DOI:
10.1016/j.ccell.2019.10.011

Document Version
Peer reviewed version

HOTTIP lncRNA promotes hematopoietic stem cell self-renewal leading to AML-like disease in mice

Huacheng Luo[1#], Ganqian Zhu[2#], Jianfeng Xu[3], Qian Lai[1,4], Bowen Yan[5], Ying Guo[6], Tsz Kan Fung[7], Bernd B. Zeisig[7], Ya Cui[3], Jie Zha[4], Christopher Cogle[8], Fei Wang[9], Bing Xu[4], Feng-Chun Yang[6], Wei Li[3], Chi Wai Eric So[7*], Yi Qiu[5*], Mingjiang Xu[2,10*], and Suming Huang[1,11,12*]

[1] Division of Pediatric Hematology/Oncology, Department of Pediatrics, [11]Penn State Cancer Institute, Pennsylvania State University College of Medicine, Hershey, PA 17033

[2]Department of Molecular Medicine, [6]Department of Cell System & Anatomy, [10]Mays Cancer Center, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229

[3]Department of Molecular and Cellular Biology, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030

[4]Department of hematology, The First Affiliated Hospital of Xiamen University, Xiamen, China 361003

[5]Department of Anatomy and Cell Biology, [8]Division of Hematology/Oncology, Department of Medicine, University of Florida College of Medicine, Gainesville, FL 32610

[7]School of Cancer and Pharmaceutical Science, King's college London, London SE5 9NU, UK

[9]Department of Hematology and Oncology, The Affiliated Zhongda Hospital, Southeast University Medical School, Nanjing, China 210009

[#]These authors contributed equally

[12]Lead Contact:
Suming Huang, PhD.
Department of Pediatrics, Pennsylvania State University College of Medicine, Hershey, PA 17033
Email: shuang4@pennstatehealth.psu.edu

* Corresponding authors
shuang4@pennstatehealth.psu.edu (S.H), qiuy@ufl.edu (Y.Q.), eric.so@kcl.ac.uk (E.S), and xum1@uthscsa.edu (M.X)

Running title: *HOTTIP* controls AML chromatin and HSC behavior

SUMMARY

Long non-coding RNAs (lncRNAs) are critical for regulating HOX genes, aberration of which is a dominant mechanism for leukemic transformation. How HOX genes-associated lncRNAs regulate hematopoietic stem cell (HSC) function and contribute to leukemogenesis remains elusive. We found that *HOTTIP* is aberrantly activated in acute myeloid leukemia (AML) to alter HOXA-driven topologically associated domain (TAD) and gene expression. *HOTTIP* loss attenuates leukemogenesis of transplanted mice, while reactivation of *HOTTIP* restores leukemic TADs, transcription, and leukemogenesis in the CTCF-boundary-attenuated AML cells. *Hottip* aberration in mice abnormally promotes HSC self-renewal leading to AML-like disease by altering homeotic/hematopoietic gene-associated chromatin signature and transcription program. Thus, *Hottip* aberration acts as an oncogenic event to perturb HSC function by reprograming leukemic-associated chromatin and gene transcription.

SIGNIFICANCE

The initiation and progression of AML have so far been studied mainly within the realm of mutations and/or dysregulation of key protein-coding genes. Whether and how lncRNA misregulation can lead to oncogeneisis remains to be explored in AML. Dysregulation of HOXA genes (e.g. *HOXA9*) is a dominant mechanism for hematopoietic deregulation and leukemogenesis. Our study demonstrates that *HOTTIP* lncRNA coordinates TAD organization of AML genome including the posterior HOXA genes and various key hematopoietic regulators loci. Expression of *HOTTIP* is required for AML-driven by MLL fusions/*NPM1* mutation and is sufficient to initiate leukemic transformation of HSC. These findings provide a framework for developing targeted therapeutics for AML.

INTRODUCTION

HOX genes, especially HOXA and HOXB families, are critical for hematopoietic lineage development (Deng et al., 2013; Deng et al., 2016; Dou et al., 2016). Activation of HOX genes is a dominant mechanism of leukemic transformation, perhaps by altering self-renewal and differentiation properties of hematopoietic stem and progenitor cells (HS/PCs) (Andreeff et al., 2008; Drabkin et al., 2002). Although overexpression of HOX genes in acute myeloid leukemia (AML) has been attributed to specific chromosomal rearrangements involved in the mixed lineage leukemia (MLL) gene (*KMT2A*) or abnormalities such as mutations in *NPM1* (Meyer et al., 2009; Rice and Licht, 2007), the molecular mechanisms that drive HOX genes activation are not fully understood.

HOX genes are critical for embryonic development and their expression patterns are temporally and spatially restricted (Deng et al., 2016; Deschamps and van Nes, 2005; Forlani et al., 2003). The lineage-restricted expression pattern of HOX genes during hematopoiesis resembles their expression in early development. Generally, anterior HOX genes are highly activated in most primitive HSCs and downregulated upon lineage commitment, while posterior HOX genes are expressed in committed lineages (Sauvageau et al., 1994; Spencer et al., 2015). The different HOX genes clusters also exhibit specific patterns of lineage-specific expression. For example, HOXA genes are expressed in immature myeloid cells that are believed to play an important role in myeloid progenitor proliferation (Crooks et al., 1999; Fuller et al., 1999; So et al., 2004; Thorsteinsdottir et al., 2002).

Furthermore, the overexpression of certain HOX genes, such as *HOXA9*, is a strong marker of poor prognosis in leukemia patients (Collins and Hess, 2016; Golub et al., 1999), while lower expression of *HOXA9* and *HOXB4* are favorable predictors for AML patient outcome (Andreeff et al., 2008; Zangenberg et al., 2009), suggesting that targeting posterior HOXA genes may provide insight into AML therapy. Recently, we identified a CTCF boundary located between *HOXA7* and *HOXA9* (*CBS7/9*) that plays a critical role in maintaining posterior HOXA genes

topologically associated domain (TAD) allowing for the aberrant HOXA genes expression (Luo et al., 2018). However, the molecular mechanism by which *CBS7/9* initiates the aberrant TAD and transcription of posterior HOXA genes remain elusive.

Several HOX genes loci associated long noncoding RNAs (lncRNAs) regulate transcription of HOX genes through their influence on the epigenetic landscape (Deng et al., 2016; Wang et al., 2011). In particular, the HOXA locus associated lncRNA *HOTTIP* acts as an epigenetic regulator that recruits WDR5/MLL complex to coordinate active chromatin modifications and HOXA gene expression (Wang et al., 2011). Although during limb development, expression of *HOTTIP* was suggested to act *in cis* and positively correlates with the formation of posterior HOXA genes TAD, whether *HOTTIP* directly binds to and regulates its chromatin targets including the HOXA locus remains unknown. It has shown that knockout (KO) of *HOTTIP* strongly inhibits the 5' tip of HOXA genes (e.g. *HOXA13* and *HOXA11*), but the inhibitory effect is gradually diminished when genes move towards the anterior end (e. g. *HOXA10-HOXA7*) (Wang et al., 2011). Interestingly, HOX genes, especially posterior *HOXA9* and *HOXA10*, are frequently activated in AML, which predicts poor prognosis and treatment responses. However, the role of *HOTTIP* in HSC function and myeloid malignancies and the mechanism by which *HOTTIP* regulates its chromatin targets in leukemogenesis remains completely unknown.

RESULTS

HOTTIP loss results in inhibition of genes critical for hematopoiesis and AML leukemogenesis

To unbiasedly uncover non-coding sequences involved in HOX gene regulation in AML, we screened all CTCF sites and lncRNAs important for *HOXA9* expression within four HOX gene loci in *MLL-AF9* rearranged MOLM13 AML cells using a CRISPR-Cas9 lentivirus screening library. Besides the *CBS7/9* boundary, *HOTTIP* lncRNA was also identified as critical for aberrant

HOXA9 expression (Luo et al., 2018). *HOTTIP* is downregulated in the *CBS7/9*-disrupted (*CBS7/9*$^{+/-}$) MOLM13 cells (Figure 1A), suggesting that *HOTTIP* acts downstream of the *CBS7/9* boundary to regulate posterior HOXA genes. To test this, *HOTTIP* was specifically deleted (*HOTTIP*$^{-/-}$) by CRISPR-Cas9 in MOLM13 cells (Figure S1A). We compared transcriptomes between WT and *HOTTIP*$^{-/-}$ MOLM13 cells by performing RNA-seq analysis. A total of 706 genes exhibited greater than 2-fold decreases whereas 513 genes had increased expression upon *HOTTIP*$^{-/-}$ (Figure 1B). *HOTTIP*$^{-/-}$ impaired the transcription of not only *HOXA13-HOXA9* genes but also many genes important for hematopoiesis and leukemogenesis (Figures 1B,1C) suggesting that *HOTTIP* may directly regulate hematopoietic genes in AML besides the posterior HOXA genes. Gene ontology (GO) analysis revealed that many pathways were affected by both *CBS7/9*$^{+/-}$ and *HOTTIP*$^{-/-}$ (Figure 1D), including cell cycle, apoptosis, myeloid/leukocyte cell differentiation, JAK-STAT signaling, and regulation of cell development. In addition, pathways regulating hematopoietic cell lineage and myeloid differentiation were specifically affected by *HOTTIP*$^{-/-}$ (Figure 1D). Furthermore, when we subjected the RNA-seq data to Gene Set Enrichment Analysis (GSEA), the top ranked pathways affected by the *HOTTIP*$^{-/-}$ are those involved in JAK-STAT, NOTCH, cell adhesion, and progression of AML (Figures 1E, S1B).

Further comparison of the expression profiles between *CBS7/9*$^{+/-}$ and *HOTTIP*$^{-/-}$ MOLM13 cells revealed that 33% of differentially regulated genes were co-regulated by both *CBS7/9* boundary and *HOTTIP* (Figure S1C, Top). Among them, 33% of genes downregulated and 24% of genes upregulated overlapped between *HOTTIP*$^{-/-}$ and *CBS7/9*$^{+/-}$ (Figure S1C, Bottom). The significant overlapping of co-regulated genes by the *HOTTIP*$^{-/-}$ and *CBS7/9*$^{+/-}$ in MOLM13 cells indicated that *HOTTIP* may act downstream of *CBS7/9* boundary to coordinate active chromatin domain and gene transcription in AML cells. The critically downregulated genes were validated by RT-qPCR (Figure S1D).

To test the role of *HOTTIP* in AML genome organization, we then carried out Hi-C analysis to assess if *HOTTIP* is required to organize the HOXA locus TAD in the AML genome by comparing WT and *HOTTIP*$^{-/-}$ MOLM13 cells (Figure 1F). KO of *HOTTIP* disrupted the posterior HOXA locus TAD, but did not affect anterior HOXA locus TAD that are demacrated by the *CBS7/9* boundary (Figure 1F). Thus, the data revealed that *HOTTIP* is involved in organization of TAD in the AML genome to drive aberrant posterior HOXA genes expression.

HOTTIP lncRNA is aberrantly expressed in a subset of AML patients and cells

We then analyzed TCGA-LAML and TARGET-AML RNA-seq datasets to examine *HOTTIP* expression patterns. Compared to the *NPM1* WT (*NPM1*$^{C-}$) and *MLL* WT (*MLLr*$^{-}$) cases (n=245), *NPM1*-mutated (*NPM1*$^{C+}$) or *MLL*-rearranged (*MLLr*$^{+}$) AML cases (n=76) exhibited elevated levels of *HOTTIP* expression (Figure 2A). Overall survival was significantly longer in patients having AML with low *HOTTIP* expression (bottom 30th percentile) than those having AML with higher *HOTTIP* expression (top 30 percentile, Figure 2B). Given that *HOTTIP* is aberrantly expressed in the *MLLr*$^{+}$ and *NPM1*$^{C+}$ AML (Figures S2A-S2B), we further analyzed RNA-seq data obtained from the TCGA-LAML and TARGET-AML datasets for the correlations between expression levels of *HOTTIP* and of posterior HOXA genes and leukemogenic genes. *HOTTIP* expression positively correlated with expression of posterior HOXA genes, *TWIST1*, and oncogenes *MEIS1* and *PBX3* in AML (Figure 2C). Thus, *HOTTIP* plays an important role in the pathogenesis and prognosis of AML patients.

HOTTIP establishes aberrant chromatin signature to drive AML specific transcription profile

To further investigate how *HOTTIP* regulates aberrant posterior HOXA locus TAD and gene expression in AML, we carried out ChIRP-seq to examine global *HOTTIP* binding in the *MLLr*$^{+}$ AML genome of WT and *HOTTIP*$^{-/-}$ MOLM13 cells. *HOTTIP* lncRNA bound to *HOXA9-HOXA13 in cis*, but not to the anterior HOXA genes or the HOXB locus (Figures 3A, S3A-S3C).

HOTTIP$^{-/-}$ greatly reduced the binding of *HOTTIP* to the posterior HOXA genes (Figures 3A, S3B) supporting that *HOTTIP* is a regulator of posterior HOXA genes. The global binding site distribution of *HOTTIP* in AML genome revealed that *HOTTIP* mainly binds to noncoding regions (Figure 3B). Although *HOTTIP* bound to 3,767 genomic sites, it only directly bound to the promoters of 259 annotated genes that are mainly involved in hematopoiesis, myeloid cell differentiation, cell cycle progression, JAK-STAT, and WNT signaling pathways (Figure 3C), concomitant with GO enriched pathways obtained from the changed transcriptomic profiles upon *HOTTIP*$^{-/-}$ (Figure 1D). In addition, the GO analysis of *HOTTIP* bound intergenic regions also revealed that *HOTTIP* targets are consistently involved in chromatin organization, myeloid cell differentiation, and hematopoiesis (Figure S3D). Interestingly, *HOTTIP* also bound *in trans* in *PBX3*, *MYC*, *KIT*, *CD33*, *MEIS2,* and *RUNX1* promoters (Figures 3D, S3E). To examine if *HOTTIP* is indeed regulating hematopoietic transcription program, we performed *de novo* motif analysis of the *HOTTIP* binding sites from ChIRP-seq (Figure 3E, Table S1). Consistently, the top transcription factor (TF) motifs bound by *HOTTIP* are those involved in HS/PC function, such as RUNX1, MYC, E-box, and STAT5 motifs (Figure 3E, Top), suggesting that these factors may interact with *HOTTIP* to mediate its function in hematopoiesis. To confirm this notion, we carried out RNA-immunoprecipitation (RIP) and showed that *HOTTIP* physically interacts with phosphorylated STAT5A, MYC, RUNX1, DOT1L, and MLL1 complex, but not HDAC1 and IKAROS controls (Figures 3F, S3F). Thus, *HOTTIP* regulates hematopoietic chromatin landscape and transcription program by interacting with hematopoietic specific TFs and epigenetic regulators.

To test whether *HOTTIP* controls chromatin signature of its targets, we performed ChIP-seq and ATAC-seq assays in WT and *HOTTIP*$^{-/-}$ MOLM13 cells. Consistent with transcriptional changes of the posterior HOXA genes, *HOTTIP*$^{-/-}$ resulted in marked decreases in H3K4me3 and H3K79me2, while expanded and elevated H3K27me3 levels in the posterior HOXA genes domain

(Figure 3G). In contrast, *HOTTIP*$^{-/-}$ affected neither the anterior HOXA genes domain nor the HOXB genes locus (Figures 3G, S3G). In *HOTTIP*$^{-/-}$ cells, significantly gained or lost chromatin accessibility in subset of genomic regulatory regions was observed (Figure 3H). *HOTTIP*$^{-/-}$ led to a significant decrease in chromatin accessibility only in the posterior HOXA genes domain, but not anterior HOXA genes or HOXB genes clusters (Figures 3I, S3H, and Table S2). In addition, the MLL1 recruitment and chromatin modification/accessibility of a subset of non-HOXA genes were impaired upon loss of *HOTTIP* binding, consistent with the binding of *HOTTIP* to these genes (Figures 3D, S3E, and Table S2). The *de novo* motif analysis of ATAC-seq altered peaks confirmed that the top *HOTTIP* bound motifs also exhibits significant chromatin accessibility alteration (Figure 3E, Bottom; Table S1). Thus, *HOTTIP* establishes aberrant HOXA genes associated chromatin signature to drive ectopic transcription profile in the *MLLr*$^{+}$ AML.

HOTTIP loss perturbs AML cell proliferation and prolongs survival of the transplanted AML mouse models

We next assessed the effects of *HOTTIP*$^{-/-}$ on leukemic cell growth and viability. Compared to WT MOLM13 cells, *HOTTIP*$^{-/-}$ showed consistent inhibitions of the posterior HOXA genes (Figure 4A) and cell proliferation (Figure 4B). Cell cycle analysis revealed that *HOTTIP*$^{-/-}$ blocked MOLM13 cells in the G_1 phase and significantly reduced the G_2/M phases (Figure 4C), suggesting that *HOTTIP* controls AML cell proliferation by regulating cell cycle progression consistent with RNA-seq analysis (Figure 1D). To exclude any effect of possible regulatory elements presented in the genomic *HOTTIP* region, we created CRISPR-dCas9-KRAB mediated *HOTTIP* epigenetic silenced clones using the KRAB repressive domain that recruits H3K9 methyltransferase Suv39H1 (Gilbert et al., 2013) to specifically target the promoter region of *HOTTIP* gene in *MLLr*$^{+}$ MOLM13 and *NPM1*$^{C+}$ OCI-AML3 cells. dCas9-KRAB targeted at the *HOTTIP* promoter elevated H3K9me2 levels, while inhibited H3K4me3 levels across *HOXA9-A13* genes (Figures S4A-S4B). Inhibition of *HOTTIP* significantly reduced *HOTTIP* binding to posterior HOXA and non-HOXA

genes in both cell lines (Figures S4C-S4E). As a result, *HOTTIP* target gene expression, cell proliferation, and cell cycle progression were specifically inhibited by *HOTTIP* inhibition (Figures S4E-S4K), consistent to the phenotypes of *HOTTIP*$^{-/-}$.

To test whether *HOTTIP* loss affects AML leukemogenesis *in vivo*, we transplanted 5x10^5 WT or *HOTTIP*$^{-/-}$ MOLM13 cells into irradiated NSG mice. All mice transplanted with WT MOLM13 cells died around 40 days after transplantation, while mice receiving *HOTTIP*$^{-/-}$ cells survived more than 70 days (Figure 4D). Indeed, FACS analysis of recipients at 35 days after transplantation revealed that the human CD45$^+$ (hCD45$^+$) cell chimerism was significantly reduced in mice receiving *HOTTIP*$^{-/-}$ MOLM13 cells (Figure 4E). Consistently, immunostaining of femur sections showed that mice transplanted with *HOTTIP*$^{-/-}$ cells had decreased infiltration of hCD45$^+$ AML blasts in bone marrow (BM, Figure 4F). Thus, deletion of *HOTTIP* reduces AML leukemic burden *in vivo*.

Furthermore, *HOTTIP* was deleted in primary AML cells with *MLLr*$^+$ (#LPP4), *NPM1*$^{C+}$;*FLT3-ITD*$^+$ (#974) or *FLT3-ITD*$^+$ (#886) obtained from patients by the CRISPR-Cas9 editing. Both #LPP4 and #974 exhibited elevated expression of *HOTTIP* and posterior *HOXA9-A13* genes, while #886 had low *HOTTIP* and posterior HOXA genes expression (Figure S4L). We then transplanted 2×10^5 control or *HOTTIP*$^{-/-}$ primary AML cells into NSG mice. Interestingly, mice receiving control *MLLr*$^+$ (#LPP4) or *NPM1*$^{C+}$/*FLT3-ITD*$^+$ (#974) AML cells all died around 31 days after transplantation, while mice transplanted with corresponding *HOTTIP*$^{-/-}$ AML cells survived up to 45 days (Figure 4G). FACS analysis revealed that *HOTTIP*$^{-/-}$ dramatically decreased the hCD45$^+$ cell chimerism in BM, spleen and peripheral blood (PB) of recipients (Figure 4H). In contrast, *HOTTIP*$^{-/-}$ neither prolonged the survival nor decreased hCD45$^+$ cell chimerism in mice transplanted with *FLT3-ITD*$^+$ (#886) (Figures 4G, 4H). Thus, loss of *HOTTIP* decreases tumor burden and attenuates leukemic progression *in vivo* specific for *HOTTIP* activated AML patients carrying *MLLr*$^+$ or *NPM1*$^{C+}$ mutations.

HOTTIP activation rescues AML-associated chromatin signature and transcription profile in the _CBS7/9_$^{+/-}$ AML cells

Since *CBS7/9*$^{+/-}$ strongly suppressed *HOTTIP* expression (Figure 1A), we next sought to test if *HOTTIP* acts downstream of *CBS7/9* to organize the AML-associated TAD and transcription profile using the dCAS9-VP160 mediated promoter activation of the endogenous *HOTTIP* gene in the *CBS7/9*$^{+/-}$ MOLM13 cells. Reactivation of *HOTTIP* in the *CBS7/9*$^{+/-}$ MOLM13 cells largely restored the expression of the posterior *HOXA9-HOXA13* genes (Figure 5A). *HOTTIP* reactivation partially rescued the defective cellular proliferation in the *CBS7/9*$^{+/-}$ MOLM13 cells by escaping G_1 phase blockage and restoring G_2/M phase (Figures 5B-5C).

Next, we carried out RNA-seq, ATAC-seq, ChIP-seq, and Hi-C analysis using WT, *CBS7/9*$^{+/-}$ and the *HOTTIP*-activated *CBS7/9*$^{+/-}$ MOLM13 cells. *CBS7/9*$^{+/-}$ downregulated 865 genes involved in myeloid differentiation and cell cycle controls (Luo et al., 2018). *HOTTIP* activation largely reversed the transcription profiles of the gene sets affected by the *CBS7/9*$^{+/-}$, making them closely resembling WT MOLM13 cells (Figure 5D). These genes are involved in hematopoiesis, myeloid differentiation, and cell cycle controls (Figures 5D-5E). GSEA revealed that the pathways involved in *HOXA9* regulation, AML progression, JAK-STAT signaling, and NOTCH signaling were enriched in *HOTTIP*-activated cells as compared to *CBS7/9*$^{+/-}$ MOLM13 cells (Figures 5F, S5A). In addition, chromatin accessibility and MLL1 recruitment was largely rescued in the posterior HOXA genes domain (Figures 5G-5H, Table S2) and the promoters of *RUNX1, TWIST1, STAT5A and MYC* (Figures S5B-S5C and Table S2), but not HOXB genes (Figure S5D). Consistently, the posterior HOXA locus TAD that was largely disrupted by the *CBS7/9*$^{+/-}$ was rescued by *HOTTIP* reactivation (Figure 5I). Thus, *HOTTIP* plays a critical role in establishing and maintaining leukemic specific HOXA locus TAD and gene expression profiles.

To evaluate whether of *HOTTIP* reactivation functionally rescues the *CBS7/9*$^{+/-}$ mediated anti-leukemia effect, we again transplanted $2x10^5$ WT, *CBS7/9*$^{+/-}$, or *HOTTIP* activated *CBS7/9*$^{+/-}$

MOLM13 cells into irradiated NSG mice. All mice transplanted with *CBS7/9*$^{+/-}$ cells died between 29-39 days after transplantation, while the mice receiving WT or *HOTTIP* activated *CBS7/9*$^{+/-}$ MOLM13 cells survived only 14-25 days (Figure 5J). FACS analysis of recipients at 14 days after transplantation revealed that the hCD45^{+} cell chimerism in the mice receiving *HOTTIP*-reactivated *CBS7/9*$^{+/-}$ MOLM13 cells was restored to compatible levels of mice transplanted with WT cells (Figure 5K). Thus, reactivation of *HOTTIP* in the *CBS7/9*$^{+/-}$ MOLM13 cells rescues leukemic HOXA locus TAD, chromatin signature and transcription profile to reverse the *CBS7/9*$^{+/-}$ mediated anti-leukemic effects.

Hottip **transgenic expression in the hematopoietic compartment perturbs HSC pools and leads to AML-like disease in mice**

Hottip expression is high in HSCs and early progenitors but decreases upon terminal differentiation (Figure S6A). Since *HOTTIP* is aberrantly expressed in the *MLLr*$^{+}$ and *NPM1*$^{C+}$ AMLs (Figure 2A) resulting in dysregulation of HOX genes-associated chromatin domain and gene expression (Figures 1B, 1F), enforced expression of *HOTTIP* in HS/PC may perturb HSC function *in vivo*. We then generated transgenic mice that express *Hottip* lncRNA under the control of the hematopoietic specific *Vav1* enhancer and promoter (Figure S6B). The transgene was integrated in chromosome 10qE4 (Chro10:60117191) (Figure S6C). Two transgenic mouse lines were obtained, which exhibited ~5- and ~11 fold increase in *Hottip* expression as compared to the endogenous *Hottip* levels in BM cells, respectively (Figure S6D). The *Hoxa9-a13* genes were also aberrantly elevated upon *Hottip* transgenic expression (Figure S6E). When a cohort of WT and *Hottip*-Tg mice from both lines (6-18 month of age) were analyzed, the *Hottip*-Tg mice exhibited increased white blood cells (WBC) and neutrophil counts (Figure 6A) and most of them developed splenomegaly (Figure S6F), indicating that enforced expression of *Hottip* led to perturbation of hematopoiesis. FACS analysis of BM cells revealed that the c-Kit^{+} cell population was significantly increased in *Hottip*-Tg mice (Figure 6B). Importantly, ~32% of *Hottip*-Tg mice

died by 18 months of age (Figure 6C). May-Giemsa staining of PB smears and BM cytospins revealed increases in immature myeloid cells in these *Hottip*-Tg mice (Figure 6D). Further histological analysis identified myeloid cell infiltration in the liver of these mice (Figure 6D). However, whole exome sequencing (WES) of bulk tumor (BM) and non-tumor (skin) cells from five diseased *Hottip*-Tg mice did not reveal any recurrent mutation in genes implicated in HSC regulation or leukemogenesis in the BM of *Hottip*-Tg mice (Table S3). The only recurrently mutated genes in *Hottip*-Tg BM cells is *Mroh2a*, which has unknown cellular function. Thus, *Hottip* overexpression plays an intrinsic oncogenic role in the pathogenesis of myeloid malignancies *in vivo*.

FACS analyses of the BM HSCs of young WT and *Hottip*-Tg mice (8-16 weeks old) revealed that overexpression of *Hottip* increased the frequencies and pools of lineage$^-$Sca-1$^+$c-Kit$^+$ (LSK) cell population in mice (Figures 6E, 6F). Furthermore, both the frequencies and total numbers of long-term (LT) and short-term (ST) HSCs were dramatically increased in *Hottip*-Tg mice (Figure 6F). Of note, alteration of proportions of common myeloid progenitor (CMP), megakaryocyte-erythroid progenitor (MEP), granulocyte-macrophage progenitor (GMP) and mature lineage cell populations in the BM were not evident in young *Hottip*-Tg mice (Figures S6G-S6H), but these abnormalities developed as mice aging.

To determine the impact of the *Hottip* transgene levels on the hematological phenotypes, the homozygous *Hottip*-Tg (*Hottip^{Homo}*-Tg) mice were generated with one fold increase in the levels of *Hottip* expression as compared to the *Hottip*-Tg mice (Figure S6I, Left) and had a shorter disease latency than *Hottip*-Tg mice. Half of the *Hottip^{Homo}*-Tg mice (5 of 10) developed AML-like disease within 8 months of age. These mice exhibited severe anemia, splenomegaly, >20% of blast in PB or BM, and increased CD117$^+$CD11b$^+$ immature myeloid populations in their BM (Figures S6I-S6L). When young *Hottip^{Homo}*-Tg were analyzed for composition of hematopoietic populations, the *Hottip^{Homo}*-Tg mice had higher frequencies and pools of LSK and lineage$^-$Sca-1$^-$

c-Kit^{+} (LK) cell populations than *Hottip*-Tg mice (Figure S6M). Consistently, the frequencies and total numbers of LT- and ST-HSCs, and multipotent progenitor cells (MPPs) were significantly higher in the BM of *Hottip*^{Homo}-Tg mice than those of WT or *Hottip*-Tg mice (Figure S6N). In addition, the frequencies of GMP were increased, and MEP frequencies were decreased in the BM of *Hottip*^{Homo}-Tg mice as compared to WT or *Hottip*-Tg mice that did not exhibit such alterations at such young age (Figure S6O). These data suggest that the *Hottip* overexpression alters HSC pool and homeostasis *in vivo* in a gene dosage dependent manner.

***Hottip* regulates the balance of self-renewal and differentiation of HSCs**

Next, we sought to investigate the role of *Hottip* in HSC function by assessing the frequencies of colony-forming unit cells (CFU-C) in the BM and spleen of WT and *Hottip*-Tg mice (Figure 7A). The frequencies of each type of CFU-C including CFU-GM (Granulocyte/monocyte), burst forming unit-erythrocyte (BFU-E) and CFU-GEMM (Granulocyte/erythrocyte/monocyte/megakaryocyte) were significantly higher in both BM and spleen of *Hottip*-Tg mice (Figure 7A). When replating assays were performed on WT and *Hottip*-Tg LSK cells purified from the BM of mice, a higher replating potential was observed in *Hottip*-Tg LSK cells (Figure 7B).

Both symmetric and asymmetric cell divisions are required for the preservation of normal HSC pool and continuous production of blood cells. To test if *Hottip* aberration alters cell fates of HS/PCs, we performed paired-daughter cell assays to assess the proportions of symmetric self-renewal, symmetric differentiation and asymmetric divisions using primitive CD34^{-} LSK cells isolated from WT and *Hottip*-Tg BM. Overexpression of *Hottip* increased in proportion of CD34^{-} LSK cells with symmetric self-renewal capacity (Figure 7C). In contrast, the proportion of cells that underwent symmetric differentiation were decreased, while the asymmetric division of CD34^{-} LSK cells was not affected (Figure 7C). The abnormal behavior of CD34^{-} LSK caused by *Hottip*-

Tg is similar to *ASXL2*$^{-/-}$ or *AML-ETO* expression, which enhances HSC self-renewal and blocks myeloid differentiation.

To further interrogate the effects of the ectopic *Hottip* expression on HS/PC proliferation and differentiation, liquid cultures were performed and the frequencies and total numbers of c-Kit$^+$ cells and CFU-Cs in the progenies were analyzed weekly. FACS analysis and colony assays revealed that *Hottip*-Tg LK cells gave rise to a greater number of c-Kit$^+$ progenies and a higher number of CFU-Cs than WT LK cells each time point assayed (Figures S7A-S7C). Thus, the ectopic expression of *Hottip* in HS/PCs leads to an increased expansion of c-Kit$^+$ cells and is likely accompanied by increased proliferation and impaired differentiation, two hallmarks of leukemogenesis.

Cell-autonomous effect of *Hottip* overexpression on HS/PC functions

Next, we carried out competitive transplantation assays to examine the repopulating capacity of *Hottip*-Tg and WT BM cells as compared to WT competitor BM cells in recipient mice. When the donor cell chimerism was analyzed kinetically in the PB of recipient mice for 6 months, the CD45.2 (Donor) cell population remained ~50% in mice receiving WT BM cells, whereas the CD45.2 chimerism in mice receiving *Hottip*-Tg BM cells steadily increased, reaching ~80% 6 months after transplantation (Figure 7D, Top). Strikingly, *Hottip*-Tg BM cells generated higher proportions and numbers of LSKs in recipient animals as compared to WT BM and competior BM cells (Figure S7D). The proportions of each mature lineage population produced by *Hottip*-Tg BM cells were comparable to that produced by WT and competitor BM cells up to 6 months after trasplantation (Figure S7D). However, when we transplanted *Hottip*^{Homo}-Tg BM cells into recipients (Figure S7E), 40% of the *Hottip*^{Homo}-Tg recipients developed AML-like disease within 6 months post-transplantation (Figure S7F). The *Hottip*^{Homo}-Tg BM cells exhibited a significantly enhanced repopulation capacity and generated higher proportions and numbers of LSK, LK, LT-HSC, CMP and immature myeloid cells (CD117$^+$CD11b$^+$ and CD117$^+$Gr1$^+$) cells as compared to

WT and competior BM cells (Figures S7G), indicating that *Hottip* transgene perturbs HSC function and hematopoiesis in a gene dosage dependent manner.

To enforce long-term HSC self-renewal and disease development, secondary transplantation was carried out by transplanting WT or *Hottip*-Tg BM cells from primary recipients (1 X 10^6 cells) into 2nd recipients. In this context, 80% of the *Hottip*-Tg 2nd transplanted recipients developed AML-like diseases as evident by splenomegaly, significant increases of immature myeloid cells in the PB or BM (Figures 7E-7G). The donor cell chimerism in the PB of the 2nd recipients transplanted with *Hottip*-Tg BM cells continuously increased, from 80% reaching 93% 6 months after 2nd transplantation (Figure 7D, Bottom). *Hottip*-Tg BM cells generated much higher proportions and numbers of LSKs and immature myeloid cells in 2nd recipient mice when compared to WT BM cells and competior BM cells (Figure 7G). Thus, ectopic expression of *Hottip* in hematopoiesis increases the repopulating capacity of HS/PCs and promotes HSC self-renewal leading to AML-like disease.

Transgenic expression of *Hottip* remodels chromatin accessibility and alters hematopoietic transcription programs

It is conceivable that transgenic expression of *Hottip* remodels the HOX genes-associated chromatin domain and drives hematopoietic transcription program to perturb HS/PC function. To examine this, we performed ATAC-seq and RNA-seq using BM LT- and ST-HSCs purified from WT and *Hottip*-Tg mice. RNA-seq revealed that a total of 535 genes exhibited >2 fold increases, whereas 275 genes had decreased expression in LT-HSCs upon *Hottip* transgenic expression (Figure 8A). Among upregulated genes included *Hoxa9-a13, Nanog*, *Sox2*, *Myc*, *Meis1, Runx1, Kit, Slamf1, Gata2, and Pbx3* (Figure 8B), many of which directly bound by *Hottip* (Figure S8A) suggesting that *Hottip* regulates posterior *Hoxa* genes and hematopoietic transcription networks. Furthermore, GO analysis revealed that *Hottip*-Tg altered transcription program is involved in pathways associated with the pluripotency of stem cells, cell fate commitment, long-term

potentials, HSC proliferation, myeloid progenitor differentiation, NOTCH, JAK-STAT and WNT signaling (Figure 8C). Similar hematopoietic pathways and genes were also altered in ST-HSCs upon *Hottip* transgenic expression (Figures S8B-S8C). Concomitantly, ATAC-seq revealed that *Hottip* aberration enhanced chromatin accessibility in the endogenous posterior Hoxa gene cluster and non-Hoxa gene targets (Figures 8D, S8D-S8F, Table S2). As a control, *Hottip* expression did not alter chromatin accessibility in anterior Hoxb genes domain (Figure S11G), which also plays an important role in HSC function. We further grouped ATAC-seq promoter gained or lost peaks and carried out GO analysis to exam the pathways associated with gain or loss of promoter accessibility upon *Hottip* aberration in BM HSCs. Gained peaks were associated with pathways involved in the pluripotency of stem cells, cell fate commitment, long-term potentials, HSC proliferation, and JAK-STAT and WNT signaling, while lost peaks were associated with hematopoietic cell lineage, the cellular differentiation program, and myeloid progenitor cell differentiation (Figures 8E, S8E). Thus, *Hottip* acts as an epigenetic regulator directing hematopoietic transcription programs.

In both LT- and ST-HSCs, WNT signaling pathway was upregulated by *Hottip* transgenic expression (Figures 8C, 8E, S8E) suggesting that canonical WNT signaling pathway may play an important role in *Hottip*-driven leukemia. The WNT signaling is required for HSC homeostasis and leukemia development (Lento et al., 2013; Reya et al., 2003). To test clinical implication of *Hottip*-driven leukemogenesis, we treated primary AML patient samples carrying *MLLr*$^{+}$ and exhibiting elevated *HOTTIP* expression with DMSO (control) or 500 nM ICG-001, a canonical WNT inhibitor (Emami et al., 2004) (Figure S11H). As a result, long-term culture initiating cell frequency of the *HOTTIP* expressed AML samples were significantly inhibited by ICG-001 while *HOTTIP* silenced/or weakly expressed *MLLr*$^{-}$ AML samples were resistant to the ICG-001 (Figures 8F, S8I-S8J). Furthermore, the primary *MLLr*$^{+}$ AML patient samples MLL7 that expressed high levels of *HOTTIP* were transplanted into NSG mice and treated with vehicle (control) or ICG-001 (50 mg/Kg). Treatment of ICG-001 significantly prolonged leukemic mice survival (Figure 8G) by

decreased CD45/CD33 human leukemic blast in BM and eradicated the human leukemic blast in spleen and liver (Figure S8K). Thus, *MLLr*$^{+}$ AML expressing high level of *HOTTIP* is indeed sensitive to WNT inhibitor and WNT signaling pathway facilitates *Hottip*-driven leukemogenesis.

Discussion

HOTTIP is known to coordinate transcription of the 5' tip of HOXA genes (Wang et al., 2011). We characterized a function of *HOTTIP* in regulating the balance of HSC self-renewal and differentiation. The *HOTTIP* action is, in part, dependent on its ability to directly bind to and regulate genes and pathways that are required for HS/PC regulation. It is particularly interesting that the WNT pathway is required for self-renewal of leukemia stem cells (LSCs) in AML that are driven by *MLLr*$^{+}$ or its targets, *MEIS1* and *HOXA9* (Wang et al., 2010; Yeung et al., 2010). *HOTTIP* is highly expressed in *MLLr*$^{+}$ or *NPM1*$^{C+}$ mutated AML patients and the *HOTTIP* expressed AML is sensitive to the WNT inhibitor. Thus, *HOTTIP* may be also involved in regulation of LSCs carrying *MLLr*$^{+}$ or *NPM1*$^{C+}$ mutations and represents a therapeutic opportunity for eradication of LSCs in part through manipulating the WNT pathway.

MLLr^{+} AML is one of the most devastating subtypes of AML, being associated with poor prognosis and chemo-resistance. Although *NPM1*$^{C+}$ mutation alone is generally associated with favorable prognosis, coexistence of *FLT3-ITD* and/or *DNMT3A* mutations predicts an increased risk of relapse and poorer outcome (Gale et al., 2015; Grimwade et al., 2016). *HOXA9* is a strong predictor of poor prognosis in AML (Collins and Hess, 2016). *MEIS1* and *PBX3* are oncogenes that co-express with many HOX genes, especially *HOXA9,* to stimulate the proliferation of HSCs (Li et al., 2016; Takeda et al., 2006). They are highly expressed in AML cases carrying *MLLr*$^{+}$ or *NPM1*$^{C+}$ mutation and depended on *HOTTIP* lncRNA. Interestingly, AMLs harboring *NPM1*$^{C+}$ with aberrant *HOXA9/A10* genes and homeotic oncogenes, *MEIS1* and *PBX3,* are synergistically required for the maintenance of *NPM1*$^{C+}$ driven AML (Brunetti et al., 2018; Dovey et al., 2017;

Kuhn et al., 2016). The reports are consistent with that *Hottip* overexpression resulted in activation of *Hoxa9*, *Meis1,* and *Pbx3* in the HSC population, promoting HSC self-renewal and leukemogenesis.

Although TADs are mostly conserved across cell types and species, TADs are indeed structural and functional chromosomal units that constrain enhancer/promoter communication for specific transcription program (Valton and Dekker, 2016). Altered TAD might result in inappropriate promoter/enhancer interactions to alter transcription of oncogenes or tumor suppressors (Groschel et al., 2014; Taberlay et al., 2016). Chromatin boundaries, CTCF binding sites in many cases, play a critical role in defining TADs and chromatin signature within the TAD (Luo et al., 2018; Narendra et al., 2015). Given that expression of *HOTTIP* restores *CBS7/9* mediated posterior HOXA locus TAD and leukemogenesis, it is likely that stratification of CTCF boundary and oncogenic TAD in the HOXA locus by *HOTTIP* lncRNA may be exploited by *MLLr*$^{+}$ or *NPM1*$^{C+}$ AML cells to promote leukemogenesis.

Apart from the HOXA genes, *HOTTIP* lncRNA also bound and regulated a subset of non-HOX genes. RUNX1 is required for definitive hematopoiesis and could act to promote the survival of *MLLr*$^{+}$ leukemia cells (Goyama et al., 2013). Intriguingly, epithelial-to-mesenchymal transition (EMT) genes are recently shown to control AML blast migration and invasion and to link to aggressiveness and poor prognostic outcomes of MLL-AF9 mediated AML (Stavropoulou et al., 2016). The question remains as to whether *HOTTIP* directly regulates these non-HOXA genes or modulates them through an indirect mechanism. Although KO of *HOTTIP* decreased more than 80% of *HOTTIP* transcript levels, it is interesting that there is only approximately 50% reduction in *HOTTIP* binding to its putative targets, posterior HOXA genes. In contrast, the binding of *HOTTIP* in the *trans* regulated non-HOXA genes was almost completely eliminated. Of note, transgenic expression of *Hottip* from the chromosome 10 is able to activate Hoxa and non-Hox genes. It is likely that *HOTTIP* mainly acts in *cis* to regulate the posterior HOXA genes in normal development due to its low expression levels. However, when *HOTTIP* is overexpressed, such

as in AML or in transgenic mice, the overexpressed *HOTTIP* may go to other chromatin sites besides the posterior HOXA genes. Thus, *HOTTIP* could act in cis and/or in *trans* in a context dependent manner. The mechanism of *HOTTIP* lncRNA in gene regulation warrants further investigation.

Acknowledgments:

We thank Penn State College of Medicine Genome Science Facility. This work was supported by grants from National Institute of Health (S.H., R01DK110108, R01CA204044; HL141950; M.X., R01CA172408, R01HL145883, R01HL141950; Y.Q., R01HL144712), Programme grants by Cancer Research UK and Bloodwise (C.W.E.S) and the Four Diamonds Fund (SH).

Author Contribution

H.L., G.Z., Y.Q., Z.J., B.Y., T.F.K., B.B.Z., C.W.E.S., Q.L., M.X., and S. H. designed and performed experiments. G.Z., Y.G., F-C.Y., and M.X. established transgenic and transplantation mouse models. H.L., J.X., Y.C., and W.L. performed bioinformatics and statistical analysis. C.C., X.B., and F.W. provided human patient samples and reagents. S.H., Y.Q., and M.X. wrote the manuscript.

Declaration of Interests:

The authors have declared that no competing interests exist.

Figure Legends

Figure 1. *HOTTIP*$^{-/-}$ perturbs HOXA genes-mediated oncogenic transcription program. (A) *HOTTIP* levels in WT vs. *CBS7/9*$^{+/-}$ MOLM13 cells by RNA-seq. FPKM: Fragments Per Kilobase of transcript per Million mapped reads. **(B)** Heat map of >2 folds up- and downregulated genes upon *HOTTIP* KO by RNA-seq. **(C)** mRNA levels of HOXA genes in WT and *HOTTIP* KO MOLM13 cells. **(D)** GO analysis of genes whose expression was altered by *HOTTIP* KO. **(E)** Enrichment of decreased genes involved in JAK-STAT (Top) and AML (Bottom) pathways upon *HOTTIP* KO by GSEA. **(F)** Hi-C interacting maps in part of the human chromosome 7p15 region containing the HOXA locus compared WT and *HOTTIP* KO MOLM13 cells. See also Figure S1.

Figure 2. *HOTTIP* lncRNA is aberrantly expressed in a specific subset of AML. (A) The *HOTTIP* levels in *NPM1*$^{C-}$ and *MLLr*$^{-}$ AML cases and in *NPM1*$^{C+}$ and *MLLr*$^{+}$ AML cases obtained from the TCGA LAML and TARGET datasets. Violin plots show mean, interquartile and 1.5 × interquartile. The width shows the probability density. **(B)** Kaplan-Meier curve of overall survival probabilities of the patients having AML with high or low *HOTTIP* levels from the TCGA-LAML and TARGET-AML datasets. **(C)** Significant correlation between the expression of *HOTTIP* and posterior *HOXA* genes, *MEIS1*, *TWIST1*, *PBX3* in the TCGA LAML and TARGET-AML datasets. Pearson correlation and corresponding p value is calculated by the cor.test of R. See also Figure S2.

Figure 3. *HOTTIP* reprograms AML chromatin and regulates leukemic specific transcription networks. (A) ChIRP-seq analysis of *HOTTIP* binding in WT and *HOTTIP* KO MOLM13 cells. **(B)** Pie chart shows global *HOTTIP* binding distribution in the human AML genome. **(C)** GO analysis of *HOTTIP* regulated transcription and signal pathways. **(D)** Changes in *HOTTIP* binding (ChIRP-seq), chromatin accessibility (ATAC-seq), MLL1 recruitment and H3K4me3 enrichment (ChIP-seq) upon *HOTTIP* KO. **(E)** Shared top TF binding motifs enriched by the *HOTTIP* ChIRP-seq (Top) and the ATAC-seq altered peaks (Bottom) upon *HOTTIP*$^{-/-}$ by the *de novo* motif analysis.

(F) *HOTTIP* and protein interactions detected by RIP. **(G)** ChIP-seq analysis of changes in H3K4me3, H3K27me3, and H3K79me2 modification levels upon *HOTTIP*$^{-/-}$ in MOLM13 cells. **(H)** Altered ATAC-seq accessibility of genomic regions upon *HOTTIP*$^{-/-}$. Box plots show horizontal line (zero z-score, mean), box indicating the median with upper and lower quartiles, and whiskers indicating the highest and lowest values. p value is calculated by Kolmagorov-Smirnov (KS) test. **(I)** ATAC-seq analysis of altered chromatin accessibility upon *HOTTIP*$^{-/-}$ in MOLM13 cells. See also Figure S3, Tables S1, S2.

Figure 4. *HOTTIP*$^{-/-}$ perturbs cell proliferation and prolongs survival of the transplanted AML mouse models. (A) RT-qPCR analysis of HOXA genes expression in WT and *HOTTIP*$^{-/-}$ MOLM13 clones. **(B)** Proliferation curves of WT and *HOTTIP*$^{-/-}$ MOLM13 cells. **(C)** Cell cycle analysis of WT and *HOTTIP*$^{-/-}$ MOLM13 clones. **(D)** Kaplan-Meier curves of NSG mice transplanted with WT and *HOTTIP*$^{-/-}$ MOLM13 cells. **(E)** hCD45^{+} cell chimerism in BM, spleen (SP), and PB of NSG mice receiving WT (n=4) or *HOTTIP*$^{-/-}$ (n=4) MOLM13 cells. **(F)** Hematoxylin and Eosin (H&E) and anti-hCD45 immuno-staining (brown) of femur sections from mice transplanted with WT or *HOTTIP*$^{-/-}$ MOLM13 cells for 16 days. **(G)** Kaplan-Meier curves of NSG mice transplanted with WT or *HOTTIP*$^{-/-}$ primary AML patient BM cells carrying *MLLr*$^{+}$ (LPP4), *NPM1*$^{C+}$*Flt3-ITD*$^{+}$ (974), or *NPM1*$^{C-}$*FLT3-ITD*$^{+}$ (886) mutations. **(H)** hCD45^{+} cell chimerism in BM, SP, and PB of NSG mice receiving WT (n=4) or *HOTTIP*$^{-/-}$ (n=4) primary AML cells. Data (Panels A-C, E, H) is presented as mean ± SD. See also Figure S4.

Figure 5. Activation of *HOTTIP* rescues the HOXA genes chromatin defects in the *CBS7/9*$^{+/-}$ AML cells. (A) RT-qPCR analysis of HOXA genes expression in WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 clones **(B)** Proliferation curves of the WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 cells. **(C)** Cell cycle analysis of the WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 clones. **(D)** Heat map of RNA-seq analysis shows up- and downregulated genes of WT and the

dCas9-VP-160 mediated *HOTTIP* activated clones as compared to the *CBS7/9*$^{+/-}$ clone. **(E)** RT-qPCR validation of the key altered genes identified by RNA-seq compared WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated clones. **(F)** Enrichment of upregulated genes involved in *HOXA9* (Top) and AML (Bottom) pathways upon *HOTTIP* activation in the *CBS7/9*$^{+/-}$ MOLM13 by GSEA. **(G)** ATAC-seq analysis of chromatin accessibility in WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 cells. **(H)** ChIP-seq analysis of MLL1 recruitment in WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 cells. **(I)** Hi-C interacting maps in part of the human chromosome 7p15 region containing the HOXA locus compared WT, *CBS7/9*$^{+/-}$, and the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 cells. **(J)** Kaplan-Meier curve of NSG mice transplanted with WT, *CBS7/9*$^{+/-}$, or the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 cells. **(K)** FACS analysis of hCD45^{+} cell chimerism in BM, spleen (SP), and PB of NSG mice 14 days after transplantation of WT (n=4), *CBS7/9*$^{+/-}$ (n=4), or the dCas9-VP-160 mediated *HOTTIP* activated MOLM13 cells (n=4). Data (Panels A-C, E, and K) is presented as mean ± SD. See also Figure S5.

Figure 6. *Hottip* transgenic expression in hematopoiesis perturbs HSC pools and led to AML-like disease. (**A**) Parameters of blood counts were summarized from 6-20 months old *Hottip*-Tg (n = 15) and age matched WT (n = 8) mice. WBC: white blood cells; NE: neutrophils; RBC: red blood cells. (**B**) FACS analysis and quantitation of c-Kit (CD117^{+}) cells within total BM cells of 6-20 months old WT (n = 7) and *Hottip*-Tg (n = 15) mice. Data shows all dots as mean ± SD by Student's t test. Horizontal bars represent mean. (**C**) Kaplan-Meier curve of WT (n=21) and *Hottip*-Tg (n=31) mice over 20 months. **(D)** Images of PB smears, BM cytospins and liver sections prepared from representative WT and *Hottip*-Tg mice. Bar, 20 μm. **(E)** FACS analysis of LSK and LK populations in the BM Lin^{-} cells (Top) as well as LT-HSC, ST-HSC and MPP populations in the BM LSK cells (Bottom) of representative young (8-16 weeks) WT and age matched *Hottip*-Tg mice. (**F**) Quantitation of the total LSK, LK, LT-HSC, ST-HSC and MPP

populations per femur of young WT (n=7) and *Hottip*-Tg (n=10) mice. Quantitation data is presented as mean ± SD. See also Figure S6, Table S3.

Figure 7. *Hottip* transgenic expression perturbs HSC function leading to AML-like disease. (**A**) Frequencies of CFU-Cs in the BM and spleen cells from WT and *Hottip*-Tg mice. GM: granulocytes/macrophages; BFU-E: burst forming unit-erythrocyte; GEMM: granulocytes/erythrocyte/monocyte/megakaryocyte. (**B**) Frequencies of colonies per 100 LSK cells in WT and *Hottip*-Tg BM cells are shown (1st). Colonies were replated every 7 days for 4 times (2nd-5th). (**C**) Paired-daughter cell assays were performed on $CD34^-$ LSK cells from WT, *Hottip*-Tg mice, *Asxl2*$^{-/-}$ mice, and *AML-ETO* mice. **(D)** FACS analyses of CD45.2 (Donor) chimerisms in the PB of recipients (CD45.1) receiving WT or *Hottip*-Tg BM cells in 1st transplantation (Top) and 2nd transplantation (Bottom)**. (E)** Kaplan-Meier curve of 2nd transplantation receiving WT (n=5) and *Hottip*-Tg (n=5) BM cells (Top) and appearance of spleens and femur of representative WT and moribund *Hottip*-Tg mice receiving 2nd transplantation (Bottom). (**F**) Images of PB smears (Top) and BM cytospins (Bottom) prepared from representative WT and moribund *Hottip*-Tg mice receiving 2nd transplantation. Scale bar, 20 μm. **(G)** FACS analyses showing CD45.2 vs. CD45.1 chimerism as well as their respective lineage distribution and LSK/LK cell populations (within Lin^- cells) in the BM of representative WT or *Hottip*-Tg mice receiving 2nd transplantation. Data (Panels A, B, D) is presented as mean ± SD. See also Figure S7.

Figure 8. Transgenic expression of *Hottip* alters HSC chromatin signature and hematopoietic transcription programs. (A) Scatter plot of RNA-seq analysis of >2 folds of differentially expressed genes upon overexpression of *Hottip* in BM LT-HSCs. **(B)** Heat map showing changed expression of representative genes upon *Hottip* overexpression. **(C)** GO analysis of the *HOTTIP* affected genes. **(D)** ATAC-seq analysis of chromatin accessibility in WT and *Hottip*-Tg BM LT-HSCs. **(E)** ATAC-seq promoter density map of LT-HSCs sorted from WT

and *Hottip*-Tg BM. Up- (Top) or downregulated (Bottom) ATAC-seq promoter peaks correlate with GO enriched pathways annotated by GREAT analysis. **(F)** Primary *MLLr*$^{+}$ AML patient samples with elevated *HOTTIP* expression were treated with DMSO or ICG-001 (500 nM) and the LTC-IC frequency of each group were determined. The black bar represents the mean expression of each group. Data is presented as mean ± SD. **(G)** Primary *MLLr*$^{+}$ AML patient samples MLL7 (2.5 million cells) were transplanted into NSG mice. The mice were treated with vehicle (n=4) or ICG-001 (50 mg/Kg; n=5) and sacrificed when they showed the signs of illness. See also Figure S8.

STAR ★ METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to and will be fulfilled by the lead contact, Dr. Suming Huang (shuang4@pennstatehealth.psu.edu)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

AML patient samples

The primary AML patient samples, #886, *FLT3-ITD*$^{+}$/*NPM1* WT; #974, *MLLr*$^{-}$/*NPM1*$^{C+}$/*FLT3-ITD*$^{+}$; #LPP4, *MLLr*$^{+}$/*NPM1* WT, were obtained with informed consent and studies were approved by the Institutional Review Board (IRB) of the University of Florida.

AML Cell lines

HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% bovine serum (FBS). MOLM13 and OCI-AML3 cells were cultured in RPMI 1640 medium with 10% FBS and alpha-MEM with 15-20% FBS, respectively.

Generation of the *Hottip* transgenic (Tg) mouse model

The complete coding region sequence of mouse *Hottip* was cloned into downstream of a *Vav1* promoter (HS321/45-vav vector) followed by *Vav1* enhancer to ensure transgene expression solely in hematopoiesis (Yang et al., 2018) (Ogilvy et al., 1999). The plasmid DNA was digested with SacII to remove the pBlueScript II SK backbone and was used for injection into pronuclei of eggs from C57BL/6 mice. Two *Hottip*-Tg founder mice were obtained by PCR screening of the tail genomic DNAs with P1 (to detect both endogenous and transgenic *Hottip* gene) and P2 primer sets (to specifically recognize the transgenic *Hottip* gene). Transgenic founder mice were crossed with WT C57BL/6 mice. *Hottip* negative siblings of the *Hottip*-Tg mice were used as controls throughout the study. Two *Hottip*-Tg lines were used for this study. The primer sets of P3 was used for real-time PCR to recognize both endogenous *Hottip* gene and *Hottip* transgenes. The levels of transgenic expression of *Hottip* were confirmed by RNA-seq analysis. All animal experiments were approved by and performed in compliance with the regulatory guidelines by the University of Texas Health Science Center at San Antonio.

METHOD DETAILS

CRISPR-Cas9 mediated HOTTIP lncRNA knock-out and lentivirus production

HOTTIP knockout (KO) MOLM13 leukemia cell was generated according to the Neon Transfection User Guide and the Alt-R CRISPR-Cas9 System User Guide (2014) Neon® Transfection System for transfecting mammalian cells, including primary and stem cells, with high transfection efficiency (Integrated DNA Technologies, available at https://tools.thermofisher.com/content/sfs/manuals/neon_device_man.pdf). Briefly, CRISPR-RNA (crRNA) and tracrRNA were mixed and annealed in 95 °C for 5 min and then cooled down to room temperature. The crRNA:tracrRNA duplex and S.p. Cas9 Nuclease components were

combined together and then mixed with 500,000 AML cells for electroporation with Neon® System. After 24 hr or 96 hr, 100 µL of cell suspension was used for DNA extraction using Qiagen quick Extract kit and the mutation was verified through Sanger sequence. *HOTTIP*$^{-/-}$-#1 targeted region is Chr7: 27241953-27241985; *HOTTIP*$^{-/-}$-#2 targeted region is Chr7: 27240098-27240123.

dCas9-mediated inactivation and overexpression of *HOTTIP* in AML cells

Two guide RNAs plasmid targeting the promoter regions of *HOTTIP* (Table S4) were designed using the Zhang laboratory web tool (http://crispr.mit.edu), and cloned into the pLKO5.sgRNA.EFS.tRFP vector (Addgene #57824). The gRNA plasmids encoding mCherry and puromycin resistance were co-transfected with a plasmid encoding dCas9-KRAB (pHR-SFFV-dCas9-BFP-KRAB, addgene plasmid #46911) or dCas9-VP160 (pAC94-pmax-dCas9VP160-2A-puro, addgene plasmid number #48226) in MOLM13 and OCI-AML3 cells. After 24 hr post-transfection, MOLM13 or OCI-AML3 cells were selected with 2 µg/mL puromycin for another 48 hr, and then FACS sorted for RFP^{+} cells. RNA was extracted from RFP^{+} cells, and RT-qPCR was performed according to the primers list (Table S4).

Transgenic integrating location identification

The PCR-based method-TAIL-PCR (Thermal Asymmetric Interlaced PCR) which relies on a series of PCR amplifications with gene specific and degenerate primers to reliably amplify the integration sites was performed according to previous report (Pillai et al., 2008). In briefly, the primary PCR reaction was performed by mixing 50-100 ng genomic DNA, 2.5 mM dNTPs, 10 uM SP1 primer, 10 mM AD primer, and 1 U Taq polymerase in 20 µL 1X reaction buffer. In the secondary or tertiary PCR primers amplification, 1 µL first or secondary 10-20 fold diluted PCR products were used as templates in the reaction, respectively. Finally, the tertiary PCR products were verified through Sanger sequencing.

Cell cycle analysis

WT and modified MOLM13 cells were harvested and washed with phosphate buffered saline (PBS). The washed cells were fixed by adding 70% ethanol drop wise to the pellet with vortexing and incubated overnight at 4 °C. After fixation, cells were washed with PBS twice. The PBS washed cells were treated with the staining buffer (RNase A, Triton X-100, propidium iodide) and then incubated at 37 °C for 30 min in the dark. Stained samples were proceeded on the BD Accuri^{TM} C6 plus flow cytometry (BD Biosciences), and cell cycle data analysis was performed using FlowJo program. Triplicate experiments were performed for each sample.

Hematopoietic stem / progenitor cells (HS/PCs) sorting, analysis and colony assay

Flow cytometric analysis of *Hottip* transgenic mice was performed as previously described (Wang et al., 2014) using a BD LSRII flow cytometer. All data were analyzed by FlowJo-V10 software. Briefly, LK (Lin^- Kit^+), LSK (Lin^- $Sca1^+$ Kit^+), Long Term (LT) and Short-Term (ST)-HSC cells were derived from total bone marrow (BM) cells, and Lin^+ BM cells was pre-depleted by Miltenyi Biotec magnetic beads (130-110-470), then the leftover Lin^- BM cells were stained with Lin, Sca-1, c-kit, CD34, CD135 and CD16/32 antibody and sorted by BD FUSION flow cytometer. The purity of selected cells were over 95%.

Total white blood cells were obtained after lysis of Peripheral blood (PB) with red cell lysis buffer (Thermo fisher). Single-cell suspensions from bone marrow (BM), spleen and PB were stained with panels of fluorochrome-conjugated antibodies. Flow cytometric analysis of hematopoietic stem / progenitor cells (HS/PCs) was performed as previously described (Li et al., 2011). The analyses were performed using the FACS Canto II or LSR Fortessa flow cytometer (BD LSRFortessa™). All data was analyzed by FlowJo.V10 software. For colony and replating assays, bone marrow (BM) cells from the femurs and tibias of 6-8 week-old mice stained with anti-murine cKit-APC, Sca1-PE antibodies and a panel of antibody-conjugated goat anti-Rat IgG BioMag beads (Qiagen) for lineages (Lin), then cells were sorted on FACS ARIA II (Becton Dickinson), thus the enriched Lin^- $Sca1^+$ Kit^+ population (LSK cells) were obtained. For colony-

forming unit (CFU) assays, BM ($1x10^4$ cells/plate) or spleen cells ($5x10^4$ cells/plate) were plated in triplicate in methylcellulose medium (Methocult M3231) supplemented with mIL-3 (mouse interleukin 3, 10 ng/mL), hIL-6 (human interleukin 6, 100 ng/mL), hEpo (human erythropoietin, 4 U/mL) and mSCF (mouse stem cell factor, 100 ng/mL), and scored in 8-10 days. For replating assays, CFU assays were performed with LSK cells in methylcellulose medium supplemented with the same cytokine cocktails. Colonies were passaged every 7 days for 4 sequential plating.

Competitive repopulation assay

1×10^6 BM cells ($CD45.2^+$) from Ctrl or *Hottip*-Tg mice were mixed with 1×10^6 competitor BM cells (CD45.1) from B6.SJL mice and then transplanted into lethally irradiated (950 centigray) recipients (B6.SJL) by tail vein injection. Transplanted mice were monitored daily for signs of disease development.

Suspension culture assay

The LK cells were incubated in suspension culture containing 30% FBS, 2% BSA, and a combination of cytokines (mouse interleukin-3, human interkeukin-6, human erythropoietin, and mouse stem cell factor). At weekly intervals, cultures were mixed by pipetting and half of the culture media were removed, which was then replaced by newly prepared medium with the same combinations of cytokines. Cells in the collected media were counted and used for flow cytometric analysis. Total CFUs generated at each time point in the suspension culture were evaluated by culturing a fraction of the expanded cells in the colony assay as described above.

RNA immunoprecipitation (RIP) assay

The RNA-IP protocol was modified from previous reported (Deng et al., 2016; Tsai et al., 2010). The MOLM13 AML cells were collected and washed with PBS (e.g. 10^7 cells in 2 mL PBS), resuspended in freshly prepared nuclear isolation buffer (1.28 M sucrose, 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 4% Triton X-100), and then kept on ice for 20 min (with frequent mixing). Nuclei

were precipitated by centrifugation at 2,500 g for 15 min, and then resuspended in freshly prepared lysis buffer (10 mM HEPES-KOH pH7, 150 mM KCl, 5 mM $MgCl_2$, 5 mM EDTA, 0.5% IGEPAL-CA-630, 0.5 mM dithiothreitol, 0.2 mg/mL Heparin, 100 U/mL RNase OUT, 100 U/mL Superase IN, protease inhibitor tablet). Nuclei were sonicated with the Bioruptor™ UCD200. The suspension was centrifuged three times at 14,000 g at 4 °C for 10 min, and supernatant was collected and precipitated with antibody (2-10 µg) overnight at 4 °C with rotation. The precipitant was captured by the equilibrated Protein A/G magnetic beads followed by washing four times in ice-cold NT2 buffer buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% IGEPAL-CA-630) supplemented with 0.02 mg/mL heparin. The RNA-protein complexes were eluted twice with 500 mL SDS-EDTA (50 mM Tris pH 8.0, 100 mM NaCl, 10 mM EDTA, 1% SDS) for 10 min at 65°C. Coprecipitated *HOTTIP* RNA was isolated by resuspending beads in TRIzol RNA extraction reagent, eluted with nuclease-free water, treated with TURBO DNase, and then detected by RT-qPCR.

RNA isolation, quantitative RT-PCR, as well as RNA-sequencing and data analysis

Total RNAs from MOLM13 cells, OCI-AML3 AML cells, or primary *Hottip*-Tg mice long-term (LT) and short-term (ST)-HSCs were purified with the RNeasy mini-isolation kit according to manufacturer's instructions (Qiagen, MD, USA). A total of 2 µg RNA was subjected to reverse-transcription with Superscript II reverse Transcriptase (Invitrogen) and analyzed by a real-time polymerase chain reaction (PCR) Detection System (Bio-Rad). Primer sequences are listed in the supplemental Information (Table S4).

Paired end RNA-Seq was performed by Pennsylvania State University College of Medicine Genome Science Facility according to standard protocols. All of sequencing reads were processed and aligned to the mouse or human genome assembly (mm9 or hg19) using TopHat (version 2.0) and Bowtie2 (Langmead et al., 2009; Trapnell et al., 2009; Trapnell et al., 2012). To prevent false positives, a stringent approach was taken to identify differentially expressed genes.

First, FPKM (paired-end fragments per kilobase of exon model per million mapped reads) was calculated for each gene and further normalized (RMS-FPKM). Second, to prevent false positives due to the fluctuation of detection among genes with low expression levels, only genes with 50 or more reads in one of the conditions (WT control or *HOTTIP* manipulations) were included in the analysis. Differential expression was determined according to abundance estimations (FPKMs) processed with Cufflinks v2.2.1 and Cuffdiff (Trapnell et al., 2010). Differentially expressed genes were identified if the ratio of RMS-FPKM in the two conditions was greater than 2.0 fold, or undetectable in one condition but detectable by more than 50 reads in the other. The heatmaps and scatter plots were based on $\log_2$ transformation of the RMS-FPKM values. Expression level increased or decreased genes were marked with red or blue, respectively. The GO mapping of differentially expressed genes were performed with Gorilla (Eden et al., 2009). The normalized expression data was uploaded to Integrated Genomic Viewer (IGV) for visulization. The sequence reads have been deposited in the NCBI GEO under accession number (GSE114981).

Chromatin immunoprecipitation (ChIP) assay

ChIP were performed as described previously (Deng et al., 2013). Briefly, Nuclei were sonicated with the Bioruptor™ UCD200. Chromatin samples prepared from 5×10^6 cells of MOLM13 cells were immunoprecipitated with antibodies against MLL1, H3K4me3, H3K9me2, H3K27me3 and H3K79me2, separately. The immunoprecipitates were subjected to a series of washing steps to remove non-specific binding materials. After reverse-crosslinking, the DNA samples were purified and then analyzed by real-time quantitative PCR. Final results represent percentage of input chromatin and error bars indicate standard deviations (S.D.) through triplicate experiments. The MLL1, H3K4me3, H3K79me2 and H3K27me3 ChIP-DNA libraries were prepared using Illumina's TruSeq ChIP Sample Preparation Kit according to the manufacturer's instructions (Cat #IP-202-1012). The quality of the library was checked with Agilent TapeStation.

Final libraries were submitted to paired-end sequencing of 100 bp length on an Illumina HiSeq 3000.

Chromatin Isolation by RNA Immunoprecipitation (CHIRP) assay

The Chromatin Isolation by RNA Immunoprecipitation (CHIRP) assay was carried out based on the protocol described in a previous studies (Chu et al., 2011) with some modifications. Briefly, 20 million cells were collected and cross-linked in 20 mL of PBS buffer containing 1% formaldehyde at room temperature for 10 mins. Cross-linked cells were washed by chilled PBS 2 times, lysed in 1mL per 100 mg of cell pellet in cell lysis buffer (50 mM Tris-Cl pH 7.0, 10 mM EDTA, 1% SDS, supplemented with PMSF, DTT, proteinase inhibitors (P.I.) and Superase-in in fresh), and sonicated using a Bioruptor™ UCD200 (Diagenode) to prepare chromatin. Chromatin was diluted 2 times using hybridization buffer (750 mM NaCl, 1% SDS, 50 mM Tris 7.0, 1 mM EDTA, 15% formamide, add DTT, PMSF, P.I, and Superase-in fresh) and hybridized with 100 pmole of biotinylated DNA probes targeting *HOTTIP* or *LacZ* containing 100 µL of Streptavidin-magnetic C1 beads (Invitrogen). RNA and DNA hybrids were purified, washed 5 times with washing buffer (2x SSC, 0.5% SDS), and subjected to analysis by RT-qPCR. RNA binding proteins were subjected to analysis by western blotting with antibodies. Probes and primers are listed in the Table S4. CHIRP libraries were prepared using Illumina's TruSeq ChIP Sample Preparation Kit according to the manufacturer's instructions (Catalog: #IP-202-1012). The quality of the library was checked with Agilent TapeStation. Final libraries were submitted to paired-end sequencing of 100 bp length on an Illumina HiSeq 2500. All genomics datasets were deposited in the NCBI GEO under accession number (GSE114981).

ChIP-seq and ChIRP-seq data analysis

The ChIP-seq or ChIRP-seq raw data were processed through cutadapt (http://cutadapt.readthedocs.io, version 1.2.0) to remove adaptors and low quality reads (Martin, 2011). Cutadapt-filtered reads aligned to human reference genome (hg19) using Bowtie2 with

default parameters (Langmead et al., 2009), and the quality of these trimmed data was evaluated by FastQC program (Wingett and Andrews, 2018). After alignment, SAM files were converted to BAM files and sorted using Samtools (Li et al., 2009). Peak calling was performed using peak calling algorithm MACS2 (Zhang et al., 2008). The bedGraphToBigWig program was employed to generate the bigWig file of fragment or read coverages, including control and experimental datasets (https://www.encodeproject.org/software/bedgraphtobigwig/). All sequencing tracks were viewed using the Integrated Genomic Viewer (Robinson et al., 2011). Peaks annotation was carried out with the command "annotatePeaks.pl" from HOMER package (Heinz et al., 2010). For ChIRP-seq motif analysis, the *de novo* motif analysis was performed by the "findmotifsgenome.pl" from the HOMER motif discovery algorithm (Heinz et al., 2010). The genes and pathways regulated by the *HOTTIP* bound promoters or intergenic regions were analyzed and annotated by the Gene Ontology analysis with the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (https://david.ncifcrf.gov/, Version 6.8) (Huang da et al., 2009a; Huang da et al., 2009b). Each GO term with a p value more than 1 × 10^-3 is used for cutoff (threshold: 10^-3). All genomics datasets were deposited in the NCBI GEO under accession number (GSE114981).

Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq)

ATAC-seq was performed using the Nextera DNA library preparation kit as described previously (Buenrostro et al., 2015). In Brief, 5 $\times 10^4$ cells in single cell suspension were used for library preparation. Washed cells were re-suspended in lysis buffer containing 10 mM Tris-HCL (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1 % NP-40. After washing with cold 1x phosphate buffered saline (PBS) buffer, cells were treated with Tn5 Transposes for transposition reaction at 37°C for 30 min. DNA was purified using the MinElute Kit (QIAGEN). Library fragments were amplified using 1x NEB next PCR master mix and 1.25 μM indexed Nextra PCR primers (Ad1_noMX and Ad2.1-2.4 barcoded primers) with following PCR conditions: 72 °C for 5 min, 98 °C for 30 s,

followed by thermocycling at 98 °C for 10 s, 63 °C for 30 s and 72 °C for 1 min. The eluted DNA was used in a quantitative PCR (qPCR) reaction to estimate the optimum number of amplification cycles. Libraries were quantified using qPCR (Kapa Library Quantification Kit for Illumina, Roche), and libraries were purified with AMPure beads (Beckman Coulter), and the quality of the DNA library was examined by Agilent Bioanalyzer 2100 prior to sequencing with 2x100 bp paired-end reads on an Illumina NextSeq 500. Each sample includes two replicates for statistical analysis.

ATAC-seq analysis

We normally carried out two biological replicates for all of the ATAC-seq experiments (Luo et al., 2018). For quality control, first, each replicate should have 50 million reads for paired-end sequencing. Second, the alignment rate of each replicate is more than 95%. Third, we also removed the mitochondrial related reads from total reads after alignment and PCR duplicates were also removed. Finally, non-uniquely aligned reads were filtered based on MAPQ scores with samtools (MAPQ > 30), and plotPCA from BiocGenerics package in R package (R/3.6.1) was carried out to identify the variance between control and treatment groups. Moreover, fragSizeDist from ATACseqQC package in R package was carried out to show the fragment size distribution for control and treatment groups. In additional, the library complexity was analyzed including nucleosome free region signals (NFRs), mono-nucleosome, di-nucleosome and tri-nucleosome signals according to previously report (Tarbell and Liu, 2019).

Briefly, all of the raw fastq files were processed through cutadapt (http://cutadapt.readthedocs.io, version 1.2.0) to remove adaptors and low quality reads(Martin, 2011). Cutadapt-filtered reads aligned to human or mouse genome (hg19 or mm9) using Bowtie2 with default parameters (version Bowtie 2/2.2.6)(Langmead et al., 2009), and the quality of these trimmed data was evaluated by FastQC program (version 0.11.8)(Wingett and Andrews, 2018). After alignment, SAM files were converted to BAM files and sorted using Samtools (version 1.8.0) (Li et al., 2009). PCR duplicates were removed using Picard MarkDuplicates (version 2.0.1), and

mitochondrial reads were removed with Samtools (Corces et al., 2017). ENCODE blacklist regions were filtered (https://sites.google.com/site/anshulkundaje/projects/blacklists). Peak calling was performed using peak calling algorithm MACS2 with parameters ("-g mm -p 1e-9 –nolambda -f BAMPE –nomodel –shiftsize=100 --extsize 200")(Zhang et al., 2008). bedGraphToBigWig program was employed to generate the bigWig file of fragment or read coverages, including control and experimental datasets (https://www.encodeproject.org/software/bedgraphtobigwig/). All sequencing tracks were viewed using the Integrated Genomic Viewer (IGV/2.4.19)(Robinson et al., 2011). Peaks annotation was carried out with the command "annotatePeaks.pl" from HOMER package (version 4.8) (Heinz et al., 2010) and GREAT(McLean et al., 2010). DEseq2 (Benjamini-Hochberg adjusted $p< 0.05$; FoldChange≥2) were also performed to find the differential binding sites between two peak files, including control and treatment groups with C+G normalized and "reads in peaks" normalized data (Ross-Innes et al., 2012). The *de novo* motif analysis was performed by the "findmotifsgenome.pl" from the HOMER package (Heinz et al., 2010). For each genomic feature (peaks or chromVAR annotation), we calculated the chromatin accessibility median deviation z-score (for chromVAR features) or fragment counts (for peaks) in control and treatment groups with chromVAR package in R language (Rubin et al., 2019; Schep et al., 2017). Pearson's correlation coefficient and Pearson's $\chi 2$-test were carried out to calculate overall similarity between the replicates of ATAC-seq global open chromatin signatures. All genomics datasets were deposited in the NCBI GEO under accession number (GSE114981).

Mouse exome sequencing assay

The whole exome sequencing (WES) was carried out to identify candidate mutations in the exomes of genes. Genomic DNA was isolated from mice ear and BM cells including wildtype and the diseased *Hottip*-Tg mice, and genomic exome library was captured and constructed according to SureSelectXT Mouse All Exon kit (Agilent, Part Number:5190-4641),

and then 100 bp paired-end sequencing was performed using an Illumina NovaSeq S1. Raw sequencing reads were processed through cutadapt (http://cutadapt.readthedocs.io, version 1.2.0) to remove adaptors and low quality reads. These clean reads were mapped to the whole mouse genome (mm10) using BWA with the default settings (bwa/0.7.4) (Liang et al., 2009). The PCR duplicates were removed with Picard with default parameters (version 1.88) (http://broadinstitute.github.io/picard/), recalibrated with GATK with default setting (version 3.7) (McKenna et al., 2010), and compared the variance between wildtype and *Hottip*-Tg group with Strelka (version 2.9.2, default setting)(Kim et al., 2018) , and then variant bases were annotated with SnpEff (latest version) (http://snpeff.sourceforge.net/SnpEff_manual.html) (Cingolani et al., 2012). All genomics datasets were deposited in the NCBI GEO under accession number (GSE114981).

Xenotransplantation of human leukemic cells and Patient-Derived Xenografts (PDX)

All animals experiments were approved by and performed in compliance with the regulatory guidelines by the Institutional Animal Care and User Ethical Committees of the University of Texas Health Science Center at San Antonio and Pennsylvania State University College of Medicine. Non-obese diabetic (NOD)/LtSz-severe combined immunodeficiency (SCID) IL2Rγ_c^{null} (NSG) mice were housed in sterile conditions using high-efficiency particulate arrestance filtered micro-isolators and fed with irradiated food and acidified water. Adult mice (6-8 weeks old) were sublethally irradiated with 280 cGy of total body irradiation before injection of leukemic cells. WT control, *CBS7/9*$^{+/-}$ cells, *HOTTIP*$^{-/-}$ cells, and *HOTTIP*-VP-*CBS7/9*$^{+/-}$ MOLM13 cells (in 300 µL of PBS) were injected into the NSG mice by tail-vein injection at a dose of 0.5 x 10^6 cells/mouse. For Patient-Derived Xenografts (PDX) assay, control or *HOTTIP*$^{-/-}$ primary AML patients (#886, *FLT3-ITD*$^{+}$/*NPM1*$^{C-}$; # 974, *NPM1*$^{C+}$/*FLT3-ITD*$^{+}$; LPP4, *MLLr*$^{+}$/*NPM1*$^{C-}$) were injected into the NSG mice at 1.8 × 10^5 cells/mouse. Daily monitoring of mice for symptoms of disease (ruffled coat, hunched back, weakness and reduced motility) determined the time of killing

for injected animals with signs of distress. NSG mice were humanely killed at the time of moribund. Peripheral blood was collected by retro-orbital bleeding, bone (tibias, femurs and pelvis) and spleen were dissected. BM cells were isolated by flushing the bones. Spleens were mashed through a 70-μm mesh filter and made into single cell suspensions. PB was prepared for flow cytometry by ammonium chloride treatment to remove red cells. Human CD45 chimerism in these hematopoietic tissues was analyzed by flow cytometry (FACS LSR II–BD Biosciences, San Jose, CA, USA). All data were analyzed by FlowJo7.6 software.

For histopathology analyses, femurs were fixed in formaldehyde, decalcified, and paraffin embedded. Spleens were treated similarly except for the step of decalcification. Sections (4.5 μm) were stained with hematoxylin/eosin (H&E) or immunohistochemistry staining with anti-hCD45 antibody (Abcam, ab10559) and detected using an HRP conjugated compact polymer system with DAB as the chromogen. The slides were then observed with a conventional microscope.

Hi-C Assay

Hi-C assay was performed to generate a genome-wide interaction as described previously with Arima-HiC Kit (Cat: A410030) (https://arimagenomics.com/) with minor modifications. In brief, 5 million cells were collected and cross-linked in 10 mL of PBS buffer containing 1 % formaldehyde at room temperature for 10 min. The reaction was quenched by 0.125 M glycine solution. Cross-linked cell pellet were washed in 1x PBS buffer and collected. Cross-linked cell pellet was treated with lysis buffer and incubate at 4 °C for 15 min, and then conditioning solution was added to continue incubate at 62 °C for 10 min. Reaction was stopped by adding stop solution and incubating at 37 °C for 15 min. Cell pellet was digested with reaction buffer and restriction enzyme cocktail (Arima-HiC Kit) overnight at 37 °C with rotation. Digested DNA was purified with DNA purification beads (AMPure XP Beads), and the concentration of DNA was measured with Qubit. 750 ng of DNA per sample was sheared through sonication (Bioruptor) with default parameters (30 seconds ON, 30 seconds OFF pulse intervals). Fragmented DNA was then size-

selected to have a size distribution between 200-600 bp. 250 ng of size-selected DNA was used to generate sequence library with KAPA Hyper Prep Kit (Catalog # KK8500, KK4824 and KK8502). Final libraries were submitted to paired-end sequencing of 100 bp length on an Illumina HiSeq 2500.

Hi-C sequence data analysis

Raw sequence reads were first cleaned to remove adapter and low quality reads with bbmap and bbduk.sh (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbduk-guide/). The paired-end sequencing data was trimmed from the 3′ end of enzymatic sequences with homerTools (homerTools trim -3 Arima -mis 0 -matchStart 20 -min 20) from Homer software (version 4.8). Trimmed reads were aligned to human reference genome (hg19) using Bowtie2 with parameters ("-n 1 -m 1 -p 8") (Langmead et al., 2009). The remainder of the analysis was performed using juicer (version 1.5.5) (Durand et al., 2016b) and Homer (Heinz et al., 2010). Paired-end sequencing was used to make a tag directory with makeTagDirectory package from Homer software. A normalized interaction matrix was generated with the analyzeHiC program with parameters (1 Mb resolution for all chromosomes, and 100 kb resolution for specific chromosome), and the intra-chromatin interactions within specific loci were also generated with the analyzeHiC program in Homer software via parameters (-res 10,000 -superRes 20,000 -pos chromosome location). In depth explanations of normalization, generation of Hi-C correlation matrices, principal component analysis (PCA) and identifying significant interactions were performed as previously described (Lin et al., 2012). These interaction matrices for Hi-C heatmap were visualized with Juicebox (Durand et al., 2016a) and Java Treeview (Saldanha, 2004). All genomics datasets were deposited in the NCBI GEO under accession number (GSE114981).

Long term culture- Initiating cells (LTC-IC) assay

Primary MLL rearranged or non-MLL-AML primary samples were seeded into 96 wells plate with MS5 stroma cells in 200 µL medium (IMEM + 10 % FBS + 20 ng/mL human IL-3, IL-6,

TPO, SCF and FLT3 ligand). Different cell doses were put for 10 or 20 wells per cell dose. Cells were either vehicle DMSO treated or treated with 500 nM ICG001 for one week. After one week, all medium were replenished with fresh medium without drugs and further cultured for other two weeks with medium replenished every week. Cell clusters (cobberstone) were scored from each well and the initiating frequency were calculated using online resources-WEHI Extreme Limiting Dilution Analysis (http://bioinf.wehi.edu.au/software/elda/). Primary *MLLr*$^{+}$ AML primary samples were either mock-treated or treated 500 nM ICG-001 for 5 days in liquid medium. Cells were then harvested and stained with 0.1% Nitro Blue Tetrazolium chloride (NBT) which is converted into a dark deposit in myeloid differentiated cells and were scored out of the total cell counted.

In vivo **and** ***in vitro*** **drug treatment**

For *in vivo* experiment, mice were distributed into their respective groups randomly. *MLLr*$^{+}$ AML cells were transplanted into 6-12 weeks old sub-leathally irradiated (250 cGy) NSG mice (male or female) via intra-femoral route. Two weeks after transplantation, the mice were control-treated or treated with ICG-001 (50 mg/kg; 5 days a week for 2 weeks) in PEG300/D5W (3:1) via intra-peritoneal route. When animals showed signs of sickness, the mice were suffocated in the CO_2 chamber and confirmed dead by cervical dislocation. The leukemic mouse is defined by >20% human AML engraftment with CD45 and CD33 positive cells in the bone marrow.

QUANTIFICATION AND STATISTICAL ANALYSIS

Differences between experimental groups were determined by the Student's *t*-test or analysis of variance (ANOVA) followed by Newman-Keuls multiple comparison tests. p value <0.05 is considered significant (*), p value <0.01 is considered highly significant (**), p value <0.001 is considered extremely significant (***). TCGA datasets were obtained from The Cancer Genome Atlas (TCGA) database (https://gdac.broadinstitute.org/). Pearson's χ2-test also was

applied to determining significance of the enrichment of prognostic data from published TCGA human de novo AML datasets (Mukaka, 2012). For *in vivo* experiment, sample size chosen was based on the generalized linear model with Bonferroni multiple comparison adjustments; with the proposed sample size of at least five mice/ group/genotype. Animals were randomly assigned to each study. For all *in vitro* experiments, at least three independent experiments with more than three biological replicates for each condition/genotype were performed to ensure adequate statistical power.

DATA AND CODE AVAILABILITY

All genomics datasets generated in this study can be accessed at GEO database (accession code GSE114981 and GSE113191).

Table S4. Primers used for quantification of mRNA and genome DNA. Related to STAR Methods.

REFERENCES

Andreeff, M., Ruvolo, V., Gadgil, S., Zeng, C., Coombes, K., Chen, W., Kornblau, S., Baron, A. E., and Drabkin, H. A. (2008). HOX expression patterns identify a common signature for favorable AML. Leukemia *22*, 2041-2047.

Brunetti, L., Gundry, M. C., Sorcini, D., Guzman, A. G., Huang, Y. H., Ramabadran, R., Gionfriddo, I., Mezzasoma, F., Milano, F., Nabet, B.*, et al.* (2018). Mutant NPM1 Maintains the Leukemic State through HOX Expression. Cancer Cell *34*, 499-512 e499.

Buenrostro, J. D., Wu, B., Chang, H. Y., and Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol *109*, 21 29 21-29.

Chu, C., Qu, K., Zhong, F. L., Artandi, S. E., and Chang, H. Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell *44*, 667-678.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., and Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide

polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80-92.

Collins, C. T., and Hess, J. L. (2016). Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. Oncogene *35*, 1090-1098.

Corces, M. R., Trevino, A. E., Hamilton, E. G., Greenside, P. G., Sinnott-Armstrong, N. A., Vesuna, S., Satpathy, A. T., Rubin, A. J., Montine, K. S., Wu, B.*, et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods *14*, 959-962.

Crooks, G. M., Fuller, J., Petersen, D., Izadi, P., Malik, P., Pattengale, P. K., Kohn, D. B., and Gasson, J. C. (1999). Constitutive HOXA5 expression inhibits erythropoiesis and increases myelopoiesis from human hematopoietic progenitors. Blood *94*, 519-528.

Deng, C., Li, Y., Liang, S., Cui, K., Salz, T., Yang, H., Tang, Z., Gallagher, P. G., Qiu, Y., Roeder, R.*, et al.* (2013). USF1 and hSET1A mediated epigenetic modifications regulate lineage differentiation and HoxB4 transcription. PLoS genetics *9*, e1003524.

Deng, C., Li, Y., Zhou, L., Cho, J., Patel, B., Terada, N., Li, Y., Bungert, J., Qiu, Y., and Huang, S. (2016). HoxBlinc RNA Recruits Set1/MLL Complexes to Activate Hox Gene Expression Patterns and Mesoderm Lineage Development. Cell Rep *14*, 103-114.

Deschamps, J., and van Nes, J. (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. Development *132*, 2931-2942.

Dou, D. R., Calvanese, V., Sierra, M. I., Nguyen, A. T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C. M., Zack, J. A., Crooks, G. M.*, et al.* (2016). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat Cell Biol *18*, 595-606.

Dovey, O. M., Cooper, J. L., Mupo, A., Grove, C. S., Lynn, C., Conte, N., Andrews, R. M., Pacharne, S., Tzelepis, K., Vijayabaskar, M. S.*, et al.* (2017). Molecular synergy underlies the co-occurrence patterns and phenotype of NPM1-mutant acute myeloid leukemia. Blood *130*, 1911-1922.

Drabkin, H. A., Parsy, C., Ferguson, K., Guilhot, F., Lacotte, L., Roy, L., Zeng, C., Baron, A., Hunger, S. P., Varella-Garcia, M.*, et al.* (2002). Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. Leukemia *16*, 186-195.

Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., and Aiden, E. L. (2016a). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst *3*, 99-101.

Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S., Huntley, M. H., Lander, E. S., and Aiden, E. L. (2016b). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst *3*, 95-98.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC bioinformatics *10*, 48.

Emami, K. H., Nguyen, C., Ma, H., Kim, D. H., Jeong, K. W., Eguchi, M., Moon, R. T., Teo, J. L., Kim, H. Y., Moon, S. H.*, et al.* (2004). A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. Proc Natl Acad Sci U S A *101*, 12682-12687.

Forlani, S., Lawson, K. A., and Deschamps, J. (2003). Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. Development *130*, 3807-3819.

Fuller, J. F., McAdara, J., Yaron, Y., Sakaguchi, M., Fraser, J. K., and Gasson, J. C. (1999). Characterization of HOX gene expression during myelopoiesis: role of HOX A5 in lineage commitment and maturation. Blood *93*, 3391-3400.

Gale, R. E., Lamb, K., Allen, C., El-Sharkawi, D., Stowe, C., Jenkinson, S., Tinsley, S., Dickson, G., Burnett, A. K., Hills, R. K., and Linch, D. C. (2015). Simpson's Paradox and the Impact of Different DNMT3A Mutations on Outcome in Younger Adults With Acute Myeloid Leukemia. J Clin Oncol *33*, 2072-2083.

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A.*, et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442-451.

Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A.*, et al.* (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531-537.

Goyama, S., Schibler, J., Cunningham, L., Zhang, Y., Rao, Y., Nishimoto, N., Nakagawa, M., Olsson, A., Wunderlich, M., Link, K. A.*, et al.* (2013). Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. J Clin Invest *123*, 3876-3888.

Grimwade, D., Ivey, A., and Huntly, B. J. (2016). Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. Blood *127*, 29-41.

Groschel, S., Sanders, M. A., Hoogenboezem, R., de Wit, E., Bouwman, B. A. M., Erpelinck, C., van der Velden, V. H. J., 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.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple combinations of lineage-determining transcription

factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1-13.

Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Kim, S., Scheffler, K., Halpern, A. L., Bekritsky, M. A., Noh, E., Kallberg, M., Chen, X., Kim, Y., Beyter, D., Krusche, P., and Saunders, C. T. (2018). Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods *15*, 591-594.

Kuhn, M. W., Song, E., Feng, Z., Sinha, A., Chen, C. W., Deshpande, A. J., Cusan, M., Farnoud, N., Mupo, A., Grove, C.*, et al.* (2016). Targeting Chromatin Regulators Inhibits Leukemogenic Gene Expression in NPM1 Mutant Leukemia. Cancer Discov *6*, 1166-1181.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lento, W., Congdon, K., Voermans, C., Kritzik, M., and Reya, T. (2013). Wnt signaling in normal and malignant hematopoiesis. Cold Spring Harb Perspect Biol *5*.

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

Li, Z., Cai, X., Cai, C. L., Wang, J., Zhang, W., Petersen, B. E., Yang, F. C., and Xu, M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509-4518.

Li, Z., Chen, P., Su, R., Hu, C., Li, Y., Elkahloun, A. G., Zuo, Z., Gurbuxani, S., Arnovitz, S., Weng, H.*, et al.* (2016). PBX3 and MEIS1 Cooperate in Hematopoietic Cells to Drive Acute Myeloid Leukemias Characterized by a Core Transcriptome of the MLL-Rearranged Disease. Cancer Res *76*, 619-629.

Liang, S. Y., Moghimi, B., Crusselle-Davis, V. J., Lin, I. J., Rosenberg, M. H., Li, X., Strouboulis, J., Huang, S., and Bungert, J. (2009). Defective erythropoiesis in transgenic mice expressing dominant-negative upstream stimulatory factor. Molecular and cellular biology *29*, 5900-5910.

Lin, Y. C., Benner, C., Mansson, R., Heinz, S., Miyazaki, K., Miyazaki, M., Chandra, V., Bossen, C., Glass, C. K., and Murre, C. (2012). Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. Nat Immunol *13*, 1196-1204.

Luo, H., Wang, F., Zha, J., Li, H., Yan, B., Du, Q., Yang, F., Sobh, A., Vulpe, C., Drusbosky, L.*, et al.* (2018). CTCF boundary remodels chromatin domain and drives aberrant HOX gene transcription in acute myeloid leukemia. Blood *132*, 837-848.

Martin, M. (2011). Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. EMBnet Journal *17*, 10-12.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297-1303.

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

Meyer, C., Kowarz, E., Hofmann, J., Renneville, A., Zuna, J., Trka, J., Ben Abdelali, R., Macintyre, E., De Braekeleer, E., De Braekeleer, M.*, et al.* (2009). New insights to the MLL recombinome of acute leukemias. Leukemia *23*, 1490-1499.

Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi medical journal : the journal of Medical Association of Malawi *24*, 69-71.

Narendra, V., Rocha, P. P., An, D., Raviram, R., Skok, J. A., Mazzoni, E. O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science *347*, 1017-1021.

Ogilvy, S., Metcalf, D., Gibson, L., Bath, M. L., Harris, A. W., and Adams, J. M. (1999). Promoter elements of vav drive transgene expression in vivo throughout the hematopoietic compartment. Blood *94*, 1855-1863.

Pillai, M. M., Venkataraman, G. M., Kosak, S., and Torok-Storb, B. (2008). Integration site analysis in transgenic mice by thermal asymmetric interlaced (TAIL)-PCR: segregating multiple-integrant founder lines and determining zygosity. Transgenic Res *17*, 749-754.

Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R., and Weissman, I. L. (2003). A role for Wnt signalling in self-renewal of haematopoietic stem cells. Nature *423*, 409-414.

Rice, K. L., and Licht, J. D. (2007). HOX deregulation in acute myeloid leukemia. J Clin Invest *117*, 865-868.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Ross-Innes, C. S., Stark, R., Teschendorff, A. E., Holmes, K. A., Ali, H. R., Dunning, M. J., Brown, G. D., Gojis, O., Ellis, I. O., Green, A. R.*, et al.* (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389.

Rubin, A. J., Parker, K. R., Satpathy, A. T., Qi, Y., Wu, B., Ong, A. J., Mumbach, M. R., Ji, A. L., Kim, D. S., Cho, S. W.*, et al.* (2019). Coupled Single-Cell CRISPR Screening and Epigenomic Profiling Reveals Causal Gene Regulatory Networks. Cell *176*, 361-376 e317.

Saldanha, A. J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.

Sauvageau, G., Lansdorp, P. M., Eaves, C. J., Hogge, D. E., Dragowska, W. H., Reid, D. S., Largman, C., Lawrence, H. J., and Humphries, R. K. (1994). Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. Proc Natl Acad Sci U S A *91*, 12223-12227.

Schep, A. N., Wu, B., Buenrostro, J. D., and Greenleaf, W. J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods *14*, 975-978.

So, C. W., Karsunky, H., Wong, P., Weissman, I. L., and Cleary, M. L. (2004). Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. Blood *103*, 3192-3199.

Spencer, D. H., Young, M. A., Lamprecht, T. L., Helton, N. M., Fulton, R., O'Laughlin, M., Fronick, C., Magrini, V., Demeter, R. T., Miller, C. A.*, et al.* (2015). Epigenomic analysis of the HOX gene loci reveals mechanisms that may control canonical expression patterns in AML and normal hematopoietic cells. Leukemia *29*, 1279-1289.

Stavropoulou, V., Kaspar, S., Brault, L., Sanders, M. A., Juge, S., Morettini, S., Tzankov, A., Iacovino, M., Lau, I. J., Milne, T. A.*, et al.* (2016). MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome. Cancer Cell *30*, 43-58.

Taberlay, P. C., Achinger-Kawecka, J., Lun, A. T., Buske, F. A., Sabir, K., Gould, C. M., Zotenko, E., Bert, S. A., Giles, K. A., Bauer, D. C.*, et al.* (2016). Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. Genome Res *26*, 719-731.

Takeda, A., Goolsby, C., and Yaseen, N. R. (2006). NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. Cancer Res *66*, 6628-6637.

Tarbell, E. D., and Liu, T. (2019). HMMRATAC: a Hidden Markov ModeleR for ATAC-seq. Nucleic Acids Res.

Thorsteinsdottir, U., Mamo, A., Kroon, E., Jerome, L., Bijl, J., Lawrence, H. J., Humphries, K., and Sauvageau, G. (2002). Overexpression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion. Blood *99*, 121-129.

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

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E., and Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

Valton, A. L., and Dekker, J. (2016). TAD disruption as oncogenic driver. Curr Opin Genet Dev *36*, 34-40.

Wang, J., Li, Z., He, Y., Pan, F., Chen, S., Rhodes, S., Nguyen, L., Yuan, J., Jiang, L., Yang, X.*, et al.* (2014). Loss of Asxl1 leads to myelodysplastic syndrome-like disease in mice. Blood *123*, 541-553.

Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A.*, et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, Y., Krivtsov, A. V., Sinha, A. U., North, T. E., Goessling, W., Feng, Z., Zon, L. I., and Armstrong, S. A. (2010). The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. Science *327*, 1650-1653.

Wingett, S. W., and Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. F1000Res *7*, 1338.

Yang, H., Kurtenbach, S., Guo, Y., Lohse, I., Durante, M. A., Li, J., Li, Z., Al-Ali, H., Li, L., Chen, Z.*, et al.* (2018). Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid malignancies. Blood *131*, 328-341.

Yeung, J., Esposito, M. T., Gandillet, A., Zeisig, B. B., Griessinger, E., Bonnet, D., and So, C. W. (2010). beta-Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. Cancer Cell *18*, 606-618.

Zangenberg, M., Grubach, L., Aggerholm, A., Silkjaer, T., Juhl-Christensen, C., Nyvold, C. G., Kjeldsen, E., Ommen, H. B., and Hokland, P. (2009). The combined expression of HOXA4 and MEIS1 is an independent prognostic factor in patients with AML. European journal of haematology *83*, 439-448.

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.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-H3K4me3 antibody, rabbit monoclonal	Millipore	Cat#04-745; RRID:AB_1163444
Anti-H3K27me3 antibody, rabbit polyclonal	Millipore	Cat#07-449; RRID:AB_310624
Anti-H3K9me2 antibody, rabbit polyclonal	Millipore	Cat#17-681, RRID:AB_1977531
Anti-H3K79me2 antibody, rabbit polyclonal	Abcam	Cat#ab3594; RRID:AB_303937
Anti-CD45 antibody, rabbit polyclonal	Abcam	Cat#ab10559, RRID:AB_442811
Anti-Ly-6A/E (Sca-1) antibody, mouse monoclonal	BioLegend	Cat#108111, RRID:AB_313348
Anti-CD117 (c-kit) antibody, mouse monoclonal	BioLegend	Cat#135135, RRID:AB_2632808
Anti-CD34 antibody, mouse monoclonal	BD Biosciences	Cat#553733, RRID:AB_395017
Anti-CD135 antibody, mouse monoclonal	BD Biosciences	Cat#558995, RRID:AB_397174
Anti-CD16 antibody, mouse monoclonal	BD Biosciences	Cat#555405, RRID:AB_395805
Anti-KMT2A/MLL antibody, rabbit polyclonal	Novus Biologicals	Cat#NB600-248, RRID:AB_2145479
Anti-WDR5 antibody, mouse monoclonal	Abcam	Cat#ab56919, RRID:AB_946146
Anti-DOT1L antibody, rabbit polyclonal	Novus Biologicals	Cat# NB100-40845, RRID:AB_789636
Anti-p-STAT5A, rabbit polyclonal	Cell Signaling Technology	Cat# 9351, RRID:AB_2315225
Anti-STAT5A, rabbit monoclonal	Cell Signaling Technology	Cat# 94205, RRID:AB_2737403
Anti-HDAC1, rabbit polyclonal	Abcam	Cat# ab7028, RRID:AB_305705
Anti-RUNX1, rabbit polyclonal	Abcam	Cat# ab23980, RRID:AB_2184205
Chemicals, Peptides, and Recombinant Proteins		
Lipofectamine 3000 reagent	Thermo Fisher Scientific	Cat#L3000-008
Proteinase K	Thermo Fisher Scientific	Cat#25530049
Alt-R® S.p. Cas9 Nuclease 3NLS	Integrated DNA Technologies	Cat#1074181
Alt-R® CRISPR-Cas9 tracrRNA	Integrated DNA Technologies	Cat#1072532
Alt-R® Cas9 Electroporation Enhancer	Integrated DNA Technologies	Cat#1075915
Protease inhibitor Cocktail	Abcam	Cat#ab65621
Dynabeads™ Protein G	Thermo Fisher Scientific	Cat#10003D
Dynabeads™ Protein A	Thermo Fisher Scientific	Cat#10001D

Dynabeads™ MyOne™ Streptavidin C1	Thermo Fisher Scientific	Cat#65001
SUPERase• In™ RNase Inhibitor	Thermo Fisher Scientific	Cat#AM2694
RNaseOUT™ Ribonuclease Inhibitor	Thermo Fisher Scientific	Cat#10777019
Pierce Protease Inhibitor Tablets	Thermo Fisher Scientific	Cat#A32963
TURBO™ DNase	Thermo Fisher Scientific	Cat#AM2238
AMPure XP beads	Beckman Coulter	Cat#A63881
Critical Commercial Assays		
RNeasy mini-isolation kit	QIAGEN	Cat#74106
Neon™ Transfection System Kit	Thermo Fisher Scientific	Cat#MPK1025
QIAquick Gel Extract kit	QIAGEN	Cat#28706
Alt-R® CRISPR-Cas9 Control Kit	Integrated DNA Technologies	Cat#1072554
mirVana PARIS kit	Thermo Fisher Scientific	Cat#AM1556
Superscript II reverse Transcriptase	Thermo Fisher Scientific	Cat#18064014
QIAquick PCR purification kit	QIAGEN	Cat#28106
QIAprep Spin Miniprep Kit	QIAGEN	Cat#27106
QIAGEN Plasmid *Plus* Maxi Kit	QIAGEN	Cat#12965
Nextera DNA Library Preparation Kit	illumina	Cat#FC-121-1030
Arima-HiC Kit	Arima	Cat#A410030
KAPA Hyper Prep Kit	KAPA	Cat # KK8500, KK4824 and KK8502
SureSelectXT Mouse All Exon	Agilent	Cat # 5190-4641
SingleShot™ SYBR® Green One-Step Kit	Bio-Rad Laboratories	Cat#1725095
Deposited Data		
RNA-seq in WT vs *HOTTIP*-KO MOLM13	This study	GEO: GSE114981
ATAC-seq in WT vs *HOTTIP*-KO MOLM13	This study	GEO: GSE114981
RNA-seq in WT vs *CBS7/9*-KO MOLM13	(Luo et al., 2018)	GEO: GSE113191
CHIRP-seq of WT vs *HOTTIP*-KO MOLM13	This study	GEO: GSE114981
ChIP-seq of WT vs *HOTTIP*-KO MOLM13	This study	GEO: GSE114981
LSK RNA-seq of WT vs *Hottip*-Tg mice	This study	GEO: GSE114981
LSK ATAC-seq of WT vs *Hottip*-Tg mice	This study	GEO: GSE114981
LT-HSC RNA-seq of WT vs *Hottip*-Tg mice	This study	GEO: GSE114981
ST-HSC RNA-seq of WT vs *Hottip*-Tg mice	This study	GEO: GSE114981
LT-HSC ATAC-seq of WT vs *Hottip*-Tg mice	This study	GEO: GSE114981
ST-HSC ATAC-seq of WT vs *Hottip*-Tg mice	This study	GEO: GSE114981
HiC-seq of WT vs *HOTTIP*-KO MOLM13	This study	GEO: GSE114981
HiC-seq of *HOTTIP*-KO MOLM13	This study	GEO: GSE114981
HiC-seq of CBS7/9-KO MOLM13	This study	GEO: GSE114981
HiC-seq of CBS7/9-KO-VP-HT MOLM13	This study	GEO: GSE114981
Mouse whole exome sequencing	This study	GEO: GSE114981
Experimental Models: Cell Lines		
MOLM-13	DSMZ	ACC 554

HEK293T	ATCC	CRL-3216
OCI-AML3	DSMZ	ACC 582
Experimental Models: Organisms/Strains		
Hottip-Transgenic mouse	Huang/Xu labs	
Xenograft AML mouse model	Huang/Xu labs	
Oligonucleotides		
sgRNAs	This study	see Table S4
RT-qPCR primers	This study	see Table S4
ChIP-PCR primers	This study	see Table S4
crRNAs	This study	see Table S4
ATAC primers	This study	see Table S4
CHIRP Probes	This study	see Table S4
Recombinant DNA		
pL-CRISPR.EFS.GFP	Addgene	Plasmid #57818
pLKO5.sgRNA.EFS.tRFP	Addgene	Plasmid #57824
lentiCRISPR v2	Addgene	Plasmid # 52961
pHR-SFFV-dCas9-BFP-KRAB	Addgene	Plasmid #46911
pAC94-pmax-dCas9VP160-2A-puro	Addgene	Plasmid #48226
pMD2.G	Addgene	Plasmid # 12259
psPAX2	Addgene	Plasmid # 12260
pGEM®-T Easy Vector Systems	Promega	Cat#A137A
Software and Algorithms		
TopHat	(Trapnell et al., 2012)	https://ccb.jhu.edu/software/tophat/
Bowtie2	(Langmead and Salzberg, 2012)	http://bowtiebio.sourceforge.net/bowtie2/
R	N/A	https://www.r-project.org/
Cufflinks	(Trapnell et al., 2010)	http://cole-trapnell-lab.github.io/cufflinks/
Cuffdiff	(Trapnell et al., 2010)	http://cole-trapnell-lab.github.io/cufflinks/
Integrated Genomic Viewer	(Robinson et al., 2011)	http://software.broadinstitute.org/
Deeptools	(Ramirez et al., 2014)	https://deeptools.github.io/
Gene Set Enrichment Analysis (GSEA)	(Subramanian et al., 2005)	http://software.broadinstitute.org/gsea/
chromVAR	(Schep et al., 2017)	https://bioconductor.org/packages/release/bioc/html/chromVAR.html
Homer	(Heinz S et al., 2010)	http://homer.ucsd.edu/homer/index.html
Juicer	(Neva C. Durand, 2016)	https://github.com/aidenlab/juicer
Juicebox	(Neva C. Durand, 2016)	http://aidenlab.org/juicebox/

Figure 1

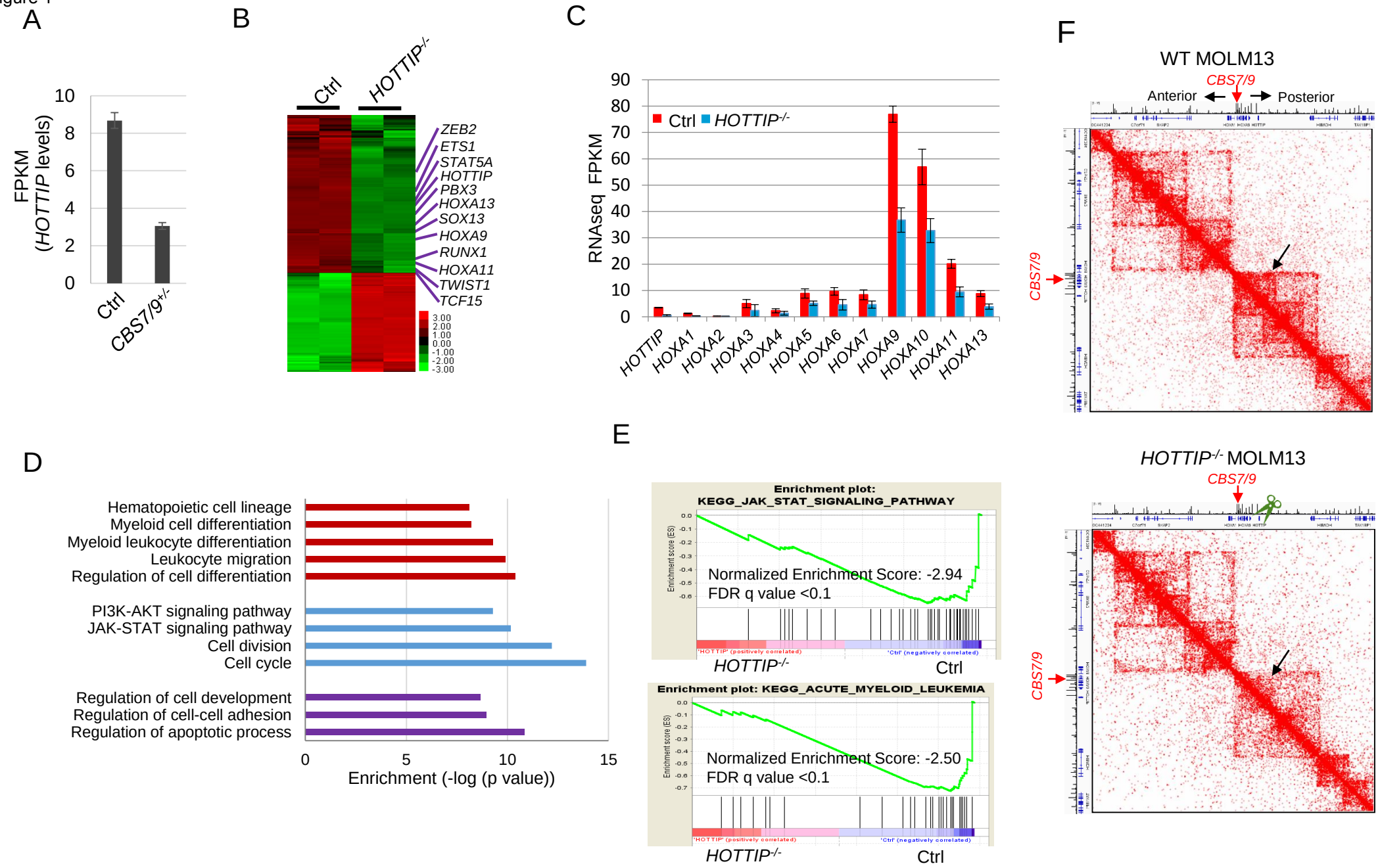

A
FPKM (HOTTIP levels)
10
8
6
4
2
0
Ctrl
CBS7/9+/-
B
Ctrl
HOTTIP-/-
ZEB2
ETS1
STAT5A
HOTTIP
PBX3
HOXA13
SOX13
HOXA9
RUNX1
HOXA11
TWIST1
TCF15
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
C
RNAseq FPKM
90
80
70
60
50
40
30
20
10
0
Ctrl
HOTTIP-/-
HOTTIP
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
HOXA13
F
WT MOLM13
CBS7/9
Anterior
Posterior
CBS7/9
D
Hematopoietic cell lineage
Myeloid cell differentiation
Myeloid leukocyte differentiation
Leukocyte migration
Regulation of cell differentiation
PI3K-AKT signaling pathway
JAK-STAT signaling pathway
Cell division
Cell cycle
Regulation of cell development
Regulation of cell-cell adhesion
Regulation of apoptotic process
0
5
10
15
Enrichment (-log (p value))
E
Enrichment plot:
KEGG_JAK_STAT_SIGNALING_PATHWAY
Enrichment score (ES)
Normalized Enrichment Score: -2.94
FDR q value <0.1
'HOTTIP' (positively correlated)
'Ctrl' (negatively correlated)
HOTTIP-/-
Ctrl
Enrichment plot: KEGG_ACUTE_MYELOID_LEUKEMIA
Enrichment score (ES)
Normalized Enrichment Score: -2.50
FDR q value <0.1
'HOTTIP' (positively correlated)
'Ctrl' (negatively correlated)
HOTTIP-/-
Ctrl
HOTTIP-/- MOLM13
CBS7/9
CBS7/9

Figure 2

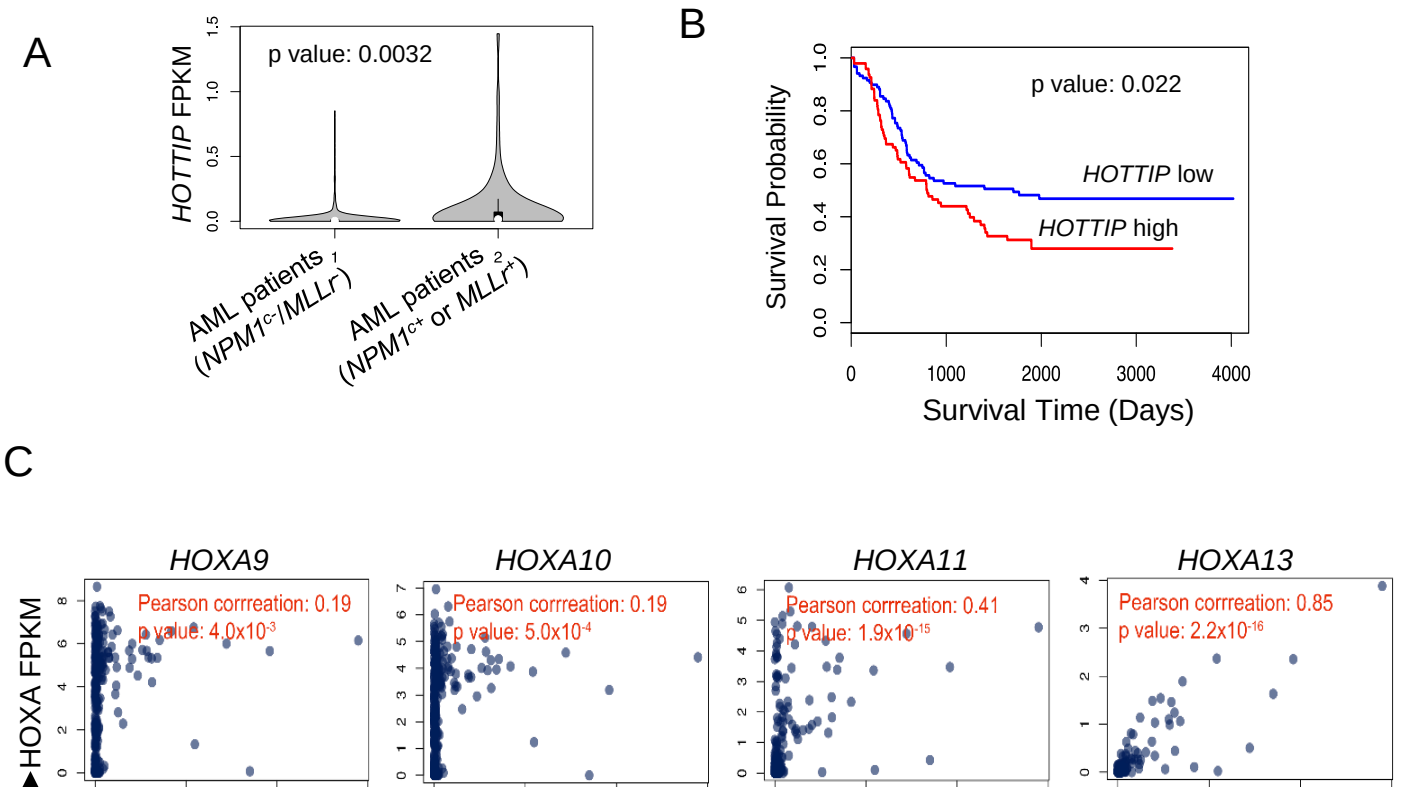
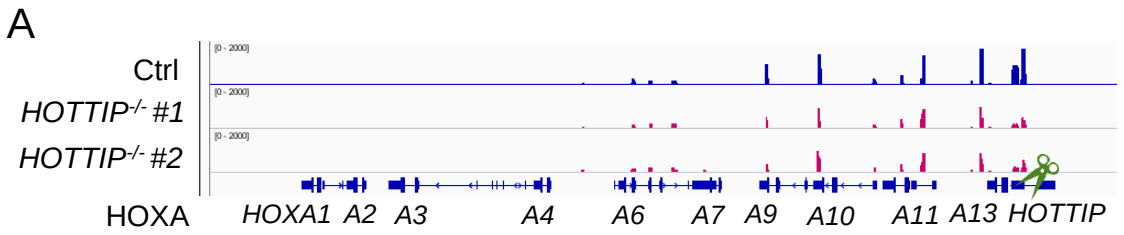
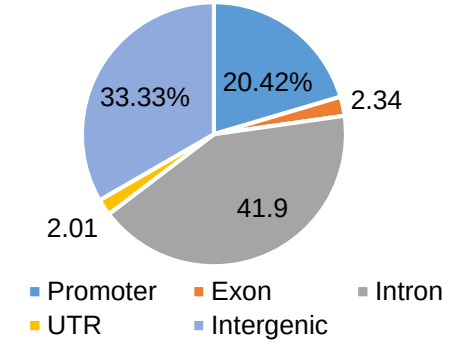
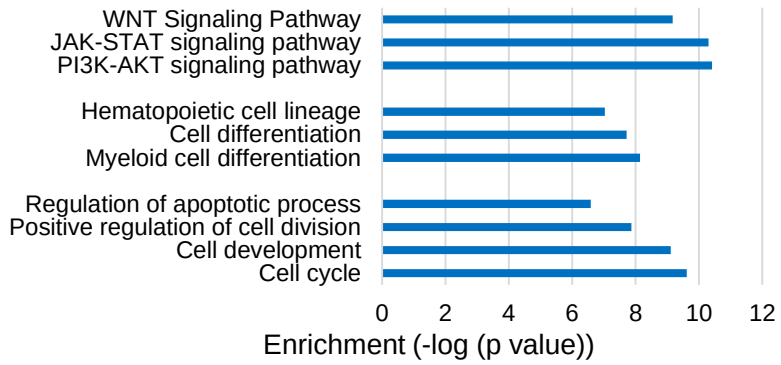
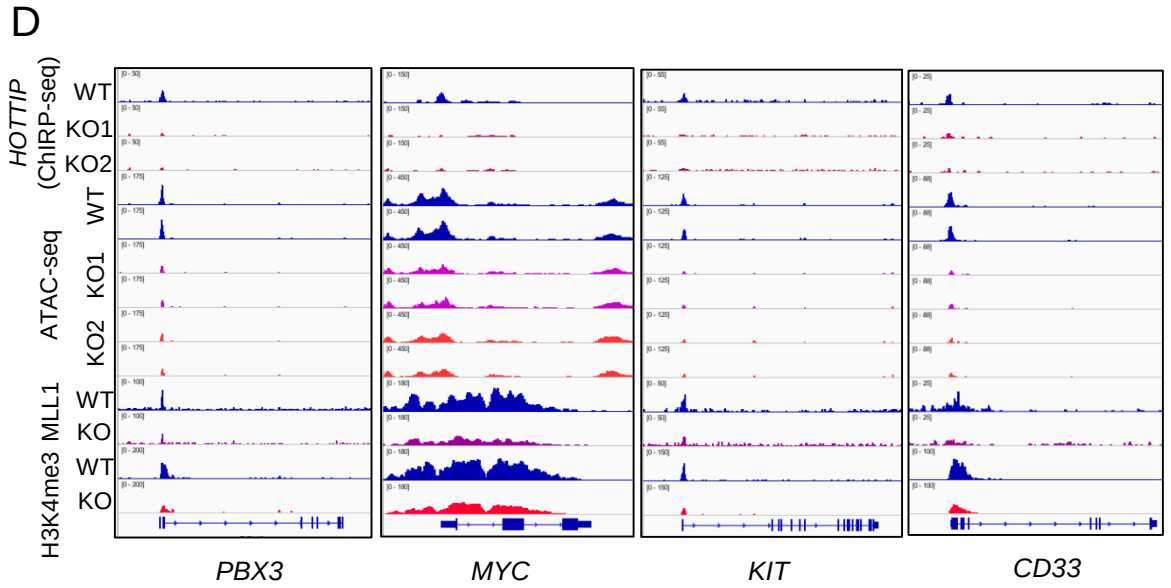

A
p value: 0.0032
HOTTIP FPKM
AML patients 1 (NPM1c-/MLLr-)
AML patients 2 (NPM1c+ or MLLr+)
B
p value: 0.022
HOTTIP low
HOTTIP high
Survival Probability
Survival Time (Days)
C
HOXA9
HOXA10
HOXA11
HOXA13
Pearson corrreation: 0.19
p value: 4.0x10-3
Pearson corrreation: 0.19
p value: 5.0x10-4
Pearson corrreation: 0.41
p value: 1.9x10-15
Pearson corrreation: 0.85
p value: 2.2x10-16
HOXA FPKM
HOTTIP FPKM
TWIST1 FPKM
MEIS1 FPKM
PBX3 FPKM
Pearson corrreation: 0.27
p value: 7.2x10-7
Pearson corrreation: 0.19
p value: 5.0x10-4
Pearson corrreation: 0.26
p value: 1.0x10-3
HOTTIP FPKM
HOTTIP FPKM
HOTTIP FPKM

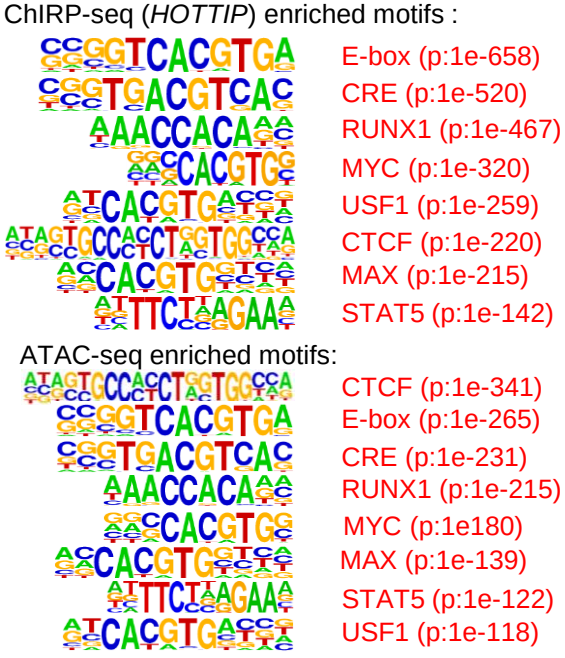
Figure 3

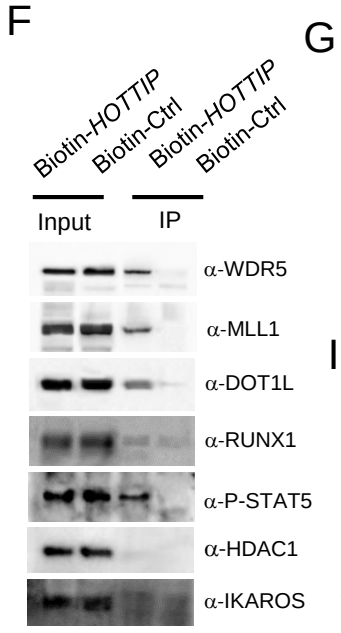

A
Ctrl
HOTTIP-/- #1
HOTTIP-/- #2
HOXA
HOXA1 A2 A3 A4 A6 A7 A9 A10 A11 A13 HOTTIP

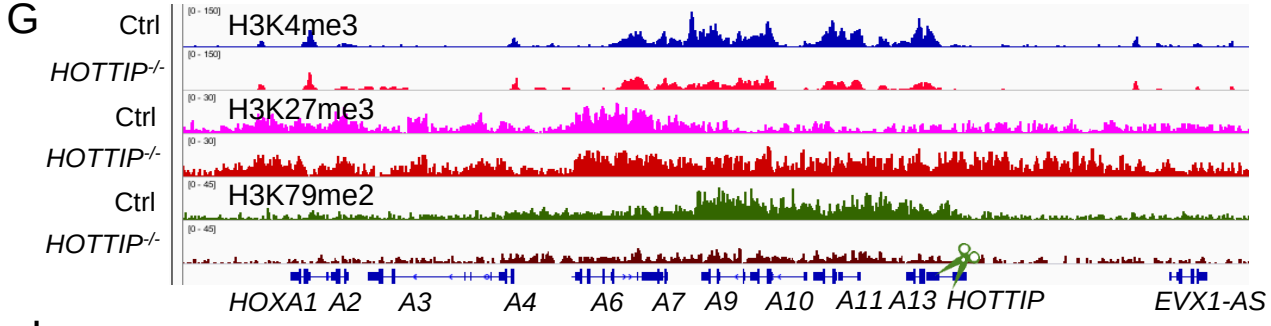

B
20.42%
2.34
33.33%
41.9
2.01
Promoter
Exon
Intron
UTR
Intergenic

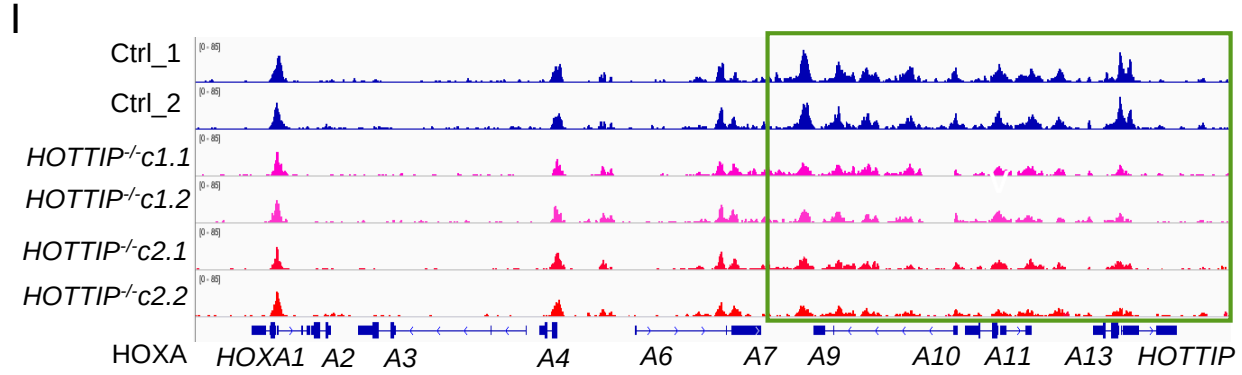

C
WNT Signaling Pathway
JAK-STAT signaling pathway
PI3K-AKT signaling pathway
Hematopoietic cell lineage
Cell differentiation
Myeloid cell differentiation
Regulation of apoptotic process
Positive regulation of cell division
Cell development
Cell cycle
0 2 4 6 8 10 12
Enrichment (-log (p value))


D
HOTTIP (ChIRP-seq)
WT
KO1
KO2
ATAC-seq
WT
KO1
KO2
H3K4me3 MLL1
WT
KO
WT
KO
PBX3
MYC
KIT
CD33


E
ChIRP-seq (HOTTIP) enriched motifs :
E-box (p:1e-658)
CRE (p:1e-520)
RUNX1 (p:1e-467)
MYC (p:1e-320)
USF1 (p:1e-259)
CTCF (p:1e-220)
MAX (p:1e-215)
STAT5 (p:1e-142)
ATAC-seq enriched motifs:
CTCF (p:1e-341)
E-box (p:1e-265)
CRE (p:1e-231)
RUNX1 (p:1e-215)
MYC (p:1e180)
MAX (p:1e-139)
STAT5 (p:1e-122)
USF1 (p:1e-118)


F
Biotin-HOTTIP
Biotin-Ctrl
Biotin-HOTTIP
Biotin-Ctrl
Input
IP
α-WDR5
α-MLL1
α-DOT1L
α-RUNX1
α-P-STAT5
α-HDAC1
α-IKAROS


G
Ctrl
H3K4me3
HOTTIP-/-
Ctrl
H3K27me3
HOTTIP-/-
Ctrl
H3K79me2
HOTTIP-/-
HOXA1 A2 A3 A4 A6 A7 A9 A10 A11 A13 HOTTIP EVX1-AS


I
Ctrl_1
Ctrl_2
HOTTIP-/-c1.1
HOTTIP-/-c1.2
HOTTIP-/-c2.1
HOTTIP-/-c2.2
HOXA
HOXA1 A2 A3 A4 A6 A7 A9 A10 A11 A13 HOTTIP

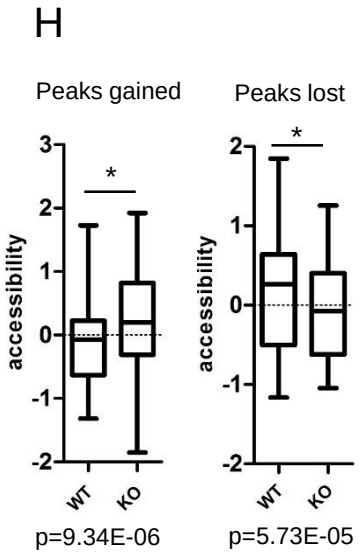

H
Peaks gained
Peaks lost
accessibility
WT
KO
p=9.34E-06
p=5.73E-05

Figure 4

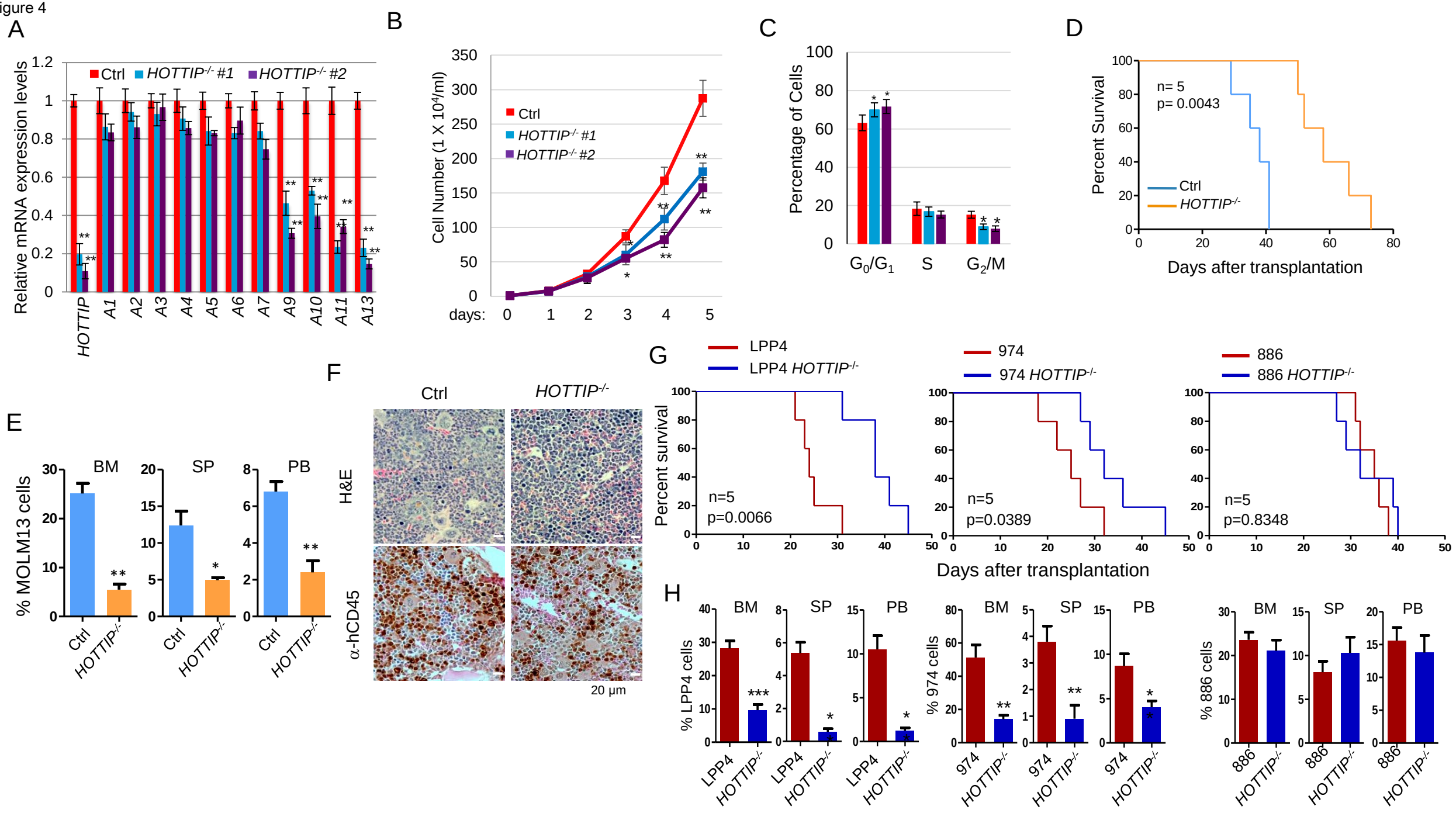

A
Relative mRNA expression levels
Ctrl
HOTTIP-/- #1
HOTTIP-/- #2
HOTTIP
A1
A2
A3
A4
A5
A6
A7
A9
A10
A11
A13
B
Cell Number (1 X 10⁴/ml)
days:
C
Percentage of Cells
G0/G1
S
G2/M
D
Percent Survival
n= 5
p= 0.0043
Days after transplantation
E
% MOLM13 cells
BM
SP
PB
F
Ctrl
HOTTIP-/-
H&E
α-hCD45
20 μm
G
LPP4
LPP4 HOTTIP-/-
974
974 HOTTIP-/-
886
886 HOTTIP-/-
Percent survival
n=5
p=0.0066
p=0.0389
p=0.8348
Days after transplantation
H
% LPP4 cells
% 974 cells
% 886 cells

Figure 5

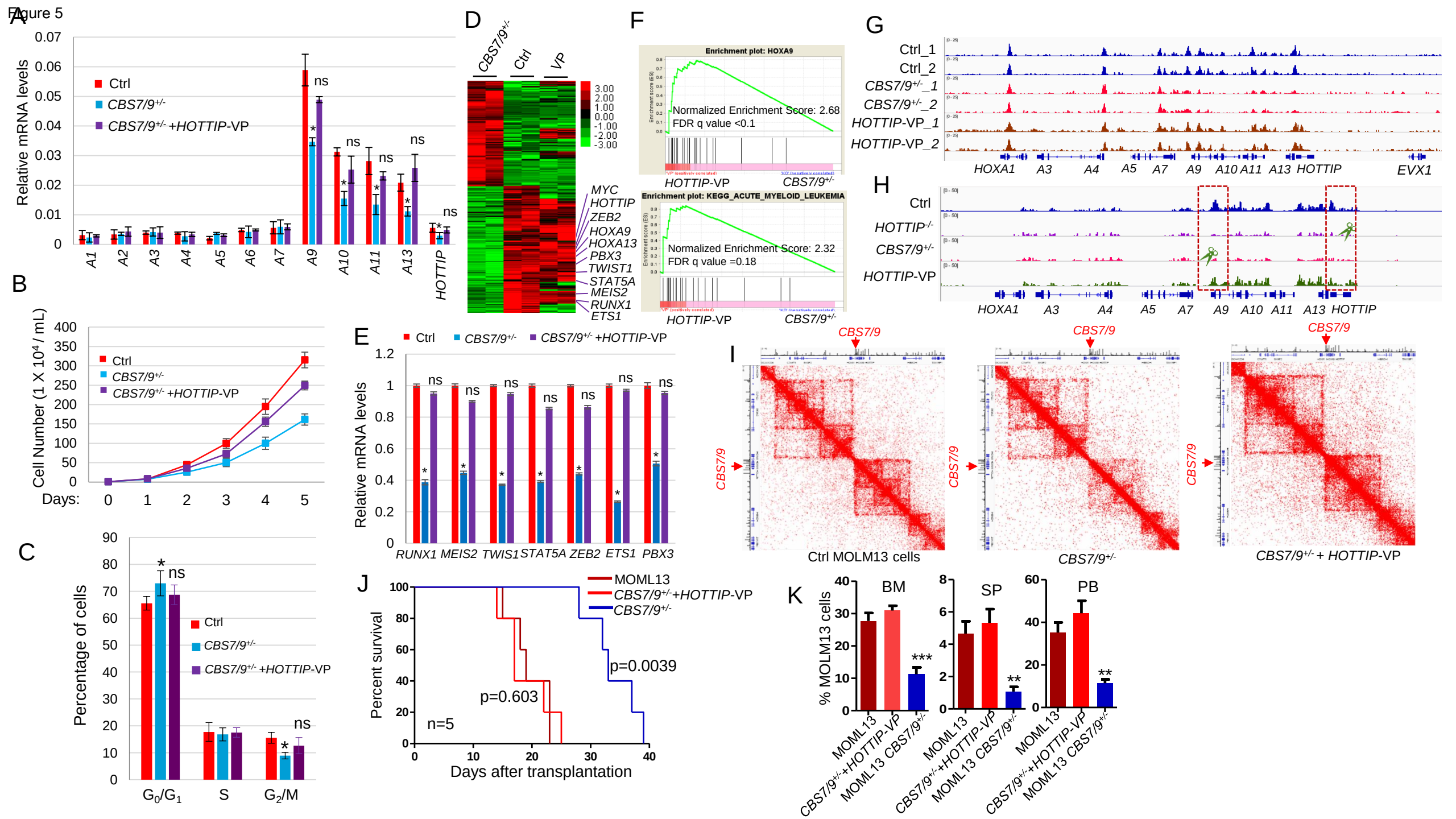

A
Relative mRNA levels
Ctrl
CBS7/9+/-
CBS7/9+/- +HOTTIP-VP
A1 A2 A3 A4 A5 A6 A7 A9 A10 A11 A13 HOTTIP
ns
B
Cell Number (1 X 10^4 / mL)
Days:
C
Percentage of cells
G0/G1 S G2/M
D
CBS7/9+/- Ctrl VP
MYC HOTTIP ZEB2 HOXA9 HOXA13 PBX3 TWIST1 STAT5A MEIS2 RUNX1 ETS1
E
Relative mRNA levels
RUNX1 MEIS2 TWIS1 STAT5A ZEB2 ETS1 PBX3
F
Enrichment plot: HOXA9
Normalized Enrichment Score: 2.68
FDR q value <0.1
HOTTIP-VP CBS7/9+/-
Enrichment plot: KEGG_ACUTE_MYELOID_LEUKEMIA
Normalized Enrichment Score: 2.32
FDR q value =0.18
G
Ctrl_1
Ctrl_2
CBS7/9+/-_1
CBS7/9+/-_2
HOTTIP-VP_1
HOTTIP-VP_2
HOXA1 A3 A4 A5 A7 A9 A10 A11 A13 HOTTIP EVX1
H
Ctrl
HOTTIP-/-
CBS7/9+/-
HOTTIP-VP
HOXA1 A3 A4 A5 A7 A9 A10 A11 A13 HOTTIP
I
CBS7/9
Ctrl MOLM13 cells
CBS7/9+/-
CBS7/9+/- + HOTTIP-VP
J
MOML13
CBS7/9+/-+HOTTIP-VP
CBS7/9+/-
p=0.0039
p=0.603
n=5
Percent survival
Days after transplantation
K
% MOLM13 cells
BM SP PB
MOML13
CBS7/9+/-+HOTTIP-VP
MOML13 CBS7/9+/-

Figure 6

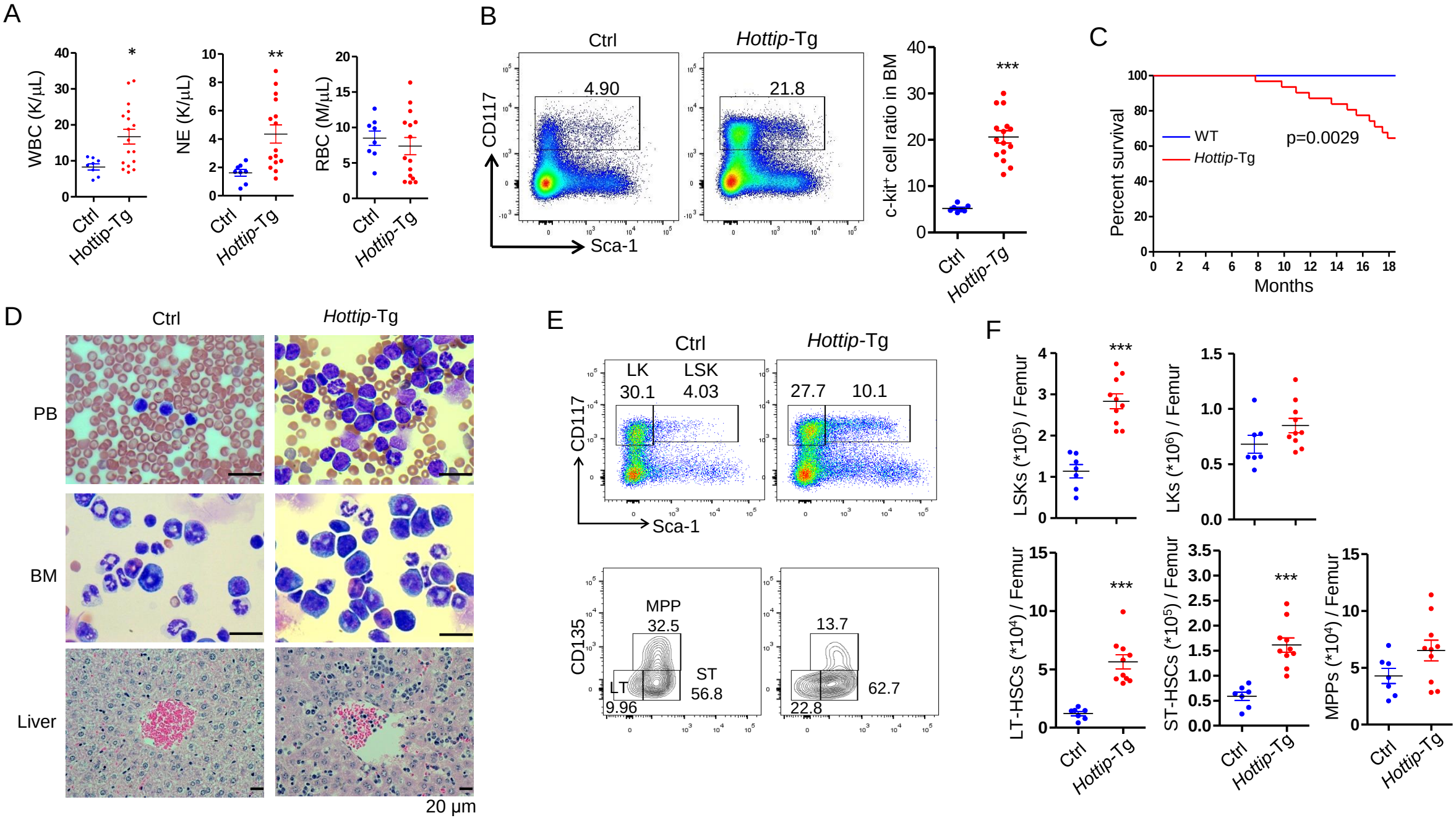
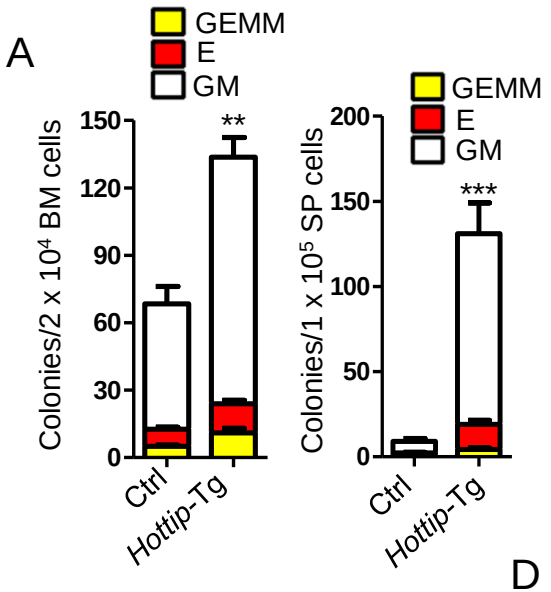
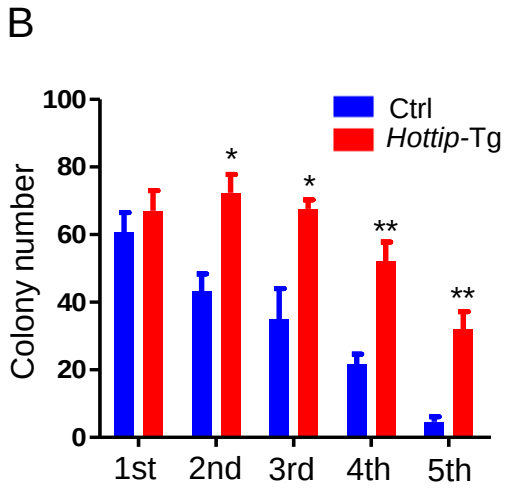
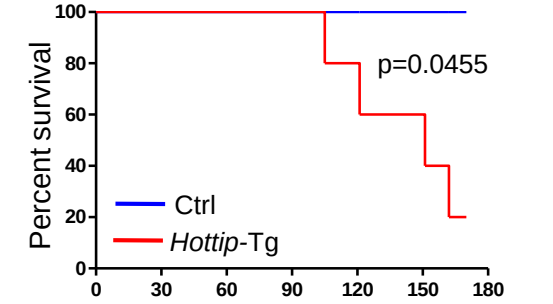
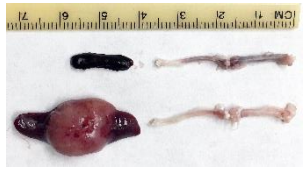

A
WBC (K/μL)
NE (K/μL)
RBC (M/μL)
Ctrl
Hottip-Tg
B
Ctrl
Hottip-Tg
4.90
21.8
CD117
Sca-1
c-kit+ cell ratio in BM
C
Percent survival
WT
Hottip-Tg
p=0.0029
Months
D
Ctrl
Hottip-Tg
PB
BM
Liver
20 μm
E
Ctrl
Hottip-Tg
LK
LSK
30.1
4.03
27.7
10.1
CD117
Sca-1
MPP
32.5
ST
56.8
LT
9.96
13.7
62.7
22.8
CD135
F
LSKs (*10^5) / Femur
LKs (*10^6) / Femur
LT-HSCs (*10^4) / Femur
ST-HSCs (*10^5) / Femur
MPPs (*10^4) / Femur

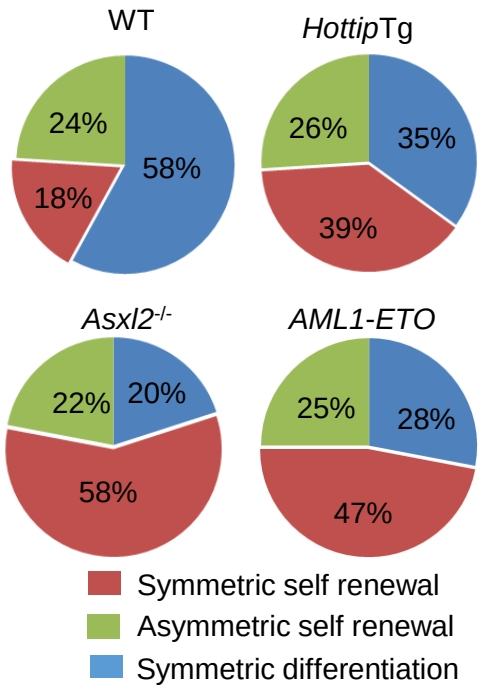
Figure 7

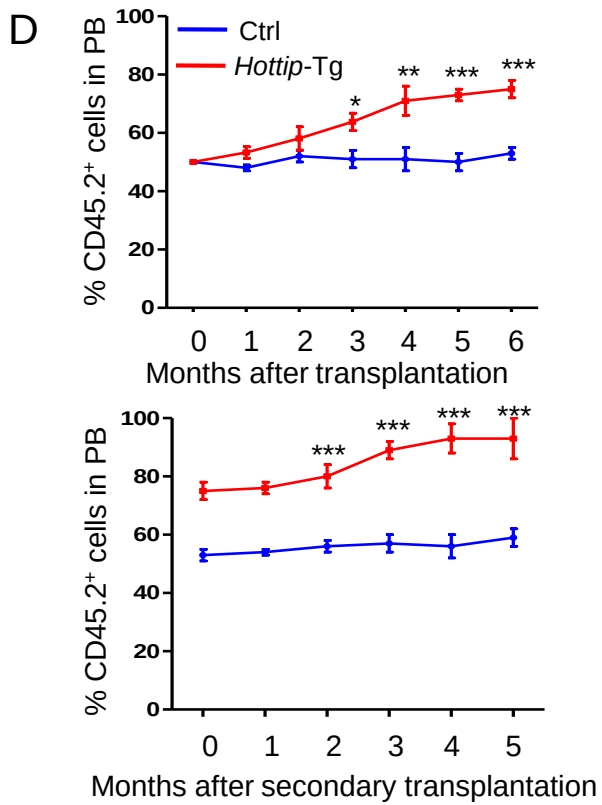

A
GEMM
E
GM
Colonies/2 x 10⁴ BM cells
Colonies/1 x 10⁵ SP cells
Ctrl
Hottip-Tg

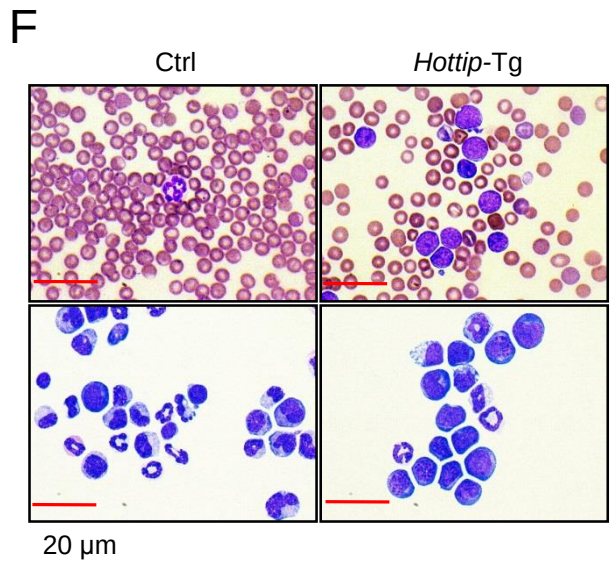

B
Ctrl
Hottip-Tg
Colony number
1st
2nd
3rd
4th
5th


E
Percent survival
p=0.0455
Ctrl
Hottip-Tg
Days after secondary transplantation


Ctrl.
Hottip-Tg


C
WT
HottipTg
Asxl2-/-
AML1-ETO
24%
58%
18%
26%
35%
39%
22%
20%
58%
25%
28%
47%
Symmetric self renewal
Asymmetric self renewal
Symmetric differentiation


D
Ctrl
Hottip-Tg
% CD45.2⁺ cells in PB
Months after transplantation
Months after secondary transplantation


F
Ctrl
Hottip-Tg
20 μm

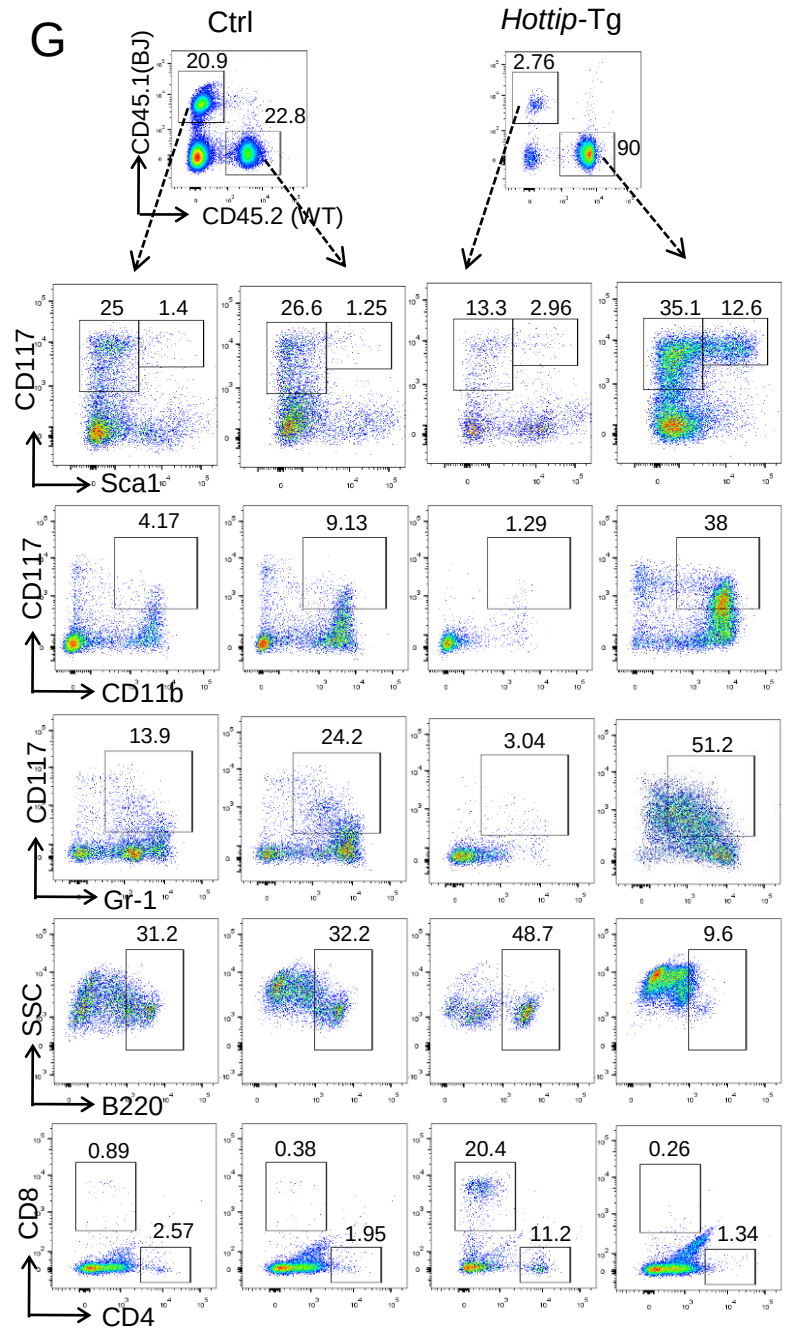
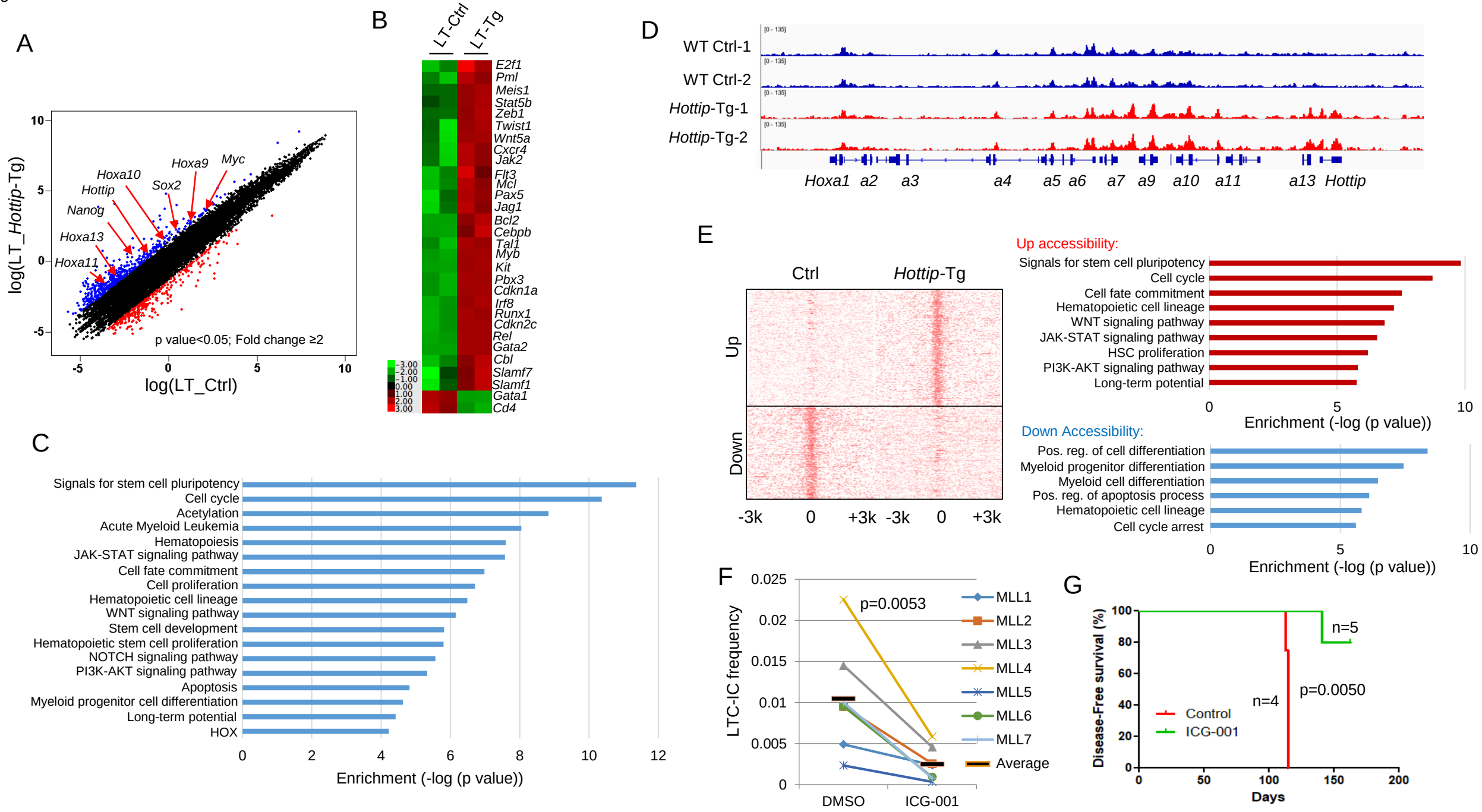
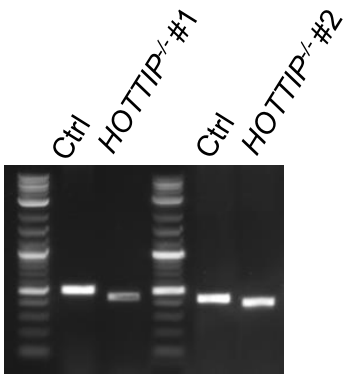
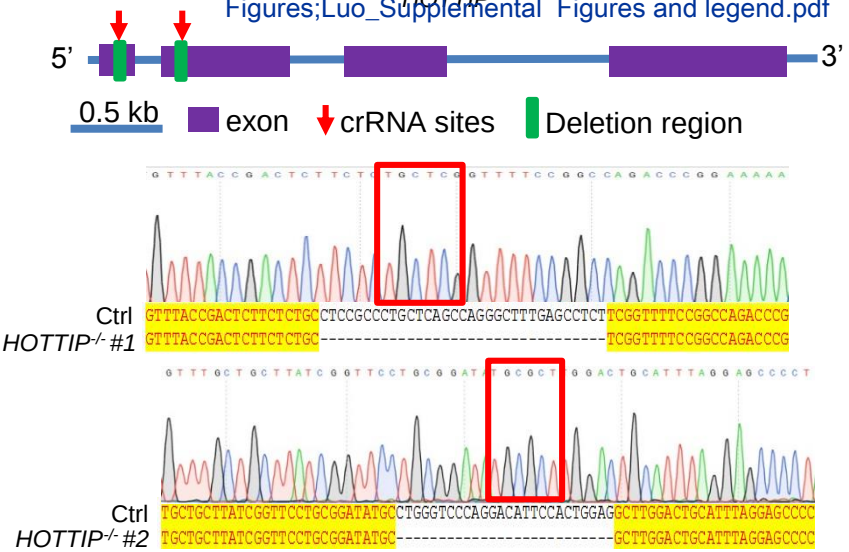

G
Ctrl
Hottip-Tg
CD45.1(BJ)
CD45.2 (WT)
CD117
Sca1
CD11b
Gr-1
SSC
B220
CD8
CD4

Figure 8

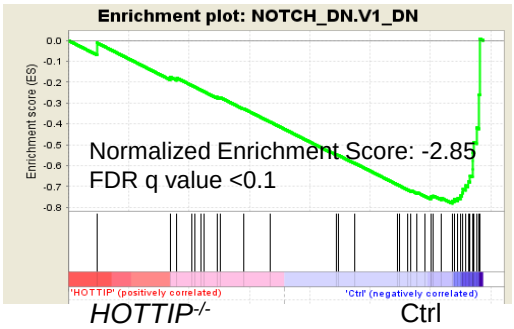

A
log(LT_Hottip-Tg)
log(LT_Ctrl)
Hoxa10
Hottip
Sox2
Hoxa9
Myc
Nanog
Hoxa13
Hoxa11
p value<0.05; Fold change ≥2
B
LT-Ctrl
LT-Tg
E2f1
Pml
Meis1
Stat5b
Zeb1
Twist1
Wnt5a
Cxcr4
Jak2
Flt3
Mcl
Pax5
Jag1
Bcl2
Cebpb
Tal1
Myb
Kit
Pbx3
Cdkn1a
Irf8
Runx1
Cdkn2c
Rel
Gata2
Cbl
Slamf7
Slamf1
Gata1
Cd4
C
Signals for stem cell pluripotency
Cell cycle
Acetylation
Acute Myeloid Leukemia
Hematopoiesis
JAK-STAT signaling pathway
Cell fate commitment
Cell proliferation
Hematopoietic cell lineage
WNT signaling pathway
Stem cell development
Hematopoietic stem cell proliferation
NOTCH signaling pathway
PI3K-AKT signaling pathway
Apoptosis
Myeloid progenitor cell differentiation
Long-term potential
HOX
Enrichment (-log (p value))
D
WT Ctrl-1
WT Ctrl-2
Hottip-Tg-1
Hottip-Tg-2
Hoxa1 a2 a3 a4 a5 a6 a7 a9 a10 a11 a13 Hottip
E
Ctrl
Hottip-Tg
Up
Down
-3k 0 +3k -3k 0 +3k
Up accessibility:
Signals for stem cell pluripotency
Cell cycle
Cell fate commitment
Hematopoietic cell lineage
WNT signaling pathway
JAK-STAT signaling pathway
HSC proliferation
PI3K-AKT signaling pathway
Long-term potential
Enrichment (-log (p value))
Down Accessibility:
Pos. reg. of cell differentiation
Myeloid progenitor differentiation
Myeloid cell differentiation
Pos. reg. of apoptosis process
Hematopoietic cell lineage
Cell cycle arrest
Enrichment (-log (p value))
F
LTC-IC frequency
p=0.0053
MLL1
MLL2
MLL3
MLL4
MLL5
MLL6
MLL7
Average
DMSO
ICG-001
G
Disease-Free survival (%)
n=5
p=0.0050
n=4
Control
ICG-001
Days

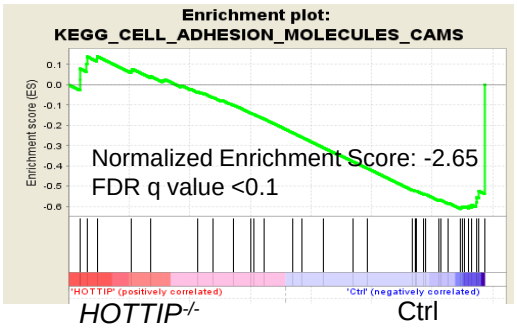
A



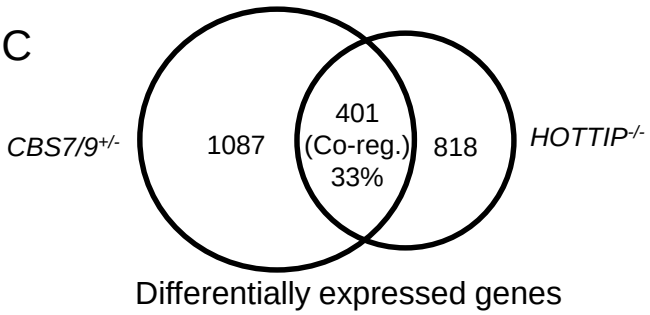


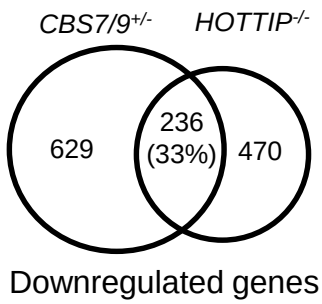
B

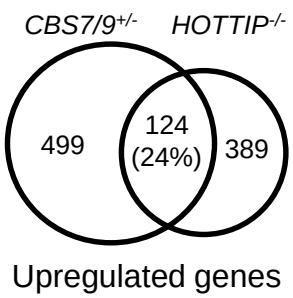




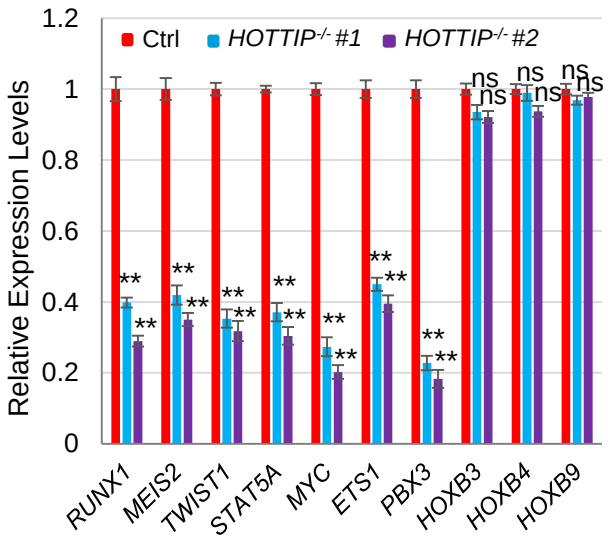
C



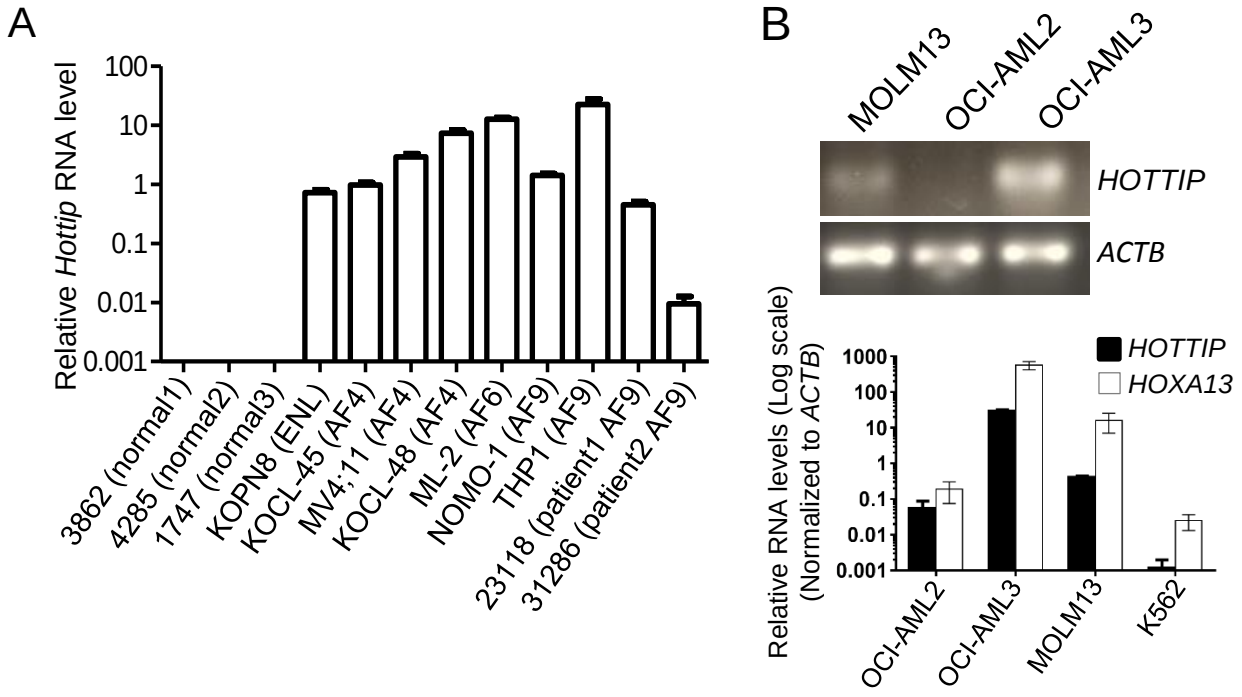




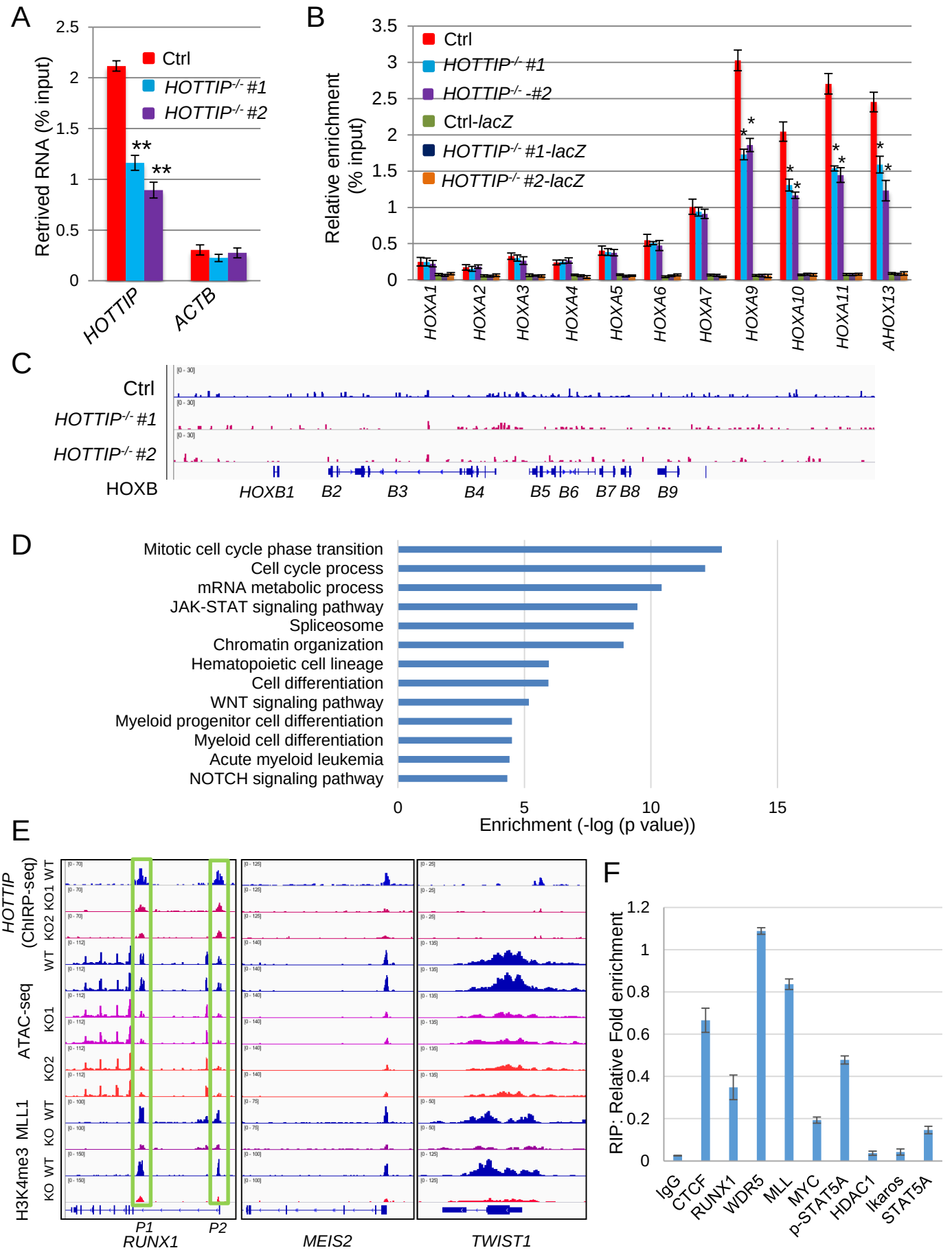
D

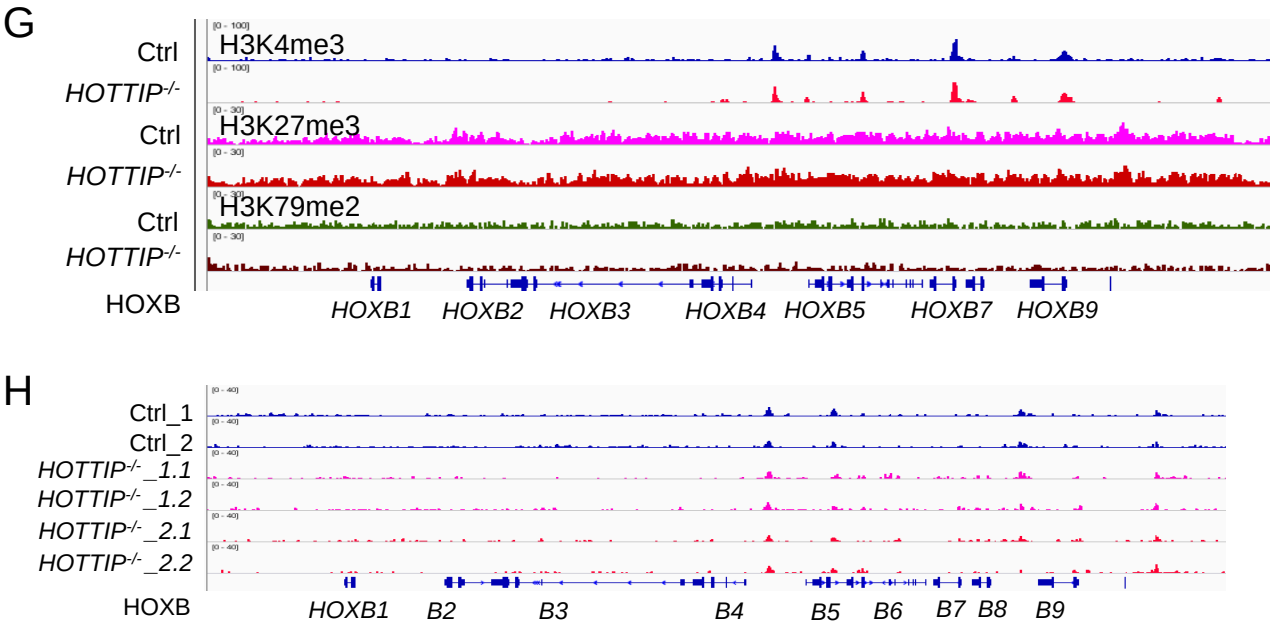


Supplementary Figure 1. Related to Figure 1; *HOTTIP* loss led to inhibition of posterior HOXA genes and other genes critical for hematopoiesis and leukemogenesis (A) PCR based genotyping and Sanger sequencing confirmation of two CRISPR-Cas9 mediated *HOTTIP*$^{-/-}$ clones in MOLM13 cells. **(B)** Enrichment of downregulated target genes involved in NOTCH signaling and cell adhesion/migration pathways in the *HOTTIP*$^{-/-}$ clones compared to WT control as shown by GSEA. **(C)** Overlap between differentially expressed genes by comparing RNA-seq data obtained from the *HOTTIP*$^{-/-}$ and the CBS7/9 boundary attenuated MOLM13 cells. **(D)** RT-qPCR validation of the key altered hematopoietic/leukemic genes identified by RNA-seq analysis of the WT control and two *HOTTIP*$^{-/-}$ clones. Data is presented as mean ± SD from three or four independent experiments; *p<0.05; **p<0.01 by Student's t-test.


A
Relative Hottip RNA level
100
10
1
0.1
0.01
0.001
3862 (normal1)
4285 (normal2)
1747 (normal3)
KOPN8 (ENL)
KOCL-45 (AF4)
MV4;11 (AF4)
KOCL-48 (AF4)
ML-2 (AF6)
NOMO-1 (AF9)
THP1 (AF9)
23118 (patient1 AF9)
31286 (patient2 AF9)
B
MOLM13
OCI-AML2
OCI-AML3
HOTTIP
ACTB
Relative RNA levels (Log scale)
(Normalized to ACTB)
HOTTIP
HOXA13
1000
100
10
1
0.1
0.01
0.001
OCI-AML2
OCI-AML3
MOLM13
K562

Supplementary Figure 2. Related to Figure 2; *HOTTIP* is aberrantly expressed in AML patients and cell lines. (A) Expression of *HOTTIP* in normal healthy individual, MLL rearranged AML cell lines and patients by RT-qPCR. Data represents mean ± SD from three independent experiments. **(B)** Expression of *HOTTIP* in MOLM13, OCI-AML3 and OCI-AML2 cells determined by RT-PCR or RT-qPCR.


A
Retrived RNA (% input)
Ctrl
HOTTIP-/- #1
HOTTIP-/- #2
HOTTIP
ACTB
B
Relative enrichment (% input)
Ctrl
HOTTIP-/- #1
HOTTIP-/- -#2
Ctrl-lacZ
HOTTIP-/- #1-lacZ
HOTTIP-/- #2-lacZ
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
AHOX13
C
Ctrl
HOTTIP-/- #1
HOTTIP-/- #2
HOXB
HOXB1
B2
B3
B4
B5
B6
B7
B8
B9
D
Mitotic cell cycle phase transition
Cell cycle process
mRNA metabolic process
JAK-STAT signaling pathway
Spliceosome
Chromatin organization
Hematopoietic cell lineage
Cell differentiation
WNT signaling pathway
Myeloid progenitor cell differentiation
Myeloid cell differentiation
Acute myeloid leukemia
NOTCH signaling pathway
Enrichment (-log (p value))
E
HOTTIP (ChIRP-seq)
ATAC-seq
H3K4me3 MLL1
WT
KO1
KO2
KO
P1
P2
RUNX1
MEIS2
TWIST1
F
RIP: Relative Fold enrichment
IgG
CTCF
RUNX1
WDR5
MLL
MYC
p-STAT5A
HDAC1
Ikaros
STAT5A



Supplementary Figure 3. Related to Figure 3; Deletion of *HOTTIP* in AML cells perturbs its binding to chromatin as well as target gene chromatin structure and accessibility. (A) RT-qPCR analysis of RNA retrieved by the complementary *HOTTIP*-tiling probes and *lacZ* probes compared WT and *HOTTIP*$^{-/-}$ MOLM13 cells. Error bars show mean ± SD from three to four independent experiments; *p<0.05; **p<0.01 by Student's t-test. **(B)** ChIRP-qPCR analysis of the *HOTTIP* RNA enrichment at the *HOXA* locus compared WT and *HOTTIP*$^{-/-}$ MOLM13 cells. Error bars show mean ± SD from three to four independent experiments; *p<0.05; **p<0.01 by Student's t-test. **(C)** ChIRP-seq analysis of *HOTTIP* lncRNA binding in the *HOXB* locus compared WT and *HOTTIP*$^{-/-}$ MOLM13 cells. **(D)** The genes and pathways regulated by the *HOTTIP* bound intergenic regions were analyzed and annotated by the Gene Ontology analysis. **(E)** Alterations in *HOTTIP* binding (ChIRP-seq), chromatin accessibility (ATAC-seq), MLL1 recruitment and H3K4me3 enrichment (ChIP-seq) in the *HOTTIP trans* regulated genes, *RUNX1*, *MEIS2*, and *TWIST1*, compared WT and *HOTTIP* KO MOLM13 cells. P1: promoter 1 of *RUNX1*; P2: promoter 2 of *RUNX1.* **(F)** RT-qPCR analysis of *HOTTIP* RNA retrieved by antibodies against RUNX1, WDR5, MLL1, MYC, and p-STAT5A precipitated from the MOLM13 nuclear extract. **(G)** ChIP-seq analysis of changes in H3K4me3, H3K27me3, and H3K79me2 modification levels in the *HOXB* locus in the MOLM13 cells compared with the WT and the *HOTTIP*$^{-/-}$. **(H)** ATAC-seq analysis of changes in chromatin accessibility upon *HOTTIP*$^{-/-}$ in MOLM13 cells. Shown are altered promoter chromatin accessibility in the *HOXB* locus.

Table S1: Transcription factor bound Motifs analyzed and annotated using ChIRP-seq and ATAC-seq datasets. Related to Figure 3.

ATAC associated			***HOTTIP*-CHIRP**	
Motif	**p value**		**Motif**	**p value**
ETS1	1.00E-465		E-box	1.00E-658
CTCF	1.00E-341		CRE	1.00E-520
E-Box	1.00E-265		RUNX1	1.00E-467
CRE	1.00E-231		HOXA13	1.00E-360
RUNX1	1.00E-215		c-Myc	1.00E-320
c-Myc	1.00E-180		MEIS1	1.00E-280
Max	1.00E-139		USF1	1.00E-259
STAT5	1.00E-122		CTCF	1.00E-220
USF1	1.00E-118		Max	1.00E-215
E2F	1.00E-98		STAT5	1.00E-142
STAT1	1.00E-87		TEAD4	1.00E-120
PU.1	1.00E-81		Smad3	1.00E-101
STAT3	1.00E-70		SOX9	1.00E-91
Nanog	1.00E-49		HLF	1.00E-78
YY1	1.00E-35		MYB	1.00E-69
ATF3	1.00E-22		CDX4	1.00E-60
FOXA1	1.00E-18		ELF4	1.00E-48
SP2	1.00E-14		USF2	1.00E-36
ATF1	1.00E-12		STAT1	1.00E-27
FLI1	1.00E-10		PU.1	1.00E-16
KLF6	1.00E-08		TCF12	1.00E-10
MYB	1.00E-08		Tbx20	1.00E-09
NFAT	1.00E-07		ERG	1.00E-08
ELF4	1.00E-06		TCF4	1.00E-07
TBX5	1.00E-05		AP1	1.00E-07
E2A	1.00E-05		LHX1	1.00E-06
TCF4	1.00E-05		ATF1	1.00E-05
			FRA1	1.00E-05

Table S2: Statistical analysis of ATAC differential peaks and p value. Related to Figures 3, 5 and 8.

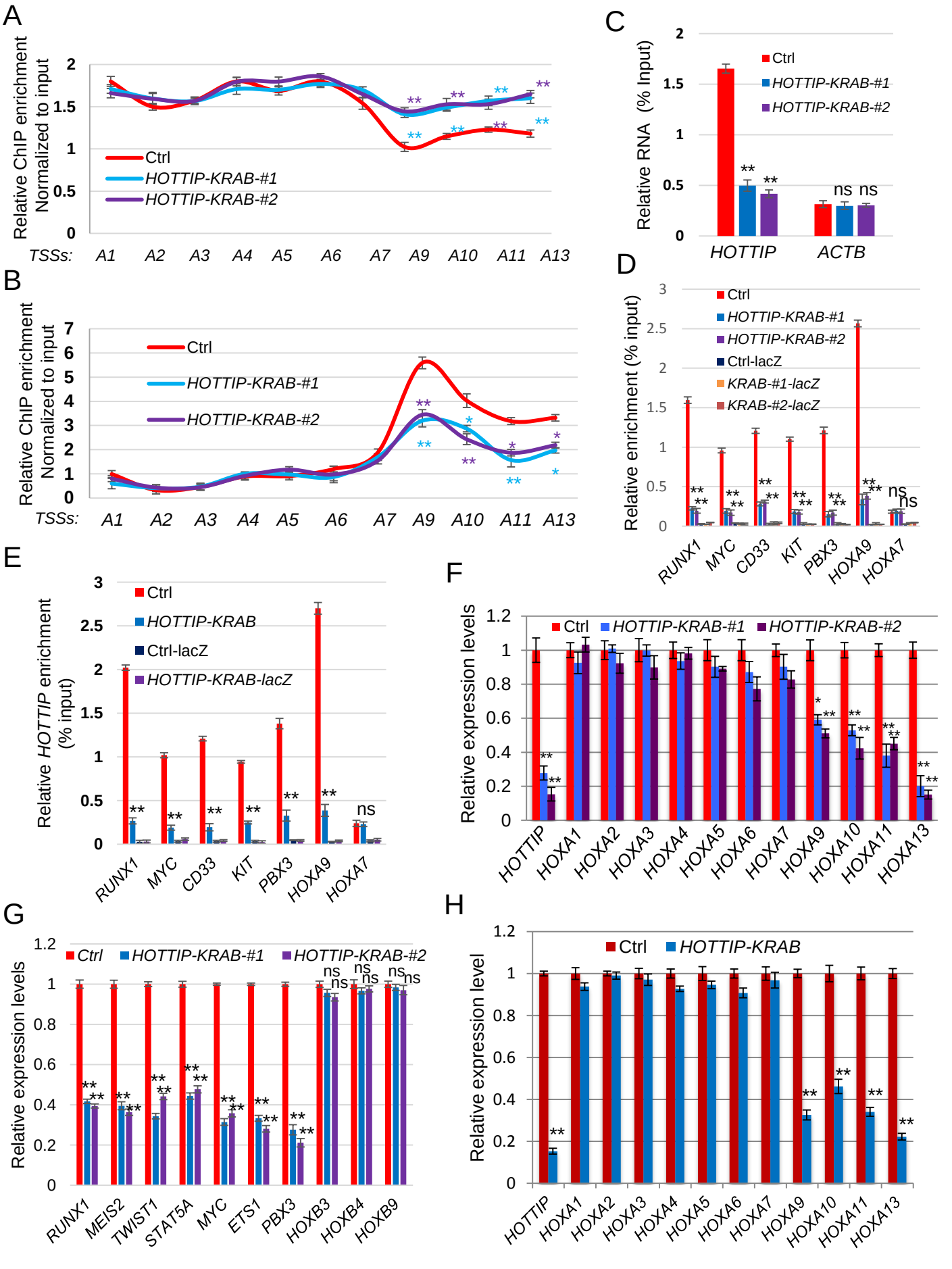
Ctrl vs *HOTTIP*$^{-/-}$					
Chr	**start**	**end**	**log2FoldChange (Ctrl vs KO)**	**Adj p value**	**Genes**
chr7	27239961	27240137	2.653	3.28E-13	*HOTTIP*
chr7	27200117	27201441	2.385	1.45E-08	*HOXA9*
chr7	27239593	27240001	1.917	1.58E-10	*HOXA13*
chr21	36260643	36261680	1.614	5.63E-06	*RUNX1*
chr17	40440299	40440741	1.445	1.72E-08	*STAT5A*
chr8	128747648	128748815	1.411	3.61E-06	*MYC*
chr7	27219354	27220019	1.395	4.17E-05	*HOXA10*
chr7	27224552	27225091	1.378	3.63E-05	*HOXA11*
chr19	51728058	51728855	1.356	4.11E-06	*CD33*
chr7	19156515	19157393	1.346	5.20E-08	*TWIST1*
chr15	37393235	37393784	1.338	5.19E-05	*MEIS2*
chr9	128509129	128509991	1.327	3.45E-06	*PBX3*
chr4	55523944	55524288	1.214	7.64E-07	*KIT*

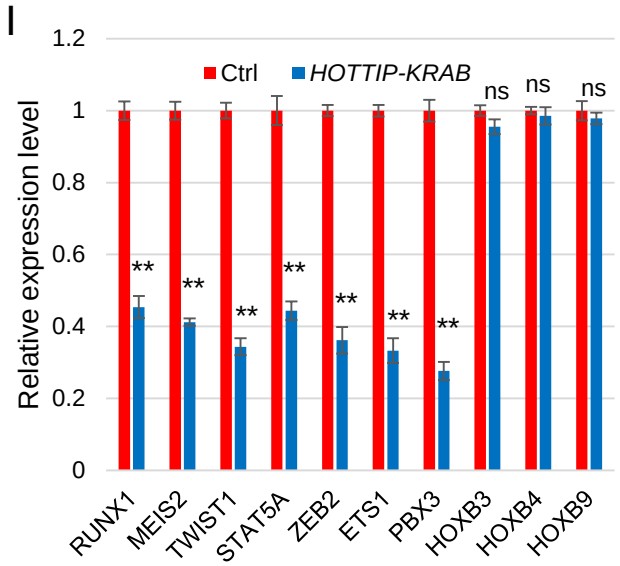
HOTTIP-VP vs *CBS7/9*$^{+/-}$					
Chr	**start**	**end**	**log2FoldChange (VP vs KO)**	**Adj p value**	**Genes**
chr7	27239612	27239999	2.882	5.56E-17	*HOXA13*
chr7	27239945	27240143	2.265	3.51E-16	*HOTTIP*
chr21	36260648	36261668	1.831	8.35E-10	*RUNX1*
chr15	37393239	37393779	1.776	6.44E-09	*MESI2*
chr7	19156523	19157390	1.482	1.08E-07	*TWIST1*
chr2	239756363	239757156	1.387	1.01E-06	*TWIST2*
chr2	66662301	66663195	1.352	1.98E-05	*MEIS1*
chr7	27200112	27201446	1.323	2.04E-09	*HOXA9*
chr17	40440313	40440733	1.302	5.33E-07	*STAT5A*
chr8	128747661	128748802	1.281	6.43E-07	*MYC*
chr2	145277521	145278173	1.256	4.11E-06	*ZEB2*
chr9	128509121	128510008	1.245	8.97E-05	*PBX3*
chr7	27219332	27220028	1.230	5.34E-06	*HOXA10*
chr7	27224556	27225095	1.202	7.95E-05	*HOXA11*
chr11	128391555	128392392	1.198	5.49E-06	*ETS1*

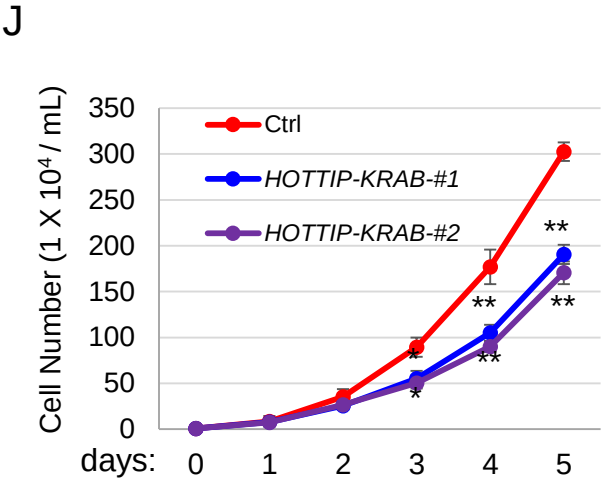
Continued

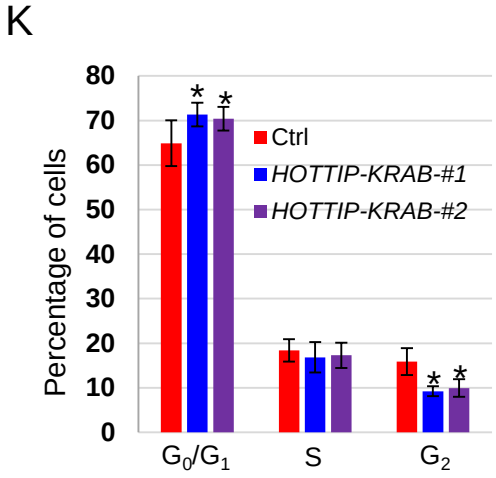
***Hottip*-Tg vs Ctrl LT-HSC**					
Chr	**start**	**end**	**log2FoldChange (HT vs Ctrl)**	**Adj p value**	**Genes**
chr6	52212340	52213115	2.268	6.89E-21	*Hottip*
chr6	52172626	52173303	2.090	1.21E-20	*Hoxa9*
chr6	52209155	52210718	1.971	3.35E-19	*Hoxa13*
chr16	92826082	92826426	1.789	1.21E-16	*Runx1*
chr6	52189793	52191345	1.747	2.26E-18	*Hoxa10*
chr2	34227091	34227906	1.521	6.48E-12	*Pbx3*
chr11	18918453	18919407	1.342	6.48E-10	*Meis1*
chr15	61816553	61817375	1.267	3.26E-13	*Myc*
chr12	34642150	34643195	1.204	3.42E-09	*Twist1*
chr6	52195231	52196121	1.145	4.98E-07	*Hoxa11*

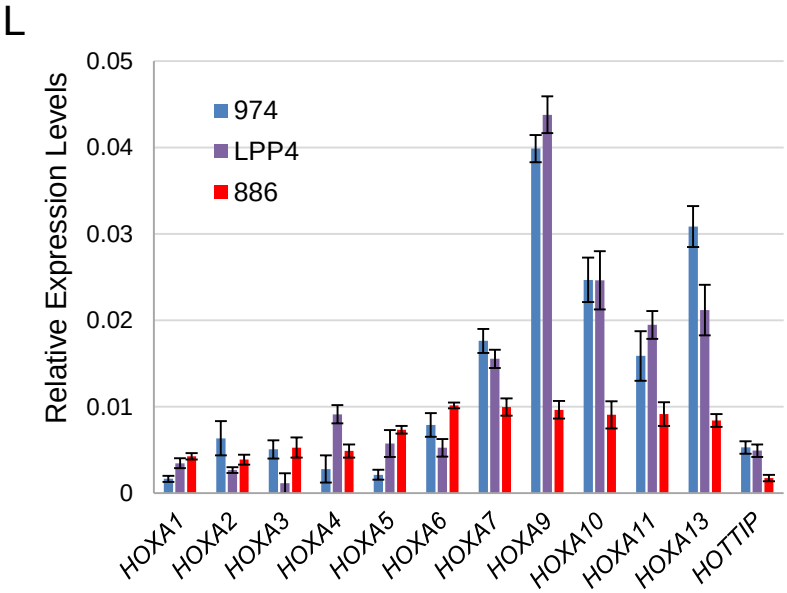
***Hottip*-Tg vs Ctrl ST-HSC**					
Chr	**start**	**end**	**log2FoldChange(HT vs Ctrl)**	**Adj p value**	**Genes**
chr6	52212348	52213124	2.395	1.18E-28	*Hottip*
chr6	52172626	52173303	1.805	4.33E-17	*Hoxa9*
chr6	52209148	52210720	1.744	2.13E-11	*Hoxa13*
chr2	34227085	34227911	1.673	1.06E-15	*Pbx3*
chr16	92826076	92826431	1.512	9.67E-13	*Runx1*
chr6	52189788	52191356	1.486	1.20E-18	*Hoxa10*
chr15	61816547	61817382	1.399	9.91E-12	*Myc*
chr6	52195236	52196129	1.376	9.21E-08	*Hoxa11*
chr11	18918478	18919416	1.246	6.43E-12	*Meis1*
chr12	34642143	34643190	1.213	8.64E-07	*Twist1*


A
Relative ChIP enrichment Normalized to input
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
TSSs: A1 A2 A3 A4 A5 A6 A7 A9 A10 A11 A13
B
Relative ChIP enrichment Normalized to input
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
TSSs: A1 A2 A3 A4 A5 A6 A7 A9 A10 A11 A13
C
Relative RNA (% Input)
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
HOTTIP
ACTB
D
Relative enrichment (% input)
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
Ctrl-lacZ
KRAB-#1-lacZ
KRAB-#2-lacZ
RUNX1 MYC CD33 KIT PBX3 HOXA9 HOXA7
E
Relative HOTTIP enrichment (% input)
Ctrl
HOTTIP-KRAB
Ctrl-lacZ
HOTTIP-KRAB-lacZ
RUNX1 MYC CD33 KIT PBX3 HOXA9 HOXA7
F
Relative expression levels
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
HOTTIP HOXA1 HOXA2 HOXA3 HOXA4 HOXA5 HOXA6 HOXA7 HOXA9 HOXA10 HOXA11 HOXA13
G
Relative expression levels
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
RUNX1 MEIS2 TWIST1 STAT5A MYC ETS1 PBX3 HOXB3 HOXB4 HOXB9
H
Relative expression level
Ctrl
HOTTIP-KRAB
HOTTIP HOXA1 HOXA2 HOXA3 HOXA4 HOXA5 HOXA6 HOXA7 HOXA9 HOXA10 HOXA11 HOXA13


I
Ctrl
HOTTIP-KRAB
Relative expression level
1.2
1
0.8
0.6
0.4
0.2
0
ns
**
RUNX1
MEIS2
TWIST1
STAT5A
ZEB2
ETS1
PBX3
HOXB3
HOXB4
HOXB9

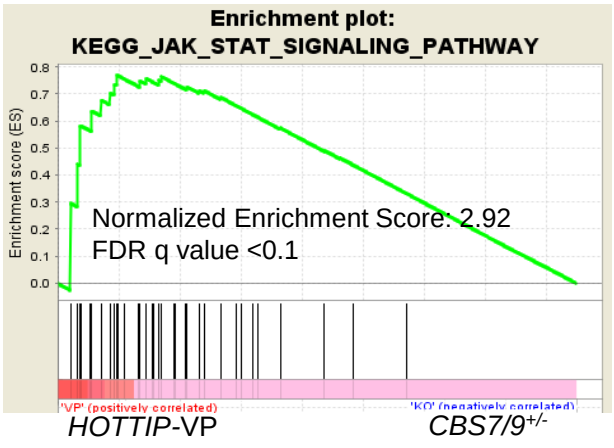

J
Cell Number (1 X 10⁴ / mL)
350
300
250
200
150
100
50
0
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
days:
0
1
2
3
4
5
*
**

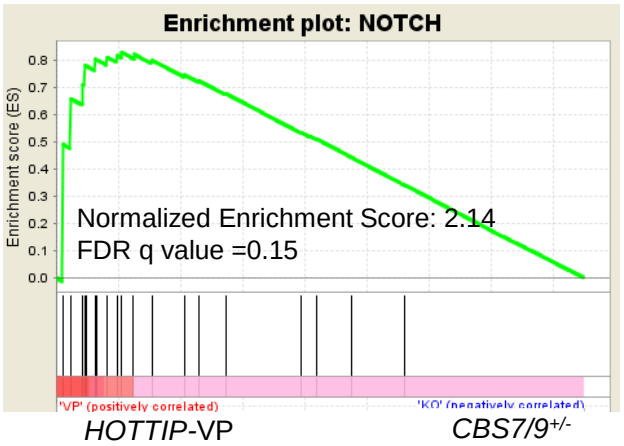

K
Percentage of cells
80
70
60
50
40
30
20
10
0
Ctrl
HOTTIP-KRAB-#1
HOTTIP-KRAB-#2
G0/G1
S
G2
*


L
Relative Expression Levels
0.05
0.04
0.03
0.02
0.01
0
974
LPP4
886
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
HOXA13
HOTTIP

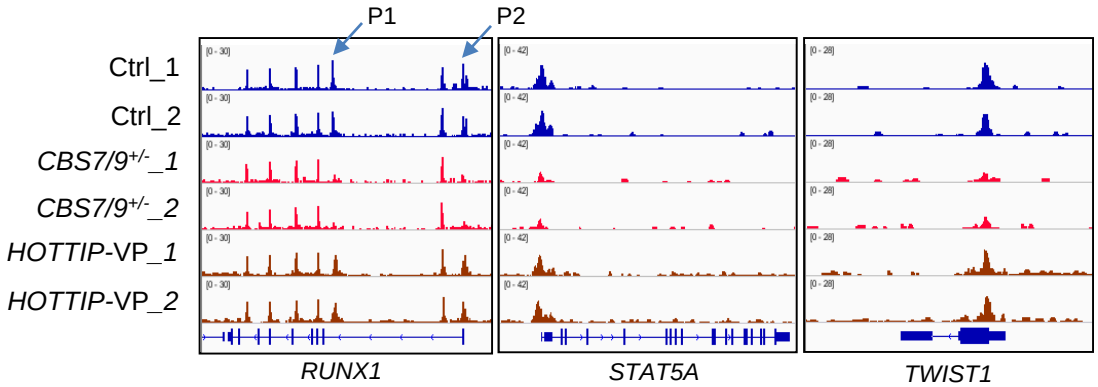
Supplementary Figure 4. Related to Figures 4; *dCas9-KRAB* mediated inhibition of *HOTTIP* lncRNA resulted in inhibition of active histone modifications and elevated repressive histone modifications in the posterior *HOXA* domain. (A, B) ChIP analysis of H3K9me2 **(A)** and H3K4me3 **(B)** enrichment at the *HOXA* locus compared WT control and *HOTTIP-dCas9-KRAB* MOLM13 clones. **(C)** RT-qPCR analysis of RNA retrieved by the complementary *HOTTIP*-tiling probes compared WT and *HOTTIP-dCas9-KRAB* inhibited MOLM13 cells. **(D)** ChIRP-qPCR analysis of the *HOTTIP* RNA enrichment at the *HOXA9* and other hematopoietic/leukemia specific genes compared WT and *HOTTIP-dCas9-KRAB* MOLM13 cells. **(E)** ChIRP-qPCR analysis of the *HOTTIP* RNA enrichment at the *HOXA9* and other hematopoietic/leukemia specific genes compared WT and *HOTTIP-dCas9-KRAB* OCI-AML3 cells carrying $NPM1^{C+}$ mutation. **(F)** RT-qPCR analysis of *HOXA* gene expression in *MLL*-rearranged MOLM13 cells compared the WT control and the *HOTTIP-dCas9-KRAB* clones. **(G)** RT-qPCR validation of the key altered hematopoietic/leukemic genes identified by RNA-seq analysis compared among the WT control and two *HOTTIP- dCas9-KRAB* clones. **(H)** RT-qPCR analysis of *HOXA* gene expression in $NPM1^{C+}$ mutated OCI-AML3 cells compared the WT control and the *HOTTIP-dCas9-KRAB* clones. **(I)** RT-qPCR validation of the key altered hematopoietic/leukemic genes identified by RNA-seq analysis compared among the WT control and *HOTTIP-KRAB* OCI-AML3 clones. **(J)** Proliferation curves of WT control and the *HOTTIP-dCas9-KRAB* MOLM13 clones were measured by cell viability count. **(K)** FACS analysis of cell cycle was carried out using propidium Iodide staining of the WT control or the *HOTTIP-dCas9-KRAB* MOLM13 clones. **(L)** RT-qPCR analysis of the expression levels of *HOTTIP* and *HOXA* genes in primary AML patient BM cells carrying $MLLr^{+}$ (LPP4), $NPM1^{C+}$/$Flt3\text{-}ITD^{+}$ (974), or $NPM1^{C-}$/$FLT3\text{-}ITD^{+}$ (886). For statistics, data is presented as mean ± SD from three to four independent experiments; *p<0.05; **p<0.01 by Student's t-test.

A

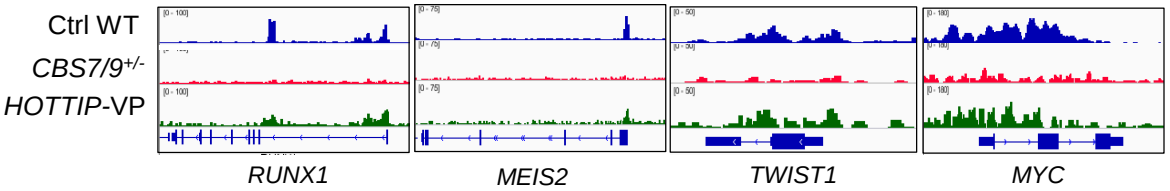

Enrichment plot:
KEGG_JAK_STAT_SIGNALING_PATHWAY
Enrichment score (ES)
Normalized Enrichment Score: 2.92
FDR q value <0.1
'VP' (positively correlated)
'KO' (negatively correlated)
HOTTIP-VP
CBS7/9+/-


Enrichment plot: NOTCH
Enrichment score (ES)
Normalized Enrichment Score: 2.14
FDR q value =0.15
'VP' (positively correlated)
'KO' (negatively correlated)
HOTTIP-VP
CBS7/9+/-

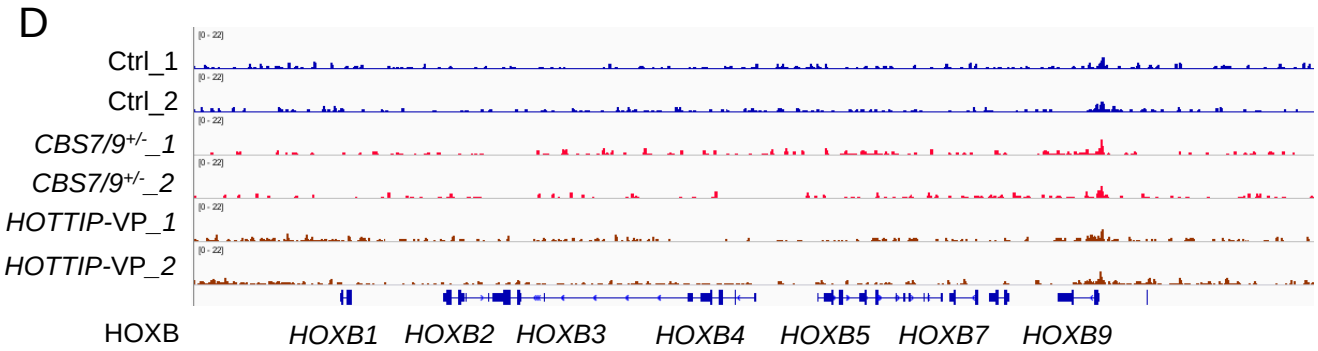
B


P1
P2
Ctrl_1
Ctrl_2
CBS7/9+/-_1
CBS7/9+/-_2
HOTTIP-VP_1
HOTTIP-VP_2
RUNX1
STAT5A
TWIST1

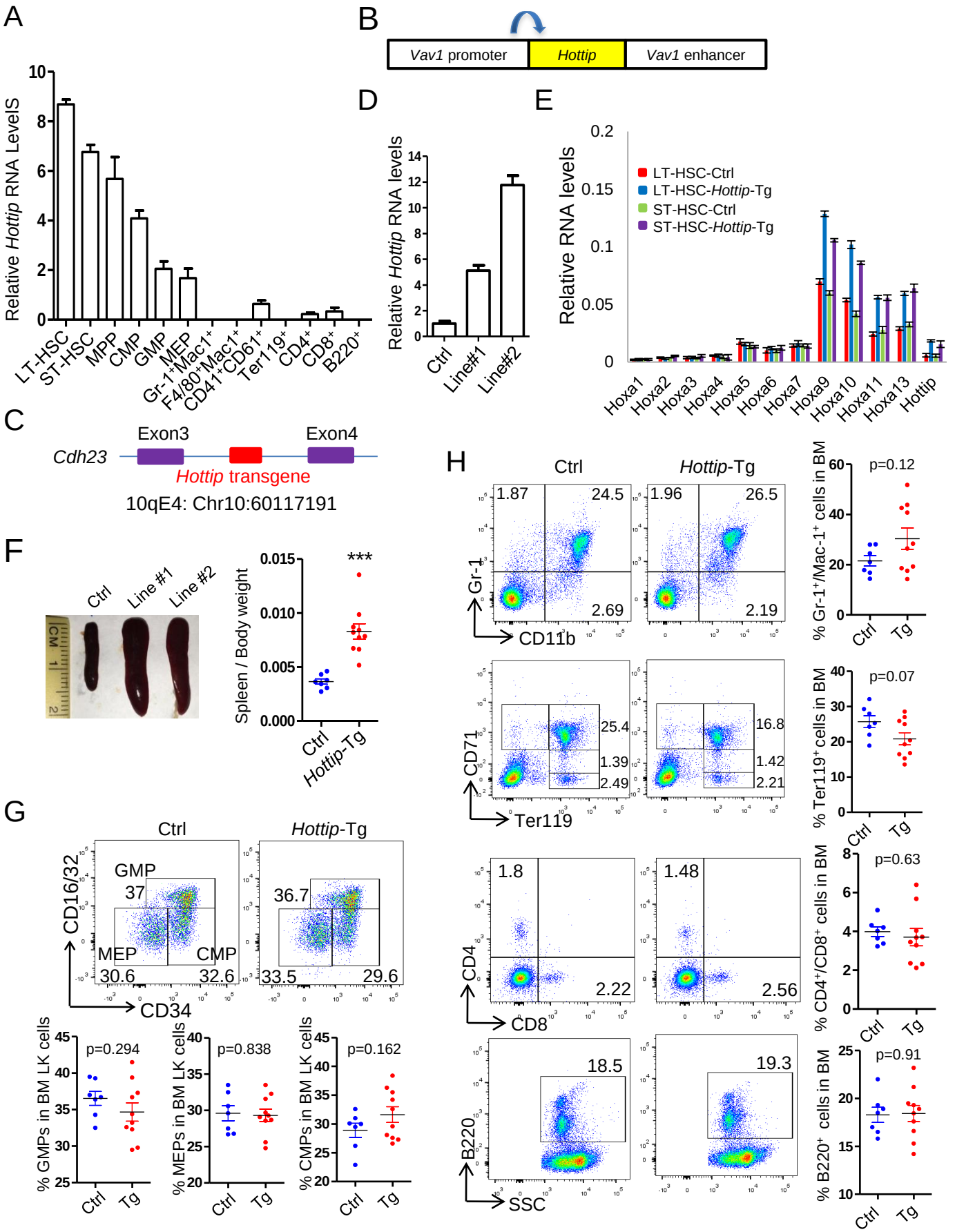
C

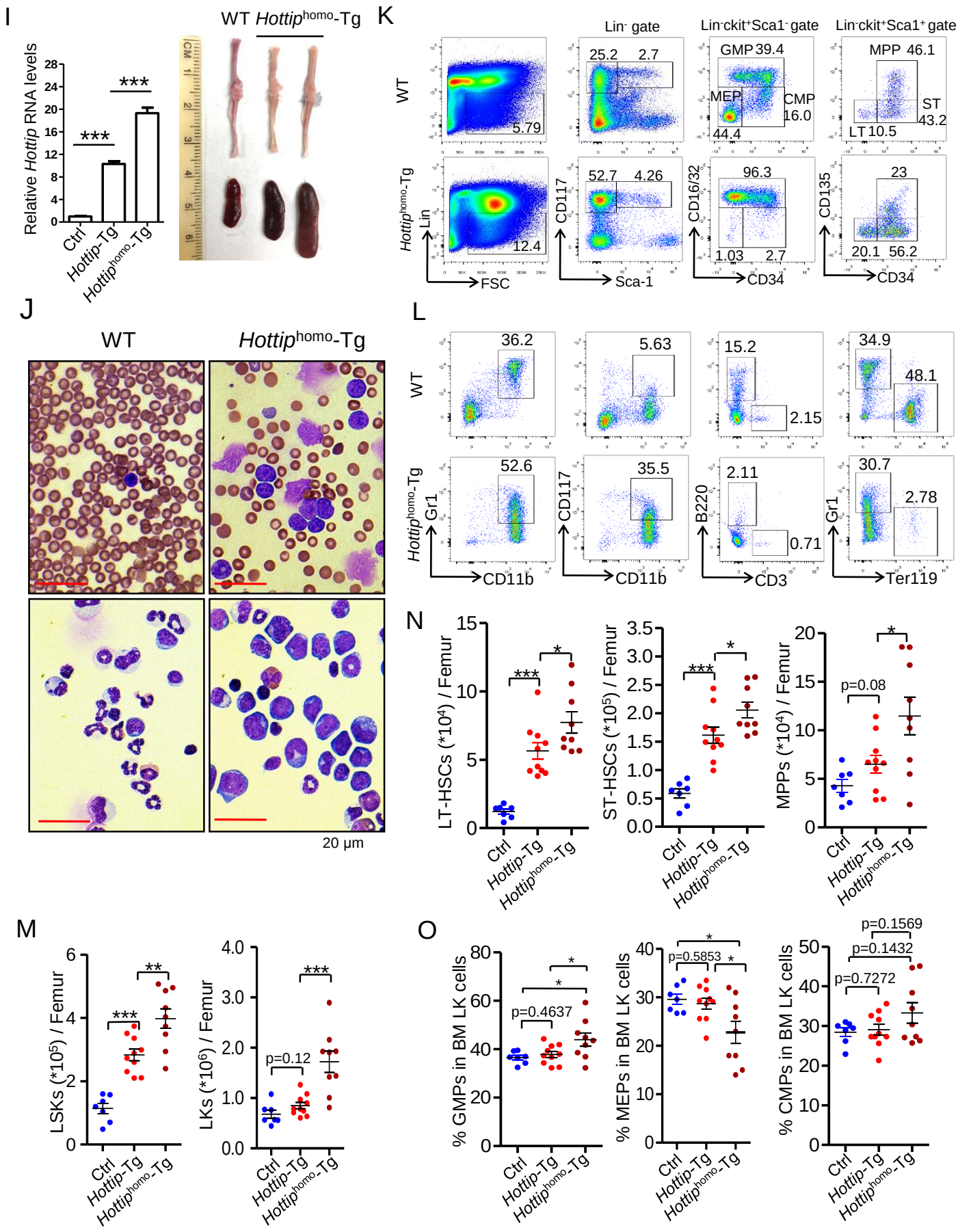

Ctrl WT
CBS7/9+/-
HOTTIP-VP
RUNX1
MEIS2
TWIST1
MYC

D


Ctrl_1
Ctrl_2
CBS7/9+/-_1
CBS7/9+/-_2
HOTTIP-VP_1
HOTTIP-VP_2
HOXB
HOXB1
HOXB2
HOXB3
HOXB4
HOXB5
HOXB7
HOXB9

Supplementary Figure 5. Related to Figure 5; Reactivation of *HOTTIP* lncRNA rescues the posterior *HOXA* chromatin defects and gene expression in the *CBS7/9* boundary-disrupted AML cells. (A) Enrichment of upregulated target genes involved in JAK-STAT signaling and NOTCH signaling pathways in the *dCas9-VP-160* mediated *HOTTIP* activated *CBS7/9*$^{+/-}$ MOLM13 clones compared to the *CBS7/9*$^{+/-}$ MOLM13 cells as shown by GSEA. **(B)** Snap shot of ATAC-seq analysis of *RUNX1*(Left), *STAT5A* (Middle), and *TWIST1* (Right) loci compared among WT MOLM13 control, *CBS7/9*$^{+/-}$ MOLM13 cells, and the *dCas9-VP-160* mediated *HOTTIP* activated *CBS7/9*$^{+/-}$ MOLM13 clones. **(C)** ChIP-seq analysis of MLL1 recruitment in the Non-*HOX* targets compared among WT MOLM13 control, *CBS7/9*$^{+/-}$ MOLM13 cells, and the *dCas9-VP-160* mediated *HOTTIP* activated *CBS7/9*$^{+/-}$ MOLM13 clones. **(D)** ATAC-seq analysis of the alteration of chromatin accessibility in the *HOXB* locus compared among WT MOLM13 control, *CBS7/9*$^{+/-}$ MOLM13 cells, and the *dCas9-VP-160* mediated *HOTTIP* activated *CBS7/9*$^{+/-}$ MOLM13 clones.

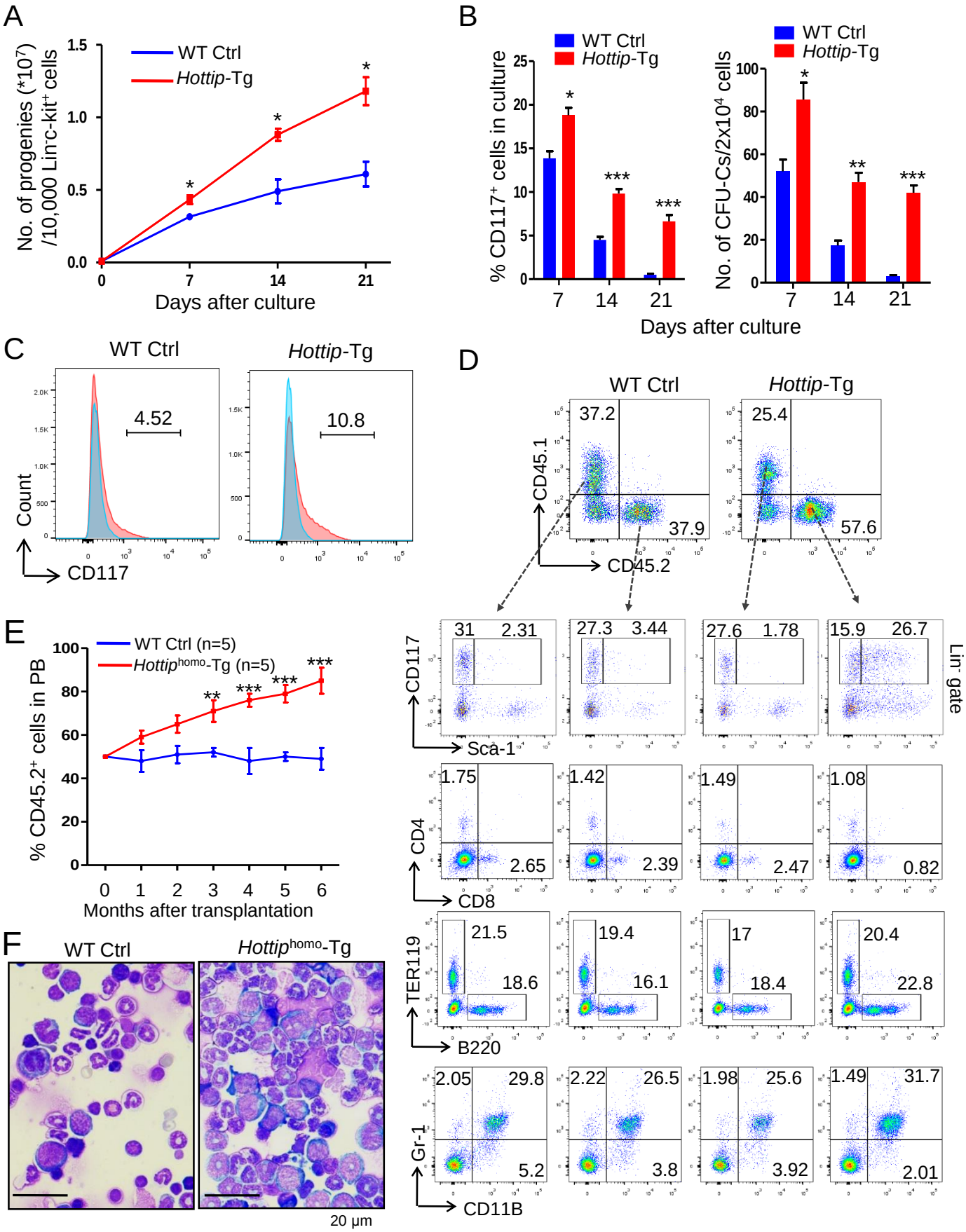

A
Relative Hottip RNA LevelS
LT-HSC
ST-HSC
MPP
CMP
GMP
MEP
Gr-1+Mac1+
F4/80+Mac1+
CD41+CD61+
Ter119+
CD4+
CD8+
B220+
B
Vav1 promoter
Hottip
Vav1 enhancer
D
Relative Hottip RNA levels
Ctrl
Line#1
Line#2
E
Relative RNA levels
LT-HSC-Ctrl
LT-HSC-Hottip-Tg
ST-HSC-Ctrl
ST-HSC-Hottip-Tg
Hoxa1
Hoxa2
Hoxa3
Hoxa4
Hoxa5
Hoxa6
Hoxa7
Hoxa9
Hoxa10
Hoxa11
Hoxa13
Hottip
C
Exon3
Exon4
Cdh23
Hottip transgene
10qE4: Chr10:60117191
F
Ctrl
Line #1
Line #2
Spleen / Body weight
***
Ctrl
Hottip-Tg
G
Ctrl
Hottip-Tg
CD16/32
CD34
GMP
37
MEP
30.6
CMP
32.6
36.7
33.5
29.6
% GMPs in BM LK cells
p=0.294
% MEPs in BM LK cells
p=0.838
% CMPs in BM LK cells
p=0.162
H
Ctrl
Hottip-Tg
Gr-1
CD11b
1.87
24.5
2.69
1.96
26.5
2.19
% Gr-1+/Mac-1+ cells in BM
p=0.12
CD71
Ter119
25.4
1.39
2.49
16.8
1.42
2.21
% Ter119+ cells in BM
p=0.07
CD4
CD8
1.8
2.22
1.48
2.56
% CD4+/CD8+ cells in BM
p=0.63
B220
SSC
18.5
19.3
% B220+ cells in BM
p=0.91
Ctrl
Tg

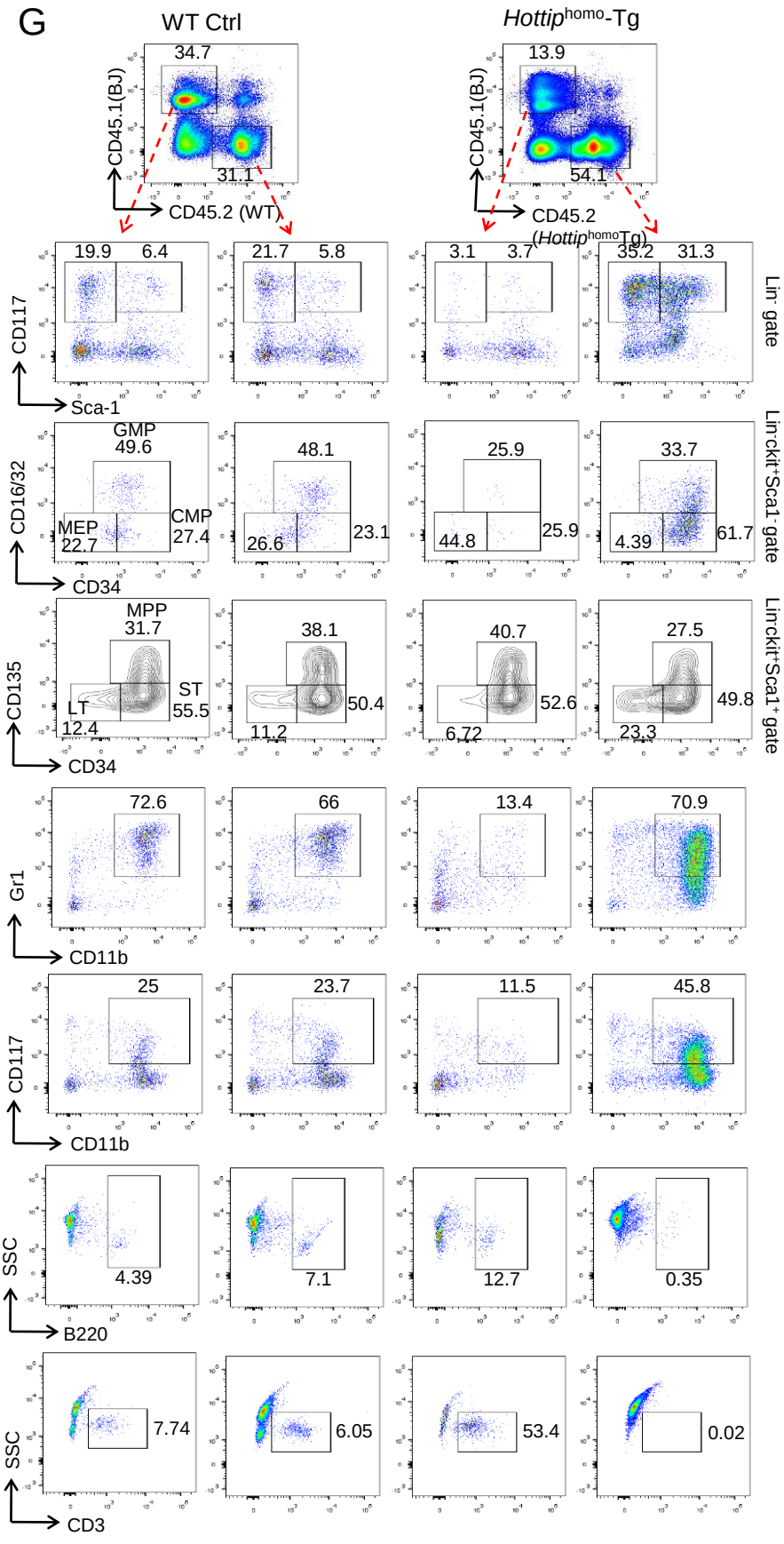

I
WT Hottip^homo-Tg
Relative Hottip RNA levels
28
24
20
16
12
8
4
0
***
***
Ctrl
Hottip-Tg
Hottip^homo-Tg
CM
1
2
3
4
5
6
K
Lin- gate
Lin-ckit+Sca1- gate
Lin-ckit+Sca1+ gate
WT
Hottip^homo-Tg
5.79
25.2
2.7
GMP 39.4
MEP
CMP
16.0
44.4
MPP 46.1
ST
43.2
LT 10.5
12.4
52.7
4.26
96.3
1.03
2.7
23
20.1
56.2
Lin
FSC
CD117
Sca-1
CD16/32
CD34
CD135
CD34
J
WT
Hottip^homo-Tg
20 μm
L
WT
Hottip^homo-Tg
36.2
5.63
15.2
2.15
34.9
48.1
52.6
35.5
2.11
0.71
30.7
2.78
Gr1
CD11b
CD117
CD11b
B220
CD3
Gr1
Ter119
N
LT-HSCs (*10^4) / Femur
ST-HSCs (*10^5) / Femur
MPPs (*10^4) / Femur
***
*
p=0.08
Ctrl
Hottip-Tg
Hottip^homo-Tg
M
LSKs (*10^5) / Femur
LKs (*10^6) / Femur
***
**
p=0.12
O
% GMPs in BM LK cells
% MEPs in BM LK cells
% CMPs in BM LK cells
p=0.4637
p=0.5853
p=0.7272
p=0.1432
p=0.1569

Supplementary Figure 6. Related to Figure 6; *Hottip* lncRNA transgenic expression in hematopoietic compartment perturbs HSC function and results in AML-like disease. (**A**) RT-qPCR analyses of *Hottip* RNA expression pattern of FACS-sorted different cell populations of WT mice. **(B)** Diagram of the *Vav1* promoter driven *Hottip* transgene strategy. **(C)** PCR based transgenic integrating location identification (TAIL) assay maps the *Vav1-Hottip* transgene integration site in the mouse chromosome 10. **(D)** RT-qPCR analysis of *Hottip* RNA expression in BM cells of WT and 2 lines of the *Hottip*-Tg mice. **(E)** RT-qPCR analysis of *Hottip* and *Hoxa* gene expression from purified LT-HSC and ST-HSC populations from WT control and *Hottip*-Tg mice. **(F)** Gross appearance of spleens of representative WT and two lines of *Hottip*-Tg mice (left). Showing is spleen/body weight ratio (right) for age matched WT (n = 7) and the *Hottip*-Tg (n = 10) mice. **(G)** FACS analysis of GMP, MEP and CMP populations within BM LK cells of representative young WT and *Hottip*-Tg mice (Top). Quantitation of the percentage of GMP, MEP and CMP cell populations in the $Lin^{-}ckit^{+}Sca1^{-}$ cells of each genotype of mice are shown (bottom). n=7-10 mice/genotype. **(H)** FACS analysis and quantification of myeloid (Gr-1/Mac-1), erythroid (Ter119/CD71), B (B220/SSC) and T (CD4/CD8) cell populations in the BM of representative young WT and *Hottip*-Tg mice (8-week old). **(I)** The expression levels of *Hottip* in WT, *Hottip*-Tg, and $Hottip^{homo}$-Tg BM mononuclear cells (Left) as well as gross appearance of spleens and femurs of representative WT and moribund $Hottip^{homo}$-Tg mice (Right). **(J)** May-Giemsa stained PB smears and BM cytospins prepared from representative WT and moribund $Hottip^{Homo}$-Tg mice. **(K)** FACS analysis of Lin^{-}, LSK/LK, GMP/CMP, MEP, as well as LT-HSC/ST-HSC/MPP cell populations in the BM of representative young WT and $Hottip^{Homo}$-Tg mice. **(L)** FACS analysis and quantification of myeloid (Gr-1/Mac-1), erythroid (Ter119), B (B220/SSC) and T (CD4/CD8) cell populations in the BM of representative young WT and *Hottip*-Tg mice (8-week old). **(M)** Quantitation of the total LSK and LK cell populations per femur of young WT (n=7), *Hottip*-Tg (n=10), and $Hottip^{Homo}$-Tg (n=9) mice are shown. **(N)** Quantitation of the total LT-HSC, ST-HSC and MPP cell numbers per femur of young WT (n=7), *Hottip*-Tg (n=10), and $Hottip^{Homo}$-Tg (n=9) mice are shown. **(O)** Quantitation of the total GMP. MEP, and CMP cell populations per femur of young WT (n=7), *Hottip*-Tg (n=10), and $Hottip^{Homo}$-Tg (n=9) mice are shown. For statistics, data (Panels F-I, M-O) is presented as mean ± SD.

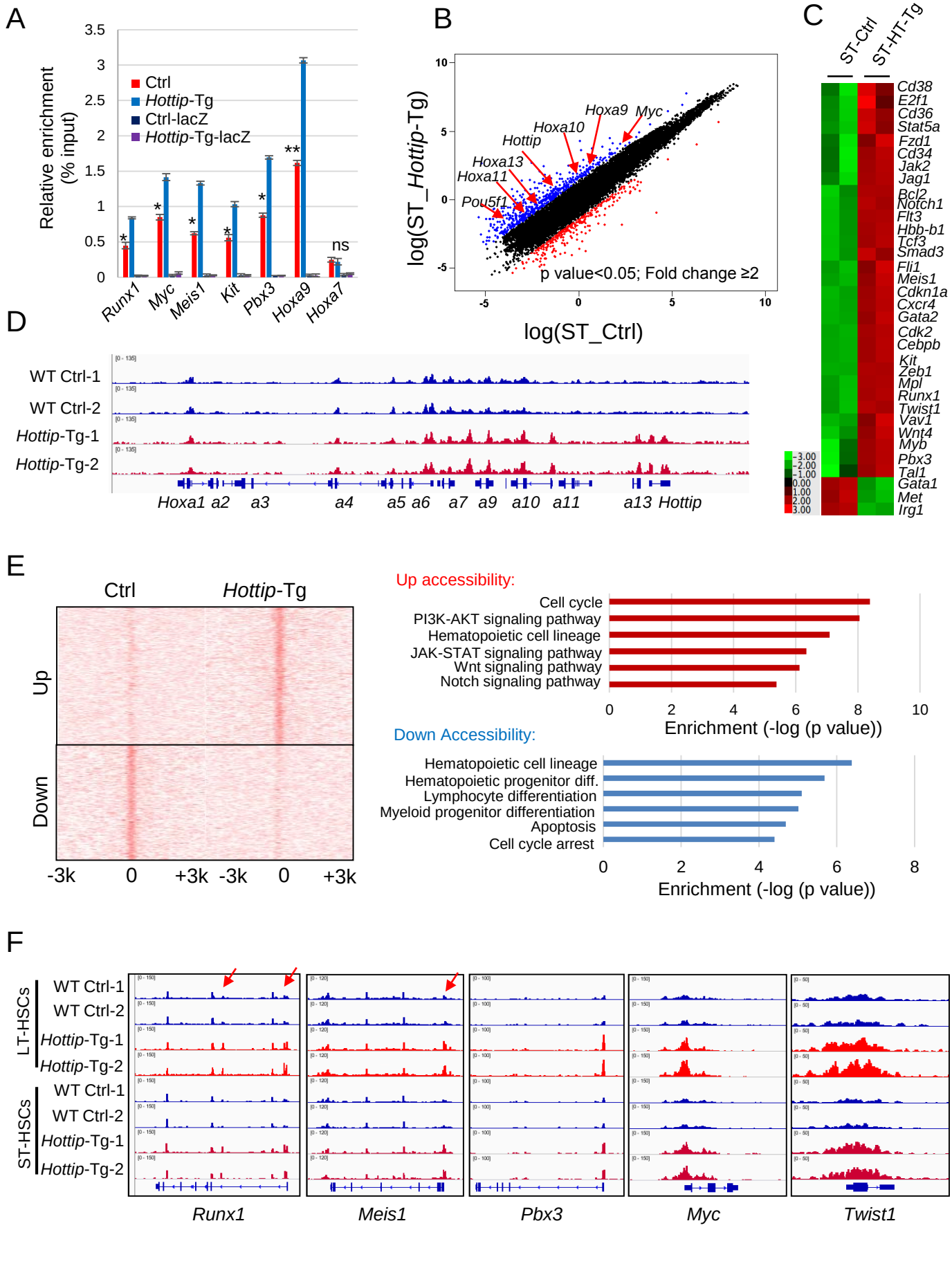
Table S3: Indel and snp variants analysis of whole exome sequencing of the *Hottip*-Tg AML mice. Related to Figure 6.

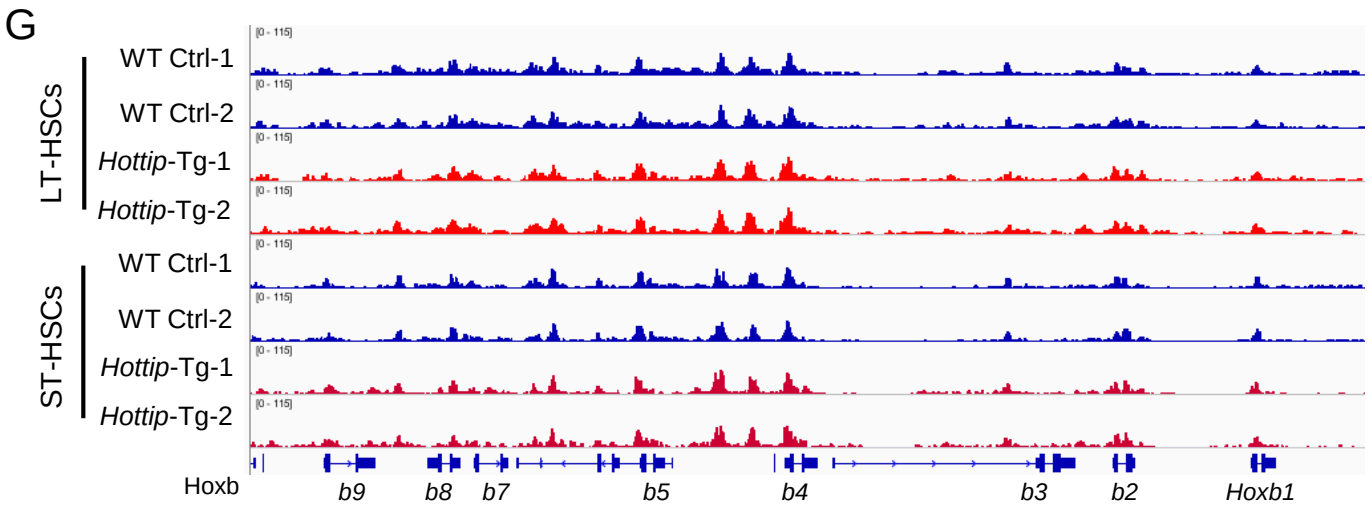
strelka method for exome-seq snp and indel analysis					
sample name #		**indel**	**snp**	**indel_AA_change**	**snp_AA_change**
Hottip-Tg mice	22	39	35	1	15
	39	81	50	3	14
	48	47	63	4	14
	51	49	76	2	27
	55	37	44	3	16
WT mice	69	54	52	4	14
	130	42	42	1	15
Recurrent mutation gene					
Mroh2a					

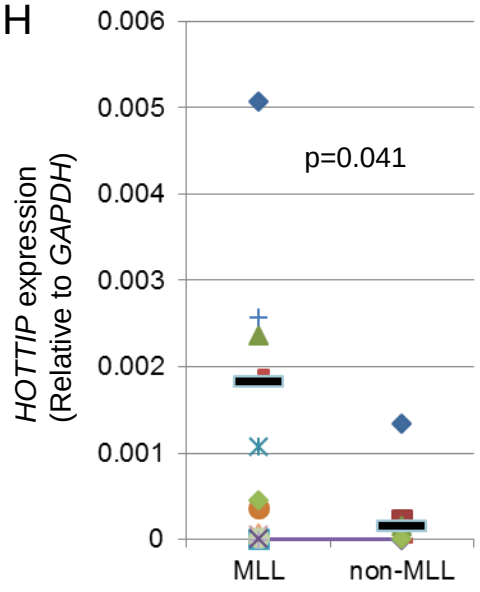

A
No. of progenies (*10^7) /10,000 Lin-c-kit+ cells
WT Ctrl
Hottip-Tg
Days after culture
B
% CD117+ cells in culture
No. of CFU-Cs/2x10^4 cells
WT Ctrl
Hottip-Tg
Days after culture
C
WT Ctrl
Hottip-Tg
4.52
10.8
Count
CD117
D
WT Ctrl
Hottip-Tg
37.2
37.9
25.4
57.6
CD45.1
CD45.2
31
2.31
27.3
3.44
27.6
1.78
15.9
26.7
CD117
Sca-1
Lin- gate
1.75
2.65
1.42
2.39
1.49
2.47
1.08
0.82
CD4
CD8
21.5
18.6
19.4
16.1
17
18.4
20.4
22.8
TER119
B220
2.05
5.2
29.8
2.22
3.8
26.5
1.98
3.92
25.6
1.49
2.01
31.7
Gr-1
CD11B
E
WT Ctrl (n=5)
Hottip homo-Tg (n=5)
% CD45.2+ cells in PB
Months after transplantation
F
WT Ctrl
Hottip homo-Tg
20 μm

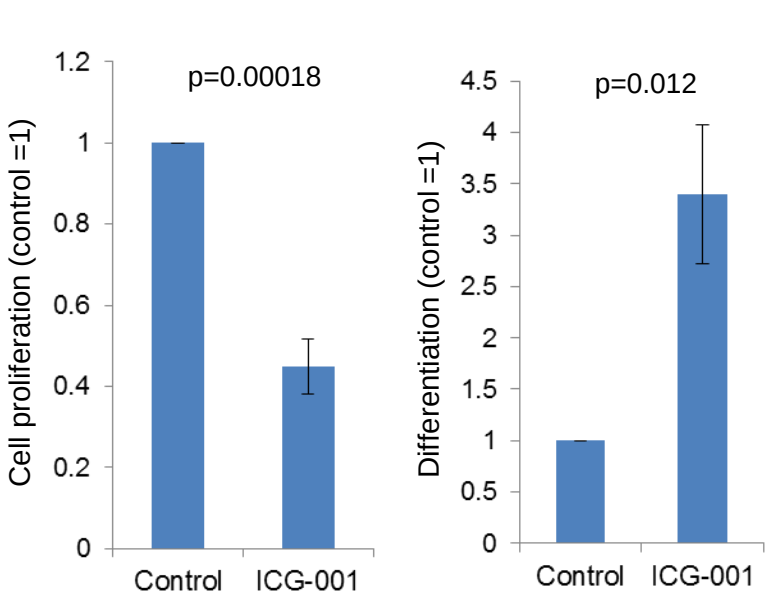

G
WT Ctrl
Hottip homo-Tg
34.7
31.1
CD45.1(BJ)
CD45.2 (WT)
13.9
54.1
CD45.2
(Hottip homoTg)
19.9
6.4
21.7
5.8
3.1
3.7
35.2
31.3
CD117
Sca-1
Lin- gate
GMP
49.6
MEP
22.7
CMP
27.4
48.1
26.6
23.1
25.9
44.8
25.9
33.7
4.39
61.7
CD16/32
CD34
Lin-ckit+Sca1- gate
MPP
31.7
LT
12.4
ST
55.5
38.1
11.2
50.4
40.7
6.72
52.6
27.5
23.3
49.8
CD135
CD34
Lin-ckit+Sca1+ gate
72.6
66
13.4
70.9
Gr1
CD11b
25
23.7
11.5
45.8
CD117
CD11b
4.39
7.1
12.7
0.35
SSC
B220
7.74
6.05
53.4
0.02
SSC
CD3

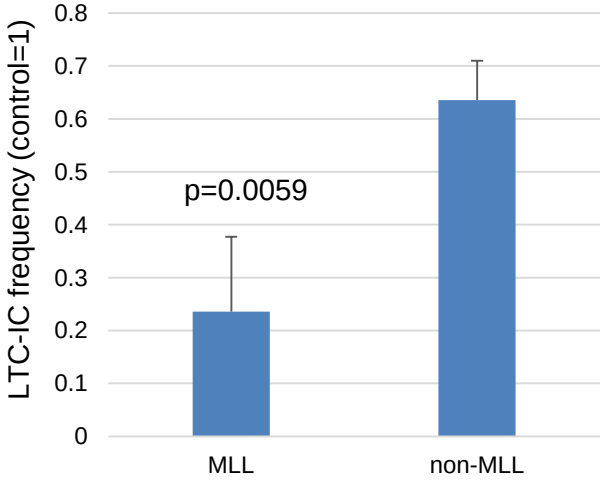
Supplementary Figure 7. Related to Figures 7; *Hottip* regulates self-renewal and proliferation of HSCs. (**A-C**) The expansion and proliferation potential of WT and *Hottip*-Tg HSC/HPCs were examined by culturing BM $Lin^{-}c\text{-}Kit^{+}$ cells in the presence of SCF, TPO, IL-3, G-CSF and EPO. No. of total cells (**A**), percent $c\text{-}Kit^{+}$ cells and No. of CFU-Cs **(B)** in the progenies of $Lin^{-}c\text{-}Kit^{+}$ cells cultured for 7, 14 and 21 days were shown; representative histogram showed the c-Kit (CD117) expression in the cultures of WT and *Hottip*-Tg $Lin^{-}c\text{-}Kit^{+}$ cells for 14 days (**C**). **(D)** FACS analysis showing CD45.2 v.s. CD45.1 chimerism as well as their respective distribution of LSK/LK cell (within Lin^{-} cells) and lineage populations in the BM of representative mice receiving WT or *Hottip*-Tg BM cells. **(E)** Kinetic FACS analyses of CD45.2 (Donor) chimerisms in the PB of recipients (CD45.1) receiving WT or $Hottip^{Homo}$-Tg BM cells. **(F)** Images of May–Grunwald–Giemsa stained cytospin preparations of BM cells from representative mice receiving WT or $Hottip^{Homo}$-Tg BM cells. Scale bar, 20 μm. **(G)** FACS analysis showing CD45.2 vs. CD45.1 chimerism and their respective lineage distribution of LSK/LK, GMP/CMP/MEP, LT-HSC/ST-HSC/MPP cell populations, as well as lineage distribution of myeloid, B and T (CD3/B220), immature myeloid in the BM of representative mice receiving WT or $Hottip^{Homo}$-Tg BM cells. For statistics, data (Panels A, B, and E) is presented as mean ± SD.

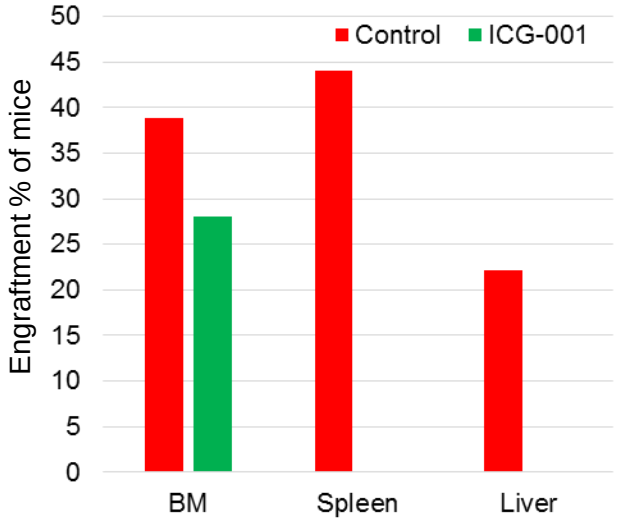

A
Relative enrichment (% input)
Ctrl
Hottip-Tg
Ctrl-lacZ
Hottip-Tg-lacZ
Runx1
Myc
Meis1
Kit
Pbx3
Hoxa9
Hoxa7
ns
B
log(ST_Hottip-Tg)
log(ST_Ctrl)
Hoxa9
Myc
Hoxa10
Hottip
Hoxa13
Hoxa11
Pou5f1
p value<0.05; Fold change ≥2
C
ST-Ctrl
ST-HT-Tg
Cd38
E2f1
Cd36
Stat5a
Fzd1
Cd34
Jak2
Jag1
Bcl2
Notch1
Flt3
Hbb-b1
Tcf3
Smad3
Fli1
Meis1
Cdkn1a
Cxcr4
Gata2
Cdk2
Cebpb
Kit
Zeb1
Mpl
Runx1
Twist1
Vav1
Wnt4
Myb
Pbx3
Tal1
Gata1
Met
Irg1
D
WT Ctrl-1
WT Ctrl-2
Hottip-Tg-1
Hottip-Tg-2
Hoxa1 a2 a3 a4 a5 a6 a7 a9 a10 a11 a13 Hottip
E
Ctrl
Hottip-Tg
Up
Down
-3k 0 +3k -3k 0 +3k
Up accessibility:
Cell cycle
PI3K-AKT signaling pathway
Hematopoietic cell lineage
JAK-STAT signaling pathway
Wnt signaling pathway
Notch signaling pathway
Enrichment (-log (p value))
Down Accessibility:
Hematopoietic cell lineage
Hematopoietic progenitor diff.
Lymphocyte differentiation
Myeloid progenitor differentiation
Apoptosis
Cell cycle arrest
Enrichment (-log (p value))
F
LT-HSCs
ST-HSCs
WT Ctrl-1
WT Ctrl-2
Hottip-Tg-1
Hottip-Tg-2
Runx1
Meis1
Pbx3
Myc
Twist1


G
LT-HSCs
WT Ctrl-1
WT Ctrl-2
Hottip-Tg-1
Hottip-Tg-2
ST-HSCs
WT Ctrl-1
WT Ctrl-2
Hottip-Tg-1
Hottip-Tg-2
Hoxb
b9
b8
b7
b5
b4
b3
b2
Hoxb1


H
HOTTIP expression (Relative to GAPDH)
0.006
0.005
0.004
0.003
0.002
0.001
0
p=0.041
MLL
non-MLL


I
Cell proliferation (control =1)
1.2
1
0.8
0.6
0.4
0.2
0
p=0.00018
Control
ICG-001
Differentiation (control =1)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
p=0.012
Control
ICG-001


J
LTC-IC frequency (control=1)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
p=0.0059
MLL
non-MLL


K
Engraftment % of mice
50
45
40
35
30
25
20
15
10
5
0
Control
ICG-001
BM
Spleen
Liver

Supplementary Figure 8. Related to Figure 8; Transgenic expression of *Hottip* lncRNA remodels chromatin structure and alters hematopoietic transcription programs. (A) ChIRP-RT-qPCR analysis of the *Hottip* RNA enrichment at the *Hoxa7, Hoxa9* and other hematopoietic/leukemia specific genes in the purified BM HSCs (Including LT-HSCs and ST-HSCs) compared WT and *Hottip*-Tg mice. Data is presented as mean ± SD from three to four independent experiments; *p<0.05; **p<0.01 by Student's t-test. **(B)** Scatter blot of RNA-seq analysis of more than two folds of differentially expressed genes upon overexpression of *Hottip* lincRNA in BM ST-HSC populations of the *Hottip*-Tg mice as compared to WT control. **(C)** Heat map analysis for changed expression of representative genes associated with hematopoiesis and leukemogenesis upon *Hottip* overexpression in BM ST-HSC populations of the *Hottip* Tg mice as compared to WT control (Cutoff: Fold change ≥ 2; q value < 0.05). **(D)** ATAC-seq analysis of chromatin accessibility in the *Hoxa* locus compared the WT control and *Hottip* Tg mouse BM ST-HSCs. **(E)** ATAC-seq promoter density map of ST-HSCs sorted from WT and *Hottip* transgenic BM. Significant upregulated ATAC-seq promoter peaks correlate with GO terms of enriched pathways annotated by GREAT analysis compared between WT control and *Hottip* transgenic ST-HSCs (Top). Significant downregulated ATAC-seq promoter peaks correlate with GO terms of enriched pathways annotated by GREAT analysis compared between WT control and *Hottip* transgenic ST-HSCs (Bottom). **(F)** ATAC-seq analysis of *Hottip trans*-regulated genes important for hematopoiesis and leukemogenesis compared between WT control and the *Hottip*-Tg BM LT-HSC (Top) and ST-HSC (Bottom) populations. **(G)** ATAC-seq analysis of *Hoxb* locus compared between WT control and the *Hottip*-Tg BM LT-HSC (Top) and ST-HSC (Bottom) populations. **(H)** Expression of *HOTTIP* by RT-qPCR. 22 MLL-AML cell lines and primary samples and 7 non-MLL-AML primary samples were subjected to RT-qPCR, the *HOTTIP* expression of each sample were normalized to house-keeping gene GAPDH. The black bar represents the mean expression of each group. p value is determined by Students' t-test. **(I)** MLL-AML samples with high expression levels of *HOTTIP* were either treated with DMSO or ICG-001 (500 nM) and the cells were counted 5 days after treatment. The cell number of each sample was normalized to DMSO control (Left) or the cells were stained with 0.1% NBT and the differentiated cells were NBT-positive. The percentage of NBT-positive cells of each sample was normalized to DMSO control (Right). p value is determined by Students' t-test. **(J)** Primary MLL-AML samples that exhibit high levels of *HOTTIP* expression and Non-MLL samples with very low or silent HOTTIP expression levels were either treated with DMSO or ICG-001 (500 nM) and the long-term culture initiating cells (LTC-IC) frequency of each group were determined. The data was normalized with DMSO control which is set to 1. p value is determined by Students' t-test. **(K)** The human engraftment of bone marrow, spleen and liver of leukemic mice were determined by positive on human CD45 and CD33 comparing vehicle (Control; red; n=4) and ICG-001 (50 mg/Kg; green; n=5).