (Supplementary Figures 6e,f), supporting our proposal that DNMT3A and TET2 can oppose each other at DNA methylation canyons[40]. Across most DMRs, methylation changes were primarily enriched at the edges of UMRs, except for the *Tet2*$^{-/-}$-specific class (Type VI) in which hypermethylation occurred exclusively in the center of DMR, indicating that TET2 normally maintains their low methylation (Supplementary Figure 6 e,g).

We were particularly interested in synergistic (Type III) DMRs, in which methylation levels decreased in the *Dnmt3a*$^{-/-}$ from WT and then further dropped in the DKO (average 55% to 26%) (Figure 4b, Supplementary Figure 6b). This behavior suggests that TET2 contributes to maintenance of the methylation signal; this is likely hydroxymethylation, since bisulfite treatment cannot distinguish mC and hmC. Once both proteins are gone, mC as well as hmC are completely depleted from the region, resulting in starkly absent methylation in the DKO at these loci.

Enhancer regions are increasingly recognized as epigenetically dynamic and impacted by DNMTs and TET2[41,42]. Enhancer methylation across hematopoietic lineages[43] was not strongly altered. However, in DMRs that overlapped with enhancers, most were hypomethylated in *Dnmt3a*$^{-/-}$ and DKO HSCs (Supplementary Figure 7a–c). Interestingly, the long-term-HSC (LT-HSC) and short-term-HSC (ST-HSC) enhancers have low methylation levels in HSCs while B cell- and RBC-specific enhancer sites became hypomethylated in both *Dnmt3a*$^{-/-}$ and DKO HSCs (Figure 4e). This suggests that *Dnmt3a* loss-of-function “primes” these lineages in HSCs by depleting methylation at relevant enhancers. Enhancer methylation changes correlated poorly with gene expression, likely due to imprecision in linking enhancers with specific transcription units[43].

Synergistic loss of mC and hmC in Double KO HSCs

In order to determine how loss of DNMT3A and TET2 affected hydroxymethylation, we mapped hmC using the CMS-IP method[44,45]. DNMT3A and TET2 both contributed to hmC maintenance, with *Tet2*$^{-/-}$ resulting in the greatest hmC redistribution (Figure 5a).

We also identified genotype-specific hmC patterns by applying the strategy described above to find differentially hydroxymethylated regions (DhMR), then classifying their dynamic hydroxymethylation alteration patterns (DhMAPs). This revealed 5 DhMAPs clusters (Figure 5b). Despite global loss of hmC in *Tet2*$^{-/-}$ and DKO (Clusters 4 and 5), some loci showed increased hmC specifically in the *Tet2*$^{-/-}$ and DKO HSCs (Clusters 1 and 2).

Increased hmC may be due to ectopic TET1 or TET3 action in absence of TET2 (Figure 5b). This shows that TET2 loss not only reduces global hmC, but leads to hmC redistribution. Finally, we found that both DNMT3A and TET2 contribute to global maintenance of hmC. Hydroxymethylation decreased in *Dnmt3a* and *Tet2*$^{-/-}$ HSCs (Figure 5c,d), with the lowest hmC signal in DKO HSCs, as expected because oxidation to 5hmC requires the presence of 5mC.

We next asked how dynamic methylation (DMAPs) and hydroxymethylation were correlated genome wide. We examined hmC enrichment in the 6 major DMAPs, finding that Type III (synergistic) DMRs were more enriched for hmC than other DMRs (Figure 5e,f). This confirms that in the absence of DNMT3A, TET2-mediated hydroxymethylation of persisting mC accounts for the apparent maintenance of methylation in Type III DMRs. In other words, the methylation signal remaining at these DMRs is actually hmC, which is read as a methylation mark after bisulfite treatment. The hmC at these DMRs disappeared in the *Tet2*$^{-/-}$, and neither mC nor hmC is present at these regions in DKO HSCs (Figure 5g,h).

Since hmC can be a step in DNA demethylation, we investigated how hmC and mC alterations correlated by examining the overlap between DMRs and DhMRs. We found the DMRs and DhMRs significantly overlap (Figure 6a). We then examined the trends in overlap and exclusion of DMR and DhMR groups (Figure 6b). As expected, Type III DMAPs tended to overlap with Cluster 3 DhMR, compared with random overlap between DMRs of all DMAPs and DhMRs of all clusters (Figure 6b), confirming that synergistic DMRs are sites of DNMT3A and TET2 competition. Additionally, Type V DMR and Cluster 4 DhMRs are observed together at Canyon edges (Figure 6b and Supplementary Figure 8a). Finally, genes associated with Type III DMRs and Cluster 3 overlapped with statistical significance (Supplementary Figure 8b). Together, these data show that DMAPs and DhMAPs are correlated in the epigenome and demonstrate the competition (Type V DMR and Cluster 4 DhMRs at canyon edges) and cooperation (Type III DMRs and Cluster 3 DhMRs) of DNMT3A and TET2 in the modifying the epigenome.

5hmC-linked HSC gene activation and lineage-TF repression

With the global hmC and mC dynamics across all 4 genotypes, we have shown that various interactions between DNMT3A and TET2 activity occur in different regions. We next asked whether hmC and mC could globally affect gene expression. By assigning genes expressed in HSCs of one genotype into quantiles by their expression level, we found that the mC distribution correlated only poorly with gene expression (Figure 6c), while the hmC displays a “bi-modal” pattern where the highly expressed genes (>median FPKM) show enriched hmC distribution in the promoter and gene body regions while the genes in the lower quartiles shows hmC enrichment in their TSS (Figure 6d). Our analysis is consistent with previous studies and reveals that genic distribution of hmC is more associated with gene expression than mC.

We next focused on the hmC distribution in the genes that changed the most in mutant HSCs. Consistent with previous reports[7,46–49], the hmC signal was associated with both repressive and active gene expression (Figure 6e–h, Figure 7a, Supplementary Figure 8c). At HSC fingerprint genes[32] that are down-regulated in DKO HSCs (Figure 6e and Figure 2b),

hmC decreased in the gene body and promoter regions (Figure 6f). These observations support a gene-activating role for hmC on genes highly expressed in HSCs. A specific example is the *Mpl* locus, encoding the thrombopoietin receptor critical to HSCs, where loss of hmC at the promoter and gene body correlated with decreased gene DKO expression (Figure 7a).

In contrast, at the transcriptional start sites (TSS) of RBC-associated genes that were up-regulated in DKO HSCs (Figure 6g and 2b), DKO hmC was depleted (Figure 6h). Importantly, RBC gene expression was generally inversely proportional with hmC content in the TSS region: higher hmC correlated with lower relative gene expression (vs WT) and vice versa (Figure 6f,h). The loss of TSS hmC in conjunction with marked gene up-regulation in DKO HSCs strongly suggests that hmC plays a repressive role to constrain RBC gene expression in HSCs. This is consistent with a previous study that showed genes at the bottom 20% of all expressed transcripts are highly enriched for hmC signal in their TSS region[46,48]. Together, these data support the view that hmC in the TSS region maintains gene expression at a low “poised” level.

In addition to the RBC genes, hmC in the TSS region of the myeloid transcription factor *Cebpe* was associated with gene repression (Supplementary Figure 4e and Supplementary Figure 8d). Except for the RBC and HSC fingerprint gene sets (Supplementary Figure 8b), there was not a significant correlation of gene expression with hmC distribution in specific genomic regions (Supplementary Figure 9a,b). Together, these data indicate distinct functions for gene body vs promoter hmC.

DNMT3A and TET2 competitively repress *Klf1* in HSCs

We then examined genes within Cluster 3 for expression alterations, identifying *Klf1* as well as the master lymphoid regulator *Ikzf1,* as members that exhibit both synergistic loss of methylation in the DKO and increased hmC in the *Dnmt3a*$^{-/-}$. Both of these genes are overexpressed in DKO HSCs (Figure 7b, Supplementary Figure 9c). *Klf1* is within the RBC gene group that gains hmC in the *Dnmt3a*$^{-/-}$ and loses hmC in the TSS region when TET2 is lost. *Klf1* also harbors a Type III (synergistic) DMR in its upstream promoter region (Figure 7b), which exhibits high hmC in the *Dnmt3a*$^{-/-}$ (Figure 7b). Consistent with the low expression in WT and *Dnmt3a*$^{-/-}$ HSCs, the Type III DMR in the *Klf1* promoter region also exhibits H3K27me3 in *Dnmt3a*$^{-/-}$ HSCs, which is absent in *Tet2*$^{-/-}$ and DKO HSCs (Figure 7b). Thus, combining *Klf1* and target gene expression, we conclude that DNA methylation and elevated hmC in the TSS region of *Klf1* are both associated with repression of *Klf1* expression in HSCs. Conversely, loss of hmC in *Tet2*$^{-/-}$ and DKO HSCs is associated with increased *Klf1* gene expression.

The promoter region of *Ikzf1* undergoes similar changes, resembling *Klf1* in mC-hmC (Supplementary Figure 9c), in contrast to the *Mpl* locus in which decreased gene-body hmC is associated with lower gene expression in absence of TET2 (Figure 7a). Based on our analysis we propose that DNA methylation at regulatory regions of lineage-specific TFs decreases in *Dnmt3a*$^{-/-}$ HSCs, but TET2 represses their expression by maintaining hmC marks in the TSS region. The loss of both TET2 and DNMT3A is required for full activation of *Klf1* and *Ikzf1.* Therefore, we conclude that in HSCs, both TET2 and DNMT3A repress

differentiation-associated transcription factors by hydroxymethylation and methylation activity in regulatory regions (Figure 7c).

Discussion

DNMT3A and TET2 modify cytosine at CpG dinucleotides, with DNMT3A adding a methyl group whilst TET2 oxidizes it to hydroxymethylcytosine. Classical epistasis theory predicts that combining null mutations in *DNMT3A* and *TET2* should resemble either single mutation. However, we show that double null mice displayed a phenotypes and accelerated disease progression. Our data, together with the fact that *DNMT3A* and *TET2* mutations co-occur in T cell lymphoma and AML, strongly suggests an unconventional interplay between DNMT3A and TET2.

Here, we substantiate these complex interactions, as DNA methylation and hydroxymethylation profiles demonstrate the proteins have both independent and interdependent roles (Figures 4,5). With regard to demethylation instigated by TET2, its deficiency would be redundant with loss of DNMT3A if the critical mC that TET2 acts upon is generated by DNMT3A. Indeed, enhancers display dominant hypomethylation in *Dnmt3a*$^{-/-}$ HSCs, which is maintained in the DKO, particularly at HSC-associated enhancers; this is consistent with participation of DNMT3A in repression of the stem cell program[17,50].

We also observe contribution of TET2 to HSC maintenance: gene bodies of highly expressed HSC-specific genes are hmC laden, and loss of *Tet2* correlates with their diminished expression. This supports previously reported observations in embryonic stem (ES) cells and erythropoiesisfor 5hmC as a stable epigenetic mark[7,26,47–49,51].

Importantly, as compound knockout mice (*Dnmt3a*/*Dnmt3b; Tet1*/*Tet2; Tet2*/*Tet3*) of DNA methylation regulators have been developed, the layered regulatory nature of DNA (hydroxy) methylation has been suggested[17,50,52,53]. The loss of one regulator results in a redistribution of other regulators and DNA modifications, contributing to the mutant phenotypes, as shown here.

Finally, we observe that TET2 is associated with repression of lineage-differentiation genes, particularly in relation to hmC at the TSS. With loss of *Tet2*, accentuated in the DKO, RBC– and myeloid–associated transcription factors were up-regulated. This is consistent with ES cells where a repressive role for hmC is suggested[7,46,48]. Furthermore, *Tet1*-null mice develop a lymphoid- gene signature associated with bivalent marks[54], and hmC derivatives can inhibit transcription elongation[55], linking TETs to poised gene repression.

Overall, we find that TET-mediated hmC appears to contribute to both differentiation and self-renewal through gene activation and repression. DNMT3A may also contribute to both activities by repressing HSC genes and generating mC for TET action. *Dnmt3a* loss exacerbates the *Tet2*$^{-/-}$ expression changes, indicating they both contribute to repression. Indeed, TET2 represses lineage-specific factors as effectively as DNMT3A, explaining dramatic up-regulation of *Klf1* in the DKO. Such parallel and sometimes redundant roles

likely lead to the synergistic development of malignancies with *Tet2* and *Dnmt3a* co-mutation in mouse and man.

Thus, we propose a dual-function model for hmC in HSCs (Figure 7c). Hydroxymethylation, added by TET2 in the gene body and some promoter regions, maintains HSC gene expression (Figure 7c). Concomitantly, mC (generated by DNMT3A) and hmC (oxidized by TET2) both maintain the repression or "primed" state of lineage-specific transcription factors. Once both are lost and these genes are fully de-repressed, their ectopic expression promotes abnormal self-renewal and primes HSCs for transformation (Figure 7c).

The increase of KLF1-regulated genes in both mice and humans is striking. Priming of lineage-specific genes is thought to enhance differentiation with stress[38,56,57]. Paradoxically, erythroid differentiation is abrogated in the DKO, suggesting instead self-renewal, potentially contributing to leukemogenesis. Lymphoid master TFs are similarly upregulated. A recent study on induced pluripotency demonstrated that the combination of ectodermal and mesendodermal inducer genes can substitute for *Oct4* and *Sox2* in the reprograming process[37] suggesting that concurrent expression of normally opposing transcription factors can drive an undifferentiated state. Our observations provide a similar example, in which dysregulated lineage programs in HSCs contribute to disease development. Targeting these pathways may provide therapeutic avenues for patients with malignancies who carry both *DNMT3A* and *TET2* mutations.

Online Methods

Mice

Tet2$^{-/-}$ *and Dnmt3a*$^{f/f}$ mice were previously reported[17,20]. *Tet2*$^{-/-}$ mice were crossed with Mx1-cre*; Dnmt3a*$^{f/f}$ to obtain *Mx1-cre; Dnmt3a*$^{f/f}$; *Tet2*$^{-/-}$. Deletion of *Dnmt3a* in Mx1-cre*; Dnmt3a*$^{f/f}$; *Tet2*$^{-/-}$ was achieved by i.p. injection of poly(I:C) every 2 days for 12 days. *Tet2*$^{-/-}$ mice are referred to as *Tet2*$^{-/-}$, and Cre-activated Dnmt3a$^{f/f}$ mice are referred to as *Dnmt3a*$^{-/-}$. Four weeks after the last injection, bone marrow transplantation was performed. All mice were maintained under pathogen-free conditions according to the Institutional Animal Care and Use Committee guidelines.

Bone Marrow Transplantation and HSCs

Bone marrow (BM) transplantation was performed with 2.5×10^5 CD45.2 mononuclear whole bone marrow cells from *Dnmt3a*$^{-/-}$*, Tet2*$^{-/-}$ or DKO mice co-transplanted with an equal number of CD45.1 WT cells into lethally irradiated C57BL/6 mice. Monthly, mice were checked for disease development and donor BM engraftment using CD45.2 and other lineage marker as described previously.[58] Bone marrow transplantation has been duplicated independently and similar result has been achieved and representative data has been shown. Unless specified, HSCs were purified from mice transplanted with BM of the indicated genotypes as CD45.2+ Lineage^{neg}, Sca1+ cKit+ CD150+ CD48− (LKS CD150+ CD48−)

Phenotyping

Tissues, –femur, tibia, sternum, spleen, liver, lung and tumor tissue — were harvested and a single-cell suspension was made by trituration. Peripheral blood was obtained retro-orbitally and stained with antibodies, analyzed or sorted by flow cytometry (LSRII or FACSAriaII, BD Biosciences). For histology, blood smears and touch preps from tibia, spleen or tumors were made for Giemsa Wright Staining (Hema Diff, Stat Lab). Fresh organs were fixed in formalin for HE staining.

In vitro Colony Forming Assay

1×10^4 whole bone marrow cells were plated in MethoCult M3434 medium (Stemcell Technologies); colonies were counted one week later. *In vitro* cultured cells were then washed in PBS and replated at $1x10^4$ in MethoCult M3434; the procedure was repeated 4 times. For the colony-forming assay with shRNA knockdown, 5-FU was injected to DKO recipient mice. One week later, mice were sacrificed and BM harvested. CD45.1 cells were magnetically depleted and Sca1+ cells enriched. 0.5×10^5 Sca1+ DKO HSPCs were then transduced with shRNA-retrovirus. Two days after transduction, 250 GFP+ Lin− Sca1+ cKit +(LSK) cells were sorted and plated in 1mL M3434. Colonies were counted one week later.

shRNA Cloning

pGIPZ shRNA clones (Openbio Systems, GE Dharmacon) against *Epor* and *Klf1* were obtained from a Baylor College of Medicine Core. The Mir-30-based hairpin sequence was removed and cloned into to the LMP vector (Openbio Systems). shRNA clones (GE Dharmacon pGIPZ): sh*Epor*: V3LMM_489985. sh*Klf1*: V3LMM_489109. Knockdown efficiency of *Epor* and targets was confirmed by immunoblot with antibody targeting EPOR, β-ACTIN (Santa Cruz), JAK2, Phos-JAK2 (Y1007/1008), BCLXL (all Cell Signaling).

RNA-sequencing (RNA-seq)

RNAseq was performed on HSCs (LKS $CD150^+CD48^-$) purified after 8 weeks from mice transplanted with $Tet2^{-/-}$ and DKO bone marrow for second time (all in duplicates),and compared with data from WT and $Dnmt3a^{-/-}$ HSCs[40,50]. HSCs were FACS sorted and RNA was isolated with the RNeasy Micro kit (Qiagen, Valencia, CA), including the DNase I (Qiagen) on- column digestion. Paired end libraries were generated by using Illumina TruSeq RNA sample preparation kit. Illumina HiSeq 2000 was used for sequencing with a paired-end sequencing length of 100bp. RNA-sequencing reads were mapped using TopHat 2.0.10. The differentially expressed genes were identified by pairwise comparison using the software Cufflinks 2.1.1 with FDR adjusted p-value <0.01 and relative fold changes of mean FPKM>2[59]. The software MEV version 4.8.1 was used to perform the unsupervised hierarchical clusters of differentially expressed genes and plot the heatmap.

ChIP-sequencing (ChIP-seq)

HSCs purified as indicated above were sorted and crosslinked with 1% formaldehyde at room temperature (RT) for 10 min, and the reaction was stopped by 0.125M glycine at RT for 5 min. Cross-linked cells were lysed and sonicated to 200–500 bp fragments (Bioruptor, Diagenode). ChIP-qualified antibodies (H3K4me3 Millipore 07-473, H3K27me3 Millipore

07-449) were added to the sonicated chromatin and incubated at 4°C overnight. Following this, 10 μl of protein A magnetic beads (Dynal, Invitrogen) previously washed in RIPA buffer were added and incubated for an additional 2 hours at 4°C. The bead: protein complexes were washed three times with RIPA buffer and twice with TE buffer. Following transfer into new 1.5 ml collection tube, genomic DNA was eluted for 2 hours at 68°C in 100 μl Complete Elution Buffer (20 mM Tris pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 μg/ml proteinase K), and combined with a second elution of 100 μl Elution Buffer (20 mM Tris pH 7.5, 5 mM EDTA, 50 mM NaCl) for 10 min at 68°C. ChIPed DNA was purified by MinElute Purification Kit (Qiagen) and eluted in 12 μl elution buffer. ChIPed DNA was successfully made into a library using ThruPLEX-FD preparation kit (Rubicon, Ann Arbor, MI). Sequencing was performed according to the manufacturer's protocol on a HiSeq 2000 (Illumina). Sequenced reads were mapped to the mm9 mouse genome and peaks were identified by model-based analysis of ChIP-seq data (MACS).[60]

Whole-Genome Bisulfite Sequencing (WGBS)

For WGBS library construction, 300ng genomic DNA was isolated from FACS-sorted HSCs (as indicated above, and performed in duplicates) and fragmented using a Covaris sonication system (Covaris S2). DNA libraries were constructed using the Illumina TruSeq DNA sample preparation kit. After ligation, libraries were bisulfite-treated using the EpiTect Bisulfite Kit (Qiagen, Valencia, CA). Ligation efficiency was tested by PCR using TrueSeq primers and *Pfu* TurboCx hotstart DNA polymerase (Stratagene). After determining the optimized PCR cycle number for each sample, a large scale PCR reaction (100 ul) was performed as described previously[40]. PCR products were sequenced with Illumina HiSeq sequencing systems. Paired- end bisulfite-treated sequencing reads were aligned to the mouse genome mm9. The adapters and the low-quality reads were trimmed using BSMAP with default. Sequencing depth of CpGs at least 5 reads of each sample were extract to calculate the methylation ratio calling [0,1] by the software MOABS[61]. A two-state first-order hidden Markov model was used to detect the undermethylated regions and larger undermethylated canyons (≥ 3.5 Kb) as described previously[40].

Identification of Differentially Methylated Regions and Clustering Analysis

To identify differentially methylated regions of multiple samples, we adopted a combinational strategy as described in Figure 5A. First, the differentially methylated regions (DMRs) were identified by pairwise comparison of WGBS profiles (Dnmt3aKO vs WT, Tet2KO vs WT and DKO vs WT) using the software MOABS with at least 5 differentially methylated CpG sites. Second, the three sets of DMRs included were integrated into the union of DMRs among the four genotypes. Next, the mean methylation levels of CpGs covered with at least 5 reads in DMRs were computed in each genotype. Finally, the DMRs were further assessed by a quantitative method QDMR based on shannon entropy among the four genotypes[62]. The DMRs with default entropy value <2 remained. To study the patterns of dynamic change of DNA methylation and their coordinate regulation relationship among the different genotypes of *Dnmt3a* and *Tet2*$^{-/-}$, we used the method of k-means clustering to classify the DMR types among four genotypes with the optimal number of clusters estimated by R package 'fpc'. The six types of DMRs were identified from 22,475 DMRs in Supplementary Table S2.

Mouse annotation information of Refseq genes and CpG islands was downloaded from UCSC The mouse genome was segmented into seven categories: Upstream 2kb regions of transcription start sites (Up2kb), 5′UTR, gene body (including encoding exons and introns), 3′UTR and Downstream 2kb regions of transcription terminal sites (Down2kb) and Intergenic regions. The CGI shores were defined as 0–2kb regions from CGI and CGI shelves as 2–4kb regions from CGI. The overlapping numbers of different types of DMRs with various annotation regions were computed using Bedtools[63].

CMS-Immunoprecipitation

Genomic DNA was isolated using Qiagen DNeasy blood and tissue kit. Purified genomic DNA was sonicated into ~300 bp using Covaris focused ultrasonicators. Sheared DNA was ligated with methylated adaptors using illumine TruSeq DNA library preparation kit followed by sodium bisulfite treatment (Life Technologies, Methylcode bisulfite conversion kit) to convert 5hmC to cytosine methyl sulfonate (CMS). CMS fragments were enriched using anti-CMS antibody that binds to protein A/G dynabeads. Enriched fragments were cleaned up using phenol/ chloroform/ isoamyl-alcohol method and then amplified using KAPA HiFi Uracil+ (Kapa Biosystems) polymerase with 12 PCR cycles. Amplified libraries were sequenced using Illumina HiSeq2500. The paired-end 100-bp 5hmc CMS reads of four genotypes (HSC, *Dnmt3a*$^{-/-}$, *Tet2*$^{-/-}$ and *Tet2_3a*DKO) were aligned to the mouse reference genome (mm9) using BSMAP alignment algorithm[64] with at most 4 mismatches as described previously[40]. For each sample, the sum read number was normalized to 30 million. To compare the 5hmc CMS distribution among four genotypes, quantile normalization for 5hmc CMS density of each was preformed using Wiq function. And the 5hmc density of DMRs and flanking regions (2kb) were computed using the Profile function in DANPOS 2.2.0.[65] The paired identification of DhMRs were performed using Dregion function of DANPOS 2.2.0 between wild type (WT) and three knock- out genotypes (*Dnmt3a*$^{-/-}$, *Tet2*$^{-/-}$ and DKO) of HSC with FDR cutoff 1x10^{-5}, respectively. The three sets of DhMRs were merged into reference DhMRs. The lengths of reference DhMRs more than 200bp were retained. The average CMS signals of reference DhMRs were computed using Bedtools.[63] The method of k-means clustering was used to classify the clusters of DhMRs among four genotypes with the optimal number of clusters estimated by R package ‘fpc’. The five clusters of DhMRs were identified from 40,969 DhMRs that show different 5hmc enrichment among the four genotypes (See Supplementary Table 3).

Functional Enrichment Analysis

For the RNA-seq functional enrichment analysis, the GSEA pre-rank test has been used to test the enrichment of LOLLIPOP lineage specific gene set for differentially expressed genes in mutant HSC vs WT HSC (DKO vs WT).[32,66] DAVID v6.7 is also used to perform function enrichment analysis for differentially expressed genes (DKO vs Tet2KO) for the target pathway analysis. For the DMR functional enrichment analysis, we used DAVID v6.7 to perform the function enrichment analysis for differentially expressed genes (DKO vs Tet2KO) and the gene set of different DMR types[67]. The significant p-values were adjusted by Bonferroni method and visualized by MeV. The DMR associated genes were extracted if DMRs located within 2kb of the genes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank M. Kampmann for discussions at AACR, C. Gillespie for editing, R. Nitsal, S. Hexige, and for Y. Zheng for assistance with histology, pathology, and bone marrow transplantation respectively. X.Z. is supported by the Welcome Trust Gene-Environment training program. This work was supported by the NIH (DK092883, CA183252, HG007538, CA193466, CA125123, P50CA126752, CA151535), the Samual Waxman Medical Research Foundation, the Edward P Evans Foundation, the Adrienne Helis Malvin Medical Research Foundation, the Leukemia and Lymphoma Society (Translational Research Program to AR and Special Fellow Award to MK), and by CPRIT (RP140001) and a CPRIT Scholar to YH (RP140053).

References for main text

1. Tahiliani M, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324:930–5. [PubMed: 19372391]
2. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science. 2011; 334:194. [PubMed: 21940858]
3. Ley TJ, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010; 363:2424–33. [PubMed: 21067377]
4. Walter MJ, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia. 2011; 25:1153–8. [PubMed: 21415852]
5. Grossmann V, et al. The molecular profile of adult T-cell acute lymphoblastic leukemia: mutations in RUNX1 and DNMT3A are associated with poor prognosis in T-ALL. Genes Chromosomes Cancer. 2013; 52:410–22. [PubMed: 23341344]
6. Van Vlierberghe P, et al. ETV6 mutations in early immature human T cell leukemias. J Exp Med. 2011; 208:2571–9. [PubMed: 22162831]
7. Pastor WA, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature. 2011; 473:394–7. [PubMed: 21552279]
8. Abdel-Wahab O, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood. 2009; 114:144–7. [PubMed: 19420352]
9. Delhommeau F, et al. Mutation in TET2 in myeloid cancers. N Engl J Med. 2009; 360:2289–301. [PubMed: 19474426]
10. Cancer Genome Atlas Research N. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013; 368:2059–74. [PubMed: 23634996]
11. Genovese G, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014; 371:2477–87. [PubMed: 25426838]
12. Jaiswal S, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014; 371:2488–98. [PubMed: 25426837]
13. Xie M, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nat Med. 2014; 20:1472–8. [PubMed: 25326804]
14. Welch JS, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012; 150:264–78. [PubMed: 22817890]
15. Shlush LI, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014; 506:328–33. [PubMed: 24522528]
16. Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc Natl Acad Sci U S A. 2014; 111:2548–53. [PubMed: 24550281]
17. Challen GA, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet. 2012; 44:23–31. [PubMed: 22138693]
18. Mayle A, et al. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. Blood. 2015; 125:629–38. [PubMed: 25416277]

19. Celik H, et al. Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation. Blood. 2015; 125:619–28. [PubMed: 25416276]
20. Ko M, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc Natl Acad Sci U S A. 2011; 108:14566–71. [PubMed: 21873190]
21. Moran-Crusio K, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell. 2011; 20:11–24. [PubMed: 21723200]
22. Li Z, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood. 2011; 118:4509–18. [PubMed: 21803851]
23. Quivoron C, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell. 2011; 20:25–38. [PubMed: 21723201]
24. Ko M, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature. 2010; 468:839–43. [PubMed: 21057493]
25. Figueroa ME, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell. 2010; 18:553–67. [PubMed: 21130701]
26. Madzo J, et al. Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. Cell Rep. 2014; 6:231–44. [PubMed: 24373966]
27. Russler-Germain DA, et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. Cancer Cell. 2014; 25:442–54. [PubMed: 24656771]
28. Palomero T, et al. Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. Nat Genet. 2014; 46:166–70. [PubMed: 24413734]
29. Couronne L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. N Engl J Med. 2012; 366:95–6. [PubMed: 22216861]
30. Sakata-Yanagimoto M, et al. Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. Nat Genet. 2014; 46:171–5. [PubMed: 24413737]
31. Scourzic L, et al. DNMT3A mutant and Tet2 inactivation cooperate in the deregulation of DNA methylation control to induce lymphoid malignancies in mice. Leukemia. 2016
32. Chambers SM, et al. Hematopoietic fingerprints: an expression database of stem cells and their progeny. Cell Stem Cell. 2007; 1:578–91. [PubMed: 18371395]
33. Tallack MR, et al. Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. Genome Res. 2012; 22:2385–98. [PubMed: 22835905]
34. Tallack MR, et al. A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. Genome Res. 2010; 20:1052–63. [PubMed: 20508144]
35. Hasemann MS, et al. C/EBPalpha is required for long-term self-renewal and lineage priming of hematopoietic stem cells and for the maintenance of epigenetic configurations in multipotent progenitors. PLoS Genet. 2014; 10:e1004079. [PubMed: 24415956]
36. Schwickert TA, et al. Stage-specific control of early B cell development by the transcription factor Ikaros. Nat Immunol. 2014; 15:283–93. [PubMed: 24509509]
37. Shu J, et al. Induction of pluripotency in mouse somatic cells with lineage specifiers. Cell. 2013; 153:963–75. [PubMed: 23706735]
38. Orkin SH. Priming the hematopoietic pump. Immunity. 2003; 19:633–4. [PubMed: 14614848]
39. Lessene G, et al. Structure-guided design of a selective BCL-X(L) inhibitor. Nat Chem Biol. 2013; 9:390–7. [PubMed: 23603658]
40. Jeong M, et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat Genet. 2014; 46:17–23. [PubMed: 24270360]
41. Ziller MJ, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013; 500:477–81. [PubMed: 23925113]
42. Rasmussen KD, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev. 2015; 29:910–22. [PubMed: 25886910]

43. Lara-Astiaso D, et al. Immunogenetics. Chromatin state dynamics during blood formation. Science. 2014; 345:943–9. [PubMed: 25103404]
44. Huang Y, et al. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One. 2010; 5:e8888. [PubMed: 20126651]
45. Huang Y, Pastor WA, Zepeda-Martinez JA, Rao A. The anti-CMS technique for genome-wide mapping of 5-hydroxymethylcytosine. Nat Protoc. 2012; 7:1897–908. [PubMed: 23018193]
46. Shen L, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell. 2013; 153:692–706. [PubMed: 23602152]
47. Williams K, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature. 2011; 473:343–8. [PubMed: 21490601]
48. Wu H, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature. 2011; 473:389–93. [PubMed: 21451524]
49. Ficz G, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011; 473:398–402. [PubMed: 21460836]
50. Challen GA, et al. Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. Cell Stem Cell. 2014; 15:350–64. [PubMed: 25130491]
51. Tsagaratou A, et al. Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. Proc Natl Acad Sci U S A. 2014; 111:E3306–15. [PubMed: 25071199]
52. Zhao Z, et al. Combined Loss of Tet1 and Tet2 Promotes B Cell, but Not Myeloid Malignancies, in Mice. Cell Rep. 2015; 13:1692–704. [PubMed: 26586431]
53. An J, et al. Acute loss of TET function results in aggressive myeloid cancer in mice. Nat Commun. 2015; 6:10071. [PubMed: 26607761]
54. Cimmino L, et al. TET1 is a tumor suppressor of hematopoietic malignancy. Nat Immunol. 2015; 16:653–62. [PubMed: 25867473]
55. Wang L, et al. Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. Nature. 2015; 523:621–5. [PubMed: 26123024]
56. Guo G, et al. Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. Cell Stem Cell. 2013; 13:492–505. [PubMed: 24035353]
57. Nimmo RA, May GE, Enver T. Primed and ready: understanding lineage commitment through single cell analysis. Trends Cell Biol. 2015; 25:459–67. [PubMed: 26004869]
58. Mayle A, Luo M, Jeong M, Goodell MA. Flow cytometry analysis of murine hematopoietic stem cells. Cytometry A. 2013; 83:27–37. [PubMed: 22736515]
59. Pollier J, Rombauts S, Goossens A. Analysis of RNA-Seq data with TopHat and Cufflinks for genome-wide expression analysis of jasmonate-treated plants and plant cultures. Methods Mol Biol. 2013; 1011:305–15. [PubMed: 23616006]
60. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]
61. Sun D, et al. MOABS: model based analysis of bisulfite sequencing data. Genome Biol. 2014; 15:R38. [PubMed: 24565500]
62. Zhang Y, et al. QDMR: a quantitative method for identification of differentially methylated regions by entropy. Nucleic Acids Res. 2011; 39:e58. [PubMed: 21306990]
63. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics. 2014; 47:11 12 1–11 12 34. [PubMed: 25199790]
64. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics. 2009; 10:232. [PubMed: 19635165]
65. Chen K, et al. DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res. 2013; 23:341–51. [PubMed: 23193179]
66. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–50. [PubMed: 16199517]

67. Huang DW, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007; 8:R183. [PubMed: 17784955]

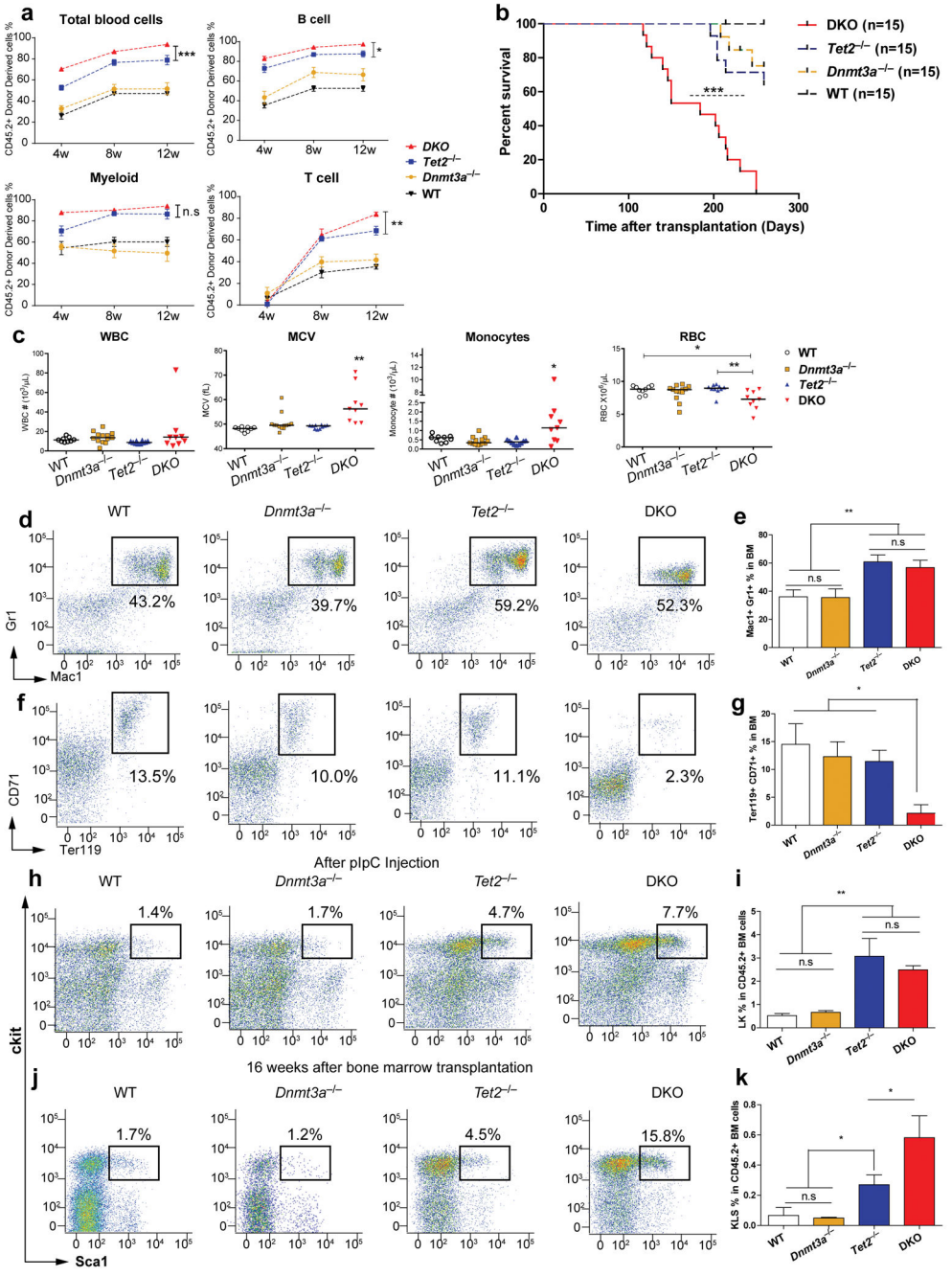

Figure 1. Phenotype of Dnmt3a-Tet2 DKO mice

a) Engraftment of donor-derived bone marrow cells after competitive bone marrow transplantation (BMT) of wild type, *Dnmt3a*$^{-/-}$, *Tet2*$^{-/-}$ and DKO in WBC, myeloid, B cell and T cell compartments in the first 12 weeks. n=15 for each group. **b)** Kaplan-Meier survival of BMT recipients ***, p<0.001, Log-rank test. **c)** Complete blood counts of BMT recipients after 4 months. WBC, white blood cell. RBC, Red blood cell, MCV, Mean corpuscular volume. n=10 for each group. **d–e)** Donor-derived cells stained with myeloid markers Mac1 and Gr1 6 months after BMT. **f–g**) Donor-derived cells stained with the

erythroid markers CD71 and Ter119 6 months after BMT. n=3 for each group. **h**) Hematopoietic progenitor analysis (Lineage^{neg} Scal^{+}cKit^{+}) in representative mice 1 month after Poly(I:C) injection; (**i**) Quantification of myeloid progenitors (LK) in donor-derived cells from (**h**) n=3 for each group. (**j,k**) Percentage of hematopoietic progenitors (LSK) in donor-derived cells 4 months after BMT. n=3 for each group. All error bars show the mean ±S.E.M. n.s, not significant; * $p<0.05$; ** $p<0.01$. Two tailed student T test.

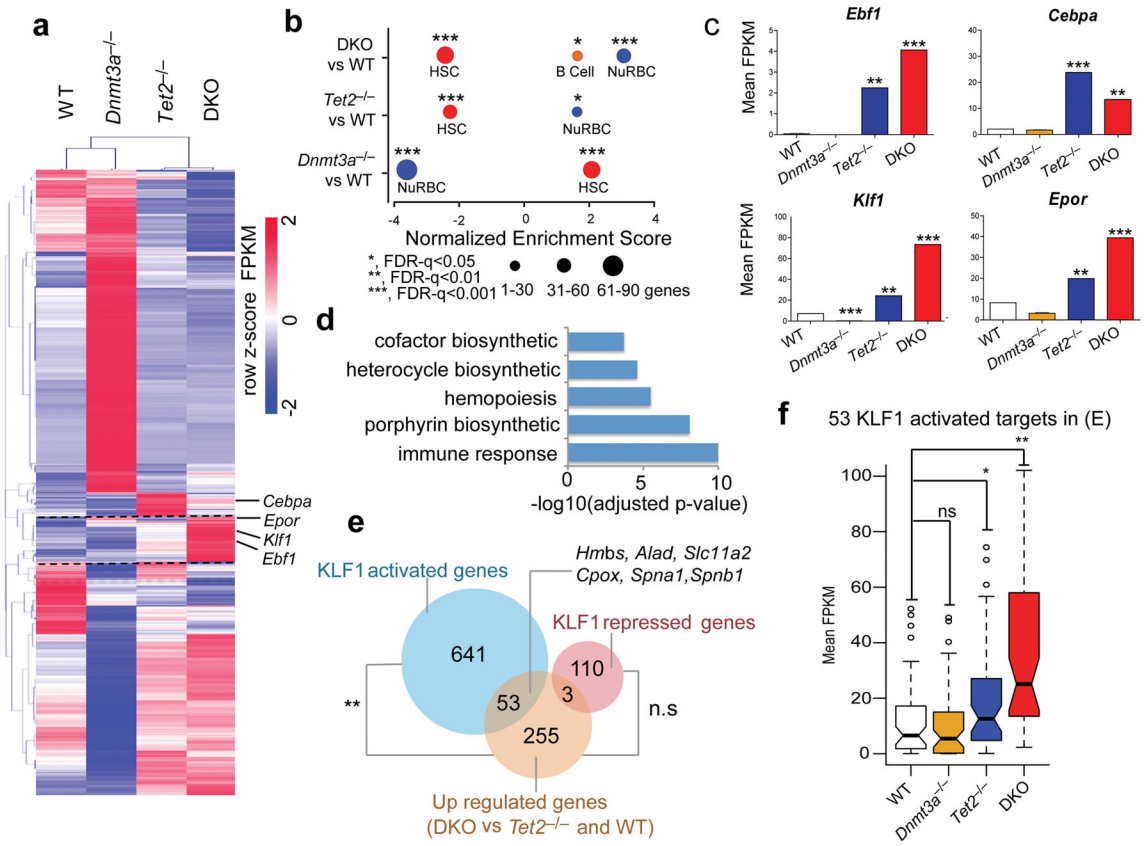

Figure 2. Synergistic dysregulation of HSC- and RBC-associated genes in DKO HSCs

a) Heatmap displaying all differentially expressed genes in each genotype relative to WT. **b**) Plot representing GSEA enrichment of Fingerprint lineage-specific genes among differentially expressed genes between the indicated genotypes. Normalized enrichment scores are plotted with FDR q value. HSC: Hematopoietic Stem Cell. NuRBC: nucleated Red Blood Cell. **c**) Expression levels of the indicated genes in HSCs. $p<0.05$. **, $p<0.01$. ***, $p<0.001$. Each group consist of 2 biological replicates. p value was calculated by cuffdiff in the RNA-seq analysis pipeline. **d**) Functional pathway enrichment analysis of genes up-regulated in DKO vs. *Tet2*$^{-/-}$ HSCs. **e**) Venn diagram depicting the overlap between genes upregulated in DKO vs. *Tet2*$^{-/-}$ HSCs compared with activated and repressed KLF1 targets. **, $p<0.01$. Fisher exact test. The indicated genes are members of the heme synthesis pathway. **f**) Box plot depicting the range of expression of KLF1 target genes from (**e**) which are activated in the DKO relative to the *Tet2*$^{-/-}$. Boxplots represent the interquartile range (25% to 75%), with the median; whiskers correspond to 1.5 times the interquartile range. Each group consist of 2 biological replicates. Each * $p<0.05$. **, $p<0.01$. ns, not significant. p value was calculated by cuffdiff in RNA-seq analysis pipeline.

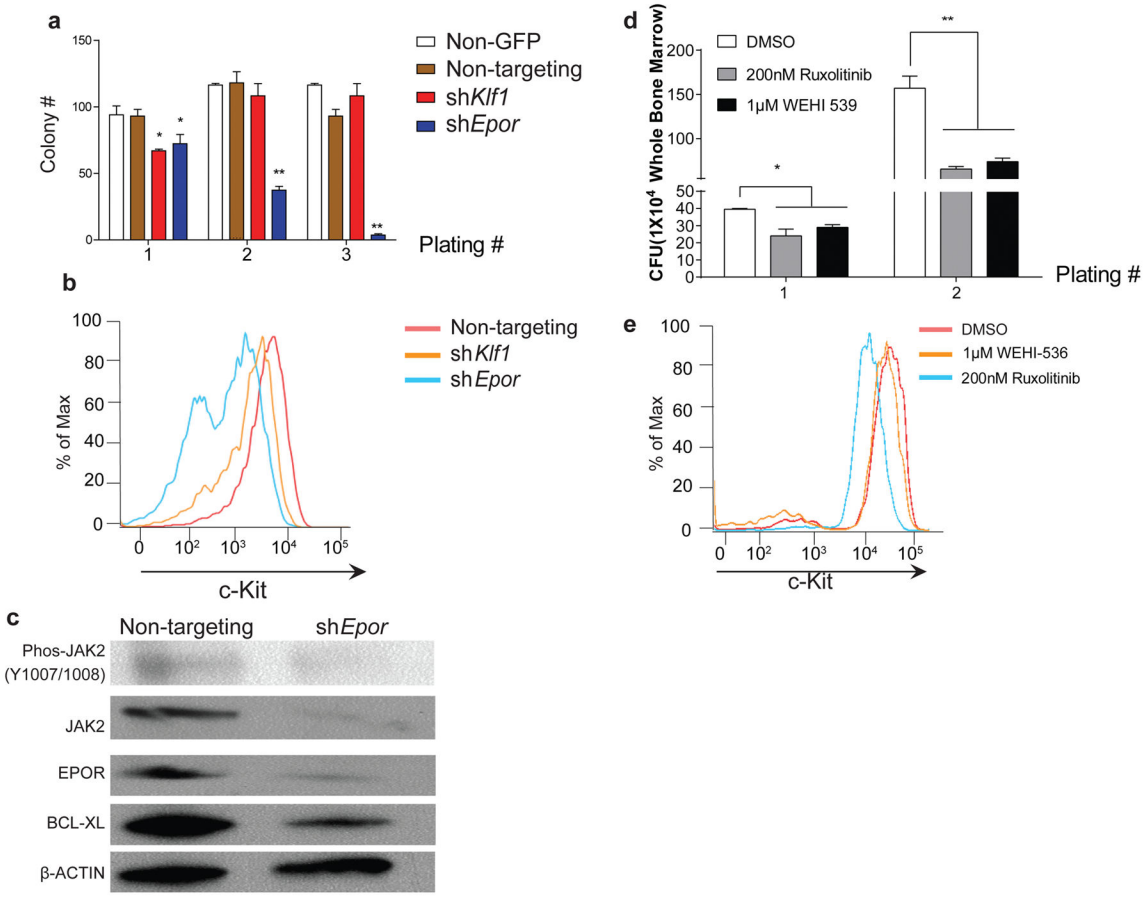

Figure 3. Knockdown of *Klf1* and *Epor* reverses the abnormal self-renewal of DKO HSPCs *in vitro*

a) Replating assay with DKO Lin− cKit+ Sca1+ (LSK) HSPC cells transfected with shScramble, sh*Epor*, and sh*Klf1* using M3434 methylcellulose medium. n=3 for each treatment group. **b)** ckit expression on plated DKO LSK HSPC cells transfected with shScramble, sh*Epor*, and sh*Klf1* cells. **c)** Immunoblot shows the knockdown of *Epor* and the expression level of downstream targets of *Epor* after knockdown of *Epor* in DKO HSPCs. **d)** Replating assay with DKO whole bone marrow cell after treatment with the JAK2 inhibitor Ruxolitinib and Bcl-xL specific inhibitor WEHI-569. n=3 for each treatment group. **e**) ckit expression on plated DKO whole bone marrow cells treated with Jak2 inhibitor Ruxolitinib and Bcl-xL specific inhibitor WEHI-569. All error bars show the mean ± S.E.M. *, $p<0.05$, Student T-test.

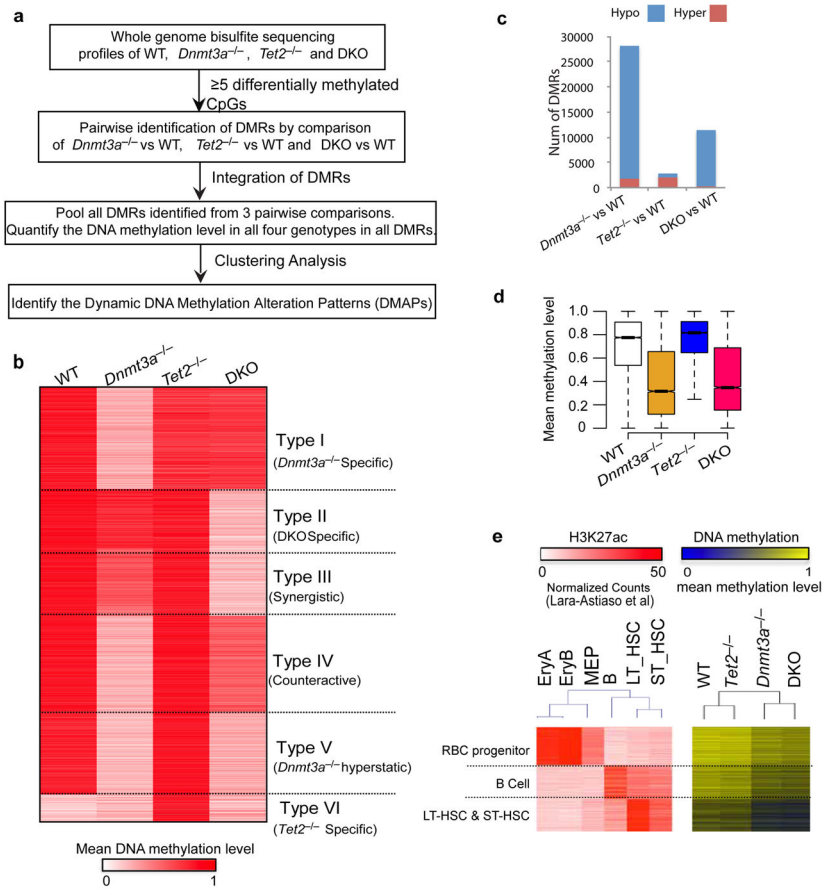

Figure 4. DNA methylation across the genotypes are highly dynamic

a) Flowchart of DNA methylation analysis strategy. **b)** Heat map depicting differentially methylated regions (DMRs) of the 6 major dynamic DNA methylation alteration patterns (DMAPs). **c)** Numbers of hypermethylated and hypomethylated DMRs in HSCs of *Dnmt3a*$^{-/-}$, *Tet2*$^{-/-}$ and DKO phenotype. **d)** Global DNA methylation levels of all DMRs in all 4 genotypes. Each group of genotype consisted of 2 biological replicates. **e)** DNA methylation levels in hematopoiesis-associated enhancers in different lineages and progenitors. Enhancers (left panel) as defined by H3K27Ac marks in B cell, RBC progenitors, ST- and LT-HSCs (from ref. 43), and (right panel) their DNA methylation levels.

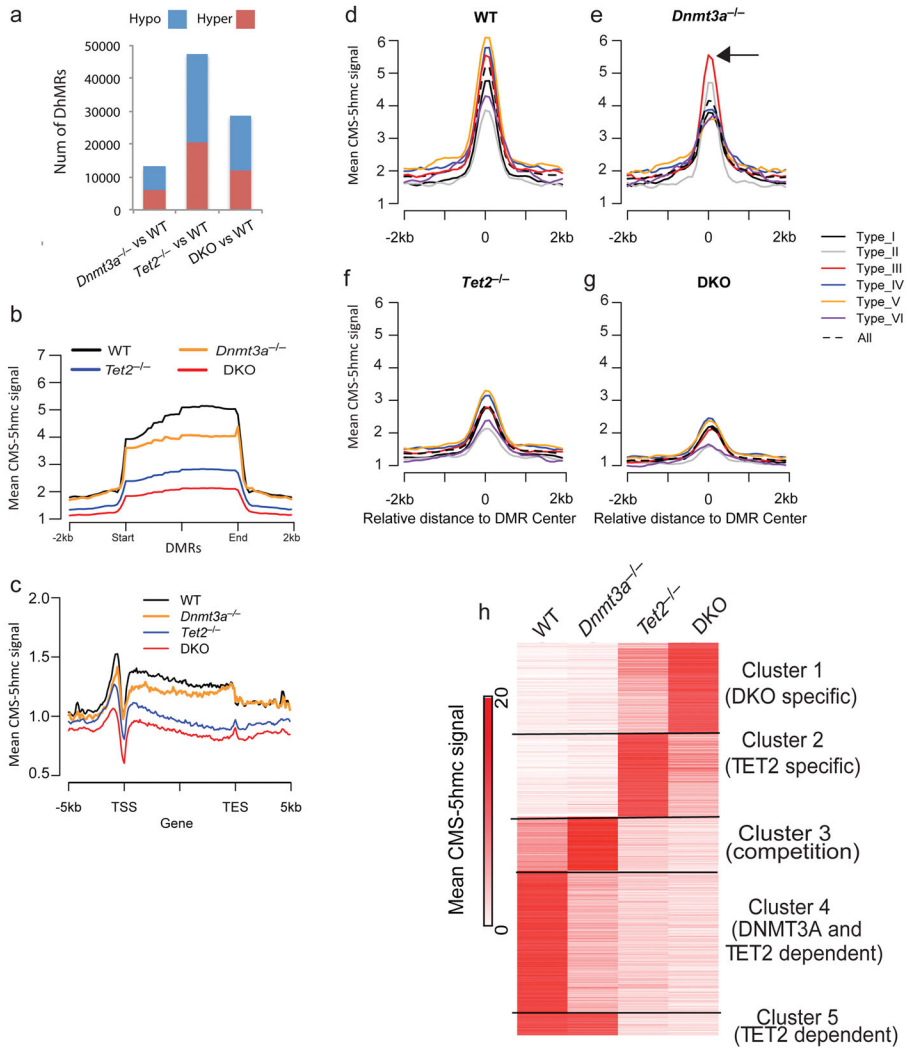

Figure 5. Hydroxymethylation (hmC) in HSCs is associated with active HSC genes and repressed RBC genes

a) Numbers of hypermethylated and hypomethylated DhMRs in HSCs of *Dnmt3a*$^{-/-}$, *Tet2*$^{-/-}$ and DKO phenotype. **b**) Heat map depicting differentially hydroxylmethylated regions (DhMRs) of the 5 major dynamic DNA hydroxylmethylation alteration patterns (DhMAPs). **c**) Cytosine-5-methylenesulfonate (CMS) signal in HSCs displayed for DMR regions. **d**) CMS hmC signal in displayed in relationship to genic features. **e–h**) Differential CMS signal distribution in regions within the 6 major DMR classes in HSCs.

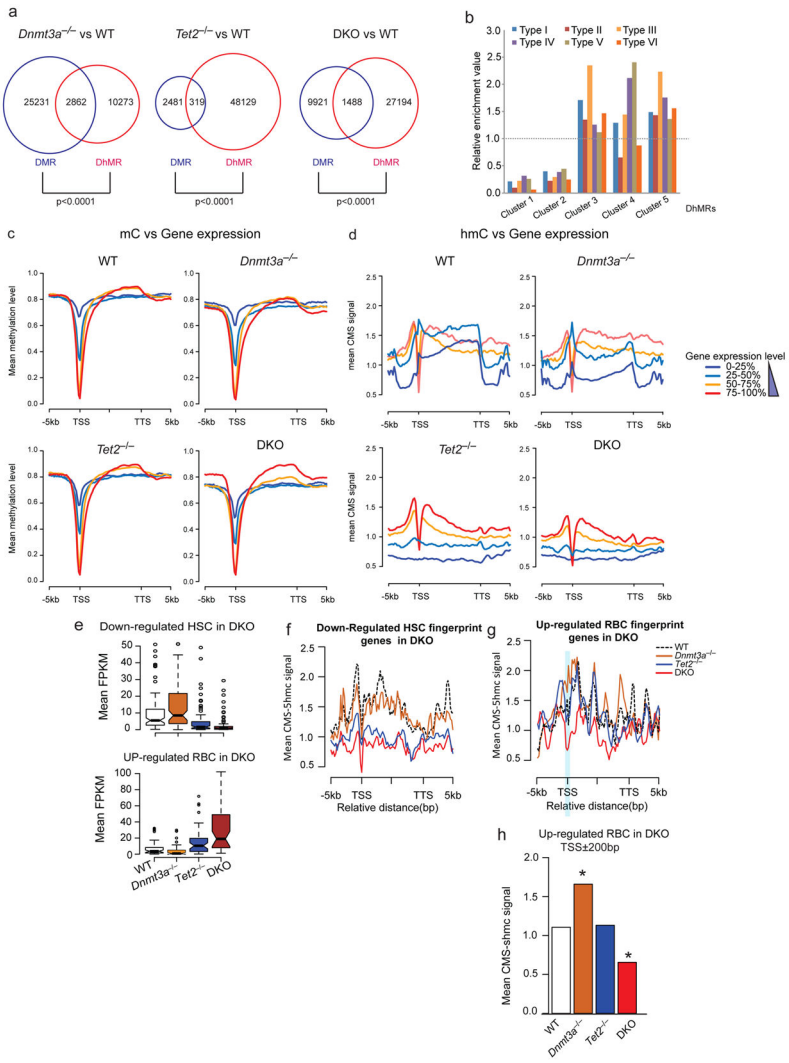

Figure 6. Hydroxymethylation (hmC) in HSCs is associated with active HSC genes and repressed RBC genes

a) Overlap between DMRs and DhMRs in *Dnmt3a*$^{-/-}$, *Tet2*$^{-/-}$ and DKO HSCs. p value is calculated with Chi-square test. **b)** Relative overlap ratio of 6 different dynamic DMR patterns with 5 clusters of DhMRs in comparison with random overlap of DMR and DhMRs. The grey line indicates random DMR and DhMR global overlap ratio. Observed/Expected is plotted. **c)** Distribution of mC in genic regions of genes from low to high expression level in HSCs of all 4 genotypes. TSS: transcriptional start site. TTS: transcriptional termination site. **d)** Distribution of mC in genic regions of genes from low to high expression level in HSCs of all 4 genotypes. TSS: transcriptional start site. TTS: transcriptional termination site. **e**) Gene expression level (FPKM) of HSC Fingerprint genes (upper panel) analyzed in **f** and RBC Fingerprint genes (lower panel) analyzed in **g**. Each genotype group consisted of 2 biological replicates. **f**) hmC signal distribution on HSC-Fingerprint genes down-regulated in DKO HSCs. Each genotype group consisted of 2 biological replicates. **g**) hmC signal on RBC-Fingerprint genes associated with up-regulation in DKO HSCs. Each genotype consisted of 2 biological replicates. **h**) The hmC signal

distribution of the blue shaded area from (**g**) containing the TSS±200bp region *, $p<0.05$. Student t-test.

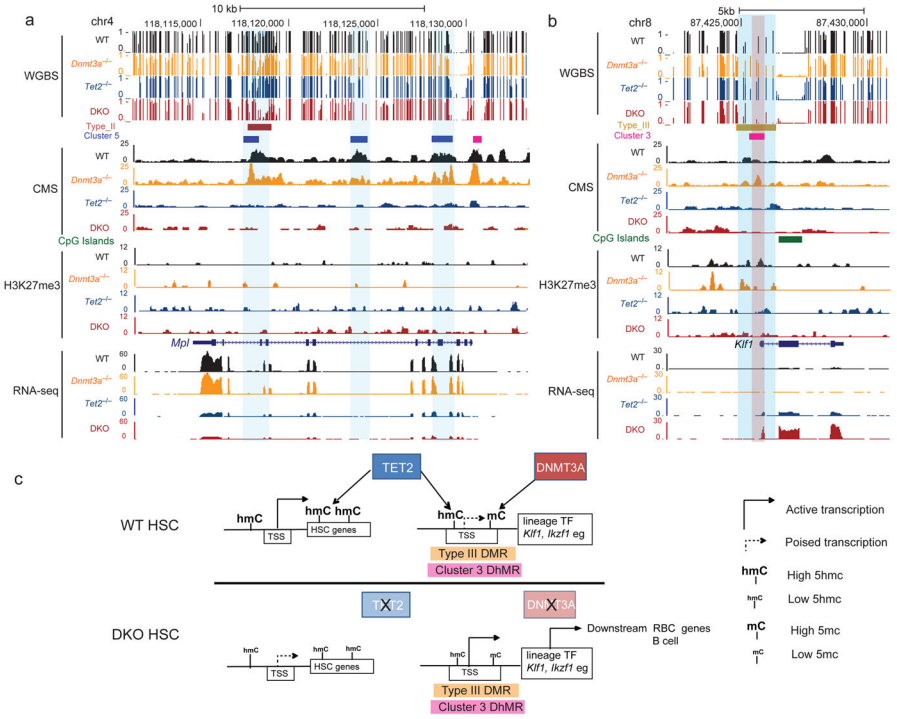

Figure 7. DNMT3A and TET2 cooperate to prevent activation of lineage-specific transcription factors in HSCs

a) Comprehensive epigenomic dynamics at the *Mpl* locus in HSCs with tracks WGBS, CMS, H3K27me3 and gene expression (RNAseq). Blue shaded area shows the cluster 5 DhMR located in the gene body of *Mpl*. **b)** Comprehensive epigenomic dynamics at the *Klf1* locus in HSCs. Blue shaded area shows the Type III DMR located in the promoter and 5′UTR regions of *Klf1*. Red shaded area shows the Cluster 4 DhMR located in the TSS and 5′UTR region in which hmC increases in *Dnmt3a*$^{-/-}$ HSC and decrease in *Tet2*$^{-/-}$ and DKO HSCs. The y axis of CMS and H3K27me3 sections shows signal of ChIP-seq occupancy in units of reads per million mapped reads per base pair (rpm/bp) in **a** and **b. c, d)** Model shows the action mode of DNMT3A and TET2 in activating HSC gene expression and repressing lineage-specific transcription factor expression.