DNMT3A loss drives enhancer hypomethylation in FLT3-ITD-associated leukemias

Liubin Yang*[1,4], Benjamin Rodriguez*[2], Allison Mayle*[1,4], Hyun Jung Park[2], Xueqiu Lin[2,3], Min Luo[4], Mira Jeong[4], Choladda V. Curry[5], Sang-Bae Kim[6], David Ruau[#9], Xiaotian Zhang[1,4], Ting Zhou[10], Michael Zhou[11], Vivienne I. Rebel[10], Grant A. Challen[8], Berthold Gottgens[9], Ju-Seog Lee[6], Rachel Rau[7], Wei Li[2,^] and Margaret A. Goodell[1,2,4,6, ^]

1 Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA
2 Dan L. Duncan Cancer Center and Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA
3 Department of Bioinformatics, School of Life sciences and Technology, Tongji University, Shanghai 20092, China.
4 Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, Texas 77030, USA
5 Department of Pathology and Immunology, Texas Children's Hospital, Baylor College of Medicine, Houston, Texas 77030, USA
6 Department of Systems Biology, Division of Cancer Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas, USA.
7 Department of Pediatrics, Baylor College of Medicine, Houston, Texas 77030, USA
8 Division of Oncology, Washington University School of Medicine, St. Louis, Missouri 63110, USA
9 Wellcome Trust/MRC Stem Cell Institute, Cambridge, UK
10 Greehey Children's Cancer Research Institute and Department of Cellular and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA
11 Rice University, Houston, Texas 77030

*equal contribution
Now at AstraZeneca, Cambridge, UK
^ Co-senior authors

Contact Information:

Margaret Goodell
Baylor College of Medicine
713-798-1265
Goodell@bcm.edu

Wei Li
Baylor College of Medicine
713-798-7854
WL1@bcm.edu

Character count 70,159

Running Title: DNMT3A in *FLT3-ITD* leukemia

SUMMARY

De novo DNA *METHYLTRANSFERASE (DNMT) 3A* is among the most frequently mutated genes in hematologic malignancies. However, the mechanisms through which DNMT3A normally suppresses malignancy development are unknown. Here, we show that DNMT3A loss synergizes with the FLT3 internal tandem duplication (ITD) in a dose-influenced fashion to generate rapid lethal lymphoid or myeloid leukemias similar to their human counterparts. Loss of DNMT3A leads to reduced DNA methylation, predominantly at hematopoietic enhancer regions in both mouse and human samples. Myeloid and lymphoid diseases arise from transformed murine HSCs. Broadly, our findings support a role for DNMT3A as a guardian of the epigenetic state at enhancer regions, critical for inhibition of leukemic transformation.

SIGNIFICANCE

Epigenetic regulators, including *DNA METHYLTRANSFERASE 3A* (*DNMT3A*), have emerged as potent tumor suppressors in many hematologic malignancies, but the mechanisms conferring this role are unknown. Here, we show that combined loss of *DNMT3A* with the leukemia-associated *FLT3*-internal tandem duplication (*FLT3-ITD*) in mice mimics human disease, with DNMT3A dosage influencing leukemia type. Both mouse and human $DNMT3A^{mut}$ leukemias are characterized by diminished DNA methylation in regulatory regions, particularly enhancers, suggesting a central role for DNMT3A in maintaining epigenetic integrity to prevent transformation. These murine models of *DNMT3A* AML and early immature T-ALL provide a venue to further study the mechanism of *DNMT3A/FLT3-ITD* leukemia and to test potential therapies.

HIGHLIGHTS

- *Dnmt3a* elimination and *FLT3-ITD* preferentially initiate early immature-like T-ALL
- Heterozygous *Dnmt3a* loss and *Flt3-ITD* initiate AML in mouse models
- *Dnmt3a*-associated myeloid and lymphoid leukemias arise from HSCs
- *Dnmt3a* loss leads to hypomethylation of active hematopoietic enhancers

INTRODUCTION

DNA methylation is an epigenetic process that influences cell fate. While aberrant DNA methylation is observed in many types of malignancies (You and Jones, 2012), a causative role for such changes has been difficult to identify. Recently, mutations in the *de novo* DNA methyltransferase *(DNMT) 3A* have been found in a variety of hematologic malignancies, suggesting a central role for this epigenetic regulator in preventing disease development (Reviewed in (Yang et al., 2015). Approximately 25% of patients who have myeloid or lymphoid malignancies, including acute myeloid leukemia (AML) (Ley et al., 2010) and T-cell acute lymphoblastic leukemia (T-ALL) (Grossmann et al., 2013; Roller et al., 2013; Van Vlierberghe et al., 2013) harbor DNMT3A mutations. In AML, a hotspot mutation at Arginine 882 (R882) is most prevalent, occurring in about 60% of $DNMT3A^{mut}$ patients (Ley et al., 2010). This alteration is thought to act as a dominant-negative, depleting most of the DNMT3A activity in the cell except for about 20% of wild-type activity (Kim et al., 2013; Russler-Germain et al., 2014). In T-ALL, the R882 mutation is less prevalent, and about 62% of $DNMT3A^{mut}$ patients harbor homozygous or compound heterozygous (biallelic) mutations (Grossmann et al., 2013; Roller et al., 2013). Together, these findings indicate that DNMT3A acts as a classic tumor suppressor, with a loss of most, or all, of its function promoting malignancy development..

Dnmt3a is highly expressed in hematopoietic stem cells (HSCs), where its loss promotes HSC self-renewal at the expense of efficient differentiation (Challen et al., 2012). In AML patients, *DNMT3A* mutations can be found in non-leukemic and leukemic cells in the peripheral blood (Corces-Zimmerman et al., 2014; Shlush et al., 2014). Moreover, blood cells with *DNMT3A* mutations can be found in healthy individuals, particularly with age (Genovese et al., 2014; Jaiswal et al., 2014; Xie et al., 2014). *DNMT3A* mutations in humans are associated with increased risk of leukemia, but alone are insufficient for transformation. The presence of *DNMT3A* mutations in HSCs that can behave relatively normally, and the latency of disease

development in individuals that harbor *DNMT3A*^{mut} HSCs, suggests that secondary mutations are key in driving the particular type of disease development.

Previous work with mice transplanted with *Dnmt3a* knock-out (KO) cells is consistent with the view that mutations in *DNMT3A* alone are insufficient for disease. Mice transplanted with *Dnmt3a* KO HSCs developed a variety of hematologic malignances, such as T-ALL, B-ALL, myelodysplastic syndrome (MDS), AML, and CMML with a long latency (3 to 14 months) (Celik et al., 2014; Mayle et al., 2014). Here we sought to combine *Dnmt3a* ablation with a specific additional mutation to investigate the mechanisms through which loss of DNMT3A promotes leukemia development. Because *DNMT3A* mutations are found in both lymphoid and myeloid leukemia, we introduced a secondary mutation that co-occurs with *DNMT3A*^{mut} in both lineages. In AML, 30% of patients with *DNMT3A* mutations harbor internal tandem duplications (ITD) in the fms-like tyrosine kinase 3 (*FLT3-ITD*) (Ley et al., 2010). *DNMT3A* and *FLT3-ITD* mutations both commonly occur in early immature T-ALL (Van Vlierberghe et al., 2013). We therefore reasoned that *FLT3-ITD* expression in *Dnmt3a*-KO mice could be used to study the role of *DNMT3A* in both lymphoid and myeloid malignancies.

RESULTS

Dnmt3a loss accelerates FLT3-ITD lymphoid leukemia

We sought to establish a model with both DNMT3A loss and *FLT3-ITD* expression. Because expression of *FLT3-ITD* via retrovirus can generate murine T-ALL (Kelly et al., 2002), we first used this strategy in *Dnmt3a*-KO cells. We deleted *Dnmt3a* in 8-week-old *Mx1-Cre*$^{+}$*;Dnmt3a*$^{fl/fl}$ mice, using polyinosinic-polycytidylic acid (pIpC) to generate animals with *Dnmt3a*$^{-/-}$ bone marrow. From these mice, we transduced 5-fluorouracil-stimulated hematopoietic stem and progenitor cells with an MSCV retrovirus containing FLT3-ITD-IRES-GFP (*3aKO/FLT3-ITD*) or IRES-GFP alone (3aKO) and transplanted the cells into lethally irradiated recipients. A separate

group of mice, *Mx1-Cre*$^{+}$*;Dnmt3a*$^{+/+}$ were also transduced with FLT3-ITD-IRES-GFP (*FLT3-ITD*), and IRES-GFP alone (WT) (Figure 1A). All mice received pIpC injections to control for possible interferon-mediated effects (Baldridge et al., 2010).

Mice transplanted with *FLT3-ITD* or *3aKO/FLT3-ITD* bone marrow cells developed leukemia. Strikingly, *3aKO/FLT3-ITD* mice had significantly shorter survival times (79 days vs. 116 days) than did *FLT3-ITD* mice (Figure 1B). Both groups showed weight loss, splenomegaly, and thymomegaly (Figures 1C and 1D) with widespread GFP+ cell infiltration in the bone marrow (Figure 1E). Notably, the *3aKO/FLT3-ITD* group had larger spleens and smaller thymuses (Figures 1C and 1D). Immunophenotyping revealed GFP+ populations of T-cells that expressed surface markers found on immature thymocytes and progenitor cells (CD4+CD8+CD25+; Figures 1E and S1A). At this time point, mice transplanted with cells from the 3aKO-alone showed no overt phenotype (Figure 1, S1). Histological examination revealed extensive infiltration of peripheral blood, bone marrow, and spleen (Figure 1F) and nonhematopoietic organs (liver, lung and kidney) by leukemic cells that were cytoplasmic $CD3^{+}$ and MPO^{-} (Figures S1B and S1C). Consistent with previous reports using the retroviral model in mixed background B6 mice (Kelly et al., 2002), we diagnosed the majority of *3aKO/FLT3-ITD* and *FLT3-ITD* mice (90% and 78%, respectively) as having a T-cell disease, specifically precursor T-cell lymphoblastic lymphoma/leukemia (similar to human T-ALL), based on the Bethesda proposal for classification of murine lymphoid neoplasms (Morse et al., 2002). The leukemic cells were capable of self-renewal as demonstrated by transplantation to sublethally irradiated WT recipients (Figure S1D). In addition, 22% of mice transplanted with *FLT3-ITD* cells and 5% with *3aKO/FLT3-ITD* died from myeloproliferative disease and 5% of *3aKO/FLT3-ITD* mice died of B-cell ALL (Figure S1E). Compared to the *FLT3-ITD* T-ALL cells, the *3aKO/FLT3-ITD* T-ALL cells were more proliferative and had higher rates of early apoptosis by Ki-67 and annexin V staining (Figures 1G and 1H). These findings indicate that loss of *Dnmt3a* promotes aggressive T-ALL in mice that acquire *FLT3-ITD*.

Dnmt3a-related lymphoid leukemia upregulates myeloid programs

To understand how loss of *Dnmt3a* contributes to lethal lymphoid leukemia in mice, we studied the global expression profiles of the *3aKO/FLT3-ITD* and *FLT3-ITD* T-ALLs by RNA-seq. We compared the two sets of leukemic cells (>95% GFP+ and $CD4^+CD8^+$) and sorted $CD4^+CD8^+$ wild-type thymocytes from transplanted mice as a control. *3aKO* mice were healthy through the time-course of this experiment, and $CD4^+CD8^+$ cells sorted from these mice were analyzed, but had negligible differences when compared to the WT controls (data not shown). Comparison of *3aKO/FLT3-ITD* with *FLT3-ITD*-only cells revealed 696 differentially expressed genes (507 upregulated and 189 downregulated in the *3aKO/FLT3-ITD* group) (Figure S2A). Gene ontology (GO) analysis showed that in the *FLT3-ITD*-only group, the upregulated genes were functionally related to the extracellular region, whereas those that were upregulated in the *3aKO/FLT3-ITD*-positive mice included genes related to inflammation and immune response (Figure 2A). Surprisingly, using Ingenuity pathway analysis we found that myeloid gene sets were upregulated in the *3aKO/FLT3-ITD* group (Figure 2B), while genes involved in mature T-cell function were downregulated (Figure 2C). Other downregulated genes include several genes that are mutated or play a role in suppression of apoptosis in hematopoietic diseases (e.g. *Cd28, Rorc, Ephb6)* (Maddigan et al., 2011; Rohr et al., 2016; Tian et al., 2015). Gene set enrichment analysis revealed that upregulated genes in *3aKO/FLT3-ITD* cells were enriched for immature gene sets, including hematopoietic stem cells (e.g. *Gata2, H19*) and mouse early thymic progenitor expression. The upregulated genes were also enriched for myeloid, AML (e.g. *G0s2, Hmox1*, *Tbxas1*), and aging gene sets (Figure S2B). The enriched gene sets were highly differentially expressed in *3aKO/FLT3-ITD* compared to *FLT3-ITD* and WT cells (Figure 2C).

To investigate the extent to which the expression profiles of the mouse model recapitulate those of human disease, we assessed the significance of gene overlap with expression signatures derived from the Microarray Innovations in Leukemia (MILE) patient

research study (Haferlach et al., 2010). Specifically, we compared genes differentially expressed in the *3aKO/FLT3-ITD* leukemia model to signatures that distinguish human disease subtypes (leukemia precursor, ALL, and AML) from all other disease patients in MILE. The mouse model profile was strongly associated with genes characterizing human AML (odds ratio 15.5, p-val = 2.23E-16) (Figure 2D and S2C). Interestingly, the strength of association between the model and human ALL was considerably less than expected by chance (odds ratio 0.24, p-val = 1.12E-08). These data suggest that in the absence of *Dnmt3a,* murine lymphoid leukemic cells maintain the expression of T-cell surface markers, but activate a myeloid gene signature at the cost of mature T-cell genes, which is reminiscent of human early immature T cell leukemias that are enriched in aberrant myeloid genes (Van Vlierberghe et al., 2011).

The T-ALL diagnosed in *3aKO/FLT3-ITD* mice is similar to early T-cell precursor T-ALL (ETP-ALL), which is associated with *FLT3* and *DNMT3A* mutations and the expression of stem and myeloid genes (Neumann et al., 2012; Neumann et al., 2013). The ETP classification is immunophenotypic, and requires the absence of the T-cell marker CD8 (and others) (Coustan-Smith et al., 2009; Neumann et al., 2012; Zhang et al., 2012). T-ALL can alternatively be classified solely on the basis of gene expression profiles. In this strategy, two distinct classes of immature vs mature, cortical T-ALL with unique survival and mutation profiles are evident (Van Vlierberghe et al., 2013). The immature T-ALL group, associated with dismal survival, is distinguished by a myeloid and stem cell gene expression signature and is enriched for *DNMT3A* mutations (Van Vlierberghe et al., 2013). While having some overlap with ETP-ALL, the immature classification includes cases with a variety of surface markers, including expression of CD4 and CD8 (Coustan-Smith et al., 2009; Van Vlierberghe et al., 2013). Considering the surface marker profile (CD4+CD8+), and myeloid and stem cell-associated gene expression signatures in the leukemia we observe, the *3aKO/FLT3-ITD* cells most closely resembles early immature T-ALL. To confirm the immature phenotype, we tested the expression of *Notch* pathway and early thymic progenitor-related genes by quantitative PCR and found

them to be downregulated and upregulated, respectively, as is expected in early immature T-ALL (Figure S2D). Therefore, these data indicate that *3aKO/FLT3-ITD* lymphoid leukemia is most similar to human early immature T-ALL.

Dnmt3a-loss initiates myeloid and lymphoid Flt3-ITD leukemia

The strategy described above was highly effective in generating early immature-like T-cell leukemia, which is also associated with *DNMT3A* and *FLT3* mutations in patients. More frequently however, this combination of mutations is seen in acute myeloid leukemia (AML) (Ley et al., 2010; Neumann et al., 2012). In an attempt to generate myeloid disease, we turned to a *Flt3-ITD* knock-in model (referred to as *Flt3-ITD^{KI}*), in which heterozygous mice develop myeloproliferative disease (Lee et al., 2007). This enabled us to compare disease development in *Flt3-ITD^{KI}* mice in the presence or absence of *Dnmt3a*.

We crossed *Flt3-ITD^{KI}* mice with our *Dnmt3a$^{fl/fl}$* mice expressing a tamoxifen inducible (ER) Cre (Hinkal et al., 2009). The *ER-Cre^{+} Dnmt3a$^{fl/fl}$ Flt3$^{+/ITD}$* and control *ER-Cre^{-} Dnmt3a$^{fl/fl}$ Flt3$^{+/ITD}$* mice were treated with tamoxifen to generate *3aKO/Flt3-ITD^{KI}, 3aKO*, and *Flt3-ITD^{KI}* mice, after which their bone marrow was transplanted into lethally irradiated recipients (Figure 3A). Consistent with previous reports, ablation of *Dnmt3a* alone led to a broad spectrum of hematologic diseases (Celik et al., 2014; Mayle et al., 2014), and the *Flt3-ITD^{KI}* group did not develop lethal disease after transplantation (Lee et al., 2007). Yet, all of the *3aKO/Flt3-ITD^{KI}* group died of acute leukemia with a median survival of 225 days, which is not significantly different than the *3aKO* alone (Figure 3B), although the spectrum of diseases is different (Figure 4G), and the median survival is significantly shorter than the *Flt3-ITD^{KI}* group, which all survive longer than 400 days. The *3aKO/Flt3-ITD^{KI}* mice exhibited leukocytosis and myeloproliferation (Figure 3C and 3D), and 8/14 mice were diagnosed with AML, as indicated by myeloid cell surface markers and leukemic cell infiltration in bone marrow and spleen (Figures 3E and S3)

(Kogan, 2002). Two mice exhibited mixed lineage leukemia, consisting of a blast-filled bone marrow (Figure 3F) that stained for either myeloid or T-cell surface markers but also had an enlarged thymus infiltrated by donor-derived T-lymphoblasts (Figure 3G). The remaining four mice were diagnosed with T-ALL.

To verify complete deletion of *Dnmt3a* in these animals, we tested a subset of AML and T-ALL samples and bone marrow and thymus tissue from mice with mixed-lineage leukemia for *Dnmt3a* deletion by PCR. Although *ER-Cre* generally recombines loxP sites with high efficiency, we noted residual floxed allele in most of the AML samples (representative samples in Figure 3H). In contrast, *Dnmt3a* deletion appeared to be complete in most T-ALL samples. We were initially surprised to find both lymphoid and myeloid leukemia in the *Flt3-ITD* knock-in mouse model known for myeloid disease but these data suggested a correlation between *Dnmt3a* dosage and disease type, with a remaining allele of *Dnmt3a* more likely found in myeloid rather than lymphoid leukemia. As discussed above, AML patients most frequently harbor mono-allelic mutation of *DNMT3A* at R882, which is thought to retain some DNMT3A activity, while T-ALL patients much more commonly harbor biallelic *DNMT3A* mutations, consistent with more complete loss-of-function.

Heterozygous Dnmt3a loss cooperates with Flt3-ITD to initiate acute myeloid leukemia

To test the hypothesis that heterozygous deletion of *Dnmt3a* is more likely to result in myeloid disease, we generated $Dnmt3a^{+/-}$;$Flt3^{ITD/+}$ (*3aHet/Flt3-ITD*^{KI}) mice by crossing *ER-Cre*$^{+}$; $Dnmt3a^{fl/fl}$ mice with $Dnmt3a^{+/+}$;$Flt3^{+/ITD}$ mice, and bone marrow from these mice was transplanted into lethally irradiated recipients. Nine out of ten of the *3aHet/Flt3-ITD*^{KI} mice analyzed died from AML with a median survival of 270 days (Figure 4A). These *3aHet/Flt3-ITD*^{KI} mice developed myelocytosis in the peripheral blood (Figure 4B and C) with increased monocytes and neutrophils (Figure 4D). Histological examination revealed leukemic blast cell

infiltration in the bone marrow (Figure 4E), peripheral blood (Figure S4A), and extramedullary organs (liver, lung, kidney) (Figure S4B). The blasts expressed myeloid cell surface markers (Mac-1^{+}Gr-1^{+}c-Kit^{+}) (Figures 4F and S4C). The disease was recapitulated when transplanted to secondary recipients (Figure S4D). Together, these features indicated a diagnosis of AML in the *3aHet/Flt3-ITD^{KI}* transplanted mice.

These studies indicate that partial *Dnmt3a* loss cooperates with *Flt3-ITD^{KI}* to instigate AML with monocytic and neutrophilic bias consistent with findings of *DNMT3A* mutations in M4/M5 AML patients (Cancer Genome Atlas Research, 2013; Yan et al., 2011). By transcriptome analysis, genes related to myeloid cell development and function were enriched in the *3aHet/Flt3-ITD^{KI}* leukemic cells compared to *Flt3-ITD^{KI}* progenitors (c-Kit^{+}Sca-1^{+}Lin-) (Figure S5).

Of the cohort of mice transplanted with *3aHet/Flt3-ITD^{KI}* cells, 90% developed AML (Figure 4G), in striking contrast to the *3aKO/Flt3-ITD^{KI}* cohort, of which only 57% developed AML with the rest exhibiting T-ALL or mixed lineage leukemia, and the *3aKO* cohort, which we have previously reported develop a variety of malignancies (Mayle et al., 2014). Together, these observations indicate a relationship between *Dnmt3a*-dosage and the type of disease development.

Dnmt3a-associated leukemia arises from HSCs

In AML patients, *DNMT3A* mutations have been found in HSCs, acting as a pre-leukemic lesion (Corces-Zimmerman et al., 2014; Ding et al., 2012; Shlush et al., 2014). However, whether this *DNMT3A* mutant cell, or other downstream populations, can serve as the leukemia stem cell (LSC) is not known. Indeed, in some leukemias, LSCs arise from committed progenitors that acquire the ability to self-renew such as *MLL-AF9* leukemia (Krivtsov et al., 2006), whereas in others, it arises from an HSC such as in *c-Cbl*-related leukemia (Rathinam et al., 2010). In our model, all cells carry the *Dnmt3a* mutation as well as the *FLT3-ITD*, so it was

not clear whether HSCs, their progeny, or both, would transmit the disease. To address this, we tested different cell populations for their ability to generate leukemia.

We sorted *Dnmt3a*-KO myeloid progenitors ($Lin^{-}Sca\text{-}1^{-}c\text{-}Kit^{+}$), lymphoid progenitors ($Lin^{-}IL7ra^{+}Sca\text{-}1^{med}c\text{-}Kit^{med}$), and HSCs (Side $population^{+}Lin^{-}Sca\text{-}1^{+}c\text{-}Kit^{+}$), and transduced them with the *FLT3-ITD* retrovirus before transplanting them into lethally irradiated recipients. We observed disease phenotype only in the mice transplanted with transduced HSCs (Figure 5A), and the T-ALL we saw previously was recapitulated (Figure 5B). These data indicate that this *Dnmt3a*-deficient lymphoid leukemia originated from a transformed HSC rather than downstream progenitor cells. To determine whether each population could propagate the disease, we also sorted various GFP+ populations from the mice after T-ALL development, and transplanted them. In this cohort, T-ALL was recapitulated in most mice (Figure 5C and 5D), suggesting that more committed populations propagated the disease, despite only the HSCs being able to initiate the disease.

In the *Flt3-ITD*^{KI} AML model, we tested for leukemia initiation by inducing *Dnmt3a* deletion by tamoxifen injection and transplanting purified cells one week later. In this short time frame, *Dnmt3a*-deleted HSCs would likely have only limited contribution to downstream populations. Thus, any contribution of sorted progenitors or differentiated cells to AML development would indicate that *Dnmt3a*-deletion may exert its effects downstream of the HSC. We sorted LT-HSC, ST-HSC, MPP, CMP, GMP, and CLP and transplanted them into lethally irradiated WT recipients (Figure 5E). Only the LT-HSC population showed long-term engraftment and development of the AML phenotype by peripheral blood analysis, whereas other populations showed only transient blood contribution, indicating more limited self-renewal (Figures 5F and 5G). These data establish that in *Dnmt3a*-related AML, the undifferentiated HSCs initiate the leukemia, in contrast to models of leukemia such as MLL-AF9 in which granulocyte/macrophage progenitors give rise to the disease (Krivtsov et al., 2006).

Loss of DNMT3A drives enhancer hypomethylation

Recent reports have shown a correlation between *DNMT3A*^{mut} and specific hypomethylation patterns in CN-AML (cytogenetically normal AML) patients (Qu et al., 2014; Russler-Germain et al., 2014). Because we were utilizing a genetically defined model, we could identify specific patterns of DNA methylation changes resulting from loss of DNMT3A in our mouse models. We first sought to assess these changes in the T-ALL model. We generated genome-wide DNA methylation maps of *3aKO/FLT3-ITD* and *FLT3-ITD* leukemic cells, and $CD4^+CD8^+$ WT thymocytes by whole genome bisulfite sequencing (WGBS; coverage 8-10x; 7.2 million CpGs with at least 5x coverage across all samples). The overall methylation differences between sample groups were minor, but all comparisons revealed loci with increased and loci with decreased DNA methylation, similar to what has been observed in human hematologic malignancies (Figueroa et al., 2013; Figueroa et al., 2010), yet the greatest magnitude of hypomethylation occurred in samples lacking *Dnmt3a* (Figure 6A).

We next focused on specific regions such as CpG islands, shores, and regulatory elements such as promoters and enhancers (Hon et al., 2013; Lara-Astiaso et al., 2014). In *3aKO/FLT3-ITD*-derived cells, nearly all of these features exhibited marked hypomethylation, with the greatest magnitude of change found in enhancer sites defined in hematopoietic tissues (bone marrow, spleen, thymus) and at the edges of large undermethylated regions (canyons) (Jeong et al., 2014); >15% of enhancers and canyons were hypomethylated, a proportion ~ 5-fold greater than observed by chance (Figure 6B and Figure S6A). In *FLT3-ITD*-only leukemic cells, by contrast, hypermethylation of enhancers and canyon edges was much more prominent (Figure 6B).

To confirm that *3aKO* hypomethylated enhancer sites were regulatory elements relevant to hematopoiesis, we examined them for enrichment of hematopoiesis-related transcription

factor-binding sites. We compared our findings to a database of 315 mouse chromatin immunoprecipitation (ChIP)-seq transcription factor studies involving over 400,000 regions (Ruau et al., 2013). The comparison revealed overrepresentation of more than 80 transcription factors at hypomethylated enhancer sites, including master regulators that determine cell fate during stem, myeloid and lymphoid cell differentiation such as *Fli1, Gfi1b*, and *Pu.1* (Figure S6B). Many of the transcription factors identified have also been implicated in hematopoietic diseases, including *Fli1*, *Lmo2*, and *Runx1* in T-ALL (Cleveland et al., 2014; Mok et al., 2014; Smeets et al., 2013; Smith et al., 2014), and *Pu.1* and *Runx1* in AML (Cancer Genome Atlas Research, 2013; Gerloff et al., 2015). These findings suggest that *Dnmt3a* loss results in the demethylation of previously methylated, functionally relevant enhancer regions during leukemogenesis, thus potentially increasing accessibility of these regions to relevant transcription factors.

To further confirm that loss of DNMT3A unveils enhancer sites, we generated maps of putative enhancers (Shen et al., 2013) based on H3K27ac and H3K4me1 histone peaks in leukemic *FLT3-ITD* and *3aKO/FLT3-ITD* cells by ChIP-seq and DNAse1 hotspots from mouse ENCODE database (Consortium et al., 2012). We identified 13,705 regions co-occupied by H3K27ac and H3K4me1 in *3aKO/FLT3-ITD* and assessed differential enrichment of H3K27ac and H3K4me1 at these regions. At these co-occupied regions, the interquartile-range distributions of both marks' signals showed greater intensity in *3aKO/FLT3-ITD* cells (Figure S6C). We observed differential enrichment of one or both marks in *3aKO/FTL3-ITD* cells at 29.4% (4031 / 13705) of regions tested (Figure S6D). We also observed an overwhelming trend towards increased occupancy of both H3K4me1 and H3K27ac, implying that many regions in *3aKO/FLT3-ITD* have acquired the chromatin marks *de novo* relative to *FLT3-ITD* alone (Figure 6C). We asked whether these regions with changes in H3K4me1 or H3K27ac also differed in DNA methylation. In regions with differential enrichment of both H3K4me1 and H3K27ac, the proportion of hypomethylated CpG sites was significantly greater in *3aKO/FLT3-ITD* than *FLT3-*

ITD (odds ratio 3.16, CI 2.89 – 3.44, P-val < 2.2e-16, Fisher's exact test) (Figure 6D). The increase in enrichment of H3K27ac and H3K4me1 peaks at hypomethylated regions in the leukemic cells that lack DNMT3A strongly suggests that DNMT3A regulates active enhancers.

Given the observations that DNA methylation of enhancers was reduced with loss of *Dnmt3a*, we considered whether exogenous expression of *Dnmt3a* would restore methylation levels at those sites. Leukemic cells derived from *3aKO/FLT3-ITD* secondary transplants were transduced with a retrovirus expressing *Dnmt3a* along with a GFP marker (MSCV-Dnmt3a-IRES-GFP) or GFP alone (MSCV-GFP) and were transplanted into sublethally irradiated tertiary recipients. Methylome maps of recipient bone marrow cells using RRBS (average of 19 million mapped reads per sample) revealed increased average DNA methylation in tumors with enforced *Dnmt3a* expression (Figure 6E and 6F). Furthermore, enhancers, CGI shores, and canyon edges all showed significant hypermethylation (Figure S6E). These data support our observations that DNMT3A is particularly active at enhancers. Expression of *Dnmt3a* did not affect the survival of the mice, consistent with *Dnmt3a* being more important for leukemic initiation than maintenance (despite an estimated 4-fold higher *Dnmt3a* expression than that in GFP-transduced samples (data not shown)).

While the *3aKO/FLT3-ITD* model showed clear methylation losses, we considered whether the heterozygous *Dnmt3a* AML model would exhibit similar effects on DNA methylation. We examined the methylome of three *3aHet/Flt3-ITD*^{KI} AML samples compared to three sorted stem and progenitor cells (KSL) samples from transplanted *Flt3-ITD*^{KI} mice by reduced representation bisulfite sequencing (RRBS) (total of 16-37 million reads with at least 13 million aligned reads). Overall mean CpG methylation was not different between the two genotypes (Figure S6F). However, of interrogated CpGs, there were more hypomethylated 1-kb regions than hypermethylated regions (Figure 6G). Further analysis of genomic regions revealed more frequent hypomethylation at enhancer regions defined as hematopoietic (bone marrow, spleen, and thymus) and more frequent hypermethylation of enhancers in lineages such as testes and

cerebellum (Figure 6H). Therefore, in the absence of one allele of *Dnmt3a*, loss of methylation was the most frequent event in the murine model of AML.

To more directly compare methylation changes in the *3aHet/Flt3-ITD*^{KI} AML model with the *3aKO/FLT3-ITD* T-ALL model, we repeated our methylation analysis on the T-ALL samples using RRBS. Even though RRBS is more biased toward CpG islands, which frequently show hypermethylation, we confirmed enhancers and canyons were dramatically undermethylated in the T-ALL leukemic cells (Figure S6G). Therefore, in both models, RRBS analysis demonstrated that canyon edges were subject to the most frequent methylation changes, with CGIs and promoters having the least frequent methylation changes (Figure 6H). Although we observe more extreme hypomethylation in the T-ALL model (possibly due to the lower dosage of DNMT3A), the shared hypomethylation of enhancers and frequent altered methylation at canyon edges strongly suggest that these regions in particular are regulated by DNMT3A.

Re-expression of DNMT3A inhibits transcription factor binding at enhancers

Having observed significant hypomethylation particularly at enhancer sites in the absence of DNMT3A, we considered whether this had functional consequences on transcription factor binding. We hypothesized that remethylation of enhancer binding sites in the presence of enforced DNMT3A expression would result in lower occupancy of some transcription factors. To test this, we utilized the T-ALL model because of the homogeneity of this leukemia and the complete absence of *Dnmt3a* expression. In addition to being implicated in T-cell development and T-ALL (Smeets et al., 2013; Smeets et al., 2014), FLI1 exhibited robust gene expression across conditions, was not differentially expressed, and the FLI1 consensus binding motif was enriched in hypomethylated enhancer sequences (P = 3.59e-15, Bonferroni correction), so we decided to test the effect of re-methylation on FLI1 binding. We performed ChIP-seq for FLI1 in *3aKO/FLT3-ITD* leukemic cells transduced with either the Dnmt3a-expressing retrovirus, or that of GFP alone, identifying 6,492 FLI1 binding sites. In the absence of DNMT3A (GFP-transduced

cells) compared to the presence of DNMT3A (Dnmt3a-transduced cells), FLI1 binding was significantly higher at 472 sites. This suggests re-expression of *Dnmt3a* exerted a mostly disruptive effect on FLI1 binding. Most differential binding occurred outside promoters, frequently in hematopoietic enhancer regions (Figure 7A). Next we investigated the genes and biological processes associated with differential occupancy of FLI1. Increased promoter binding was observed at genes involved in regulation of transcription and signal transduction (e.g., Wnt), decreased binding was associated with embryonic lethality, cell cycle, and constituents of the spliceosome (Figure 7B). Remarkably, 76% of distal sites with increased binding in absence of *Dnmt3a* (255 of 337) occurred in hematopoietic enhancer regions and were associated with genes involved in human leukemia (including AML) as well as with relevant mouse model phenotypes including increased tumor incidence, abnormal proliferation and abnormal differentiation of T cells (Figure 7C). We identified no biological processes or ontologies enriched among the limited number of distal sites with decreased binding (60 of 105) in such regions. Finally, we asked whether increased binding of FLI1 was concomitant with the hypomethylation we observed in *3aKO/FLT3-ITD* leukemia. Sites of increased FLI1 occupancy were significantly enriched for hematopoietic enhancer regions that are also hypomethylated in the absence of DNMT3A (odds ratio 2.35 P = 5.26e-06, Fisher's test). Whether FLI1 has a role in this particular leukemia is not clear, but it serves to support the concept that DNA methylation changes can affect transcription factor occupancy at enhancer sites (Figure 7D).

Mutant DNMT3A/FLT3 AML patients also exhibit enhancer DNA hypomethylation

Given the findings of DNA methylation changes at enhancer regions when *Dnmt3a* is lost and ectopically expressed, we considered whether patients with $DNMT3A^{mut}$ also exhibited DNA methylation loss at hematopoietic enhancer regions. For our analysis, we selected from the TCGA cohort patients with the $DNMT3A^{R882}$ mutation as this group is of sufficient size to allow statistical analysis, and the R882 mutation has been functionally characterized as a

dominant negative (Kim et al., 2013; Russler-Germain et al., 2014). Patients with non-R882 mutations were excluded from our analysis due to their uncharacterized nature and small numbers. Leukemic cells with the R882 mutation are thought to retain around 20% of wild-type activity, so we cannot exclude some effects peculiar to this mutant that would be distinct from the *Dnmt3a*KO mouse situation.

We mined the TCGA AML patient dataset to compare the DNA methylation maps of AML patients with co-mutations of *DNMT3A*^{R882} and *FLT3* to those with either mutation alone or to patients of normal karyotype (WT NK AML) lacking either mutation (Figure 8A) (Cancer Genome Atlas Research, 2013). To better control for other mutations that may have epigenetic effects, patients with mutations in IDH1, IDH2, TET1, or TET2 were excluded. All patients harbor cytoplasmic NPM1. We observed a dramatic increase in CpG hypomethylation (n = 35,436 sites) in patients with R882 and *FLT3* mutations compared to *FLT3* alone (Figure 8B), a unique result (among the group comparisons) in that it completely spanned the inter-quartile range of changes at sites of variable methylation in the four patient groups. The hypomethylated sites were significantly enriched for enhancers and to a lesser extent, CpG island shores (Figure 8C). While limited hypomethylation (n = 3437 sites) was observed in patients with FLT3 and R882 co-mutation compared to R882 alone (Figure 8B), the methylation states we observed in enhancer regions at sites of variable methylation clearly discriminated the two patient groups from one another, though to a lesser extent than either from the WT or FLT3 alone groups (Figure 8D). We mapped the enhancers to gene regulatory domains and assessed what biological functions most distinguished those hypomethylated in R882 + FLT3 mutant patients relative to the larger population of enhancers. They were enriched for sequence-specific (including enhancer) binding transcriptional regulators, corepressors, genes involved in cell fate specification, targets of PTEN, p53, PRC2, as well as genes with conserved homeobox sites (Figure S7A). To investigate the transcriptional profile of genes with hypomethylated enhancers, we generated linear models of RNA-seq counts for 69 AML patients in the TCGA dataset,

contrasting the *DNMT3A* mutation status among patients harboring *FLT3* mutations. Ingenuity knowledge-based functional enrichment analysis of differentially expressed, hypomethylated-enhancer associated genes showed over-representation in functions supporting hematopoietic progenitor cells (including quantity, proliferation, and maturation) as well as disease states, including myeloproliferative disorders and AML (Figure 8E). Applying a more unbiased approach, GSEA, we observed a strong, negative correlation with targets of MYC and E2F, as well as genes involved in the G2/M cell cycle checkpoint (Figure S7B). We also observed overexpression of homeobox genes, including the *HOXB* cluster (members 2 – 5) consistent with other reports (Yan et al., 2011). These data indicate that in AML patients with mutated FLT3, the loss of enhancer methylation observed with mutation of DNMT3A may contribute to deregulation of transcriptional programs key to cell identity and normal hematopoietic function, thus promoting leukemogenesis.

DISCUSSION

Here, we have generated a mouse model of human *DNMT3A*^{mut} leukemia and show that deletion of *Dnmt3a* cooperates with the *FLT3-ITD* mutation to initiate both myeloid and lymphoid leukemias, establishing a powerful model to investigate the mechanisms of *Dnmt3a*-associated transformation. Because the promoter used to drive *FLT3-ITD* expression in mice has been reported to affect disease development (Lee et al., 2007), we utilized both a retroviral transduction model, which leads primarily to lymphoid disease (Kelly et al., 2002), and a knock-in model, which leads to myeloproliferative disease (Lee et al., 2007).

We have previously reported that loss of *Dnmt3a* predisposes transplanted mice to a variety of hematologic diseases such as B- and T-ALL, MDS, AML, CMML and primary myelofibrosis, (Mayle et al., 2014). Here, we combined loss of *Dnmt3a* with *FLT3-ITD* retroviral overexpression, which led to rapid T-cell leukemia that resembled early immature lymphoid leukemia. The *Flt3-ITD* knock-in combined with *Dnmt3a* loss led primarily to myeloid leukemia

that resembled the M4/M5 subtypes of AML, which are highly associated with *DNMT3A* mutations (Yan et al., 2011). Thus, combining *FLT3-ITD* with loss of *Dnmt3a* drove development of specific diseases. This suggests that loss of *Dnmt3a* establishes a cellular milieu that is permissive for transformation down either the myeloid or lymphoid lineage, and that the secondary hits play a key role in disease specification and latency.

Importantly, the *Dnmt3a/Flt3-ITD* knock-in model also generated a number of lymphoid leukemias that were more prevalent with complete absence of *Dnmt3a*, whereas heterozygous *Dnmt3a* loss almost exclusively led to myeloid leukemia. These data establish that *Dnmt3a* dosage influences the lineage outcome of *Flt3-ITD* leukemia. This observation is consistent with the observation that the mutational spectrum of *DNMT3A* is distinct among patients with lymphoid vs. myeloid leukemias (reviewed in (Yang et al., 2015)). Yet, the mechanism behind the differential dependence on *DNMT3A* is not clear. During the natural evolution of leukemia, patients will initially have a single mutant allele, and probably sustain that state for some time. Thus, we can conjecture that a patient acquiring a dominant-negative R882 mutation would have ~20% of WT DNMT3A activity, while those with a nonsense mutation might have 50% activity remaining. Both these levels will be haploinsufficient, and may influence the differentiation of HSCs, such that myeloid differentiation may be favored in the R882 case, where lymphoid differentiation is favored when DNMT3A is present at a higher level. This suggests that a greater amount of DNMT3A maybe required for normal lymphoid differentiation. In T-ALL, the second *DNMT3A* lesion is probably acquired later and likely leads to rapid leukemia development down the lymphoid pathway.

Our data also reinforce the concept that DNMT3A functions primarily at the stem cell level. Both the myeloid and lymphoid leukemias are initiated by HSCs transformed with *FLT3-ITD.* Our findings indicate that *Dnmt3a* loss in stem cells likely primes the pre-leukemic clone for a secondary oncogenic hit at the stem cell level. These data are consistent with the highly specific effect that *Dnmt3a* loss has on expansion of the stem cell compartment (Challen et al.,

2012). In addition, the inability to slow the leukemia by re-expression of DNMT3A suggests that DNMT3A loss is important for leukemia initiation, but probably less so for maintenance. This association with initiation and stem cells is also seen in patients, as *DNMT3A*^{mut} can be found in multiple lineages with *DNMT3A*^{mut} AML, indicating it first emerges in HSC- or progenitor-generated clones (Corces-Zimmerman et al., 2014; Jan et al., 2012; Shlush et al., 2014). Furthermore, in patients with AML, *DNMT3A*^{mut} are thought to be the initiating event, as they are almost always found at the highest variant allele frequencies (Welch et al., 2012). Together, these data lead to a model in which an HSC acquires a *DNMT3A* mutation, which expands and remains as a reservoir for clonal expansion until a new genetic lesion, such as *FLT3-ITD*, is acquired and leads to full leukemic transformation.

Methylation profiling of *Dnmt3a*^{mut} leukemias showed loss of DNA methylation at hematopoietic enhancer regions, and this hypomethylation can impact transcription factor binding as we showed for FLI1. Similarly, AML patients with *DNMT3A*^{R882} and FLT3 mutation also exhibited hypomethylation at enhancer regions relative to other non-*DNMT3A/FLT3*-mutant AML. These observations are consistent with other reports of hypomethylation observed in *DNMT3A*^{mut} AML patients (Qu et al., 2014; Russler-Germain et al., 2014). In the T-cell ALL retroviral transduction model, global hypomethylation was observed in contrast to leukemia developing from *FLT3-ITD* transduction alone. In the knock-in AML model in which only one copy of *Dnmt3a* was lost, both hyper- and hypomethylation of enhancer sites was observed relative to a non-leukemic stem and progenitor population with Flt3-ITD. Moreover, expression of genes associated with myeloid function was de-repressed. The hypermethylated enhancers may be related to aberrant methylation function of the remaining DNMT3A in the presence of FLT3-ITD. Together, these experiments show that *3aKO* leads to hypomethylation at distal gene regulatory regions, especially at active enhancers together with *FLT3* mutations, indicating that DNMT3A regulates these regions by DNA methylation. Distal regulatory elements can mediate disease as demonstrated by rearrangements or removal of single enhancers that can initiate

leukemia (Groschel et al., 2014). Furthermore, during HSC differentiation, changes in DNA methylation are mapped to regulatory elements and transcription factor binding sites (Cabezas-Wallscheid et al., 2014), supporting our findings that DNMT3A regulates distal regulatory elements by DNA methylation in leukemogenesis. *Dnmt3a* loss in non-leukemic HSCs can also induce hypomethylation at distal regulatory elements (Jeong et al., 2014), indicating that our observations are *Dnmt3a*-specific. Our discoveries imply that during leukemogenesis, DNMT3A acts as a tumor suppressor, guarding stem-cell regulatory regions through methylation at enhancers (Hnisz et al., 2013; Hon et al., 2013; Ziller et al., 2013).

Overall, these experiments show that *Dnmt3a* loss drives leukemogenesis in multiple lineages, recapitulating *DNMT3A*^{mut} AML and T-ALL in a *Dnmt3a*-dosage dependent manner. These mouse models serves as a valuable tools to study and develop therapeutic targets for *DNMT3A*-related leukemia.

EXPERIMENTAL PROCEDURES

Cloning and retrovirus, peripheral blood analysis, antibody cocktails, histopathology and immunohistochemistry, bioinformatics analysis, quantitative real-time PCR, chromatin immunoprecipitation are described in Supplemental Experimental Procedures.

Vector and animal models

All animal experiments were reviewed and approved by the Institutional Animal Care and Use Committee of Baylor College of Medicine. *FLT3-ITD* was subcloned into MSCV-IRES-GFP from vector pcDNA-FLT3-ITD (a gift from Jonathan Licht). Flt3-ITD mice were obtained from The Jackson Laboratory. *Mx1-cre; Dnmt3a*$^{fl/fl}$ and *ER-T2-cre; Dnmt3a*$^{fl/fl}$ mice carrying the CD45.2 allele (Challen et al., 2012) on the C57BL/6 background were used to harvest bone marrow cells, which were then transplanted by retro-orbital injection into lethally irradiated (9.5 Gy) CD45.1 syngeneic mice. Tamoxifen treatment was performed with 5 intraperitoneal injections of 100μg each, Mx1-cre mice were induced with 6 intraperitoneal injections of pIpC (Sigma, 300 μg per mouse in PBS) on alternating days.

Mouse Phenotype analysis

Tissues—femur, tibiae, iliac crests, spleen, and thymus—were harvested and made into single-cell suspension by manual trituration. Peripheral blood was obtained retro-orbitally and cells were stained with antibodies at 1:100 dilution at 4°C, analyzed on a LSR II flow cytometer (BD Biosciences), or sorted with a BD FACSAria II cell sorter (BD Biosciences) (Mayle et al., 2013). For histologic studies, fresh tissues were made into touch preps or fixed and mounted for H&E staining.

Cell sorting

Bone marrow cells were isolated as described above. Cells were enriched for CD117 by magnetic enrichment (Miltenyi Biotec). Enriched cells were stained with the stem and progenitor antibodies described the Supplemental Experimental Procedures and sorted on a four-laser FACSAria (BD Biosciences). Sorted cells were transplanted as described above.

RNA/DNA Sequencing

RNA and DNA were isolated with AllPrep DNA/RNA mini kit (Qiagen). We made 100 bp paired-end WGBS and RNA-seq (Illumina TruSeq DNA or RNA Sample Preparation Kit) and RRBS libraries (Boyle et al., 2012) from 2 to 3 biological replicates. Briefly, 1ug of genomic DNA was fragmented and made into libraries, undergoing two rounds of bisulfite conversion (Qiagen EpiTect Bisulfite Kit) (Jeong et al., 2014). DNA libraries were sent to BCM Genomic and RNA Profiling Core for quality control and sequenced on Illumina HiSeq sequencing systems. Demultiplexing analysis was run allowing one mismatch in the index read on undetermined reads. Adapter and base quality trimming was performed with Trimgalore for all raw data files using a Phred score threshold 20. RNA and DNA sequencing analysis are described in supplemental experimental procedures. The WGBS, RRBS, and RNA-seq data will be deposited prior to publication in Gene Expression Omnibus of the National Center for Biotechnical Information and accession numbers will be provided.

TCGA DNA methylation analysis

AML patient DNA methylation data (Illumina Infinium HumanMethylation450) were obtained from the TCGA Research Network dataset http://cancergenome.nih.gov/. Patients with mutations in *IDH1*, *IDH2*, *TET1*, or *TET2* were excluded from further analysis. Patient groupings were established according to *DNMT3A*^{R882} and *FLT3* mutation status. Patients with neither mutation were limited to those with normal karyotype. Differential methylation between two

groups were defined as: FDR < 0.05 and mean (Beta-value) methylation difference > 0.15 using the R package CpGassoc (Barfield et al., 2012) with logit-transformed methylation beta values. Hypomethylated enhancers covered at least 3 CpGs and their mean methylation difference between R883+FLT3 and FLT3 was < -0.15. Gene and CpG island annotations were from the Illumina annotation file. Promoter probes includes the four categories TSS1500, TSS200, 5'UTR and first exon. Enhancer probes were assigned by overlap of H3K27ac or H3K4me1 peaks outside of Refseq promoter regions. H3K27ac peaks and H3K4me1 peaks were called by MACS2 (Zhang et al., 2008) with CD34+ primary cells (GSM772885 and GSM706845).

Statistics

All values are means ± s.e.m. Comparisons between groups were made with ANOVA, the log-rank test, or Student's *t*-test, using Graphpad Prism 5.0b.

AUTHOR CONTRIBUTIONS

L.Y. B.R., and A.M. designed and performed experiments, analyzed and interpreted the data and wrote the manuscript. M.L., M.J., S.K., J.L., X.Z., G.A.C., V.I.R., T.Z., R.R., B.G., W.L., and M.A.G. also contributed to experimental design and data interpretation. M.L., M.J., X.Z. also performed experiments. C.C., D.R., H.P., X.L. and B.G. analyzed and interpreted data. All authors discussed results and commented on the manuscript. W.L. supervised the bioinformatics analyses. M.A.G. supervised the study and wrote the manuscript.

ACKNOWLEDGEMENTS

We thank J. Licht and F. Rassool for providing the FLT3-ITD construct. We also thank members of the Goodell and King Labs for scientific advice; Y. Zheng, A. Rosen, R. Nitsal, M. Landis, K. Lin for technical support, and to J. Gilbert and C. Gillespie for critical reading of the manuscript.

L.Y. is funded by the Robert and Janice McNair Foundation as an MD/PhD McNair Scholar. This project was funded by CPRIT (RP110028, RP110471 and RP150292), the NIH (DK092883 and HG007538), and the Samuel Waxman Cancer Research Foundation. We also thank the Cytometry and Cell Sorting and Genomic and RNA Profiling Cores (NCI P30CA125123, P30 AI036211, P30 CA125123, and S10 RR024574) at Baylor College of Medicine. Authors declare no conflicts of interest.

REFERENCES

Baldridge, M. T., King, K. Y., Boles, N. C., Weksberg, D. C., and Goodell, M. A. (2010). Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection. Nature *465*, 793-797.

Barfield, R. T., Kilaru, V., Smith, A. K., and Conneely, K. N. (2012). CpGassoc: an R function for analysis of DNA methylation microarray data. Bioinformatics *28*, 1280-1281.

Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M. A., Beaudet, A. L., Ecker, J. R.*, et al.* (2010). The NIH Roadmap Epigenomics Mapping Consortium. Nature biotechnology *28*, 1045-1048.

Boyle, P., Clement, K., Gu, H., Smith, Z. D., Ziller, M., Fostel, J. L., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A., and Meissner, A. (2012). Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol *13*, R92.

Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D. B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S.*, et al.* (2014). Identification of Regulatory Networks in HSCs and Their Immediate Progeny via Integrated Proteome, Transcriptome, and DNA Methylome Analysis. Cell Stem Cell.

Cancer Genome Atlas Research, N. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. The New England journal of medicine *368*, 2059-2074.

Celik, H., Mallaney, C., Kothari, A., Ostrander, E. L., Eultgen, E., Martens, A., Miller, C. A., Hundal, J., Klco, J. M., and Challen, G. A. (2014). Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation. Blood.

Challen, G. A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J. S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y.*, et al.* (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. Nature Genetics *44*, 23-31.

Cleveland, S. M., Goodings, C., Tripathi, R. M., Elliott, N., Thompson, M. A., Guo, Y., Shyr, Y., and Davé, U. P. (2014). LMO2 induces T-cell leukemia with epigenetic deregulation of CD4. Experimental hematology *42*, 581-593.e585.

Consortium, M. E., Stamatoyannopoulos, J., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D., Groudine, M., Bender, M., Kaul, R.*, et al.* (2012). An encyclopedia of mouse DNA elements (Mouse ENCODE). Genome Biology *13*, 418.

Corces-Zimmerman, M. R., Hong, W. J., Weissman, I. L., Medeiros, B. C., and Majeti, R. (2014). Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc Natl Acad Sci U S A *111*, 2548-2553.

Coustan-Smith, E., Mullighan, C. G., Onciu, M., Behm, F. G., Raimondi, S. C., Pei, D., Cheng, C., Su, X., Rubnitz, J. E., Basso, G.*, et al.* (2009). Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncology *10*, 147-156.

Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., Ritchey, J. K., Young, M. A., Lamprecht, T., McLellan, M. D.*, et al.* (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature *481*, 506-510.

Figueroa, M. E., Chen, S. C., Andersson, A. K., Phillips, L. A., Li, Y., Sotzen, J., Kundu, M., Downing, J. R., Melnick, A., and Mullighan, C. G. (2013). Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. J Clin Invest *123*, 3099-3111.

Figueroa, M. E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P. J., Schifano, E., Booth, J., van Putten, W., Skrabanek, L.*, et al.* (2010). DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. Cancer Cell *17*, 13-27.

Genovese, G., Kahler, A. K., Handsaker, R. E., Lindberg, J., Rose, S. A., Bakhoum, S. F., Chambert, K., Mick, E., Neale, B. M., Fromer, M.*, et al.* (2014). Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. The New England journal of medicine *371*, 2477-2487.

Gerloff, D., Grundler, R., Wurm, A. A., Bräuer-Hartmann, D., Katzerke, C., Hartmann, J.-U., Madan, V., Muller-Tidow, C., Duyster, J., Tenen, D. G.*, et al.* (2015). NF-κB/STAT5/miR-155 network targets PU.1 in FLT3-ITD-driven acute myeloid leukemia. Leukemia *29*, 535-547.

Groschel, S., Sanders, M. A., Hoogenboezem, R., de Wit, E., Bouwman, B. A., Erpelinck, C., van der Velden, V. H., Havermans, M., Avellino, R., van Lom, K.*, et al.* (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. Cell *157*, 369-381.

Grossmann, V., Haferlach, C., Weissmann, S., Roller, A., Schindela, S., Poetzinger, F., Stadler, K., Bellos, F., Kern, W., Haferlach, T.*, et al.* (2013). 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 *52*, 410-422.

Haferlach, T., Kohlmann, A., Wieczorek, L., Basso, G., Kronnie, G. T., Bene, M. C., De Vos, J., Hernandez, J. M., Hofmann, W. K., Mills, K. I.*, et al.* (2010). Clinical Utility of Microarray-Based Gene Expression Profiling in the Diagnosis and Subclassification of Leukemia: Report From the International Microarray Innovations in Leukemia Study Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *28*, 2529-2537.

Hinkal, G., Parikh, N., and Donehower, L. A. (2009). Timed somatic deletion of p53 in mice reveals age-associated differences in tumor progression. PLoS One *4*, e6654.

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-Andre, V., Sigova, A. A., Hoke, H. A., and Young, R. A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934-947.

Hon, G. C., Rajagopal, N., Shen, Y., McCleary, D. F., Yue, F., Dang, M. D., and Ren, B. (2013). Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nature genetics *45*, 1198-1206.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., Lindsley, R. C., Mermel, C. H., Burtt, N., Chavez, A.*, et al.* (2014). Age-related clonal hematopoiesis associated with adverse outcomes. The New England journal of medicine *371*, 2488-2498.

Jan, M., Snyder, T. M., Corces-Zimmerman, M. R., Vyas, P., Weissman, I. L., Quake, S. R., and Majeti, R. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med *4*, 149ra118.

Jeong, M., Sun, D., Luo, M., Huang, Y., Challen, G. A., Rodriguez, B., Zhang, X., Chavez, L., Wang, H., Hannah, R.*, et al.* (2014). Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat Genet *46*, 17-23.

Kelly, L. M., Kutok, J. L., Williams, I. R., Boulton, C. L., Amaral, S. M., Curley, D. P., Ley, T. J., and Gilliland, D. G. (2002). PML/RARalpha and FLT3-ITD induce an APL-like disease in a mouse model. Proc Natl Acad Sci U S A *99*, 8283-8288.

Kim, S. J., Zhao, H., Hardikar, S., Singh, A. K., Goodell, M. A., and Chen, T. (2013). A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. Blood *122*, 4086-4089.

Kogan, S. C. (2002). Bethesda proposals for classification of nonlymphoid hematopoietic neoplasms in mice. Blood *100*, 238-245.

Krivtsov, A. V., Twomey, D., Feng, Z., Stubbs, M. C., Wang, Y., Faber, J., Levine, J. E., Wang, J., Hahn, W. C., Gilliland, D. G.*, et al.* (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL–AF9. Nature *442*, 818-822.

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D. A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S.*, et al.* (2014). Immunogenetics. Chromatin state dynamics during blood formation. Science *345*, 943-949.

Lee, B. H., Tothova, Z., Levine, R. L., Anderson, K., Buza-Vidas, N., Cullen, D. E., McDowell, E. P., Adelsperger, J., Frohling, S., Huntly, B. J.*, et al.* (2007). FLT3 mutations confer enhanced proliferation and survival properties to multipotent progenitors in a murine model of chronic myelomonocytic leukemia. Cancer Cell *12*, 367-380.

Ley, T. J., Ding, L., Walter, M. J., McLellan, M. D., Lamprecht, T., Larson, D. E., Kandoth, C., Payton, J. E., Baty, J., Welch, J.*, et al.* (2010). DNMT3A mutations in acute myeloid leukemia. The New England journal of medicine *363*, 2424-2433.

Maddigan, A., Truitt, L., Arsenault, R., Freywald, T., Allonby, O., Dean, J., Narendran, A., Xiang, J., Weng, A., Napper, S., and Freywald, A. (2011). EphB Receptors Trigger Akt Activation and Suppress Fas Receptor-Induced Apoptosis in Malignant T Lymphocytes. The Journal of Immunology *187*, 5983-5994.

Mayle, A., Luo, M., Jeong, M., and Goodell, M. A. (2013). Flow cytometry analysis of murine hematopoietic stem cells. Cytometry Part A : the journal of the International Society for Analytical Cytology *83*, 27-37.

Mayle, A., Yang, L., Rodriguez, B., Zhou, T., Chang, E., Curry, C. V., Challen, G. A., Li, W., Wheeler, D., Rebel, V. I., and Goodell, M. A. (2014). Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. Blood.

Mok, M. M. H., Du, L., Wang, C. Q., Tergaonkar, V., Liu, T. C., Kham, S. K. Y., Sanda, T., Yeoh, A. E.-J., and Osato, M. (2014). RUNX1 point mutations potentially identify a subset of early immature T-cell acute lymphoblastic leukaemia that may originate from differentiated T-cells. Gene *545*, 111-116.

Morse, H. C., 3rd, Anver, M. R., Fredrickson, T. N., Haines, D. C., Harris, A. W., Harris, N. L., Jaffe, E. S., Kogan, S. C., MacLennan, I. C., Pattengale, P. K.*, et al.* (2002). Bethesda proposals for classification of lymphoid neoplasms in mice. Blood *100*, 246-258.

Neumann, M., Heesch, S., kbuget, N. G. o., Schwartz, S., Schlee, C., Benlasfer, O., Farhadi-Sartangi, N., Thibaut, J., Burmeister, T., Hoelzer, D.*, et al.* (2012). Clinical and molecular characterization of early T-cell precursor leukemia: a high-risk subgroup in adult T-ALL with a high frequency of FLT3 mutations. Blood Cancer Journal *2*, e55-57.

Neumann, M., Heesch, S., Schlee, C., Schwartz, S., Gokbuget, N., Hoelzer, D., Konstandin, N. P., Ksienzyk, B., Vosberg, S., Graf, A.*, et al.* (2013). Whole-exome sequencing in adult ETP-ALL reveals a high rate of DNMT3A mutations. Blood *121*, 4749-4752.

Qu, Y., Lennartsson, A., Gaidzik, V. I., Deneberg, S., Karimi, M., Bengtzen, S., Hoglund, M., Bullinger, L., Dohner, K., and Lehmann, S. (2014). Differential methylation in CN-AML

preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. Epigenetics *9*.

Rathinam, C., Thien, C. B. F., Flavell, R. A., and Langdon, W. Y. (2010). Myeloid Leukemia Development in c-Cbl RING Finger Mutant Mice Is Dependent on FLT3 Signaling. Cancer Cell *18*, 341-352.

Rohr, J., Guo, S., Huo, J., Bouska, A., Lachel, C., Li, Y., Simone, P. D., Zhang, W., Gong, Q., Wang, C.*, et al.* (2016). Recurrent activating mutations of CD28 in peripheral T-cell lymphomas. 1-9.

Roller, A., Grossmann, V., Bacher, U., Poetzinger, F., Weissmann, S., Nadarajah, N., Boeck, L., Kern, W., Haferlach, C., Schnittger, S.*, et al.* (2013). Landmark analysis of DNMT3A mutations in hematological malignancies. Leukemia *27*, 1573-1578.

Ruau, D., Ng, F. S., Wilson, N. K., Hannah, R., Diamanti, E., Lombard, P., Woodhouse, S., and Gottgens, B. (2013). Building an ENCODE-style data compendium on a shoestring. Nat Methods *10*, 926.

Russler-Germain, D. A., Spencer, D. H., Young, M. A., Lamprecht, T. L., Miller, C. A., Fulton, R., Meyer, M. R., Erdmann-Gilmore, P., Townsend, R. R., Wilson, R. K., and Ley, T. J. (2014). The R882H DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers. Cancer Cell *25*, 442-454.

Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V. V., and Ren, B. (2013). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., Kennedy, J. A., Schimmer, A. D., Schuh, A. C., Yee, K. W.*, et al.* (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature *506*, 328-333.

Smeets, M. F. M. A., Chan, A. C., Dagger, S., Bradley, C. K., Wei, A., and Izon, D. J. (2013). Fli-1 Overexpression in Hematopoietic Progenitors Deregulates T Cell Development and Induces Pre-T Cell Lymphoblastic Leukaemia/Lymphoma. PLoS ONE *8*, e62346.

Smeets, M. F. M. A., Wiest, D. L., and Izon, D. J. (2014). Fli-1 regulates the DN2 to DN3 thymocyte transition and promotes γδ T-cell commitment by enhancing TCR signal strength. European Journal of Immunology *44*, 2617-2624.

Smith, S., Tripathi, R., Goodings, C., Cleveland, S., Mathias, E., Hardaway, J. A., Elliott, N., Yi, Y., Chen, X., Downing, J.*, et al.* (2014). LIM Domain Only-2 (LMO2) Induces T-Cell Leukemia by Two Distinct Pathways. PLoS ONE *9*, e85883.

Tian, T., Yu, S., Liu, L., Xue, F., Yuan, C., Wang, M., Ji, C., and Ma, D. (2015). The Profile of T Helper Subsets in Bone Marrow Microenvironment Is Distinct for Different Stages of Acute Myeloid Leukemia Patients and Chemotherapy Partly Ameliorates These Variations. PLoS ONE *10*, e0131761.

Van Vlierberghe, P., Ambesi-Impiombato, A., De Keersmaecker, K., Hadler, M., Paietta, E., Tallman, M. S., Rowe, J. M., Forne, C., Rue, M., and Ferrando, A. A. (2013). Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia. Blood *122*, 74-82.

Van Vlierberghe, P., Ambesi-Impiombato, A., Perez-Garcia, A., Haydu, J. E., Rigo, I., Hadler, M., Tosello, V., Della Gatta, G., Paietta, E., Racevskis, J.*, et al.* (2011). ETV6 mutations in early immature human T cell leukemias. The Journal of Experimental Medicine *208*, 2571-2579.

Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., Wartman, L. D., Lamprecht, T. L., Liu, F., Xia, J.*, et al.* (2012). The origin and evolution of mutations in acute myeloid leukemia. Cell *150*, 264-278.

Xie, M., Lu, C., Wang, J., McLellan, M. D., Johnson, K. J., Wendl, M. C., McMichael, J. F., Schmidt, H. K., Yellapantula, V., Miller, C. A.*, et al.* (2014). Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nature medicine *20*, 1472-1478.

Yan, X. J., Xu, J., Gu, Z. H., Pan, C. M., Lu, G., Shen, Y., Shi, J. Y., Zhu, Y. M., Tang, L., Zhang, X. W.*, et al.* (2011). Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nat Genet *43*, 309-315.

Yang, L., Rau, R., and Goodell, M. (2015). DNMT3A in haematological malignancies. Nature Reviews Cancer *In press.*

You, J. S., and Jones, P. A. (2012). Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell *22*, 9-20.

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S. L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M.*, et al.* (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature *481*, 157-163.

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

Ziller, M. J., Gu, H., Muller, F., Donaghey, J., Tsai, L. T., Kohlbacher, O., De Jager, P. L., Rosen, E. D., Bennett, D. A., Bernstein, B. E.*, et al.* (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477-481.

FIGURE LEGENDS

Figure 1. *Dnmt3a* deletion potentiates *FLT3-ITD*-mediated induction of pre T-lymphoblastic leukemia

(A) Experimental scheme showing induction of Mx1-cre, *FLT3-ITD* retroviral transduction, and experimental groups. (B) Kaplan-Meier survival plots comparing experimental groups with *WT* and *3aKO* controls of *FLT3-ITD* overexpression. n=10, ***P < 0.001 by log-rank test with Bonferroni correction, representative of six independent experiments. (C) Spleen weights of moribund and control mice normalized to body weight (n=9) representative of three independent experiments (D) Thymus weights normalized to body weights of moribund mice and control

mice (n=10 per group) for three independent experiments. (E) Flow cytometry analysis of surface markers CD45.2 (donor-derived cells), GFP, CD4 and CD8 in bone marrow (BM). Arrows indicate gating strategy. (F) Histological analysis of peripheral blood (Giemsa-Wright stain), BM (Giemsa-Wright stain), and spleen (H&E stain). Scale bars = 100μm. (G) Ki67 staining of *3aKO/FLT3-ITD* and *FLT3-ITD* (H) Analysis of apoptotic rate of *3aKO FLT3-ITD* and *FLT3-ITD* (n=5). All bars denote mean ± s.e.m values *P < 0.05 and ** P < 0.01 and *** P < 0.001 by one-way ANOVA. See also Figure S1.

Figure 2. Deletion of *Dnmt3a* in T-cell acute lymphoblastic leukemia induces aberrant HSC and myeloid gene expression

(A) Gene ontology (DAVID) analysis of three pairwise comparisons (B) Ingenuity pathway analysis of differential gene expression comparing *3aKO/FLT3-ITD* and *FLT3-ITD* only leukemic cells (FPKM > 0.5, fold change > 1.5, FDR q-value < 0.05). (C) Heat map of average log-transformed gene scaled FPKM expression values of representative enriched gene sets from GSEA in *3aKO/FLT3-ITD* leukemic cells compared to WT and *FLT3-ITD* only leukemic cells. Up and Down indicate genes that are upregulated or downregulated, respectively, in the gene set. (D) Percentage of up- and down-regulated genes in *3aKO/FLT3-ITD* murine cells found in the signatures of over- and under-represented genes that characterize human leukemia subtypes (Haferlach et al., 2010). PRE_SUBTYPE, AML_SUBTYPE, ALL_SUBTYPE represent leukemia precursor, acute myeloid leukemia, and acute lymphoblastic leukemia data, respectively. See also Figure S2.

Figure 3. *Dnmt3a* loss initiates acute myeloid and T- lymphoblastic *Flt3-ITD^{KI}* leukemia

(A) Experimental scheme showing deletion of *Dnmt3a* by *ER-Cre* and bone marrow transplantation into lethally irradiated recipients to generate *3aKO*/*Flt3-ITD^{KI}* mice and controls

(B) Kaplan-Meier survival plot of mice transplanted with cells from *3aKO*, *Flt3-ITD*^{KI} and *3aKO*/*Flt3-ITD*^{KI} mice. (C) Peripheral blood analysis at 7 months after transplantation showing white blood cell counts and (D) lineage distribution. * P<0.05 and *** P<0.001 by one-way and Two-way ANOVA. (E) Giemsa-Wright staining of *3aKO, 3aKO*/*Flt3-ITD*^{KI} and *Flt3-ITD*^{KI} bone marrow touch preps. Scale bar = 50μm. (F) Representative bone marrow touch preps of a group of mice with bone marrow leukemic cell infiltration stained with Giemsa-Wright. Scale bar = 50μm. (G) Flow cytometry analysis of myeloid cell surface markers (Mac-1/Gr-1) (red box) and T-cell surface markers (black box) in the bone marrow (left) and thymus (right). (H) PCR analysis to detect *Dnmt3a*-floxed vs fully deleted (Δ) allele of leukemic cells from the bone marrow (BM) and thymus (Th). L=molecular weight ladder. See also Figure S3.

Figure 4. Heterozygous loss of *Dnmt3a* induces acute myeloid leukemia

(A) Kaplan-Meier survival plot of mice transplanted with bone marrow of *3aHet*/*Flt3-ITD*^{KI}, *3aHet,* or *Flt3-ITD*^{KI}. (B) Peripheral blood WBC count over time with flow cytometry analysis (C) and WBC differential. Shown are the relative proportions of the indicated populations among the stained cells. (D). (E) Bone marrow touch prep stained with Giemsa-wright staining. Scale bar = 50μm. (F) Flow cytometry analysis of donor-derived CD45.2+ bone marrow cells with myeloid and progenitor surface markers (Mac-1/Gr-1/c-Kit) in *3aHet*/*Flt3-ITD*^{KI}, *3aHet,* and *Flt3-ITD*^{KI} with the red box in the middle panel indicating the myeloid population. Arrows indicate the gating strategy. (G) Table of leukemia incidence in *3aHet*/*Flt3-ITD*^{KI}, *3aKO/Flt3-ITD*^{KI}, and *3aKO* mice. See also Figure S4.

Figure 5. *Dnmt3a*-related myeloid leukemia arises from transformed HSCs

(A) Table of sorted cell types that were transduced with *FLT3-ITD* retrovirus and transplanted into lethally irradiated recipients. Representative of three independent experiments. (B) Representative flow cytometry analysis of recipient peripheral blood 10 weeks after

transplantation of *3aKO/FLT3-ITD* leukemic blast cells. (C) Leukemia incidence after transplantation of various *3aKO / FLT3-ITD* leukemic T cell populations into lethally irradiated recipients. (D) Representative flow cytometry analysis of recipient peripheral blood 10 weeks after transplantation of *3aKO / FLT3-ITD* leukemic blast cells. (E) Table of FACS sorted populations from *3aHet/Flt3-ITD*^{KI} mice that were transplanted into lethally irradiated recipients and development of leukemia. MPP=multipotent progenitor. LT-HSC= long-term HSC, ST-HSC= short-term HSC, CMP= common myeloid progenitor. GMP= granulocyte-macrophage progenitor. CLP= common lymphoid progenitor. Disease observation up to 12 months. Representative of two independent experiments. (F) Lineage distribution of donor-derived peripheral blood cells after *3aKO* and *3aHet/Flt3-ITD*^{KI} LT-HSC transplantation. (G) Donor engraftment of purified HSCs and progenitors from *3aHet/Flt3-ITD*^{KI} AML mice at 3, 8, 12 weeks after transplant, representative of two independent experiments. ***p<0.001 by two-way ANOVA.

Figure 6. *Dnmt3a* loss causes hypomethylation at enhancer sites

(A) Number of differentially methylated CpGs that were hypo- (blue) and hypermethylated (red) n=3. q-value < 0.05, DM ≥ 25%. (B) Heatmap of percent differentially methylated regions between pairwise comparisons as indicated at genomic regions. UMR, undermethylated regions in HSCs; CGI, CpG islands. Blue indicates enhancers. Heme, hematopoietic enhancers (Lara-Astiaso et al., 2014). (C) H3K4me1 and H3K27ac signal density across regions differentially enriched in *3aKO/FLT3-ITD* cells. Plotted is the subset of significant regions (n = 2,909) located greater than 1-kb away from an annotated Refseq TSS and covered across all experimental WGBS datasets. (D) Violin plot showing CpG methylation distributions at regions differentially enriched with chromatin marks in *3aKO/FLT3-ITD*. Plotted are data for 24,409 CpG sites located in the 2,909 significant regions with sufficient CpG coverage and 25,000 randomly selected control CpG sites. (E) Mean CpG methylation ratio of leukemic cells overexpressing Dnmt3a or GFP. *p<0.05 by student's t-test. (F) Number of hypermethylated 1-kb regions

relative to tumors overexpressing GFP. (G) Number of significant differentially methylated regions between *3aHet/Flt3-ITD*^{KI} AML cells and KSL *Flt3-ITD* progenitor cells at 1kb tiling windows using RRBS. Analysis compared three biological replicates per condition. (H) Percent of differential methylation of interrogated genomic regions of *3aHet*/*Flt3-ITD*^{KI} AML and *3aKO/FLT3-ITD* T-ALL cells. UMR, unmethylated regions in HSC; CGI, CpG island; heme, hematopoietic enhancers (Lara-Astiaso et al., 2014). See also Figure S6.

Figure 7. Enhancer hypomethylation is associated with increased recruitment of transcription factor FLI1.

(A) Bar plot summarizing differential binding of FLI1 in *3aKO/FLT3-ITD* leukemia cells upon res-expression of *Dnmt3a*. Data are presented relative to the absence of Dnmt3a. Prom, promoter binding (within 1-kb of annotated TSS); Distal, binding outside of promoter; K4K27 and K27ac, distal regions marked by H3K4me1 and/or H3K27ac in 3aKO/FLT3-ITD leukemia cells; Ref Hemat Enh, reference hematopoietic cell or tissue type enhancer region in Mouse Encode (Bone Marrow, Spleen, Thymus) or Lara-Astiaso datasets. (B) Functional significance of promoter regions differentially bound by FLI1 was predicted by GREAT 2.0. Top, functions associated with increased Fli1. Bottom, functions associated with decreased Fli1. (C) Functional significance of increased Fli1 binding in hematopoietic enhancer regions (consisting of reference datasets and distal regions marked by H3K27Ac and H3K4me1 in 3aKO/FLT3-ITD cells). Horizontal line separates enriched terms from the human Disease Ontology and Mouse Genome Informatics Mouse Phenotype databases. Associations in GREAT based off gene regulatory domain extension of 150-kb up and downstream of TSS. (D) Model of enforced Dnmt3a expression disrupting FLI1 binding at enhancers.

Figure 8. *DNMT3A R882* mutation in human AML is associated with enhancer hypomethylation with FLT3 mutations.

(A) Groups of TCGA AML patients organized by mutation status of the indicated gene. *WT NK-AML refers to patients with normal karyotype AML who do not harbor mutations in *DNMT3A* or *FLT3*. Patients with mutations in IDH1, IDH2, TET1, or TET2 were excluded. All patients harbor cytoplasmic NPM1. (B) Boxplot of differential methylation in indicated pair-wise comparisons. For each comparison the distribution of mean β-value differences for 51,776 sites differentially methylated in one or more pair-wise are plotted. Sites of differential methylation identified in each comparison are represented by red (hypermethylation) and blue (hypomethylation) points, respectively. Dashed lines are the β-value difference thresholds. FLT3, *FLT3* mutation; R882, *DNMT3A*^{R882} mutation; WT, normal karyotype AML patients without *DNMT3A* or *FLT3* mutations. (C) Percent hypomethylated CpGs (comparing between R882 + FLT3 vs FLT3) at different genomic regions. The expected percentage: grouped CpG number divided by total CpG number; the observed percentage: grouped number of hypomethylated-CpG sites divided by the total number of hypomethylated-CpG sites. * P-value < 2.22x10-16 by one-tailed Fisher's exact test. Enhancers were defined by the overlap of either histone H3K4me1 and/or H3K27ac peaks from human CD34+ primary cells (Bernstein et al., 2010). (D) Unsupervised principal coordinate analysis of CpG methylation states at sites of variable methylation in enhancers. Analyses were performed on the mean β-value s for each patient group at sites with evidence of differential methylation (Left) as described above and the subset of these sites which are mapped to enhancer regions (Right). Data are represented as points and labeled by patient group. (E) Functional significance of differentially expressed genes associated with enhancer regions hypomethylated in *DNMT3A*^{R882} FLT3 mutant AML patients. A linear model of gene expression contrasted TCGA AML patients with mutations of both DNMT3A and FLT3 to those with FLT3 alone. Expressed genes within 500-kb were included in Ingenuity functional enrichment analysis with thresholds for differential expression: 1.25-fold change $p < 0.05$ ($n = 171$). X-axis, log-transformed p-value of enrichment test. Size, fraction of differentially

expressed genes associated with the term. HPC, hematopoietic progenitor cells. See also Figure S7.

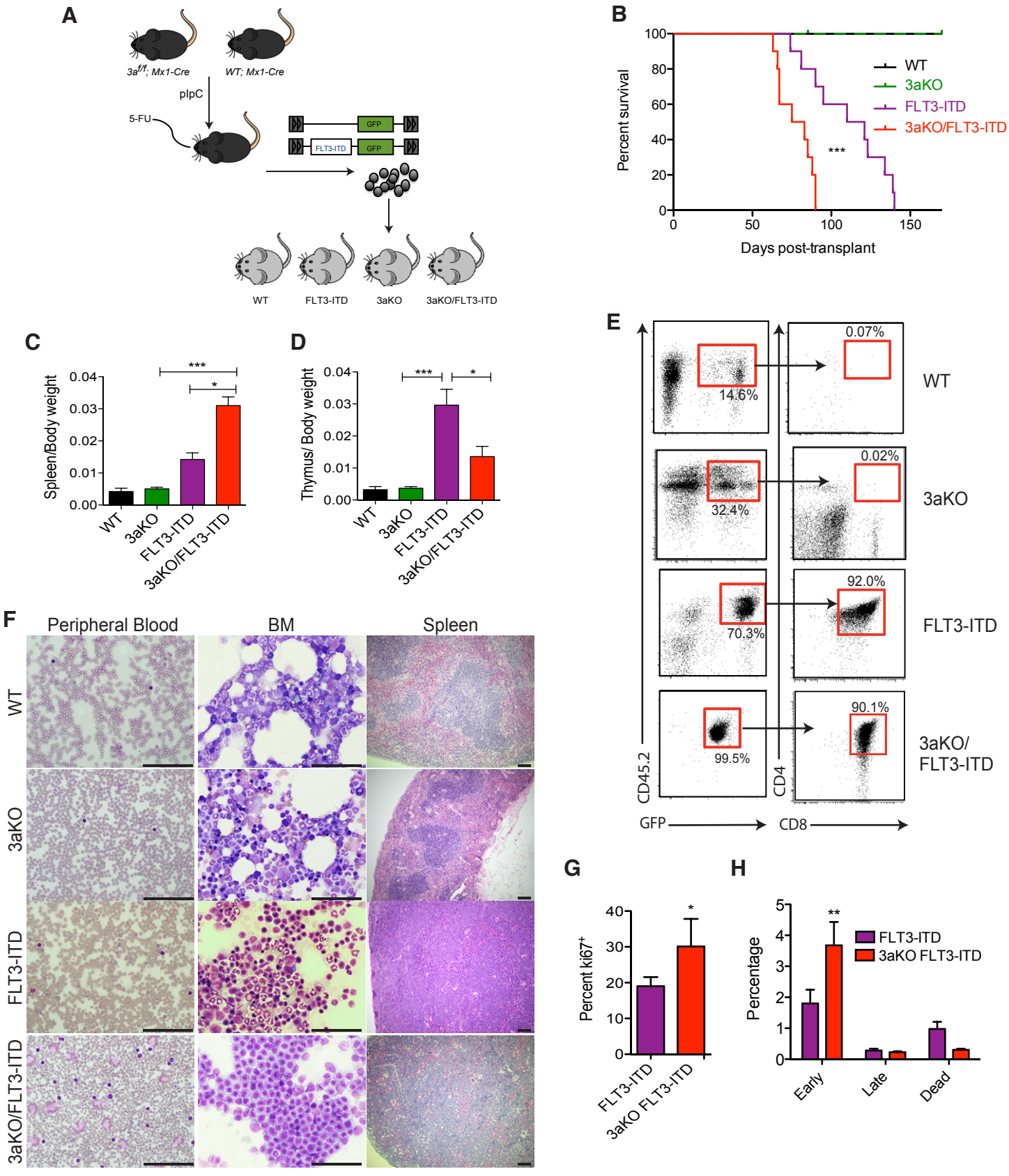

A
3a^f/f; Mx1-Cre
WT; Mx1-Cre
pIpC
5-FU
GFP
FLT3-ITD
GFP
WT
FLT3-ITD
3aKO
3aKO/FLT3-ITD
B
Percent survival
Days post-transplant
WT
3aKO
FLT3-ITD
3aKO/FLT3-ITD
***
C
Spleen/Body weight
D
Thymus/ Body weight
E
0.07%
14.6%
WT
0.02%
32.4%
3aKO
92.0%
70.3%
FLT3-ITD
90.1%
99.5%
3aKO/
FLT3-ITD
CD45.2
GFP
CD4
CD8
F
Peripheral Blood
BM
Spleen
G
Percent ki67+
H
Percentage
FLT3-ITD
3aKO FLT3-ITD
Early
Late
Dead

Figure 1

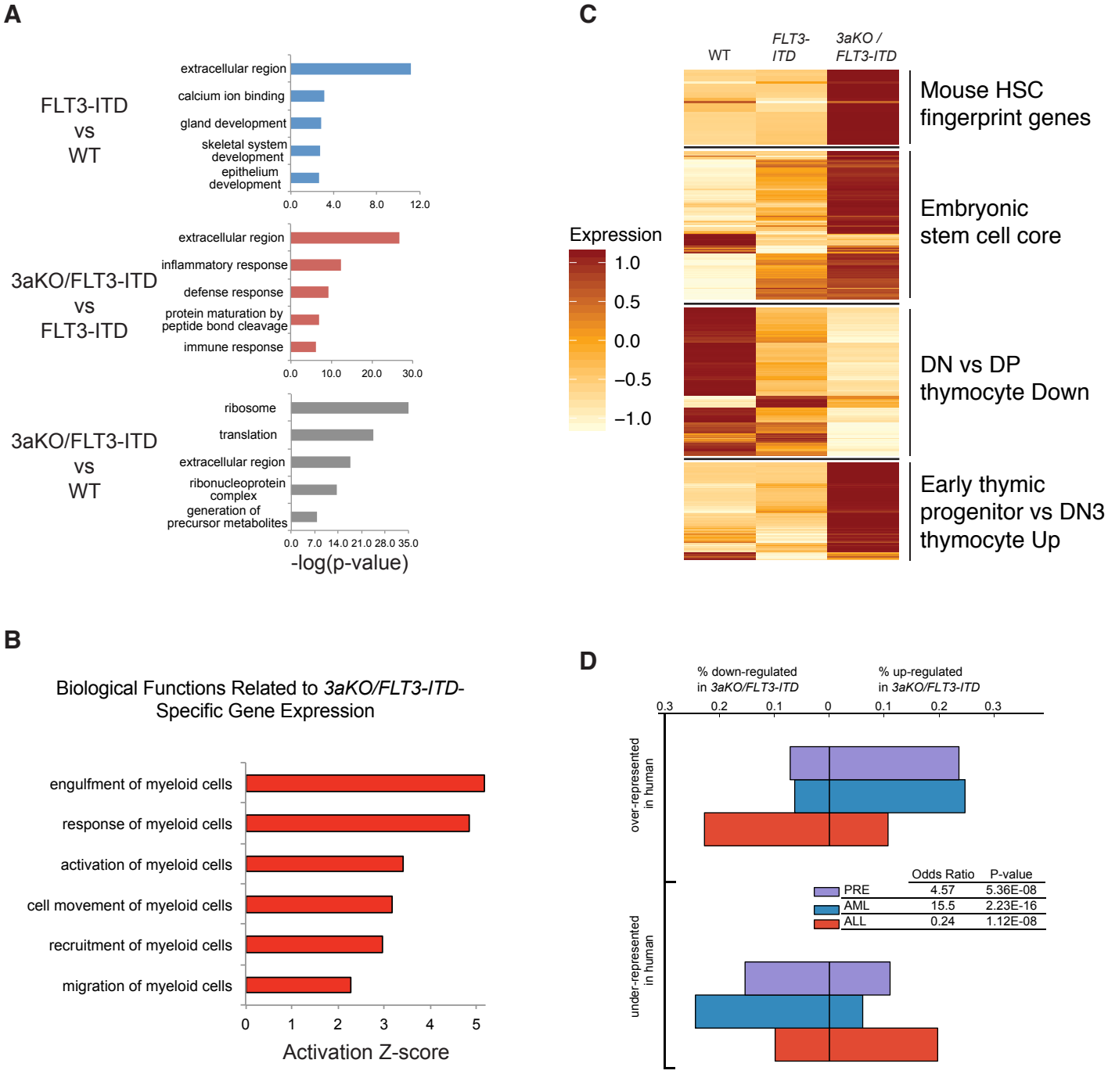

A
FLT3-ITD vs WT
extracellular region
calcium ion binding
gland development
skeletal system development
epithelium development
0.0 4.0 8.0 12.0
3aKO/FLT3-ITD vs FLT3-ITD
extracellular region
inflammatory response
defense response
protein maturation by peptide bond cleavage
immune response
0.0 10.0 20.0 30.0
3aKO/FLT3-ITD vs WT
ribosome
translation
extracellular region
ribonucleoprotein complex
generation of precursor metabolites
0.0 7.0 14.0 21.0 28.0 35.0
-log(p-value)
C
WT
FLT3-ITD
3aKO / FLT3-ITD
Expression
1.0
0.5
0.0
−0.5
−1.0
Mouse HSC fingerprint genes
Embryonic stem cell core
DN vs DP thymocyte Down
Early thymic progenitor vs DN3 thymocyte Up
B
Biological Functions Related to 3aKO/FLT3-ITD-Specific Gene Expression
engulfment of myeloid cells
response of myeloid cells
activation of myeloid cells
cell movement of myeloid cells
recruitment of myeloid cells
migration of myeloid cells
0 1 2 3 4 5
Activation Z-score
D
% down-regulated in 3aKO/FLT3-ITD
% up-regulated in 3aKO/FLT3-ITD
0.3 0.2 0.1 0 0.1 0.2 0.3
over-represented in human
under-represented in human
Odds Ratio P-value
PRE 4.57 5.36E-08
AML 15.5 2.23E-16
ALL 0.24 1.12E-08

Figure 2

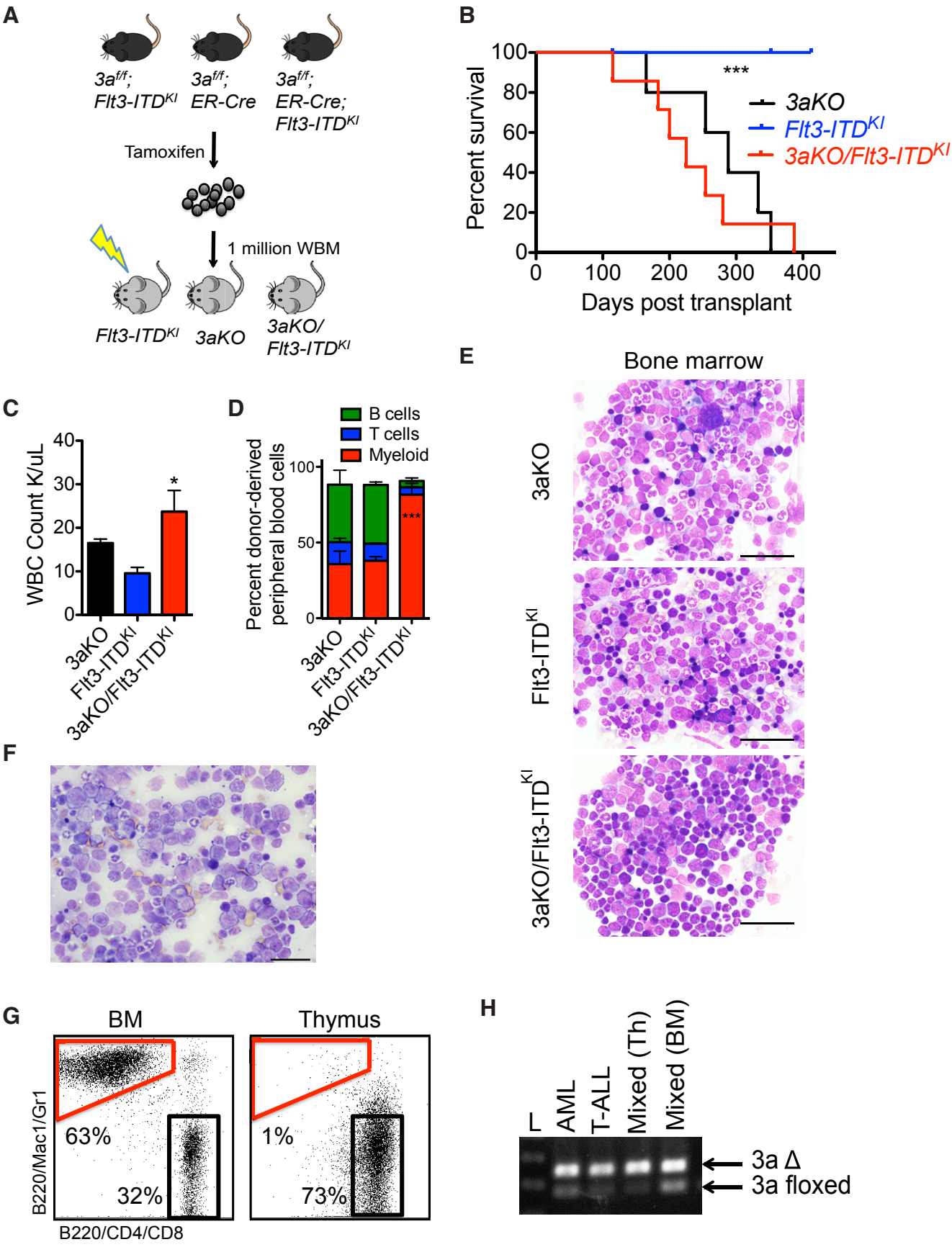
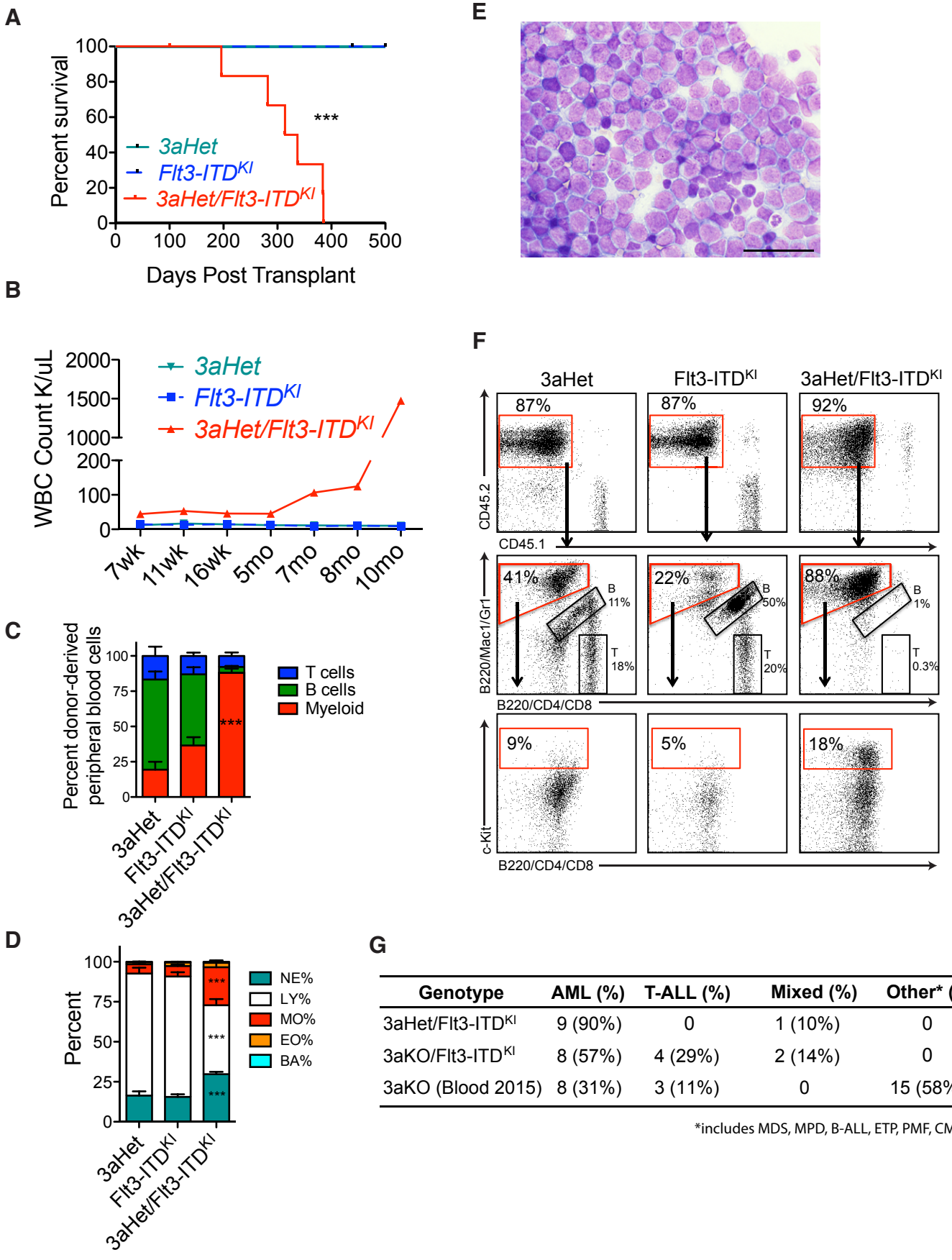

A
3a^f/f; Flt3-ITD^KI
3a^f/f; ER-Cre
3a^f/f; ER-Cre; Flt3-ITD^KI
Tamoxifen
1 million WBM
Flt3-ITD^KI
3aKO
3aKO/ Flt3-ITD^KI
B
Percent survival
Days post transplant
***
3aKO
Flt3-ITD^KI
3aKO/Flt3-ITD^KI
C
WBC Count K/uL
*
3aKO
Flt3-ITD^KI
3aKO/Flt3-ITD^KI
D
Percent donor-derived peripheral blood cells
B cells
T cells
Myeloid
***
E
Bone marrow
3aKO
Flt3-ITD^KI
3aKO/Flt3-ITD^KI
F
G
BM
Thymus
B220/Mac1/Gr1
63%
32%
1%
73%
B220/CD4/CD8
H
L
AML
T-ALL
Mixed (Th)
Mixed (BM)
3a Δ
3a floxed

Figure 3



G

Genotype	AML (%)	T-ALL (%)	Mixed (%)	Other* (%)
3aHet/Flt3-ITD^{KI}	9 (90%)	0	1 (10%)	0
3aKO/Flt3-ITD^{KI}	8 (57%)	4 (29%)	2 (14%)	0
3aKO (Blood 2015)	8 (31%)	3 (11%)	0	15 (58%)

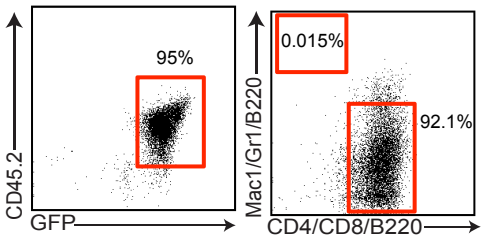
*includes MDS, MPD, B-ALL, ETP, PMF, CMML

Figure 4

A

3aKO/FLT3-ITD population	Disease incidence
Myeloid Progenitor	0/10
Lymphoid Progenitor	0/10
HSC	10/10

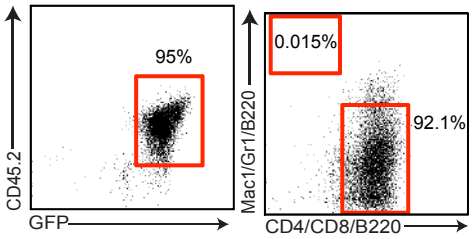
B



C

Donor cells	Number of cells	Number of helper cells	Leukemia
GFP+	1,000,000	0	10/10
GFP+	500,000	200,000	5/5
GFP+CD4+CD8+	200,000	250,000	5/5
GFP+c-kit+Lin-	2400	200,000	5/5
GFP+c-kit+Lin-Sca+	2400	200,000	3/5
GFP+c-kit+Lin-CD34-Flk2-	100	250,000	2/2

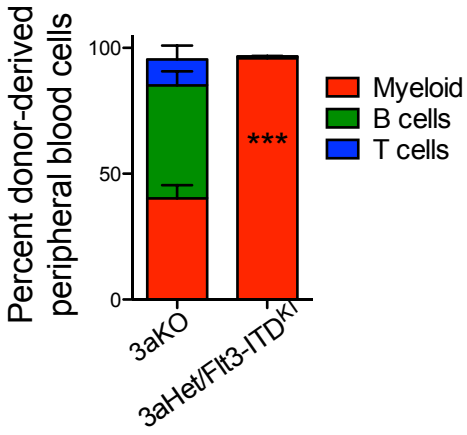
D



E

Population	Number of cells transplanted	Leukemia
LT-HSC	420	5/5
ST-HSC	2025	0/5
MPP	5389	0/5
CMP	58114	0/5
GMP	8840	0/5
MEP	40780	0/5
CLP	3258	0/4

F



G

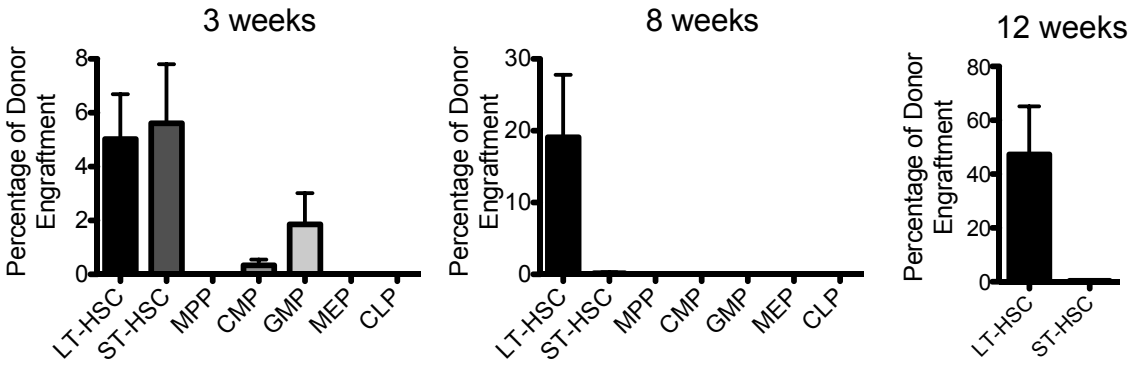


Figure 5

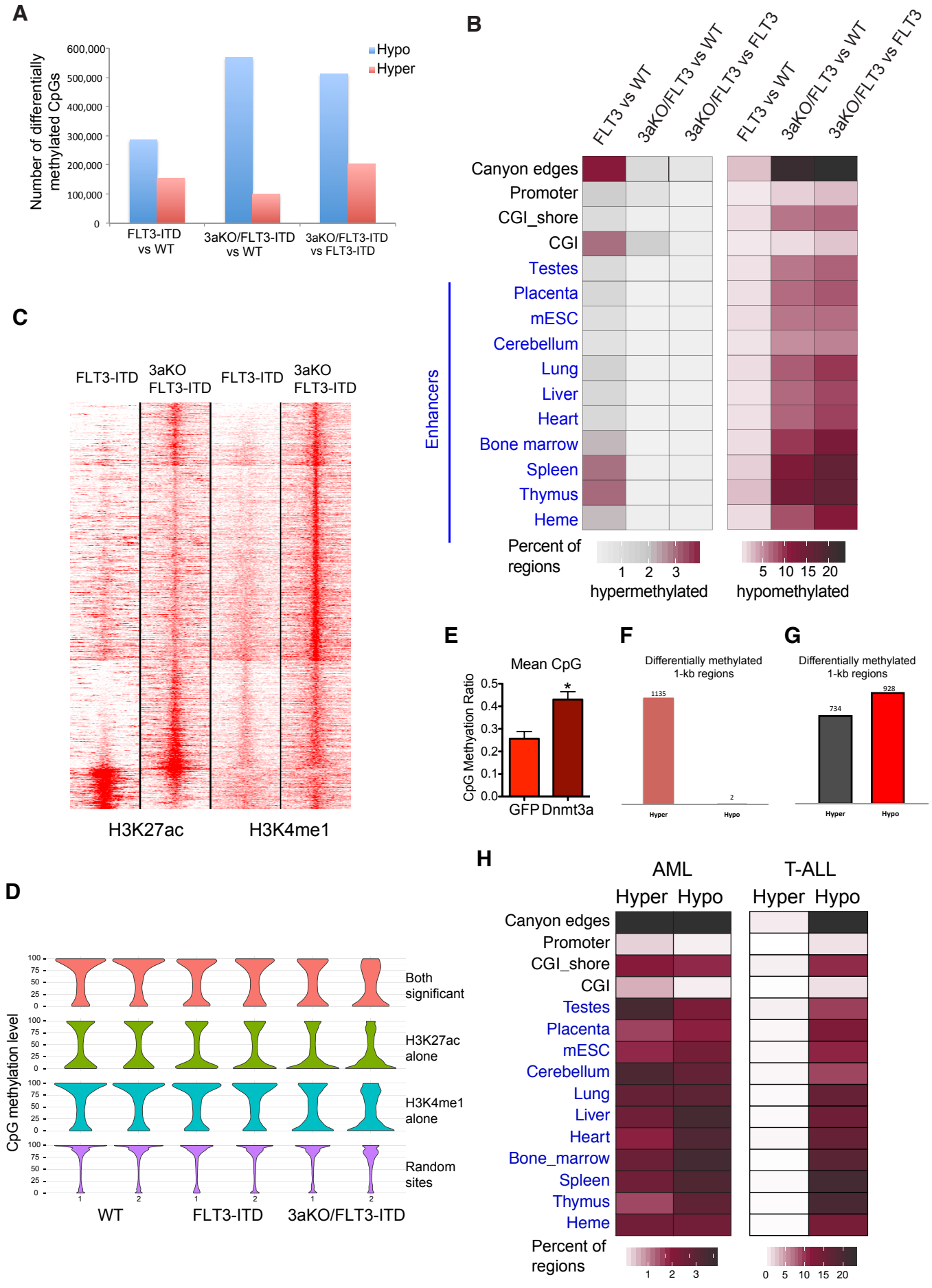

A
Number of differentially methylated CpGs
600,000
500,000
400,000
300,000
200,000
100,000
0
Hypo
Hyper
FLT3-ITD vs WT
3aKO/FLT3-ITD vs WT
3aKO/FLT3-ITD vs FLT3-ITD
B
FLT3 vs WT
3aKO/FLT3 vs WT
3aKO/FLT3 vs FLT3
FLT3 vs WT
3aKO/FLT3 vs WT
3aKO/FLT3 vs FLT3
Canyon edges
Promoter
CGI_shore
CGI
Testes
Placenta
mESC
Cerebellum
Lung
Liver
Heart
Bone marrow
Spleen
Thymus
Heme
Enhancers
Percent of regions
1 2 3
hypermethylated
5 10 15 20
hypomethylated
C
FLT3-ITD
3aKO FLT3-ITD
FLT3-ITD
3aKO FLT3-ITD
H3K27ac
H3K4me1
D
CpG methylation level
100
75
50
25
0
Both significant
H3K27ac alone
H3K4me1 alone
Random sites
1
2
1
2
1
2
WT
FLT3-ITD
3aKO/FLT3-ITD
E
Mean CpG
CpG Methyation Ratio
0.5
0.4
0.3
0.2
0.1
0.0
*
GFP Dnmt3a
F
Differentially methylated 1-kb regions
1135
2
Hyper
Hypo
G
Differentially methylated 1-kb regions
734
928
Hyper
Hypo
H
AML
Hyper Hypo
T-ALL
Hyper Hypo
Canyon edges
Promoter
CGI_shore
CGI
Testes
Placenta
mESC
Cerebellum
Lung
Liver
Heart
Bone_marrow
Spleen
Thymus
Heme
Percent of regions
1 2 3
0 5 10 15 20

Figure 6

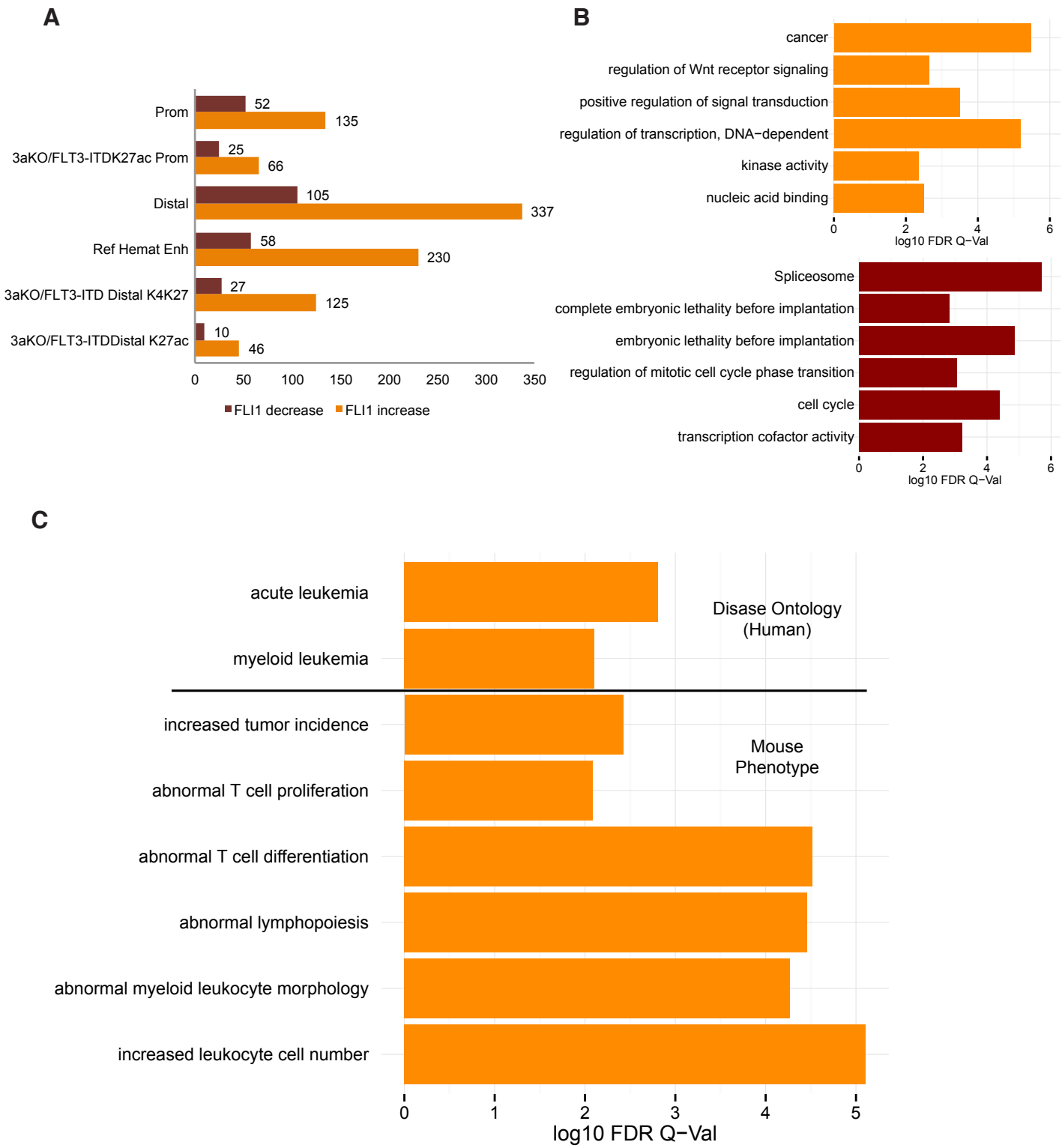
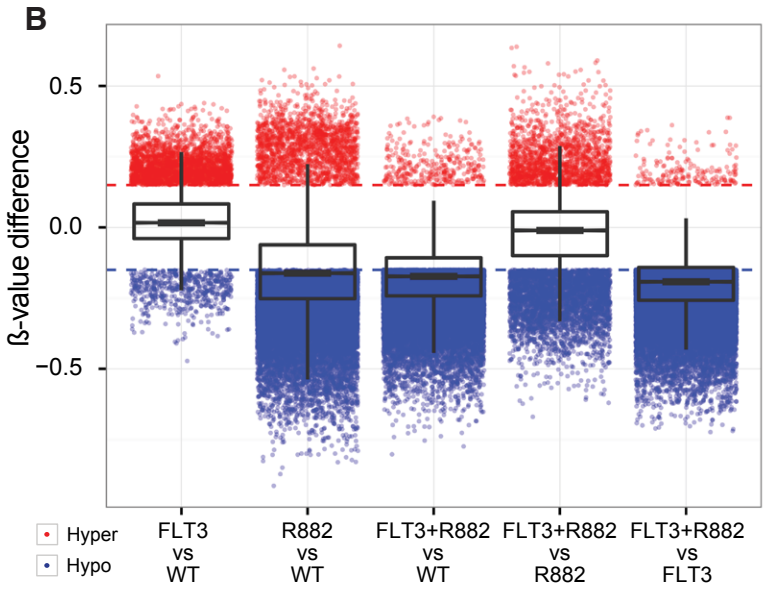
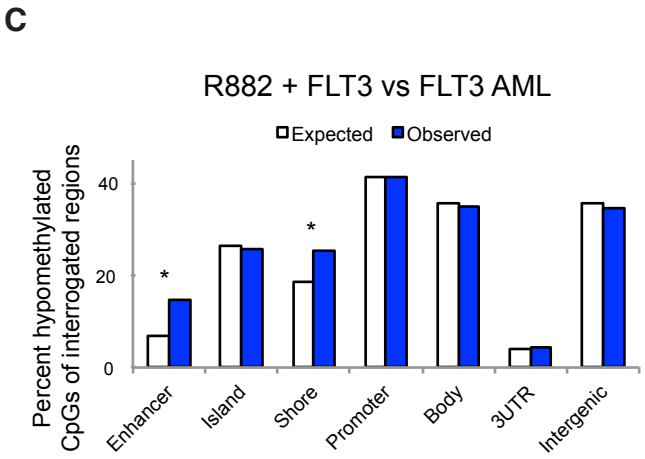

A
Prom
52
135
3aKO/FLT3-ITDK27ac Prom
25
66
Distal
105
337
Ref Hemat Enh
58
230
3aKO/FLT3-ITD Distal K4K27
27
125
3aKO/FLT3-ITDDistal K27ac
10
46
0
50
100
150
200
250
300
350
FLI1 decrease
FLI1 increase
B
cancer
regulation of Wnt receptor signaling
positive regulation of signal transduction
regulation of transcription, DNA−dependent
kinase activity
nucleic acid binding
0
2
4
6
log10 FDR Q−Val
Spliceosome
complete embryonic lethality before implantation
embryonic lethality before implantation
regulation of mitotic cell cycle phase transition
cell cycle
transcription cofactor activity
0
2
4
6
log10 FDR Q−Val
C
acute leukemia
myeloid leukemia
Disease Ontology
(Human)
increased tumor incidence
abnormal T cell proliferation
abnormal T cell differentiation
abnormal lymphopoiesis
abnormal myeloid leukocyte morphology
increased leukocyte cell number
Mouse
Phenotype
0
1
2
3
4
5
log10 FDR Q−Val

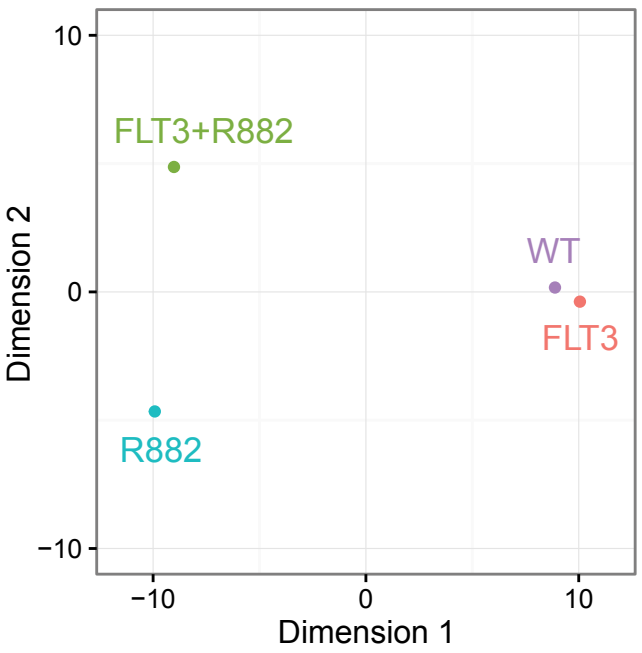
Figure 7

A

AML Group	#Samples
WT NK-AML*	8
FLT3 alone	10
R882 alone	3
R882 + FLT3	10







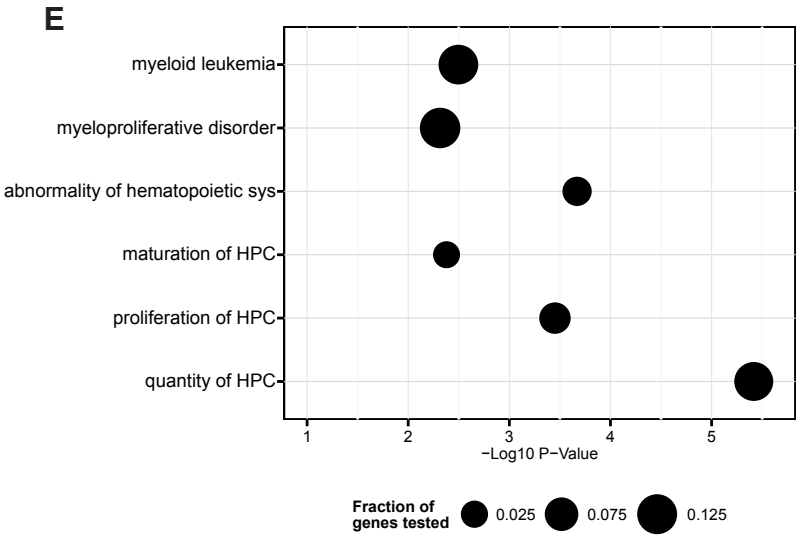


Figure 8