RESEARCH ARTICLE

PU.1-Dependent Enhancer Inhibition Separates *Tet2*-Deficient Hematopoiesis from Malignant Transformation

Maria M. Aivalioti[1,2], Boris A. Bartholdy[1], Kith Pradhan[3], Tushar D. Bhagat[3], Aliona Zintiridou[1], Jong Jin Jeong[4], Victor J. Thiruthuvanathan[1], Mario Pujato[5], Aditi Paranjpe[5], Chi Zhang[1], Ross L. Levine[6], Aaron D. Viny[7], Amittha Wickrema[4], Amit Verma[3,8], and Britta Will[1,3]

ABSTRACT Cytosine hypermethylation in and around DNA-binding sites of master transcription factors, including PU.1, occurs in aging hematopoietic stem cells following acquired loss-of-function mutations of DNA methyl-cytosine dioxygenase ten–eleven translocation-2 (TET2), albeit functional relevance has been unclear. We show that *Tet2*-deficient mouse hematopoietic stem and progenitor cells undergo malignant transformation upon compromised gene regulation through heterozygous deletion of an upstream regulatory region ($URE^{\Delta/WT}$) of the PU.1 gene. Although compatible with multilineage blood formation at young age, *Tet2*-deficient PU.1 $URE^{\Delta/WT}$ mice develop highly penetrant, transplantable acute myeloid leukemia (AML) during aging. Leukemic stem and progenitor cells show hypermethylation at putative PU.1-binding sites, fail to activate myeloid enhancers, and are hallmarked by a signature of genes with impaired expression shared with human AML. Our study demonstrates that Tet2 and PU.1 jointly suppress leukemogenesis and uncovers a methylation-sensitive PU.1-dependent gene network as a unifying molecular vulnerability associated with AML.

SIGNIFICANCE: We identify moderately impaired PU.1 mRNA expression as a biological modality predisposing *Tet2*-deficient hematopoietic stem and progenitor cells to malignant transformation. Our study furthermore uncovers a methylation-sensitive PU.1 gene network as a common feature of myeloid leukemia potentially allowing for the identification of patients at risk for malignant transformation.

See related commentary by Schleicher and Pietras, p. 378.

INTRODUCTION

Age-related clonal hematopoiesis (ARCH) is frequently found in elderly individuals and is characterized by the acquisition of recurrent DNA mutations in multipotent stem and progenitor cells that expand and form clones, often at low frequency (1). ARCH contributes to several age-related pathologic conditions, including atherosclerosis (2), cardiovascular disease (3), vascular complications (4), and increased risk for developing hematologic malignancies (5, 6).

Genetic alterations affecting key enzymes governing DNA cytosine methylation, ten–eleven translocation-2 (TET2) and DNA methyltransferase 3 alpha (DNMT3A), are among the most common ARCH-associated gene lesions (1, 7–9) and are also prevalent in acute myeloid leukemia (AML; ref. 5). AML is a molecularly and clinically highly diverse hematologic malignancy (10, 11) with increased incidence rates in the elderly (12), whose mechanisms of pathogenesis have remained incompletely resolved. This has hampered the development of curative therapies. Except for acute promylelocytic leukemia (APL), a distinct subtype of AML driven by chromosome translocation t(15;17) and the resulting *PML–RARα* fusion oncogene, several cooperating gene lesions appear to be required for AML pathogenesis (13–16). For instance, loss of Tet2 increases hematopoietic stem cell (HSC) self-renewal while progressively decreasing their ability to undergo differentiation commitment, yet is insufficient to drive leukemogenesis (17). Only combined loss of *Tet2* and *Dnmt3A* (18) or presence of AML-associated gene lesions impairing myeloid cell maturation triggers malignant transformation of *Tet2*-deficient mouse HSC (13).

Several AML-typic mutations, such as *NPM1c*, *FLT3-ITD*/TDK, or *AML1-ETO*, moderately compromise the regulation or function of PU.1 (19–22), an E26-erythroblast transformation-specific (ETS) transcription factor essential for hematopoietic multilineage commitment and differentiation (23). Past work has demonstrated that PU.1 acts in a highly dose- and context-specific manner as a master transcription factor governing hematopoiesis, but also contributes to leukemogenesis if perturbed. Several mouse models have allowed investigation of the impact of PU.1 deregulation at various dose levels using conventional gene knockout (24), enhancer (25), or autoregulatory element (26) perturbation. Although a lack of PU.1 is incompatible with multilineage hematopoiesis (24), homozygous deletion of a critical upstream regulatory element (–14 kb $URE^{\Delta/\Delta}$) of PU.1 (resulting in 80% reduced PU.1 mRNA expression) in mouse hematopoietic stem and progenitor cells allows for multilineage differentiation, but leads to an early-onset fully penetrant leukemia or lymphoma phenotype (25). PU.1 haploinsufficient (23) or heterozygous PU.1 URE deletion-bearing mice (leading to 30% reduction of PU.1 mRNA expression in hematopoietic stem and progenitor

[1]Department of Cell Biology, Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, New York. [2]Graduate Programs in the Biomedical Sciences, Albert Einstein College of Medicine, Bronx, New York. [3]Department of Medicine (Oncology), Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, New York. [4]Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Chicago, Illinois. [5]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio. [6]Center for Hematologic Malignancies, Memorial Sloan Kettering Cancer Center, New York, New York. [7]Department of Genetics and Development, Columbia University, New York, New York. [8]Department of Developmental and Molecular Biology, Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, New York.

Corresponding Author: Britta Will, Albert Einstein College of Medicine, 1300 Morris Park Ave, Chanin 401A, Bronx, NY 10461. Phone: 718-430-3786; E-mail: britta.will@einsteinmed.edu

Blood Cancer Discov 2022;3:444–67

doi: 10.1158/2643-3230.BCD-21-0226

cells (25) have no overt hematologic defects in steady state, which had been in stark contrast to the observed minimal to moderate PU.1 perturbations observed in human AML.

Our past work uncovered the loss of appropriate PU.1 gene regulation hallmarked by sub-haploinsufficient, but significant PU.1 mRNA expression decline in HSCs (27). Moreover, hypermethylation in and around PU.1 DNA-binding sites in preleukemic hematopoietic stem and progenitor cells (28) further supports that perturbed PU.1-dependent gene regulation may precede AML. Using a model of genetic stem cell aging conferred by mismatch-repair deficiency, we demonstrated that heterozygous deletion of the PU.1 URE and subsequent loss of appropriate enhancer-mediated gene regulation of PU.1 can drive the evolution of AML, which we attributed to epigenetic inactivation of Irf8, a critical PU.1 cooperating factor sustaining myelopoiesis (27). Whether impairment of PU.1 may drive the malignant transformation of hematopoietic stem and progenitor cells harboring specific ARCH lesions had been unexplored.

To model PU.1 mRNA expression impairment in the context of *Tet2* deficiency, we used mice with deletion of an enhancer element (the −14 kb URE) in the PU.1-encoding *Sfpi1* gene governing appropriate PU.1 regulation in HSCs and myeloid cells (25). We show that heterozygous URE deletion leads to highly penetrant AML in a mouse model lacking Tet2 (17). Hematopoietic stem and progenitor cells from leukemic compound mutant mice exhibited significant loss of chromatin accessibility at PU.1-associated myeloid enhancers, concomitant DNA hypermethylation, and inhibition of a set of myeloid genes shared with human AML. Importantly, expression of these signature genes was preserved in compound mutant mice not developing AML upon aging, indicating that inhibition of PU.1-associated myeloid enhancers separates healthy from malignant clonal hematopoiesis. Our study identifies PU.1 as a critical tumor suppressor in *Tet2*-deficient hematopoiesis and suggests PU.1 impairment as a molecular denominator for malignant transformation of *Tet2*-deficient aged stem and progenitor cells.

RESULTS

Progressive Inhibition of PU.1 during Leukemic Evolution

To gain insights into PU.1 expression levels and network activity during malignant transformation, we assessed PU.1 mRNA levels in phenotypically defined human healthy (*Healthy*), age-related clonal hematopoiesis stem cells (*ARCH HSCs*), and leukemic hematopoietic stem cells (*LSC*) using a previously published data set (29). ARCH HSCs, closely resembling the regulatory landscape that defines the transcriptional make-up, of non-ARCH, healthy multipotent progenitor (MPP) cells (29), showed no discernible alterations of PU.1 mRNA levels compared with age-matched healthy individual-derived control cells (Fig. 1A). However, in LSCs, mostly sharing transcriptional programs with healthy granulocytic–monocytic progenitors (GMPs; ref. 29), PU.1 mRNA expression was found reduced to sub-haploinsufficient levels [by 21% ± 4% ($P < 0.01$) compared with their healthy counterparts; Fig. 1B]. We also noted a similar extent of PU.1 reduction in leukemic blast cells (reduced by 23% ± 7% compared with healthy controls; Fig. 1C). Pathway enrichment analysis of differentially expressed genes (DEGs) in healthy ARCH HSCs and LSCs did not reveal a strong indication for these moderate expression alterations to be accompanied by PU.1 network deregulation (z score < |2|) at the stem cell level—either within the entire group of patients with AML or within genetically defined subgroups harboring common age-associated clonal hematopoiesis lesions (Fig. 1D). Similarly, differential gene-expression analysis of hematopoietic cells isolated from a mouse model, carrying hematopoietic-specific deletion of *Tet2* (17), uncovered no detectable deregulation of the PU.1 network in either *Tet2*^{HET} or *Tet2*^{KO} hematopoietic stem and progenitor cells (HSPC) [defined throughout this study as cKit$^+$ CD8a$^-$CD4$^-$CD19$^-$B220$^-$ (Lymph$^-$) Ter119$^-$ (Ery$^-$) cells; Fig. 1E]. Using published data sets, we performed pathway analysis of DEGs in *Tet2*^{KO} LSKs (vs. wild-type LSK; GSE132090) and *Tet2*^{KO} GMPs (vs. wild-type LSKs; GSE27816) cells and found no statistically significant evidence for a compromised PU.1 network in *Tet2*-deficient cells (Fig. 1F). These observations suggested that albeit PU.1 mRNA levels may decline moderately upon leukemic transformation, its global network activity is preserved at the stem cell level. However, given that sub-haploinsufficient PU.1 expression predisposes hematopoietic stem and progenitor cells to malignant transformation in the absence of intact DNA mismatch repair (27), it is possible that a subset of PU.1-regulated targets, most vulnerable to even minimal dose alterations, may be abnormally expressed and contribute to leukemogenesis.

Perturbation of PU.1 mRNA Expression Regulation Is Compatible with Normal Hematopoiesis in Young *Tet2*-Deficient Mice

Our past work showed that the malignant transformation of PU.1 sub-haploinsufficient stem and progenitor cells requires inhibition of a key PU.1 cooperating factor, interferon-responsive factor 8 (Irf8; ref. 27). To assess if epigenetic deregulation, such as seen upon acquisition of ARCH-associated loss-of-function mutations in *TET2*, may cooperate in driving leukemogenesis of HSPCs with compromised PU.1 mRNA expression regulation, we crossed mice carrying a heterozygous deletion of a regulatory element 14 kb upstream (URE) of the transcriptional start site of the PU.1-encoding *Spi1* gene (URE$^{\Delta/+}$), a critical enhancer in immature and myeloid progenitor cells (25), with *Tet2*$^{+/flox}$ (17) and Vav-iCre$^+$ mice to induce *Tet2* deletion in hematopoietic cells (Supplementary Fig. S1A). Because TET2 deficiency has been associatated with increased DNA lesions (30), we performed Sanger sequencing analysis of the key regulatory and coding regions of the PU.1 encoding gene (*Spi1*) and excluded the mice that harbored acquired DNA lesions, which may further compromise PU.1 regulation in compound mutant animals (Supplementary Fig. S1B; Supplementary Table S1). PU.1 mRNA expression was found indistinguishable from wild-type (WT) and (PU.1$^{URE\Delta/+}$ URE^{HET}) controls in young (3- to 5-month-old) Vav-iCre$^+$PU.1$^{URE\Delta/+}$*Tet2*$^{+/flox}$ (URE^{HET}*Tet2*^{HET}) and Vav-iCre$^+$PU.1$^{URE\Delta/+}$*Tet2*$^{flox/flox}$ (URE^{HET}*Tet2*^{KO}) derived CD4$^-$ CD8a$^-$ CD19$^-$ B220$^-$ CD11b$^-$ Gr-1$^+$ Ter119$^-$ cKit$^+$ Sca-1$^+$ (further referred to as LSK) and mature myeloid cell populations (Supplementary Fig. S1C–S1F). Compound mutant mice presented with normal blood cell counts in the periphery (Supplementary Fig. S1G and S1H) and showed no overt pathology in primary or secondary sites of hematopoiesis (Supplementary Fig. S1I–S1L). Albeit URE^{HET}*Tet2*^{KO} showed a minor decrease in the abundance of phenotypical immature CD4$^-$ CD8a$^-$ B220$^-$ cKit$^+$ cells in the bone marrow (Supplementary

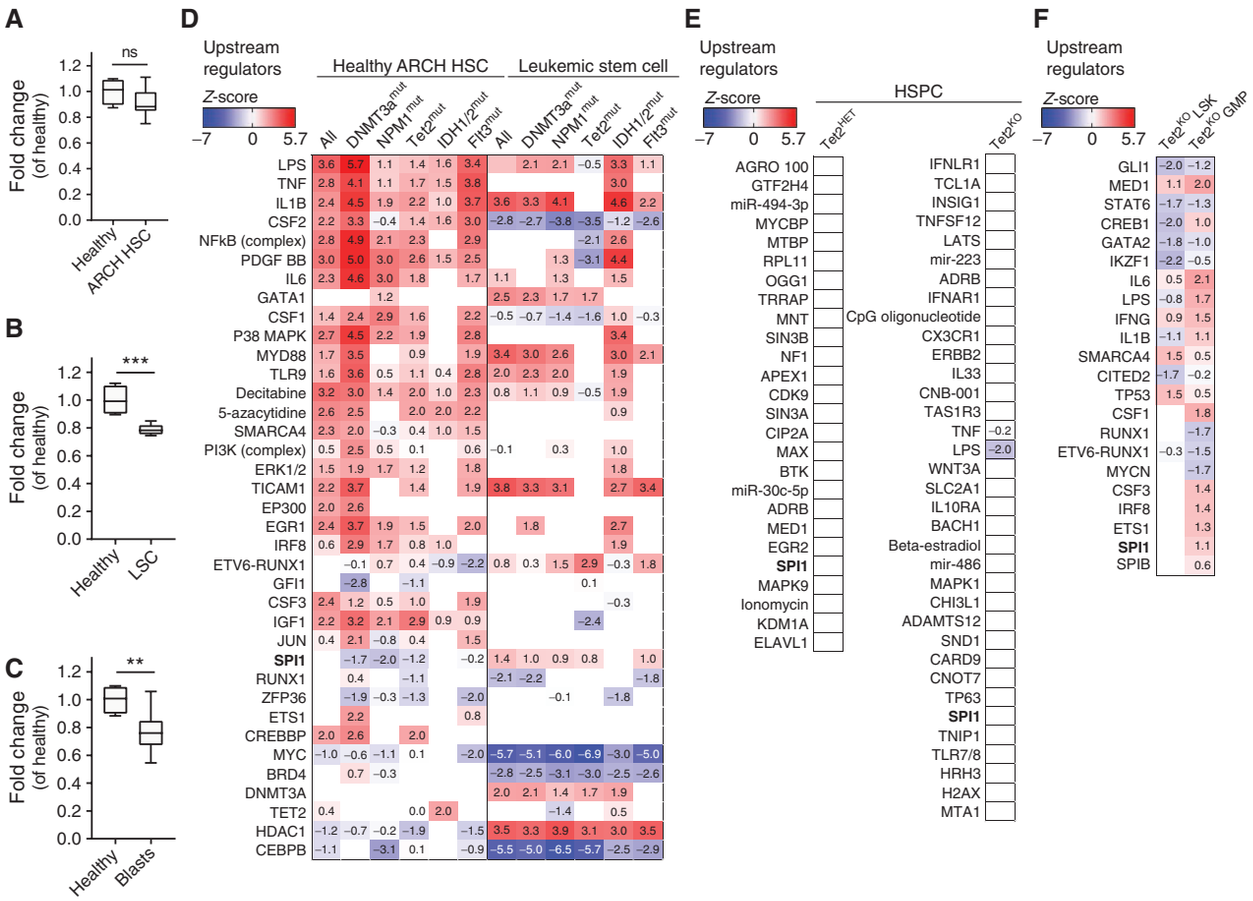


Figure 1. Progressive inhibition of PU.1 during leukemic evolution. **A,** PU.1 mRNA expression levels in preleukemic HSC relative to healthy MPP, expressed as fold change. HSC ($N = 12$ samples) and MPP ($N = 4$ samples) are human-derived samples. Plotted are fold changes, represented as min-to-max box and whiskers graph and line at the median. Fold change was calculated using counts from the GSE74246 RNA-seq data set. Significance was assessed using an unpaired Student *t* test, indicated as *ns* (not significant). **B,** PU.1 mRNA expression levels in LSCs relative to healthy GMPs, expressed as fold change. LSC ($N = 8$ samples) and GMP ($N = 4$ samples) are human-derived samples. Plotted is fold change, represented as min-to-max box and whiskers graph and line at the median. Fold change was calculated using counts from the GSE74246 RNA-sequencing data set. Significance was assessed using an unpaired Student *t* test, indicated as ***, $P \leq 0.001$. **C,** PU.1 mRNA expression levels in the bulk blast population, relative to healthy monocytes, expressed as fold change. Blast population ($N = 12$ samples) and monocytes ($N = 4$ samples) are human-derived samples. Plotted is fold change, represented as min-to-max box and whiskers graph and line at the median. Fold change was calculated using counts from the GSE74246 RNA-sequencing data set. Significance was assessed using an unpaired Student *t* test, indicated as **, $P \leq 0.01$. **D,** Comparative upstream regulator analysis using DEGs from preleukemic, ARCH HSCs (vs. MPP controls) or LSCs (vs. GMP controls), from GSE74246, that carry clonal hematopoiesis (*Dnmt3a*, *Tet2*, and *IDH1/2*)- or other AML-associated mutations (*NPM1* and *Flt3*). HSC ($N = 12$) and LSC ($N = 8$) are human-derived samples. Plotted is *z* score ($z \geq 2$ denotes activation; $z \leq -2$ denotes inhibition), calculated through the built-in function of the IPA software. **E,** Upstream regulator analysis using DESeq2-determined DEGs from a comparison of mouse *Tet2*-deficient $CD4^-CD8a^-CD19^-B220^-Ter119^-$ $cKit^+$ cells ($Tet2^{HET}$ and $Tet2^{KO}$) vs. WT $CD4^-CD8a^-CD19^-B220^-Ter119^-$ $cKit^+$ cells. Plotted is *z* score ($z \geq 2$ denotes activation; $z \leq -2$ denotes inhibition), calculated through the built-in function of the IPA software. $N = 2$ independent RNA-seq experiments. **F,** Comparative upstream regulator analysis of DEGs between $Tet2^{KO}$ LSK vs. *Tet2* WT LSK (GSE132090) and $Tet2^{KO}$ GMP vs. *Tet2* WT GMP (GSE27816). $Tet2^{KO}$ LSK ($N = 2$), *Tet2* WT LSK ($N = 2$), $Tet2^{KO}$ GMP ($N = 2$), and *Tet2* WT GMP ($N = 2$) are mouse-derived, and DEGs were determined by DESeq2 analysis of published RNA-seq data values (GSE132090 and GSE27816). Plotted is *z* score ($z \geq 2$ denotes activation; $z \leq -2$ denotes inhibition), calculated through the built-in function of the IPA software.

Fig. S1M–S1O) and a concomitant 3-fold ($P < 0.001$) expansion in the spleen (Supplementary Fig. S1M and S1N and S1P), compared with WT controls. These data support that loss of one copy of a key myeloid enhancer assuring sufficient PU.1 gene activation (25) can be compensated for in *Tet2*-deficient LSK and myeloid cells in young animals.

Compound Mutant Mice Develop AML upon Aging

We monitored aging compound mutant mice for signs of abnormal hematopoiesis and found that heterozygous deletion of the URE in combination with Tet2 deficiency had a two-pronged effect: First, we noted a *Tet2* dose-dependent reduction of median survival to 623 ($URE^{HET}Tet2^{HET}$) and 290 days ($URE^{HET}Tet2^{KO}$; Fig. 2A; Supplementary Fig. S2A). Second, we observed the emergence of a highly penetrant AML (Fig. 2B–F), as opposed to URE^{HET}, or $Tet2^{HET}$ and $Tet2^{KO}$ single mutant mice; both of the latter develop a late-onset myeloproliferative phenotype (Supplementary Fig. S2B–S2I; refs. 13, 17). Moreover, mice carrying homozygous URE deletions succumbed to an accelerated fatal myeloid malignancy within 3 to 6 months upon heterozygous *Tet2* deletion (Fig. 2A–D; Supplementary Fig. S2A).

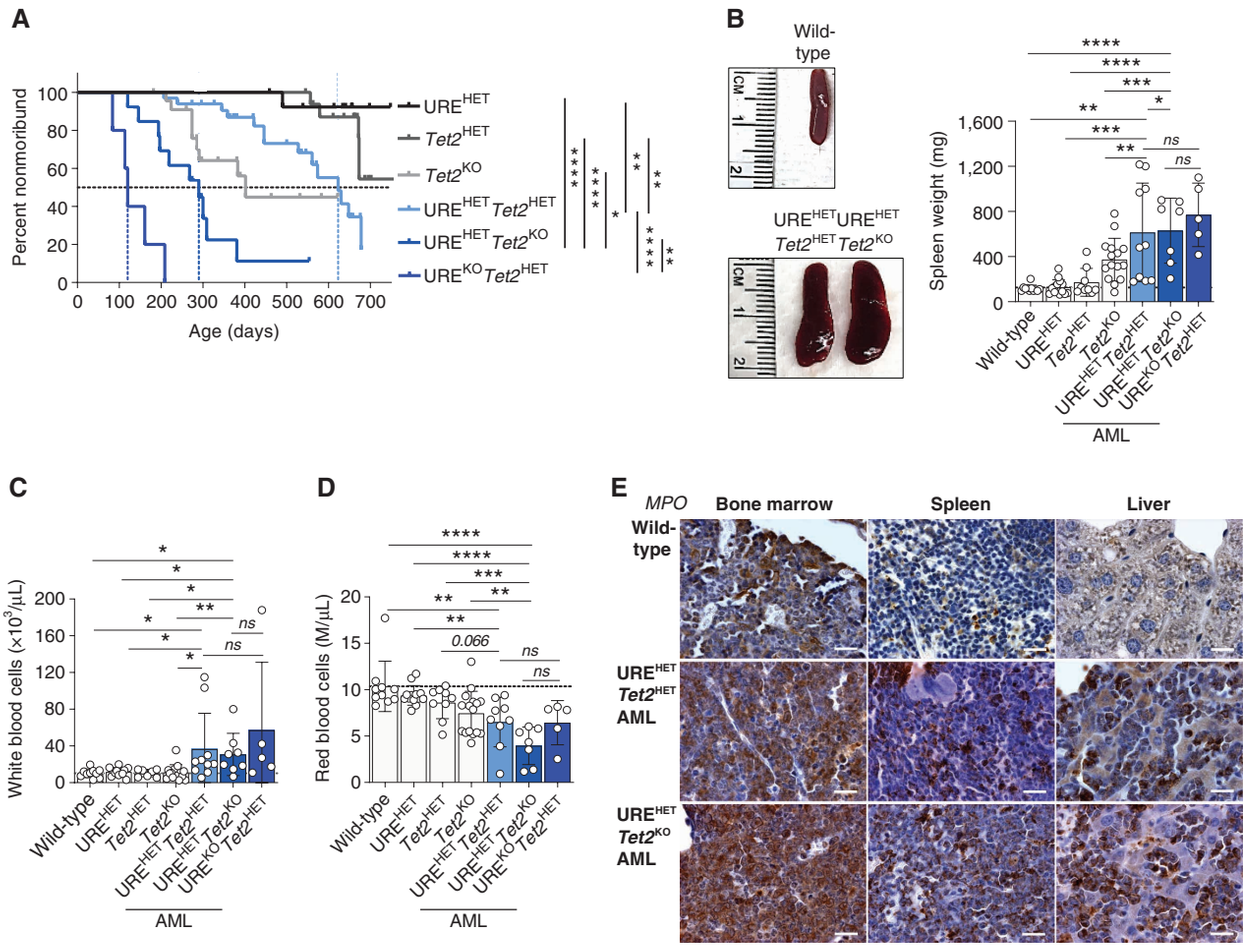
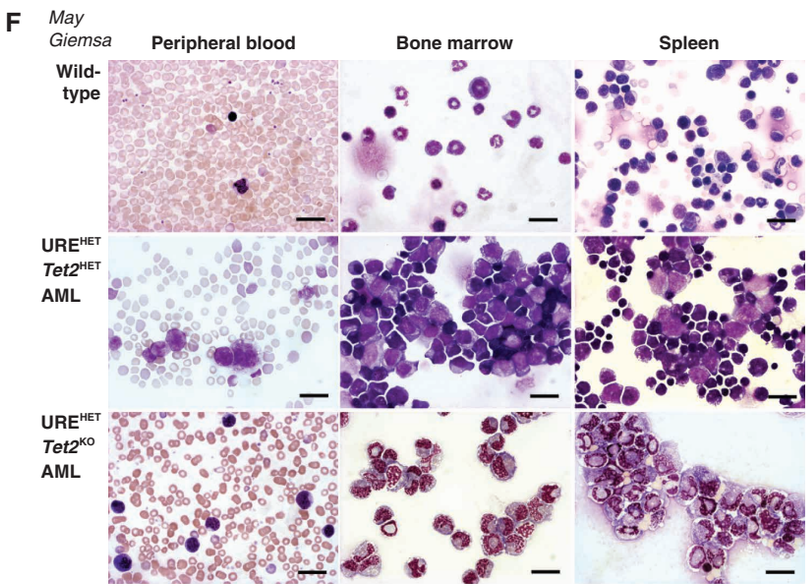
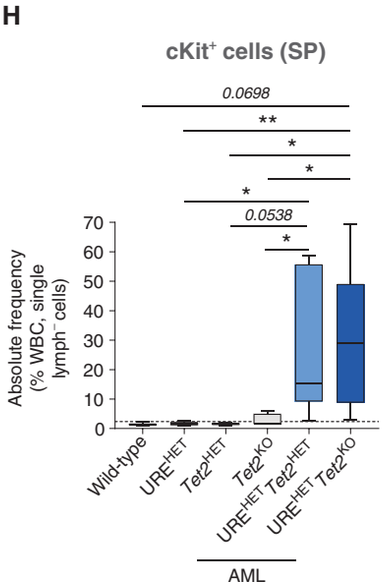


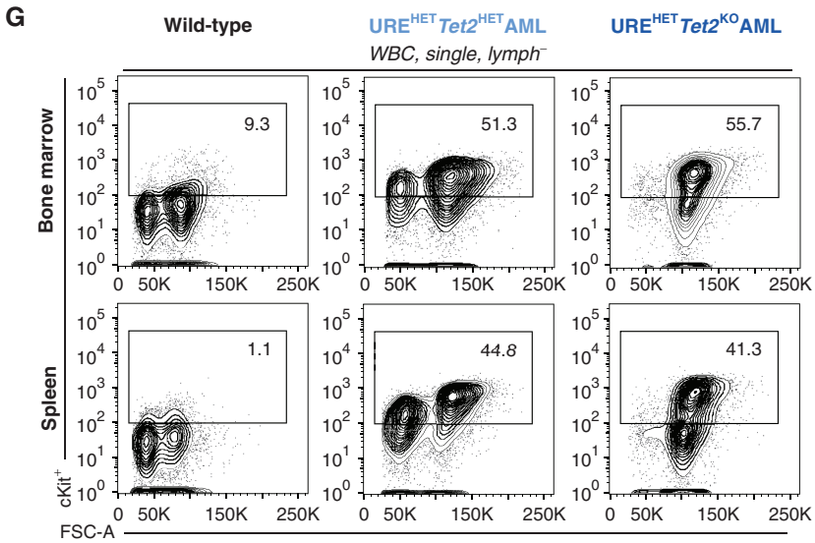
Figure 2. Tet2-deficient mice develop AML upon perturbation of PU.1 gene regulation during aging. **A,** Kaplan–Meier survival analysis of URE^{HET}*Tet2*^{HET}, URE^{HET}*Tet2*^{KO}, and URE^{KO}*Tet2*^{HET} mice (same mice as in Supplementary Fig. S2A), along with single mutant controls. Data are plotted as a percentage of nonmoribund mice at indicated days since their birth. Significance was assessed using the log-rank test (GraphPad Prism). **B,** Pictures of spleens, depicting size of age-matched WT, URE^{HET}*Tet2*^{HET}, and URE^{HET}*Tet2*^{KO} mice (left) and spleen weights of URE^{HET}*Tet2*^{HET} ($N = 10$), URE^{HET}*Tet2*^{KO} ($N = 7$), and URE^{KO}Tet2^{HET} ($N = 5$; same mice as in Supplementary Fig. S2A) and age-matched WT ($N = 11$) and single mutant control mice ($N = 9$–15; right). Individual samples in each group are indicated as circles. Mouse spleens were analyzed in independent experiments. Data represent mean ± SD. Significance was assessed using an unpaired Student *t* test and is indicated as *ns*, not significant. **C** and **D,** WBC ($N = 5$–14) and RBC levels ($N = 5$–16) of leukemic mice and age-matched WT and single mutant control animals. Individual samples in each group are indicated as circles. Mice blood counts were analyzed in independent experiments. Data represent mean ± SD. Significance was assessed using an unpaired Student *t* test and is indicated as *ns*, not significant. **E,** Myeloperoxidase (MPO) stain of bone marrow, spleen, and liver sections from WT and leukemic compound mutant mice. Scale bars indicate 20 μm.
(*continued on following page*)

Specifically, moribund URE^{HET}*Tet2*^{HET} and URE^{HET}*Tet2*^{KO} compound mutant mice presented with splenomegaly (Fig. 2B), accompanied by elevated white blood cell counts (WBC) and reduced red blood cells (RBCs), in comparison with age-matched WT and single-gene mutant (URE^{HET}, *Tet2*^{HET}, and *Tet2*^{KO}) control animals (Fig. 2C and D). Compound mutant mice exhibited an accumulation of differentiation-impaired blast cells in the peripheral blood, bone marrow, and spleen, staining positive for the myeloid marker myeloperoxidase (MPO), and disrupting the tissue architecture of these hematopoietic organs, in contrast to control mice (Fig. 2E and F; Supplementary Fig. S2B–S2D). Compared with controls, compound mutant mice with AML showed an expanded Lymph$^-$ cKit$^+$ cell population (Fig. 2G and H; Supplementary Fig. S2E and S2F; gating strategy in Supplementary Fig. S1M), as well as an increase in cell populations with detectable myeloid marker presentation (Lymph$^-$ Gr1$^+$CD11b$^-$ and Lymph$^-$ Gr1$^+$CD11b$^+$; Fig. 2I–K; Supplementary Figs. S1M and S2G–S2I).

Exome sequencing of sorted cKit$^+$ Lymph$^-$ Ery$^-$ HSPCs and paired tail tissue specimen using the Memorial Sloan Kettering Cancer Center's *Integrated Mutation Profiling of Actionable Cancer Targets* (MSK-IMPACT) mouse panel of 578 cancer-related genes (31) showed evidence for expansion of clonal cell populations in all leukemic compound mutant mice tested (Supplementary Tables S2 and S3). Notably, several of the acquired mutations (i.e., *Cux1*, *Kmt2d*, *Zfhx3*, *Daxx*, *Setd1a*, and *Vav1*) were shared by at least two mice; we also uncovered recurring nonsense insertion mutations in *Cux1* in 4 of 5 mice at subclonal levels (Supplementary Tables S2 and S3). These data provide proof of concept that perturbation of PU.1 mRNA expression regulation can functionally cooperate







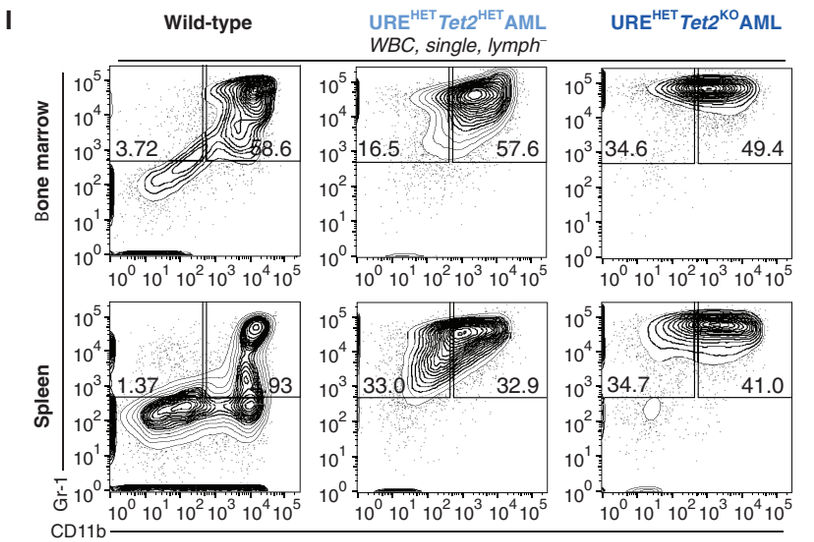


Figure 2. ***(Continued)*** **F,** May-Giemsa stain of peripheral blood, bone marrow, and spleen cell cytospins from WT and compound mutant mice with AML. Scale bars indicate 20 μm. **G,** Representative flow cytometry plots of cKit antigen presentation on bone marrow and spleen cells from leukemic compound mutant mice and age-matched WT controls. Plots are gated at single, $CD4^-CD8a^-B220^-$ ($Lymph^-$) cells, as in Supplementary Fig. S1M. **H,** Graphical representation of the calculated absolute frequency of $Lymph^-$ $cKit^+$ cells in the spleens of compound mutant mice ($N = 7$–8), and age-matched WT ($N = 3$) and single-gene mutant parental ($N = 4$–7) controls. Spleen cell-surface antigen expression was analyzed in independent experiments. Data are plotted as min-to-max box and whiskers graphs, with a line at the median. Significance was assessed using an unpaired Student *t* test and is indicated as *ns*, not significant. **I,** Representative flow cytometry plots of Gr-1 and CD11b antigen presentation on bone marrow and spleen cells from leukemic compound mutant mice and age-matched WT controls. Plots are gated at single $Lymph^-$ cells, as in Supplementary Fig. S1M. (*continued on next page*)

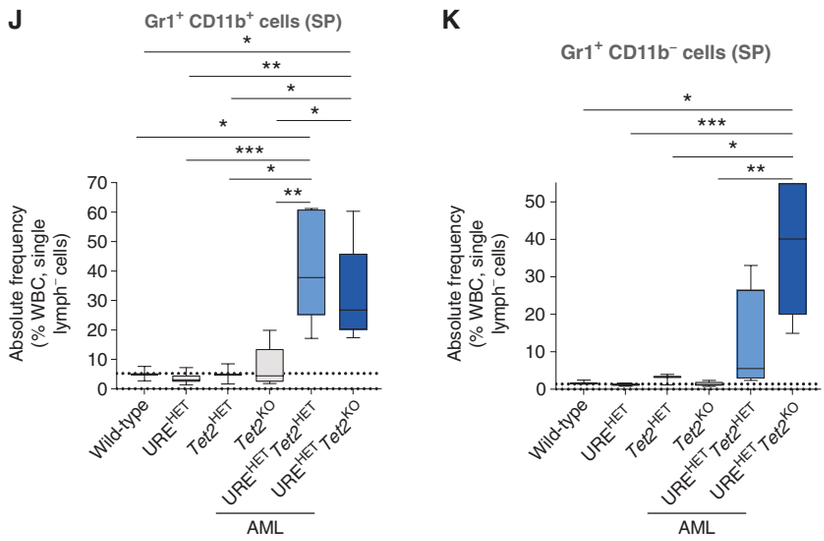


Figure 2. ***(Continued)*** **J** and **K,** Graphical representation of the calculated absolute frequency of Lymph⁻ Gr-1⁺CD11b⁺ and Lymph⁻ Gr-1⁺CD11b⁻ cells in the spleens of compound mutant mice ($N = 4$–6) and age-matched WT ($N = 3$) and single gene-mutant parental ($N = 3$–6) controls. Spleen cell-surface antigen expression was analyzed in independent experiments. Data are plotted as min-to-max box and whiskers graphs, with a line at the median. Significance was assessed using an unpaired Student *t* test and is indicated as *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$.

with heterozygous or homozygous inactivation of *Tet2* in the initiation of myeloid leukemia during aging.

Defective HSPC Self-Renewal and Differentiation in Leukemic Compound Mutant Mice

We next characterized cell-intrinsic functional aberrations of URE^{HET}*Tet2*^{HET} and URE^{HET}*Tet2*^{KO} bone marrow cells. Compared with single-gene mutant cells, bone marrow cells from compound mutant mice harbored a comparable number of colony-initiating progenitors in primary colony-forming assays, albeit lacking multipotent and erythroid potential (Fig. 3A; Supplementary Fig. S3A). Although only a minor fraction of the compound mutant cells gave rise to colonies lacking morphologic features of myeloid cell maturation in the first plating, these blastlike colonies gradually became the dominant type upon serial replating (Fig. 3A and B; Supplementary Fig. S3B). URE^{HET}*Tet2*^{KO} cells showed greater clonogenic capacity from the second, 88%, 142%, and 190% more than the URE^{HET}, *Tet2*^{HET}, and *Tet2*^{KO} controls, respectively, to the eighth plating. They continued to give rise to aberrant blastlike colonies after the URE^{HET} and *Tet2*^{HET} colony-initiating cells were exhausted. URE^{HET}*Tet2*^{HET} cells started producing higher numbers of blastlike colonies (232%, 89%, and 58% more than the URE^{HET}, *Tet2*^{HET}, and *Tet2*^{KO} controls, respectively) with the fourth plating (Fig. 3A and B; Supplementary Fig. S3B). Compound mutant cells also exhibited greater serial colony-forming activity than *Tet2*^{KO} cells, which showed extended serial replating activity, in line with past reports (17, 32). Compound mutant cell-initiated colonies consisted of cells with blast morphology, largely devoid of morphologic signs of terminal myeloid cell maturation (Fig. 3C). Cells isolated from blastlike colonies were of a myeloid progenitor-like immunophenotype, as the primary AML bulk cell population found in the moribund mice (Fig. 3D and E; Supplementary Fig. S3C–S3H).

Next, we tested whether bone marrow cells from leukemic compound mutant mice harbored disease-initiating potential. Donor cells from control URE^{HET} BM showed low engraftment (1.5%–2.3%), whereas *Tet2*-deficient cells exhibited slightly higher donor cell chimerism (9.45%–53.5% and 6.52%–98.3% for *Tet2*^{HET} and *Tet2*^{KO} cells, respectively; Supplementary Fig. S4A and S4B), consistent with previous reports (17). In contrast, transplantation of total bone marrow (TBM) cells and purified HSPC-enriched cKit⁺ cells from compound mutant donors was followed by moribund states of the recipients within 4 to 13.5 weeks (TBM: 4–13.5 weeks, cKit⁺ cells: 4–9 weeks; Fig. 3F; Supplementary Fig. S4C). Compound mutant cells, however, robustly expanded in the bone marrow of recipient mice [41.9%–92.8% (TBM) and 35.2%–77.9% (cKit⁺ cells) and 23.95%–97.9% (TBM) and 13.04%–97.71% (cKit⁺ cells) for URE^{HET}*Tet2*^{HET} and URE^{HET}*Tet2*^{KO} cell-transplanted mice, respectively; Supplementary Fig. S4D and S4E; gating strategy (1) in Supplementary Fig. S4A]. Recipients of compound mutant cells also develop splenomegaly (Fig. 3G and H), and analysis of bone marrow and spleen of sacrificed compound mutant cell recipients showed a dominant donor cell population with a striking similarity to the primary tumors [Fig. 3I–L; gating strategy (2) in Supplementary Fig. S4A and S4F]. These data demonstrate that compound mutant phenotypic HSPCs harbor cell-autonomous defects resulting in enhanced self-renewal, impaired myeloid differentiation, and leukemia-initiating capacity.

Reduced PU.1 mRNA Expression Is Compatible with Nonleukemic Hematopoiesis during Aging

Compound mutant mice presented with blast cell phenotypes of varying morphology regardless of the underlying genotype which was reminiscent of human AML, and distinct from the previously reported URE^{KO} myeloid malignancy which presents with early-onset AML homogenously

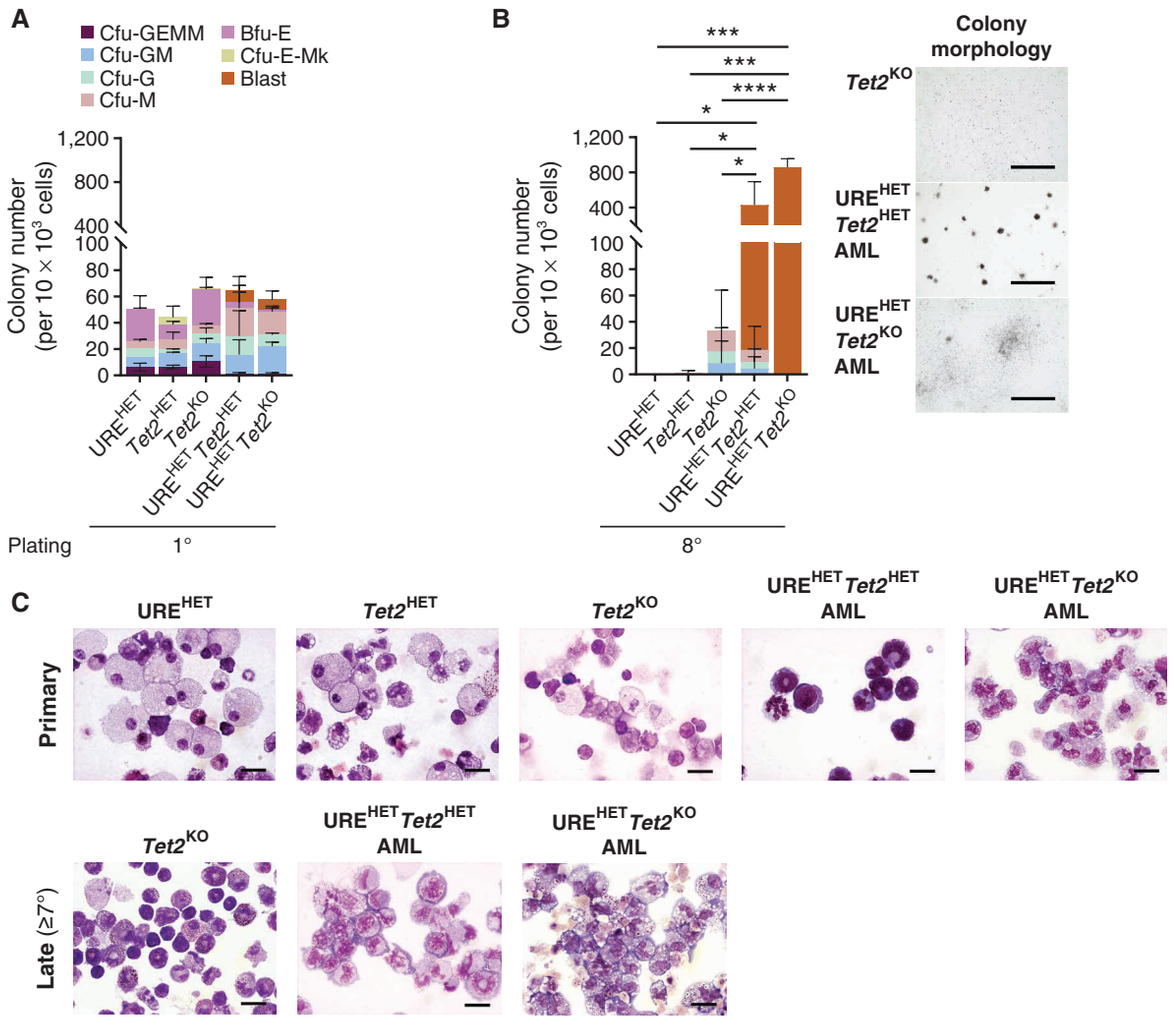


Figure 3. Cell-intrinsic defects in self-renewal and differentiation are hallmarks of compound mutant leukemia. **A–E,** Serial replating colony-forming assay (eight platings: 1°–8°) of TBM cells from URE^{HET}*Tet2*^{HET} and URE^{HET}*Tet2*^{KO} mice with AML or age-matched parental controls ($N = 3$–4 per genotype, $N = 4$ independent experiments). **A** and **B,** Number of colony-forming units (cfu)-granulocyte/erythrocyte/monocyte/megakaryocyte (Cfu-GEMM), cfu-granulocyte/monocyte (Cfu-GM), (Cfu)-granulocyte (Cfu-G), cfu-monocyte (Cfu-M), burst-forming units-erythroid (Bfu-E), cfu-erythrocyte/megakaryocyte (Cfu-E/Mk), and of colonies containing undifferentiated blastlike cells (Blast), at plating 1 (**A**) and 8 (**B**, left). Data represent mean ± SD. Significance was assessed using an unpaired Student *t* test. Colony morphology of cells after the eighth round of plating (**B**, right). Scale bars are 1,000 μm. **C,** May-Giemsa stain of leukemic compound mutant and parental control cells, at primary and late (>7th) plating. Stain was performed in independent experiments. Scale bars are 20 μm. (*continued on next page*)

hallmarked by the expansion of myeloblast-like cells (25). Approximately half (42.9% for URE^{HET}*Tet2*^{HET} and 50% for URE^{HET}*Tet2*^{KO}) of the mice harbored leukemic blasts hallmarked by immature cell morphology and a dominant cKit-expressing cell population with high Gr-1 antigen presentation in the bone marrow and spleen (classified further as *immature AML*); the remainder presented with blast cells showing some degree of myeloid differentiation (metamyelocyte-like cell morphology, and lower cKit, but high CD11b and Gr-1 abundance on their cell surface; referred further to as *mature AML*; Fig. 4A and B; gating strategy in Supplementary Fig. S1M; Supplementary Fig. S5A and S5B).

RNA-sequencing of HSPCs from leukemic mice (Fig. 4C and D) mirrored this morphologic heterogeneity; compared with phenotypical control cells from control animals, principal component analysis (PCA) showed a similar heterogeneity of the compound mutant cells, with HSPCs from mice with a mature AML phenotype more closely related with WT control-derived cells of the same immunophenotype (Fig. 4E).

PU.1 mRNA and protein abundance tracked with the myeloid differentiation stage of the AML blasts. Compared with single mutant controls, only some compound mutant mice showed reduced PU.1 mRNA (Fig. 4F). Although some of the leukemic compound mutant mice showed PU.1 mRNA expression similar to parental control mice (PU.1^{High}), others exhibited a decrease in expression of the transcription factor (PU.1^{Low}) in comparison with control-derived cells of the same immunophenotype (Supplementary Fig. S5C). At the protein level, we found that all compound mutant mice with AML had detectable PU.1 in cKit^{+} cells (Fig. 4G; Supplementary Fig. S5D). Moreover, we found that mice presenting with AML blasts harboring signs of maturation toward

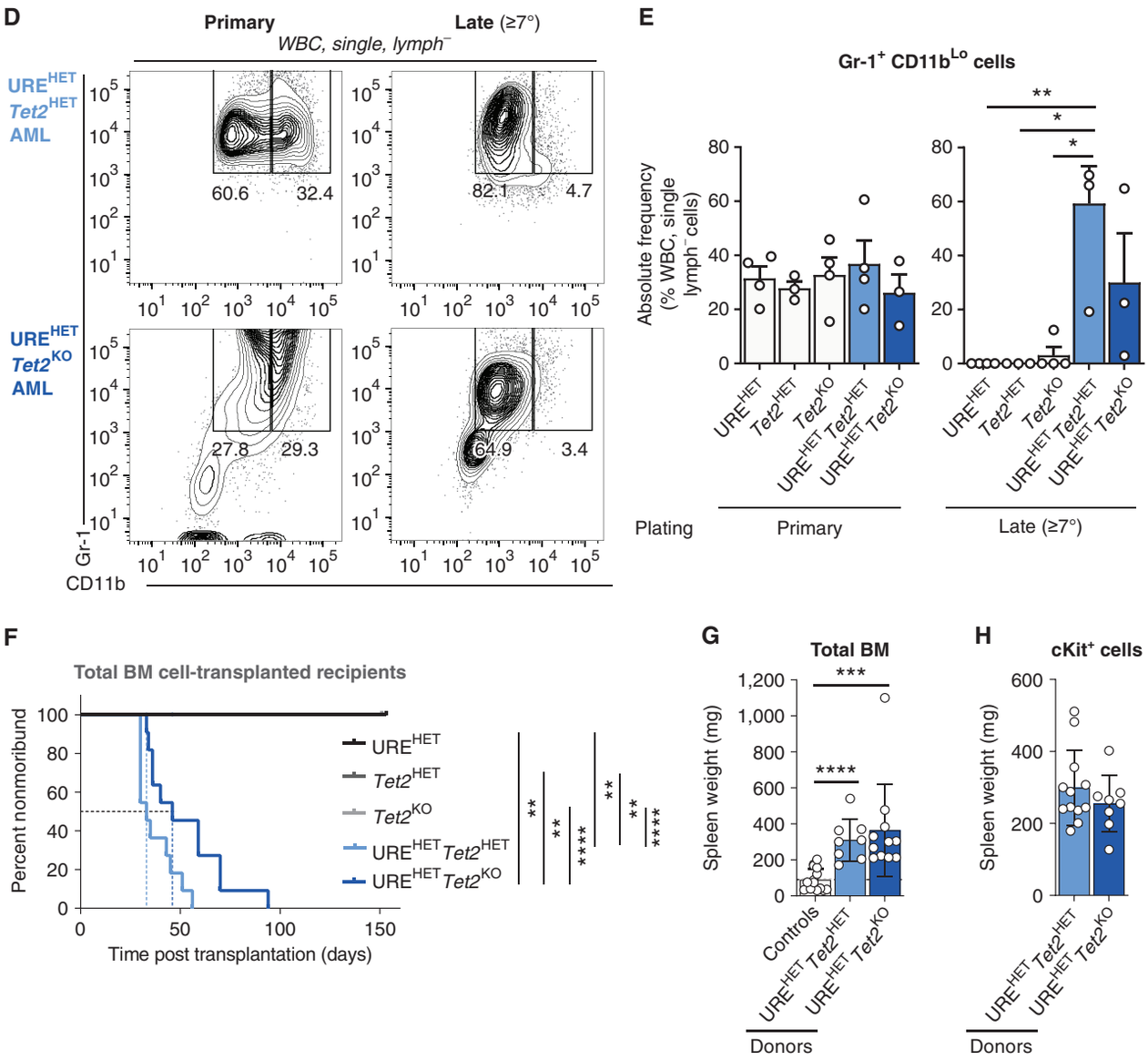


Figure 3. ***(Continued)*** **D,** Representative flow cytometry plots of Gr-1 and CD11b antigen presentation on colony assay-derived cells (primary and late (>7th) plating) from leukemic compound mutant mice, gated at single, $CD4^-CD8a^-CD19^-B220^-$ ($Lymph^-$) cells, as in Supplementary Fig. S3C (stain 1). **E,** Graphical representation of the calculated absolute frequency of $Lymph^-$ $Gr\text{-}1^+CD11b^{Lo}$ cells from colony-assay–derived cells, at primary and late (>7th) platings. Data represent mean ± SEM. Significance was assessed using ordinary one-way ANOVA. **F–L,** Bone marrow transplantation assay of total nucleated bone marrow cells (total BM) or $cKit^+$ leukemic or total BM control cells, into sublethally irradiated NSG mice ($N = 2$ donors per genotype; $N = 2$–12 NSG recipients per genotype in five independent experiments). **F,** Kaplan–Meier survival analysis of total BM-transplanted recipient mice. Data are plotted as a percentage of nonmoribund mice at indicated days, post bone marrow transplantation. Significance was assessed using the log-rank test. **G,** Spleen weights of transplanted mice with control or leukemic mice-derived total BM cells. Data represent mean ± SD. Significance was assessed using an unpaired Student *t* test. **H,** Spleen weights of transplanted mice with leukemic mice-derived $cKit^+$ BM cells. Data represent mean ± SD. (*continued on following page*)

the neutrophil lineage appeared to have lower PU.1 levels than $cKit^+$ cells from immature blast AML (Supplementary Fig. S5D), which is consistent with lower PU.1 protein abundance resulting in enhanced granulocytic lineage differentiation (33). This strongly suggests that the observed alterations in PU.1 mRNA and protein levels in compound mutant mice are to a certain extent reflective of the underlying cellular differentiation state (also observed at the global gene-expression level; Supplementary Tables S4 and S5), and not a sole feature of their malignant state. In support, we detected a reduction in PU.1 mRNA expression levels by 54% ± 5% and 64 ± 5% in aged $URE^{HET}Tet2^{HET}$ and $URE^{HET}Tet2^{KO}$ mice, respectively, which had not developed AML by age 11 to 22 months (*Healthy*), compared with young (3–5 months) mice of the same genotype (*Young*; Fig. 4H and I). Although in most leukemic $URE^{HET}Tet2^{HET}$ mice, LSK cells maintained PU.1 mRNA expression at the level found in old mice (*Healthy*; 75 ± 9% reduction of *Young*; Fig. 4H), we detected further reduced mRNA levels in some leukemic $URE^{HET}Tet2^{KO}$ LSK cells (*AML*; 84% ± 2% reduction, of *Young*; Fig. 4I), which

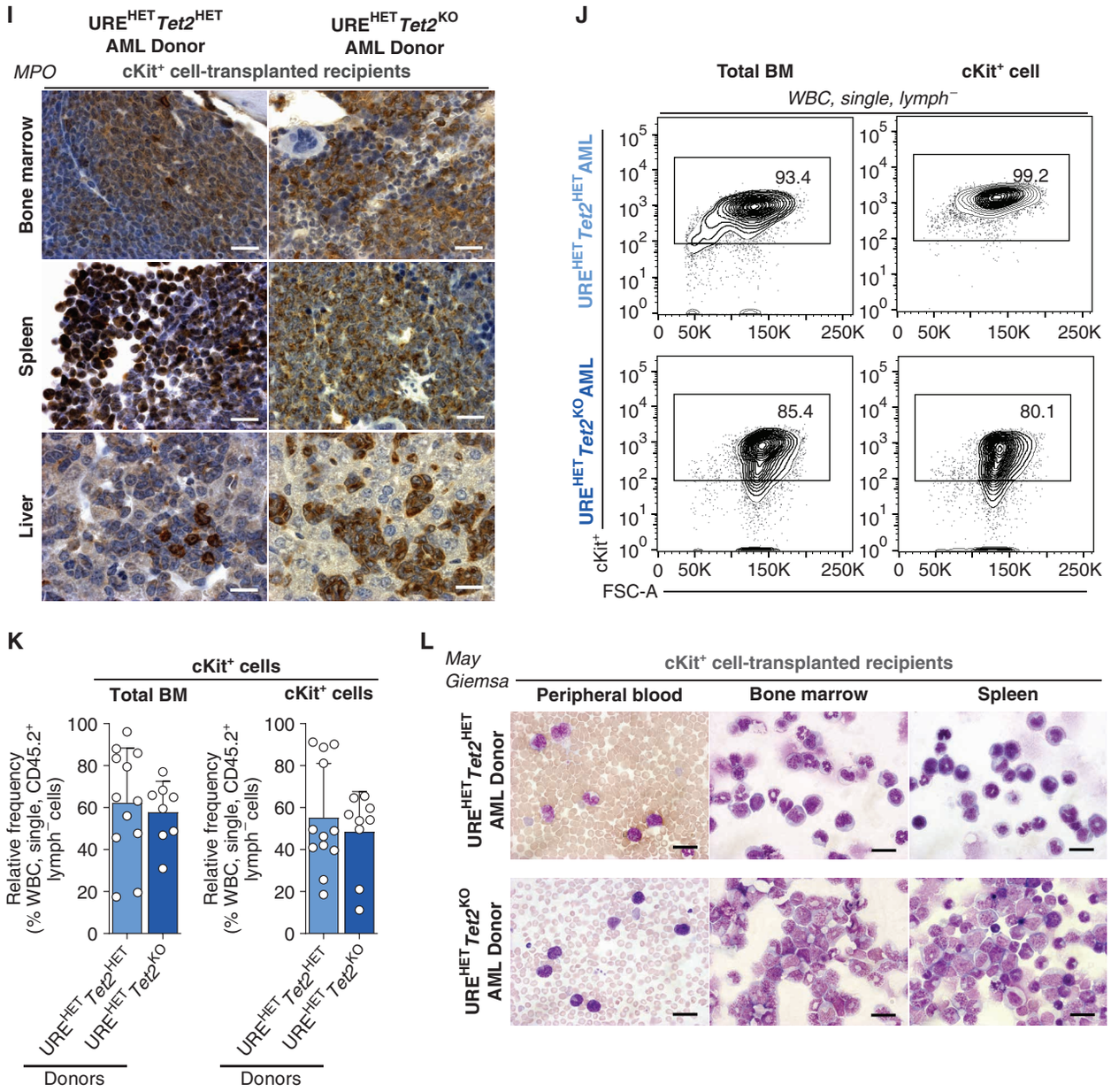


Figure 3. ***(Continued)*** **I,** MPO stain of bone marrow, spleen, and liver sections of leukemic compound mutant mice cKit+ cell-transplanted mice. Scale bars are 20 μm. **J,** Flow cytometry plots of cKit antigen presentation on bone marrow cells of recipients of total BM- and cKit+ cells, gated at single, CD45.2+ CD45.1− Lymph− cells, as in Supplementary Fig. S4A. **K,** Graphical representation of the relative frequency of CD45.2+ CD45.1− Lymph− cKit+ cells in the bone marrow of total BM- and cKit+ cell-transplanted recipients. Data represent mean ± SD. **L,** Cytospins of peripheral blood, bone marrow, and spleens of compound mutant cKit+ AML cell recipients. Scale bars are 20 μm. *, $P \le 0.05$; **, $P \le 0.01$; ***, $P \le 0.001$; ****, $P \le 0.0001$.

was inconsistent with PU.1 protein abundance, found not markedly changed in healthy, aged compound mutant mice (Supplementary Fig. S5E). Together, this strongly suggests that reduced PU.1 mRNA expression can occur in aging compound mutant mice in the absence of alterations at the protein level, and is compatible with a nonleukemic state.

Loss of PU.1 Network Activity in AML

To better understand the molecular underpinnings of compound mutant AML, we performed differential gene-expression analysis, which identified a total of 442 and 1,564 DEGs in URE^{HET}*Tet2*^{HET} and URE^{HET}*Tet2*^{KO} HSPCs, respectively compared with WT controls (Supplementary Fig. S5F and S5G; sort gating strategy as for Fig. 4C and D). We found the majority of DEGs to be downregulated in URE^{HET}*Tet2*^{HET} (263 transcripts) and URE^{HET}*Tet2*^{KO} (949 transcripts) HSPCs, as well as enriched for targets harboring a core PU.1 DNA-binding motif (Supplementary Fig. S5H and S5I). Intersection with published chromatin immunoprecipitation (ChIP) data for PU.1 and Tet2 further showed that a third of the downregulated genes are direct targets of PU.1 alone or along with Tet2 (Supplementary Fig. S5J; Supplementary Tables S6 and S7), which can physically

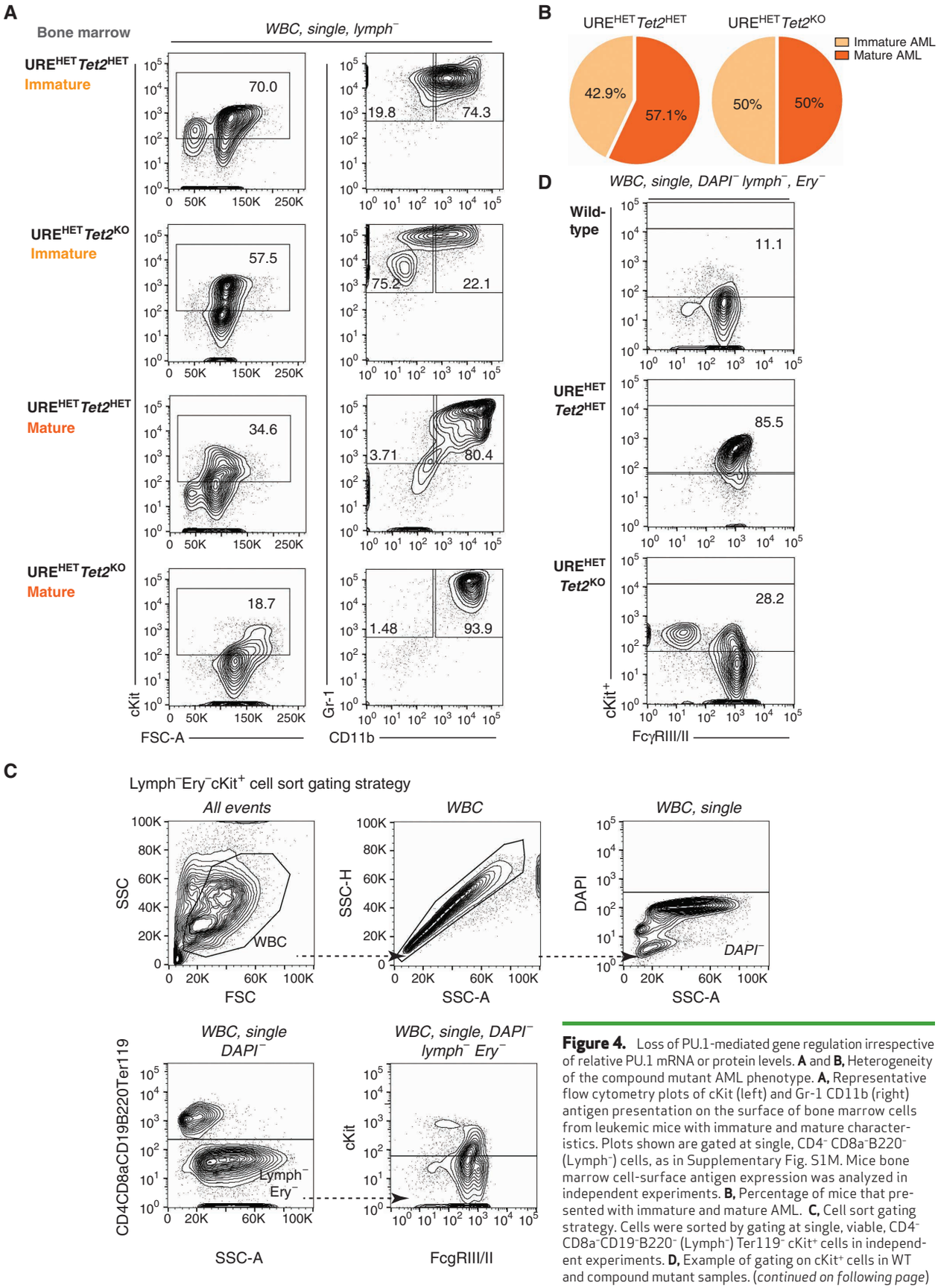


Figure 4. Loss of PU.1-mediated gene regulation irrespective of relative PU.1 mRNA or protein levels. **A** and **B,** Heterogeneity of the compound mutant AML phenotype. **A,** Representative flow cytometry plots of cKit (left) and Gr-1 CD11b (right) antigen presentation on the surface of bone marrow cells from leukemic mice with immature and mature characteristics. Plots shown are gated at single, $CD4^-$ $CD8a^-B220^-$ ($Lymph^-$) cells, as in Supplementary Fig. S1M. Mice bone marrow cell-surface antigen expression was analyzed in independent experiments. **B,** Percentage of mice that presented with immature and mature AML. **C,** Cell sort gating strategy. Cells were sorted by gating at single, viable, $CD4^-$ $CD8a^-CD19^-B220^-$ ($Lymph^-$) $Ter119^-$ $cKit^+$ cells in independent experiments. **D,** Example of gating on $cKit^+$ cells in WT and compound mutant samples. (*continued on following page*)

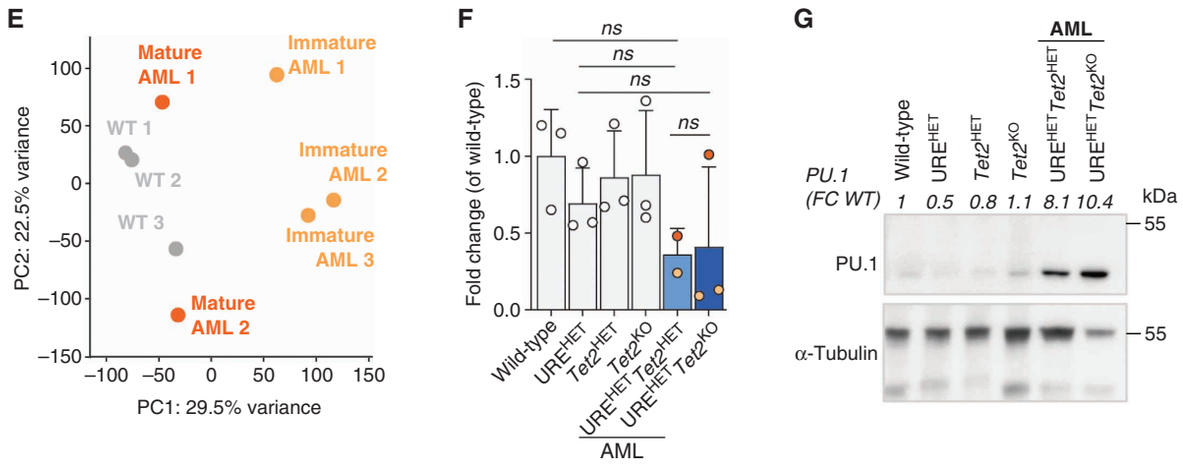


Figure 4. ***(Continued)*** **E,** PCA of gene expression of leukemic and age-matched WT mice. $N = 2$ in independent experiments. **F,** PU.1 mRNA expression levels in leukemic and parental control animals, dichotomizing mice in PU.1^{Hi} and PU.1^{Low}-expressing. $N = 2$–3 per genotype. Plotted are RNA-seq transcripts per million expressed as fold change of WT. Significance was assessed using unpaired Student *t* test and is indicated as *ns*, not significant. Data represent mean ± SD. **G,** Quantification of PU.1 protein abundance in bead-sorted cKit$^+$ cells from moribund URE^{HET}*Tet2*^{HET}, URE^{HET}*Tet2*^{KO} mice with AML or age-matched WT and single mutant control mice using Western blot analysis. Values on top of blots indicate image-assisted protein quantification, as tubulin-normalized fold changes compared with wild-type controls (FC WT). (*continued on next page*)

interact with the myeloid master transcription factor (TF; Supplementary Fig. S5K and S5L; ref. 34). Consistent with the cell morphologic and functional phenotypes, these network perturbations affected pathways involved in myeloid cell function, macrophage-mediated inflammatory response, innate immune system function, cell fate commitment and maturation, and leukemia (Fig. 4J; Supplementary Fig. S5M). Network analysis showed PU.1 and PU.1-cofactor–associated gene-expression programs to be compromised (i.e., *Jun*, *Csf1*, *Irf8*, and *Runx1*), which occurred largely irrespective of the underlying genotype or PU.1 mRNA expression level (Fig. 4K). These data implicate the loss of PU.1 network activity, rather than declining PU.1 levels alone, as a unifying mechanism underlying myeloid malignancy in compound mutant mice. We next set out to pinpoint which PU.1-dependent gene networks may confer AML-driving properties.

Inhibition of Myeloid Enhancers in Leukemic HSPC

As a key regulator of hematopoietic cell specification and differentiation, PU.1 establishes gene-expression programs, particularly through binding and engagement of cell-type–specific enhancers (35). In line, improved transposase-accessible chromatin by sequencing (ATAC-seq; ref. 36) of Lymph$^-$ Ery$^-$ cKit$^+$ populations (sorted as in Fig. 4C and D) uncovered alterations in chromatin accessibility in compound mutant cells compared with nonleukemic (combined WT, URE^{HET}, *Tet2*^{HET}, and *Tet2*^{KO}) cells (Fig. 5A).

To gain insight into the functional impact of the observed alterations in chromatin states of compound mutant HSPC, we annotated functional regulatory elements by overlapping our identified differentially accessible loci with *cis*-regulatory elements (cCREs) from the Integrative and Discriminative Epigenome Annotation System (IDEAS) database from the Validated Systematic IntegratiON (VISION) project (37). Targeted assessment of DNA accessibility at monocyte and neutrophil-specific cCREs that correspond to enhancers, showed a highly significant reduction in chromatin accessibility (P values: 3.52×10^{-53} for URE^{HET}*Tet2*^{HET} and 8.84×10^{-26} for URE^{HET}*Tet2*^{KO}) in leukemic versus WT HPSCs, whereas single mutant HSPC showed no detectable differences (Fig. 5B and C; Supplementary Fig. S6A–S6D). Loss of accessibility was also observed at monocyte- and neutrophil-specific promoters (P values: 2.71×10^{-52} for URE^{HET}*Tet2*^{HET} and 1.76×10^{-16} for URE^{HET}*Tet2*^{KO}; Supplementary Fig. S6E and S6F). Notably, combined B- and T-cell–specific enhancers (P values: 2.25×10^{-43} for URE^{HET}*Tet2*^{HET} and 2.1×10^{-24} for URE^{HET}*Tet2*^{KO}) and promoters (P values: 4.93×10^{-54} for URE^{HET}*Tet2*^{HET} and 1.89×10^{-18} for URE^{HET}*Tet2*^{KO}) also showed reduced accessibility (Supplementary Fig. S6G–S6J), in line with previous work demonstrating that PU.1-dependent enhancer regulation is also important for early stages of T-cell differentiation (38) and for B-cell maturation (39). In contrast, we were unable to observe changes in DNA accessibility at a randomly sampled set of 5,000 enhancers in compound mutant cells (vs. WT; Supplementary Fig. S6K and S6L). These findings indicate wide-ranging inhibition of hematopoietic regulatory regions in compound mutant HSPCs.

Using defined cell type–specific epigenetic states (ref. 37; Fig. 5D), we annotated the chromatin accessibility information of leukemic HSPCs and found an under-representation of nuclease-accessible, active enhancer states (states *E*, *N* and *E*, *N*, *A*), important for LSK, and immature (CMP and GMP) and mature (CFU-M, neutrophils, and monocytes) myeloid cells (Fig. 5E and F). This perturbation appeared independent of residual *Tet2* dosage. In support, we uncovered wide-ranging alterations in the abundance of enhancer RNAs, a hallmark of active enhancers (40), at sites coinciding with cCRE. We found a slightly higher number of downregulated (1,629) than upregulated (1,144) enhancer-originating RNA transcript (eRNA; Fig. 5G), demonstrating that there is a reduction in the activity of certain enhancers. Hierarchical clustering of all differential chromatin peaks that correspond to differentially expressed eRNAs (cCRE) further showed that nonleukemic mice shared most of the regulatory element patterns defining the epigenome of immature cells

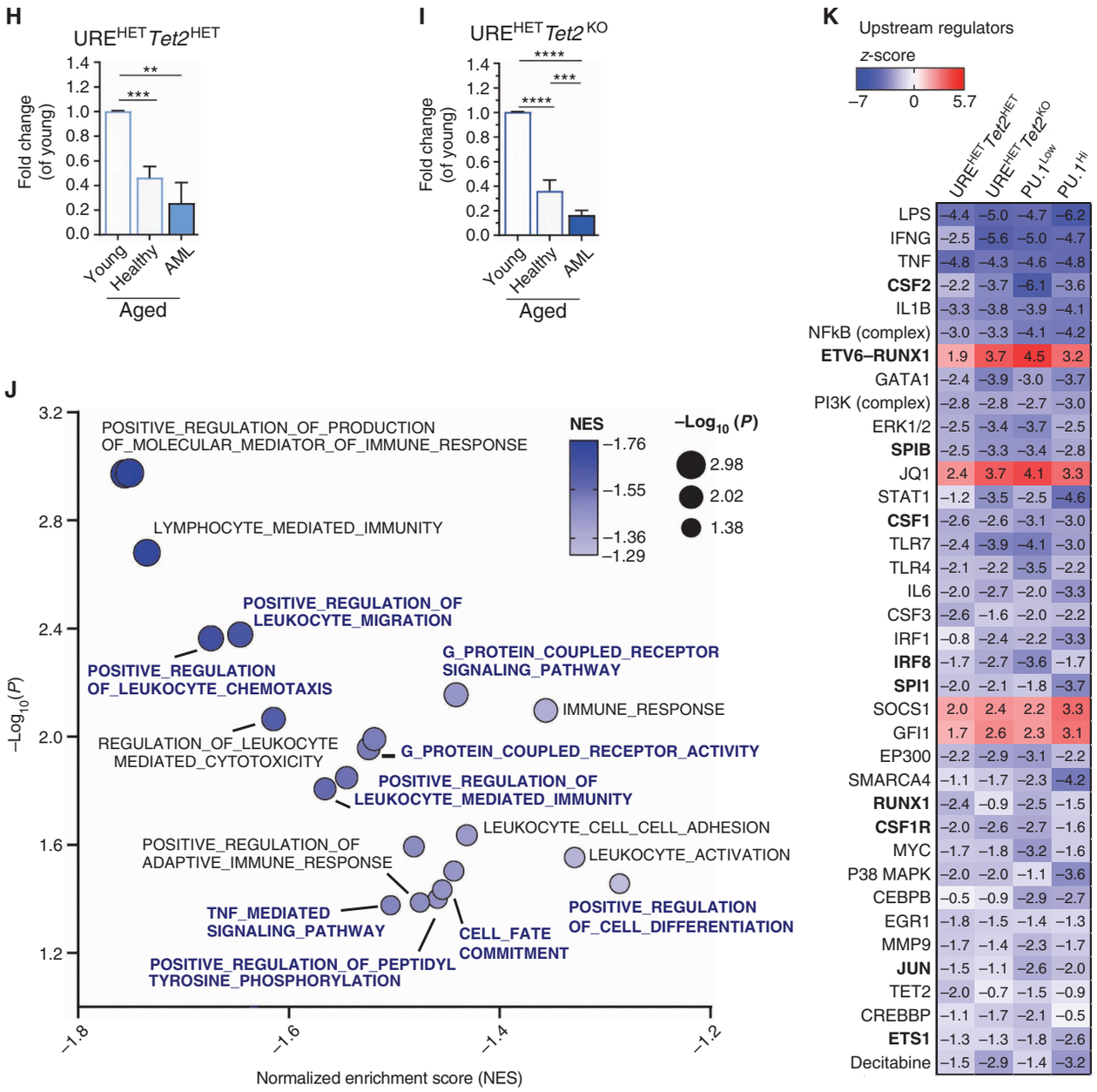


Figure 4. ***(Continued)*** **H,** qRT-PCR analysis of PU.1 mRNA expression levels in FACS-purified LSK cells of young, aged, and leukemic URE^{HET}*Tet2*^{HET} mice. Data are expressed as fold change to young compound mutant mice. $N = 2$–3 mice per age group and per genotype. $N = 2$ in independent qPCR experiments. Significance was assessed using an unpaired Student *t* test. Data represent mean ± SD. **I,** qRT-PCR analysis of PU.1 mRNA expression levels in FACS-purified LSK cells of young, aged, and leukemic URE^{HET}*Tet2*^{KO} mice. Data are expressed as fold change to young compound mutant mice. $N = 2$–3 mice per age group and per genotype. $N = 2$ independent qPCR experiments. Significance was assessed using an unpaired Student *t* test. Data represent mean ± SD. **J,** GSEA of DEGs in leukemic animals. DEGs were determined by DESeq2. Significance was assessed using the built-in function of GSEA. **K,** Comparative pathway analysis (IPA) of DEGs in compound mutant mice (vs. WT), showing similarities in upstream regulators, among the leukemic genotypes and PU.1^{Hi} and PU.1^{Low} mice. Significance was assessed using the built-in function of IPA and is indicated as *z* score. *z* score ≥ 2 denotes activation; z score ≤ −2 denotes inhibition. **, $P \le 0.01$; ***, $P \le 0.001$; ****, $P \le 0.0001$.

(HPC-7 and LSK cells) and progenitors and mature cells of most hematopoietic lineages (megakaryocytic, lymphoid, and myeloid), with only ~1/3 overlap with erythroid and mature CD8a$^+$, CD4$^+$ T cells and natural killer cells (Supplementary Fig. S6M). In contrast, compound mutant mice with AML shared many of the epigenetic state-defining regulatory element patterns with myeloid progenitors (GMP and CMP), CFU-Mk, and less with immature LSK cells, Mk, and mature monocytes, with very little representation of lymphoid and erythroid progenitor and differentiated cell states (Supplementary Fig. S6N), indicating a profound impairment of myeloid enhancers in leukemic HSPCs. Gene set enrichment analysis (GSEA) further revealed loss of gene-expression programs governing *chromatin assembly and disassembly*

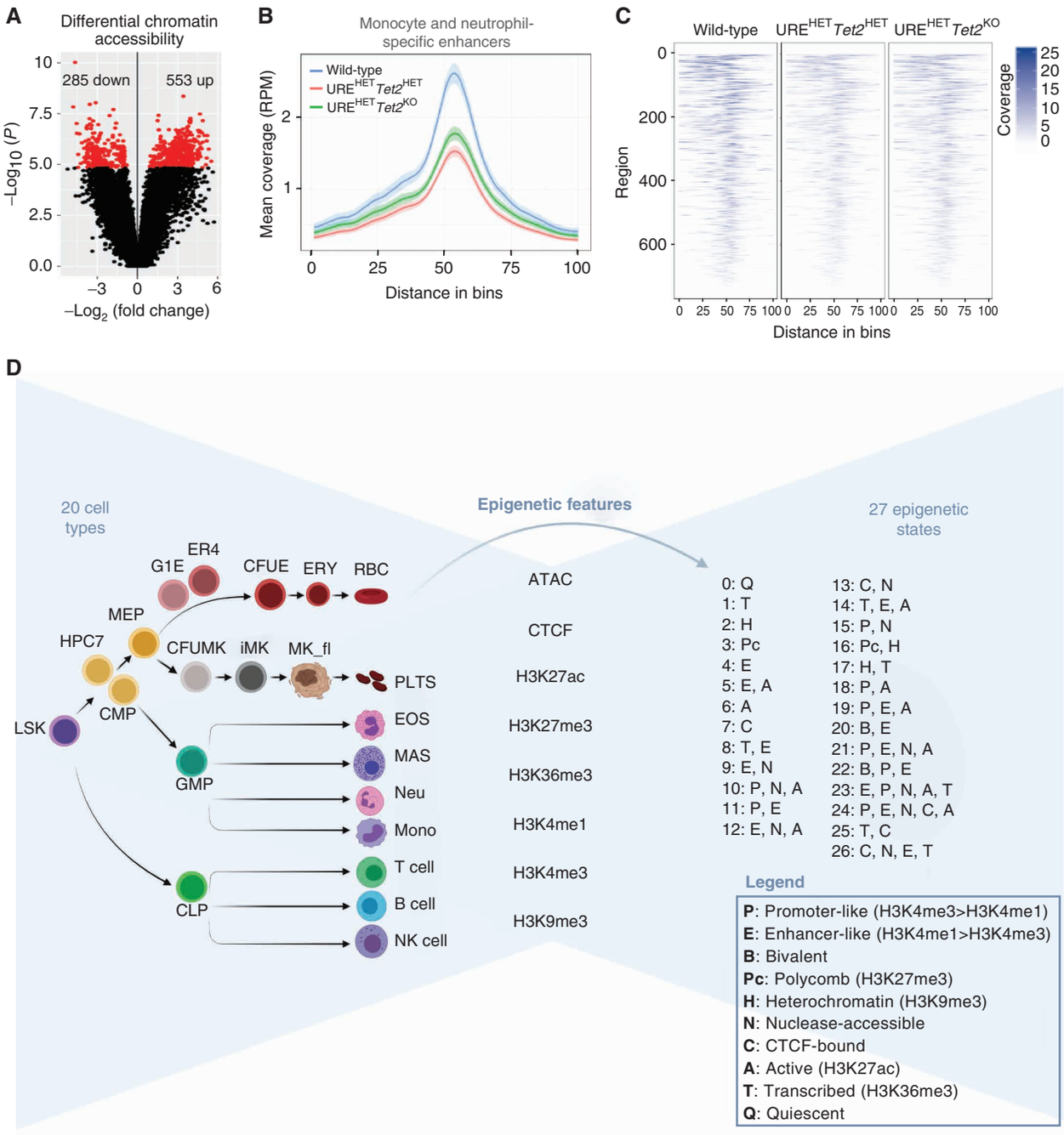


Figure 5. Loss of PU.1-mediated enhancer activity governing proper myeloid differentiation in compound mutant mice with leukemia. **A,** Volcano plot showing differential chromatin status in leukemic animals ($N = 3$ PU.1^{Low} AML vs. $N = 9$ nonleukemic). Significance was assessed using the built-in function of the respective R/Bioconductor package, described in the Methods. **B** and **C,** Quantitative enrichment analysis plot (**B**) and heat map (**C**) representing chromatin accessibility in WT ($N = 3$ pooled mice), URE^{HET}*Tet2*^{HET} ($N = 2$ biological replicates), and URE^{HET}*Tet2*^{KO} mice ($N = 2$ biological replicates). Data are expressed as mean normalized read coverage across the indicated regions (monocytic- and neutrophilic-specific enhancers), which were split into 100 bins. Significance was assessed using Wilcoxon ranked-sum tests with correction for multiple testing. **D,** Schematic showing the 20 cell types and 8 epigenetic features that were used for annotation of 27 epigenetic states and their representation in each of the cell types. Epigenetic states are defined by 1–5 capital letters that are described in the legend. (*continued on next page*)

and *protein–DNA complex subunit organization* (Supplementary Fig. S6O and S6P), suggesting a compromised differentiation-relevant chromatin rearrangement, in line with recent reports (41). Intersection with PU.1 and Tet2 ChIP-seq data uncovered that close to 90% of the cCREs localized in the closed chromatin section in compound mutant leukemic HSPCs overlapped with PU.1 and/or Tet2 occupancy, whereas less than half of cCREs within loci with increased accessibility are bound by either regulator (Fig. 5H). Collectively, these data show that compound mutant mice with AML lose

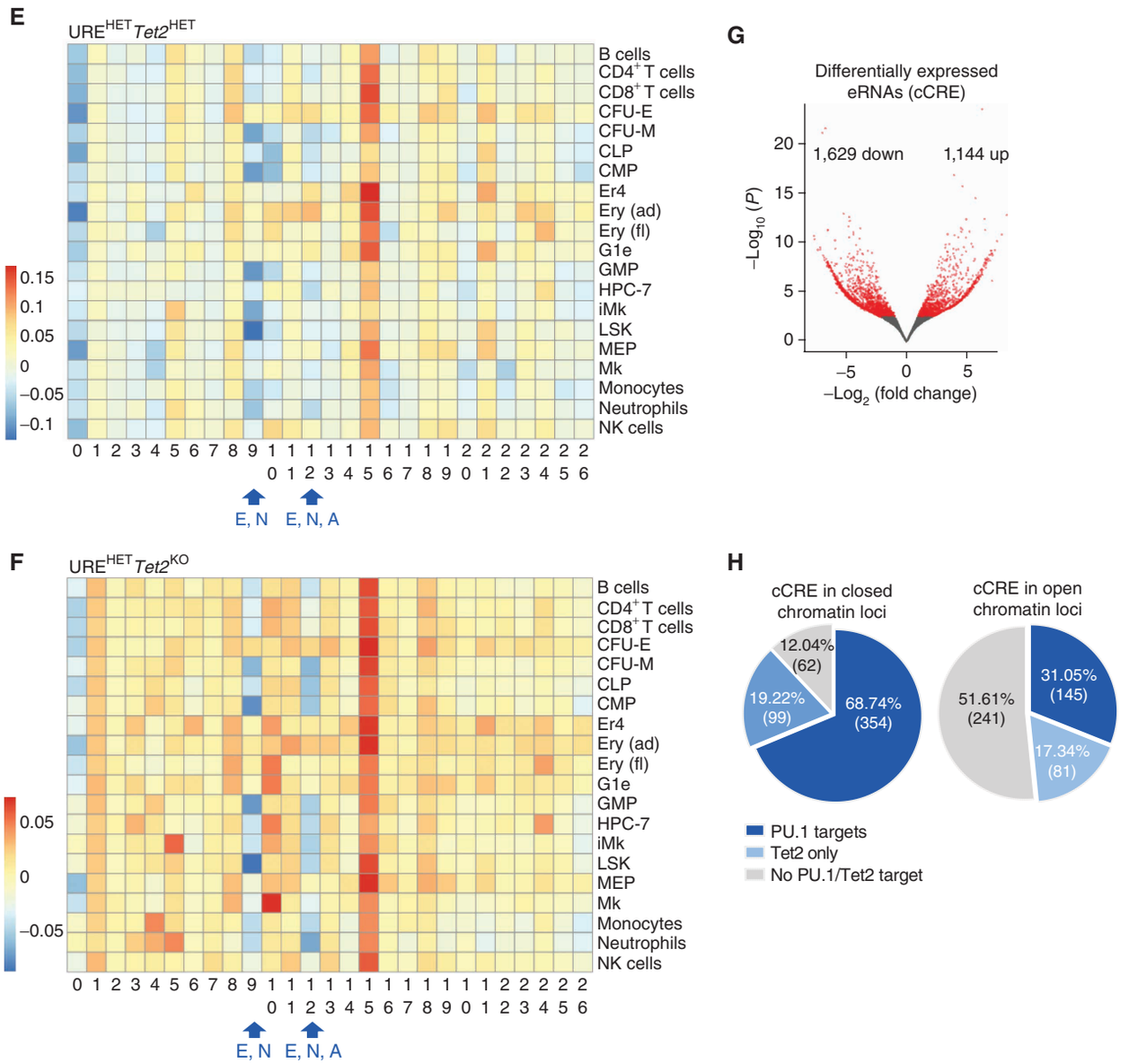


Figure 5. ***(Continued)*** **E** and **F,** Heatmaps of differential chromatin status, in URE^{HET}*Tet2*^{HET} ($N = 2$ biological replicates; vs. $N = 3$ WT; **E**) and URE^{HET}*Tet2*^{KO} ($N = 2$ biological replicates; vs. $N = 3$ WT; **F**) mice, contrasted to the IDEAS cell type–specific epigenomic states (37). Significance was assessed using the built-in function of the respective R/Bioconductor package, described in the Methods. **G,** Differentially expressed enhancer RNAs (eRNA) that correspond to *cis*-regulatory elements (cCRE) in leukemic animals (PU.1^{Low} AML vs. WT). Significance was assessed using the respective R/Bioconductor package, described in the Methods. $N = 2$ independent RNA-seq experiments. **H,** Pie charts with the percentages of PU.1 or Tet2 targets or not, present at cCREs in closed (left) or open (right) chromatin regions.

accessibility of PU.1-associated enhancer activation governing proper myeloid differentiation.

Inhibition of PU.1 Regulated DNA Methylation-Sensitive Myeloid Regulatory Regions in Leukemic HSPCs

We next predicted that loss of PU.1-associated myeloid enhancer activation is a common molecular theme in leukemic HSPCs. To test this hypothesis, we first cross-compared differentially expressed gene sets in compound mutant AML cells (grouped by either genotype or PU.1 mRNA expression levels) with publicly available gene-expression sets derived from patients with myeloid malignancies [comprising promyelocytic AML (APL), AML, and myelodysplastic syndrome (MDS)], which uncovered a striking overlap of concordantly altered gene-expression programs, many of which with known perturbation in myeloid malignancies, irrespective of PU.1 levels (Fig. 6A). The perturbation, mostly loss of expression, was accompanied by concordantly altered canonical pathways and cellular functions between mouse and human myeloid malignancies (Supplementary Fig. S7A and S7B). What caught our particular attention was a signature of differentially genes associated with the presence of oncogenic fusion gene *ETV6–RUNX1*, the most common t(12;21) originating

gene lesion found in childhood acute lymphoblastic leukemia. The translocation fuses the N terminus of the ETS family member *ETV6* (TEL) to nearly full-length *RUNX1* (*AML1*), retaining the DNA-binding domain of RUNX1 (*Runt*) and turning the oncogenic TF into a repressor via ETV6-mediated recruitment of transcriptional corepressors (NCOR or SIN3A) and epigenetic modifiers (HDAC; ref. 42). In line, we found that the vast majority of genes contributing to the ETV6–RUNX1-associated gene signature showed reduced expression in compound mutant leukemic HSPCs compared with WT controls (Fig. 6B) and concomitant loss of chromatin accessibility of a Runt DNA-binding motif (Supplementary Fig. S7C). Yet, we did not find evidence for the presence of the *ETV6–RUNX1* fusion gene in our compound mutant mice (Supplementary Table S8), in line with clinical observations of adults and patients with myeloid malignancies typically lacking this oncogene (43). Strikingly, we found that 59% of closed chromatin regions at cCRE of *ETV6–RUNX1*-associated genes could be bound by PU.1 (Fig. 6C), suggesting loss of PU.1 mediated gene activation at these sites. In support, HOMER motif analysis of cCRE sequences in chromatin with reduced accessibility in compound mutant leukemic HSPC revealed several variations of an ETS motif composed of a core PU box (*NN* GGAAGT) along with cytosine compositions (CA and CC) at the NN position as top ranking and strongly enriched DNA-binding motifs within the *ETV6–RUNX1* gene set–regulating cCRE, as well as all closed chromatin cCRE (compared with WT controls; Fig. 6D). In contrast, PU.1 DNA-binding motifs lacking cytosines immediately upstream of the core PU box, such as the functionally highly relevant PU.1:IRF composite motif (27), were not found within closed chromatin cCRE-regulating genes of the *ETV6–RUNX1* signature (Supplementary Fig. S7C), strongly suggesting inhibition of a group of PU.1-regulated enhancers that bear a highly defined cytosine containing DNA-binding motif. We predicted that this motif may render gene regulation sensitive to DNA methylation (further referred to as *methETS*).

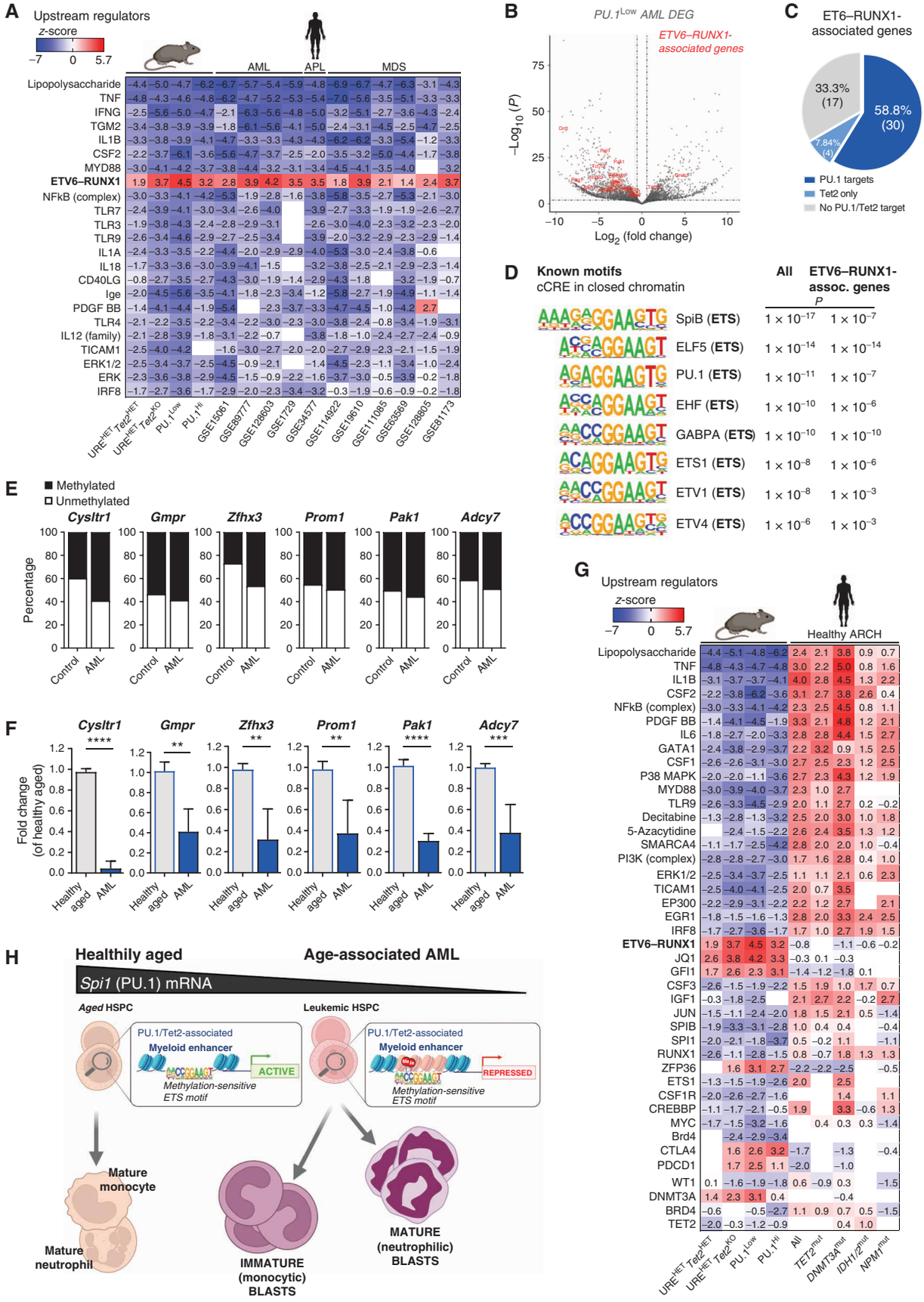
Although many TF, including several members of the ETS family, display reduced binding affinities to DNA motifs containing methylated cytosine residues (44), PU.1's ability to bind DNA is largely unperturbed in the context of DNA methylation. Although this rule likely applies to most PU.1 motifs, exceptions have been reported in the past (41). We predicted that expression loss of genes regulated by *methETS* motifs would associate with increased DNA methylation at these sites. In confirmation, bisulfite sequencing of leukemic and nonleukemic HSPCs showed significant DNA hypermethylation in or around PU.1 *methETS* motifs present at enhancers associated with *ETV6–RUNX1* signature genes (Fig. 6E) and was accompanied by reduced gene expression in leukemic GMPs compared with age-matched control cells from healthy aged compound mutant or WT mice (Fig. 6F; Supplementary Fig. S7D and S7E). Together, these observations strongly suggest epigenetic inhibition of a subset of PU.1-controlled genes harboring DNA methylation-sensitive PU.1-binding motifs in leukemic HSPCs. Lastly, we compared mRNA levels of *ETV6–RUNX1* signature genes in HSPCs from nonleukemic compound mutant mice of the same or higher age as their leukemic counterparts, which uncovered that healthy aged compound mutant mice had retained high expression levels (Fig. 6F; Supplementary Fig. S7D). Comparative pathway analyses of human healthy ARCH HSCs and leukemic stem cells further showed significant alterations in this gene set exclusively in *TET2*^{mut} leukemic HSPCs (Fig. 1D) and dissimilar alterations in canonical pathways and upstream regulators between our compound mutant mice and human healthy ARCH HSCs, including in the *ETV6–RUNX1* signature (Fig. 6G; Supplementary Fig. S7F). Together, these findings show that expression loss of genes containing PU.1-associated *methETS* motifs is restricted to leukemic HSPCs.

DISCUSSION

Altered DNA methylation is prevalent in the elderly, such as observed and well characterized in ARCH (28), which often precedes AML (5). Our study uncovered transcriptional master regulator PU.1 as a key suppressor for the malignant transformation of TET2-deficient hematopoietic stem and progenitor cells during aging and uncovers a set of downregulated genes with concomitant DNA hypermethylation, demarcating healthy from malignant hematopoiesis (Fig. 6H). Past work has demonstrated that ARCH-associated mutations leading to perturbation of key regulators of DNA methylation, such as TET2 and DNMT3A, alter DNA methylation in a highly locus-specific fashion (18, 45). Sites with aberrant DNA methylation in *TET2* or *DNMT3A* mutant cells are enriched for several key hematopoiesis regulating TF, including PU.1 (28, 45), suggesting that alterations in the PU.1-regulated network may be a common feature of ARCH. Here we show that even moderate perturbation of PU.1 leads to the evolution of myeloid malignancies in the context of *Tet2* deficiency.

PU.1 mRNA levels moderately decline aging HSCs (27). PU.1 is also inhibited in most patients with AML through various molecular pathways (19, 20, 22, 46, 47). Using published data sets, we found that expression levels of the hematopoietic master regulator are comparable in aged HSCs with and without ARCH mutations; yet they appear to drop moderately in LSCs, suggesting compromised gene regulation of the TF. Despite moderately reduced PU.1 mRNA levels, its broader transcriptional network appears intact, in line with the essentiality of the transcription factor in maintaining normal and malignant hematopoiesis (20, 22, 23, 46, 48).

We previously uncovered that perturbed PU.1 gene regulation and impaired expression drives malignant transformation of mismatch-repair–deficient HSCs (27), demonstrating a role for PU.1 in suppressing leukemogenesis in the context of accumulating genetic lesions, such as those seen in aging stem cells or treatment-associated myeloid malignancies (49, 50). Here we found that PU.1 also safeguards HSPCs in the context of loss-of-function mutations in the DNA demethylating enzyme *TET2*. In mice, loss of *Tet2* triggers a late-onset myeloproliferative disease with penetrance inversely correlated with residual *Tet2* gene dosage (17). Tet2 binds and governs site-specific DNA demethylation upon being recruited to the chromatin by DNA-binding TF, such as RUNX1 (51) or PU.1 (34), which we find to interact via Tet2's catalytic domain. Upon DNA recruitment, Tet2 also provides a hub for subsequent lineage-instructing transcription factor binding at cell type–specific enhancers residing in open chromatin (52). Consistently, *TET2*-mutant immature and myeloid cells


A
Upstream regulators
z-score
–7 0 5.7
AML
APL
MDS
Lipopolysaccharide
TNF
IFNG
TGM2
IL1B
CSF2
MYD88
ETV6–RUNX1
NFkB (complex)
TLR7
TLR3
TLR9
IL1A
IL18
CD40LG
Ige
PDGF BB
TLR4
IL12 (family)
TICAM1
ERK1/2
ERK
IRF8
URE^HET Tet2^HET
URE^HET Tet2^KO
PU.1^Low
PU.1^Hi
GSE15061
GSE89777
GSE128603
GSE1729
GSE34577
GSE114922
GSE19610
GSE111085
GSE63569
GSE128805
GSE81173
B
PU.1^Low AML DEG
ETV6–RUNX1-associated genes
–Log10 (P)
Log2 (fold change)
C
ET6–RUNX1-associated genes
33.3% (17)
58.8% (30)
7.84% (4)
PU.1 targets
Tet2 only
No PU.1/Tet2 target
D
Known motifs
cCRE in closed chromatin
All
ETV6–RUNX1-assoc. genes
P
SpiB (ETS) 1 × 10^–17 1 × 10^–7
ELF5 (ETS) 1 × 10^–14 1 × 10^–14
PU.1 (ETS) 1 × 10^–11 1 × 10^–7
EHF (ETS) 1 × 10^–10 1 × 10^–6
GABPA (ETS) 1 × 10^–10 1 × 10^–10
ETS1 (ETS) 1 × 10^–8 1 × 10^–6
ETV1 (ETS) 1 × 10^–8 1 × 10^–3
ETV4 (ETS) 1 × 10^–6 1 × 10^–3
E
Methylated
Unmethylated
Cysltr1
Gmpr
Zfhx3
Prom1
Pak1
Adcy7
Percentage
Control
AML
F
Fold change (of healthy aged)
Healthy aged
AML
G
Upstream regulators
z-score
–7 0 5.7
Healthy ARCH
Lipopolysaccharide
TNF
IL1B
CSF2
NFkB (complex)
PDGF BB
IL6
GATA1
CSF1
P38 MAPK
MYD88
TLR9
Decitabine
5-Azacytidine
SMARCA4
PI3K (complex)
ERK1/2
TICAM1
EP300
EGR1
IRF8
ETV6–RUNX1
JQ1
GFI1
CSF3
IGF1
JUN
SPIB
SPI1
RUNX1
ZFP36
ETS1
CSF1R
CREBBP
MYC
Brd4
CTLA4
PDCD1
WT1
DNMT3A
BRD4
TET2
URE^HET Tet2^HET
URE^HET Tet2^KO
PU.1^Low
PU.1^Hi
All
TET2^mut
DNMT3A^mut
IDH1/2^mut
NPM1^mut
H
Healthily aged
Age-associated AML
Spi1 (PU.1) mRNA
Aged HSPC
Leukemic HSPC
PU.1/Tet2-associated
Myeloid enhancer
ACTIVE
REPRESSED
Methylation-sensitive ETS motif
Mature monocyte
Mature neutrophil
IMMATURE (monocytic) BLASTS
MATURE (neutrophilic) BLASTS

present with hypermethylation of active enhancers and perturbed TF activity (28, 32, 53).

Here, we uncovered that moderate perturbation of PU.1 gene regulation through deletion of an URE critical for high expression of PU.1 in HSCs and differentiating myeloid cells (54) causes malignant transformation of *Tet2*-deficient HSPC, whereas not sufficient to induce malignancy in *Tet2* WT cells. Compound mutant mice lacking one copy of the PU.1 URE, along with either heterozygous or homozygous deletion of *Tet2*, developed aggressive AML hallmarked by splenomegaly, anemia, leukocytosis, and the accumulation of differentiation-impaired blast cell populations with varying phenotypes. Although displaying mostly no discernible differences in PU.1 mRNA expression at a young age, some compound mutant mice showed reduced PU.1 expression in accumulating leukemic HPSCs and blast cell populations when compared with their phenotypical WT and single mutant counterparts, which we attributed to epigenetic inactivation in at least some of the animals. Strikingly, mRNA and protein expression did not correlate in cKit$^+$ progenitor cells from compound mutant mice (moribund as well as healthy aged), indicating differential regulation at the post-transcriptional level. Reduced PU.1 mRNA expression levels were unlikely the sole trigger of leukemic progression as healthy aged compound mutant mice showed a similar drop in PU.1 expression levels as the majority of mice with leukemia, consistent with past studies (55).

Mutational analysis of transplantable leukemic HSPCs did not reveal any prevalent AML driver mutations, but uncovered highly recurrent and defined nonsense insertion mutations in *Cux1* at subclonal levels. *CUX1*-inactivating mutations are frequently found in human myeloid malignancies and associated with poor survival and trigger an MDS/MPN phenotype in mice (56). Although only found at low subclonal levels and ruling out *Cux1* lesions as AML drivers in our model, the high occurrence of defined loss-of-function mutations in our compound mutant mice suggests this gene as a mutational hot spot in Tet2 and PU.1-deficient HSPCs.

Although we cannot rule out other genetic drivers of the myeloid disease developing in our compound mutant mouse model, our molecular analyses into gene regulatory circuits support a role for wider-ranging epigenetic alterations, especially perturbing key enhancer landscapes with high similarity to human disease (29). In support, we found epigenetic inhibition of hematopoietic enhancers and loss of gene activation at loci harboring core PU.1 DNA-binding motifs which are bound by the TF in hematopoietic stem and differentiating myeloid cells, consistent with the known tumor-suppressive function of PU.1 in myeloid malignancies (25, 27). Our study provided a core network of epigenetically inactivated genes commonly inhibited by the transcriptional repressor *ETV6–RUNX1* fusion oncogene (57), which suggests a role for RUNX1 in the activation of these loci, consistent with the known function of PU.1 and RUNX1 to jointly exclude transcriptional corepressors from multiprotein complexes driving the expression of key myeloid genes (58). We identified a set of highly enriched and similar ETS motifs harboring a core PU-box motif and 5′ flanking cytosines that have been shown to exhibit a reduced binding affinity for PU.1 when methylated (referred to as *methETS*; ref. 41). Consistently, we uncovered hypermethylation at or close to the identified *methETS* motifs in leukemic compound mutant HSPCs, which was accompanied by reduced gene-expression in leukemic GMPs. Notably, this was seen only in leukemic but not in healthy aged compound mutant mice.

In contrast to the remarkable enrichment of PU.1-associated enhancers in aberrantly closed chromatin, we found only a fraction of PU.1-controlled enhancers with an aberrant gain of accessibility and activity. This may be attributable to bystander effects of a compromised core PU.1 network, transcription factor repositioning (59), or a reflection of the oncogenic activity, which has recently been demonstrated for PU.1 in a highly context- and cell-specific fashion (60).

Our study provides proof of concept that PU.1 and TET2 cooperate in sustaining appropriate leukemia suppressive gene-expression programs and proposes a methylation-sensitive PU.1 network as an indispensable and common

Figure 6. Suppression of a distinct set of direct PU.1 targets is shared between mouse and human myeloid leukemia. **A,** Comparative upstream regulator analysis of DEGs in compound mutant HPSC from leukemic mice and blast cells from patients with myeloid malignancies (AML, APL, MDS; accession numbers of publicly available data sets are shown at the bottom of the heatmap), showing significant similarities. Significance was assessed using the built-in function of the IPA software and is indicated as *z* score. *z* score ≥ 2 denotes activation; *z* score ≤ –2 denotes inhibition. **B,** Volcano plot of DEGs (PU.1^{Low} AML vs. WT), showing ETV6–RUNX1-associated targets (in red) with significant expression deregulation. Significance was assessed using the built-in function of the respective R/Bioconductor package. **C,** Pie chart illustrating the fraction of PU.1 and Tet2 direct targets at cCRE in closed chromatin regions, within the group of ETV6–RUNX1-associated downstream genes. PU.1 and Tet2 targets are defined as overlapping cCRE with PU.1/Tet2 binding, derived from published ChIP-seq data (GSE50762 and GSE115965). **D,** HOMER motif enrichment analysis of all cCRE-associated closed chromatin loci (*P* value on the left column) or ETV6–RUNX1-associated genes (*P* value on the right column) in leukemic animals (PU.1^{Low} AML vs. WT). Significance was assessed using the built-in function of the HOMER software. **E,** Percentage of methylation states in regions at and around PU.1 motifs at enhancers of PU.1/ETV6–RUNX1-associated targets in compound mutant mice ($N = 2$ biological replicates) compared with WT controls ($N = 3$ pooled mice). CD4$^-$CD8a$^-$CD19$^-$B220$^-$ (Lymph$^-$) Ter119$^-$ cKit$^+$ cells were sorted as in Fig. 4C and D in independent experiments. Significance was assessed using the Fisher exact test. **F,** qRT-PCR analysis of expression of PU.1/ETV6–RUNX1-associated targets in FACS-purified GMP cells of leukemic mice, relative to healthy age-matched compound mutant animals (healthy aged; $N = 3$–5 biological replicates). qRT-PCR analysis was performed in four independent experiments. GMP were sorted as indicated in the gating strategy in Supplementary Fig. S7D. Significance was assessed using an unpaired Student *t* test and is indicated as **, $P \le 0.01$; ***, $P \le 0.001$; ****, $P \le 0.0001$. **G,** Comparative upstream regulator analysis of DEGs from leukemic compound mutant mice and human healthy ARCH HSC (all combined or *Tet2*-, *Dnmt3a*-, *IDH1/2*-, *NPM1*-mutated) from a published data set (GSE74246). Significance was assessed using the built-in function of the IPA software and is indicated as *z* score. *z* score ≥ 2 denotes activation; *z* score ≤ –2 denotes inhibition. **H,** Proposed mechanistic model (created with BioRender.com). Enhancer-mediated PU.1 dysregulation (modeled herein through heterozygous deletion of the –14 kb URE of the PU.1 encoding gene) predisposes *Tet2*-deficient HSPC to AML. Reduced PU.1 mRNA expression is compatible with nonleukemic hematopoiesis in compound mutant mice. Healthy aged compound mutant URE$^{\Delta/WT}$ *Tet2*-deficient HSPC (left) gave rise to myeloid cells with full maturation capacity and expression of genes regulated by PU.1-associated methylation-sensitive ETS sites (*methETS*). In contrast, leukemic compound mutant HSPC (right) showed a signature of PU.1-associated *methETS* motif bearing, hypermethylated genes with loss of expression. Functionally, leukemic HPSC gave rise to myeloid precursor cells with a stage-specific differentiation block at immature monocytic and more mature neutrophilic blast levels.

feature of myeloid leukemia. This may comprise new functional tumor suppressors, as well as novel biomarkers, which may allow for the identification of patients at risk for malignant transformation.

Limitations of the Study

There are a few limitations to this work that warrant noting. We did not directly focus on PU.1 and Tet2 targets that lost chromatin accessibility but analyzed published ChIP-seq data sets of PU.1 or Tet2 binding to their target loci, in macrophages or myeloid cells, respectively. However, as these commonly lost accessibility in compound mutant mice, accompanied by recurrent enrichment of PU.1/ETS motifs in HOMER motif analysis and hypermethylation of a subset of targets genes, we believe that we look at representative PU.1/Tet2 target loci. Further work will functionally characterize the importance of all or part of the ETV6–RUNX1-associated genes in malignant transformation in mice, as well as human ARCH and hematologic disease. Furthermore, the determination of specificity of loss of PU.1 binding to *meth*ETS motif-containing loci during malignant transformation in mice and human AML will be needed to validate the importance of these in myeloid leukemia, as well as their prognostic value.

METHODS

Mice

Vav-iCre$^+$PU.1$^{URE\Delta/+}$*Tet2*$^{+/flox}$ and Vav-iCre$^+$PU.1$^{URE\Delta/+}$*Tet2*$^{flox/flox}$ mice were generated by mating PU.1 URE$^{\Delta/+}$ mice (C57BL/6, 129SV, Balbc; CD45.2$^+$; our lab; ref, 25) to*Tet2*$^{+/flox}$ mice (C57BL/6, 129SV; CD45.2$^+$; Stock No 017573; ref. 17) and Vav-iCre$^+$ mice (C57BL/6NJ; CD45.2$^+$; JAX stock #008610). NOD-*scid* IL2Rγ^{null} (NSG) mice (CD45.1$^+$; JAX stock #005557; refs. 61, 62) were purchased from Jackson Labs. All mice were housed in a special pathogen-free barrier facility at the Albert Einstein College of Medicine. All experimental procedures were approved by and performed in accordance with the guidelines from the Institutional Animal Care and Use Committee (IACUC) of the Albert Einstein College of Medicine (Protocol# 0000-1015). Mice with the indicated genotypes were included in the study and consisted of balanced numbers from both genders; we used age-matched mice. Investigators were not blinded to genotype group allocations.

Histology

Femoral or sternum bones, spleens, and livers were fixed in neutrally buffered formalin at room temperature for >24 hours, embedded in paraffin, cut into sections using a microtome, and stained with hematoxylin and eosin or anti-myeloperoxidase (MPO), according to standard protocols. Cell morphology was evaluated using an Axiovert 200M microscope (Zeiss) and analyzed pictures through ImageJ (ImageJ, RRID:SCR_003070; ref. 63).

Complete Blood Counts

Peripheral blood was obtained through the mouse submandibular vein using standard techniques and analyzed using the Forcyte Hematology Analyzer (Oxford Science Inc) according to the manufacturer's instructions.

May-Grünwald-Giemsa Stain

Peripheral blood smears and erythrocyte-lysed single-cell suspensions from bone marrow and spleen cells were used to prepare cytospins on polylysine-coated slides and were stained using the May-Grünwald-Giemsa coloration method (64). Slides were incubated in the solutions described below, for the indicated amount of time, in the following order: (i) May-Grünwald pure (2 minutes); (ii) May-Grünwald (diluted 1:2 in phosphate buffer, pH 6.8; 3 minutes); (iii) Giemsa stain diluted (30 mL Giemsa pure plus 250 mL phosphate buffer, pH 6.8; 18 minutes); (iv) Phosphate buffer, pH 6.8 (3 minutes). Then slides were washed with $_{dd}H_2O$ until the slide background was decolored (~1–2 minutes) and were let dry at room temperature. Cell morphology was evaluated using an Axiovert 200M microscope (Zeiss), and pictures were analyzed through ImageJ (ImageJ, RRID:SCR_003070; ref. 63).

Mutational Analysis of the Spi-1 Locus

Genomic DNA and RNA from TBM cells were prepared using the AllPrep DNA/RNA Micro Kit. RNA was reverse transcribed using the iScript cDNA synthesis kit. Genomic DNA and cDNA were subsequently used for PCR amplification of the URE, promoter, and coding regions of the *Spi-1* locus, using the Platinum SuperFi PCR Master Mix or Phusion High-Fidelity PCR Master Mix. Amplified bands were run on and excised from a 1% agarose gel (in 1× Tris-Acetate-EDTA buffer). DNA was purified using the MinElute Gel Extraction Kit or E.Z.N.A. MicroElute Gel Extraction Kit, concentration was measured by NanoDrop 1,000 full spectrum spectrophotometer and appropriate volume was sent for Sanger sequencing. Snapgene viewer, MultAlign, or Clustal Omega (RRID:SCR_001591) were used for sequencing quality analysis and alignment with the expected sequences.

Mutational Analysis by the Mouse MSK-IMPACT Panel

FACS-purified cKit$^+$ cell (tumor) and tail (normal) samples derived from compound mutant mice with AML were used for genomic DNA isolation, using the Gentra Puregene Tissue kit/Puregene Core kit B (Qiagen), according to the manufacturer's instructions. Genomic DNA (100–200 ng) from each sample was sequenced using the mouse MSK-IMPACT platform. Sample processing was performed as previously described (31).

Immunoprecipitation and Western Blot Analysis

HA-tagged PU.1 plasmid was constructed using the pcDNA3.1$^-$ (RRID:Addgene_124144) vector, C-terminal HA tag and full-length PU.1 (813 bp) that was cloned from UT7 cell cDNA and ligated to pcDNA3.1$^-$ vector, with the HA tag at its C-terminus. The pCMV6-Entry vector (RRID:Addgene_58324), with full-length Tet2, was purchased from Origene and was used to clone out the Tet2 catalytic domain (aa 1,129–2002; DNA size: 2,619 bp), which was subsequently cloned in pCMV6-Entry (C-terminal FLAG tag) again, where an N-terminal GFP was added. HEK293T cells were cotransfected with HA-PU.1 (pcDNA3.1$^-$, C-terminal HA tag) and GFP/FLAG-Tet2-Catalytic domain cDNA plasmid (pCMV6-Entry, C-terminal FLAG tag), using MIDSCI, Turbo DNAfectin 3000. Twenty-four hours after transfection, cells were harvested and lysed as previously, and protein concentration was determined by the standard Bradford assay. Cell lysis, immunoprecipitation, and Western blot analysis were performed according to standard protocols from Pierce, as previously described (65). Cell lysis and immunoprecipitation were carried out in a buffer containing 25 mmol/L Tris-HCl pH 7.4, 150 mmol/L NaCl, 1% NP-40, 1 mmol/L EDTA, 5% glycerol, and a protease inhibitor cocktail. Lysates were immunoprecipitated using anti-FLAG M2 magnetic beads to pull down FLAG-Tet2-Catalytic. Following SDS-PAGE and blotting of equal amounts of the immunoprecipitate, the membranes were incubated with either an anti-PU.1 Rabbit mAb (9G7; 1:1,000) or an anti-Flag (DYKDDDDK) Tag Rabbit mAb (D6W5B; 1:1,000 dilution) antibody. IgG and input samples were used as controls. Infrared

dye–labeled secondary antibodies (IRDye 800CW Goat anti-Rabbit IgG antibody; 1:10,000 dilution) and an Odyssey Infrared Imaging System (LI-COR) were used for detection and visualization. Bead-sorted primary cKit$^+$ cells were lysed in a buffer containing 50 mmol/L Tris-HCl pH 7.4, 150 mmol/L NaCl, 1% NP-40, 5 mmol/L EDTA, 0.25% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail, 1 mmol/L Na_3VO_4, 20 mmol/L β-glycerophosphate, and 1 mmol/L PMSF. Following SDS-PAGE electrophoresis and protein transfer, pvdf membranes were incubated with either PU.1 Rabbit mAb (9G7; 1:500) or α-tubulin mouse mAb (B-7; 1:1,000) antibody. Anti-rabbit IgG HRP (#ab6721; 1:3,000), anti-mouse IgG HRP (1:2,500), and an Odyssey Infrared Imaging System (LI-COR) were used for detection and visualization.

Detection of ETV6–RUNX1 Fusion

The RNA-sequencing data from cKit$^+$ cells from leukemic and non-leukemic animals were used to detect the potential of acquired gene fusions. The software Arriba (66) was used and specifically focused on detecting the *Etv6-Runx1* fusion.

Analysis and Purification of Hematopoietic Cells

TBM cells were isolated from the femurs, tibiae, pelvic bones, and spines of mice and RBC were lysed with 1× ACK lysing buffer pH 7.4. After two washes with 1× phosphate-buffered saline—containing 2% FBS (1× PBS/2% FBS). Where indicated, BM cells were further enriched using bead sorting as per the manufacturer's recommendation (CD117 MicroBeads, mouse, Miltenyi); post-sort enrichment was quantified by FACS analysis (using the anti-CD117-PE antibody, Miltenyi #130-091-730). Cells were stained with the following antibody cocktails for isolation of (**a**) cKit$^+$ CD4$^-$CD8a$^-$CD19$^-$ B220$^-$ (Lymph$^-$) Ter119$^-$ (Ery$^-$) cells, (**b**) HSPC (LSK, GMP), or (**c**) neutrophils and monocytes (all diluted at 1:100 in PBS/2%FBS) for 30 minutes on ice. (**a**) CD4-PE-Cy5, CD8a-PE-Cy5, CD19-PE-Cy5, B220-PE-Cy5, Ter119-PE-Cy5, CD117-APC, CD16/CD32-PE-Cy7, CD34-eFluor450, CD11b-FITC, Gr-1-PE. (**b**) CD4-PE-Cy5, CD8a-PE-Cy5, CD19-PE-Cy5, B220-PE-Cy5, Ter119-PE-Cy5, Gr1-PE-Cy5, CD11b-PE-Cy5, CD117-PE, Sca-1-BV421, CD16/CD32-PE-Cy7, CD34-FITC. (**c**) CD4-PE-Cy5, CD8a-PE-Cy5, CD19-PE-Cy5, B220-PE-Cy5, Ter119-PE-Cy5, Ly6G-PerCPCy5.5, CD11b FITC. Cells were washed once, filtered, and immediately sorted using a MoFlo Astrios EQ (Beckman Coulter). For mouse phenotyping, cells were stained with the following antibody cocktail: CD11b PE, Gr-1 PE-Cy7, CD44-eFluor450, CD4 PE-Cy5, CD8a FITC, B220 APC-eFluor780, CD117 APC. Cells were washed once and fluorescent signals were acquired with BD FACSAria I system (Becton Dickinson). Data were analyzed with FlowJo V10.7.2 (RRID:SCR_008520).

Transplantation of Bone Marrow–Derived Cells

TBM cells were isolated from the femurs, tibiae, pelvic bones, and spines of mice, RBC were lysed with 1× ACK lysing buffer, washed twice with 1× PBS, and stained with antibodies. Using a MoFlo Astrios EQ (Beckman Coulter), cKit$^+$Lymph$^-$Ery$^-$ cells were sorted in Iscove's DMEM, supplemented with 2% FBS, washed twice, counted and resuspended in HBSS (Thermo Fisher). Both sorted (40,000 cells/mouse) and 100,000 TBM cells (per mouse) were transplanted into sublethally irradiated 4- to 6-week-old NSG recipient animals via retro-orbital injection 4 to 5 hours after irradiation. Total body irradiation was delivered in a single dose of 250 cGy using a Shepherd 6,810 sealed-source ^{137}Cs irradiator. Engraftment of donor cells was monitored through CD45.2 and CD45.1 expression on peripheral blood and bone marrow cells. We stained ACK-treated cells with the antibodies CD45.1-BV510, CD45.2-APC-eFluor780, CD4-PE-Cy5, CD8a-PE-Cy5, CD19-PE-Cy5, B220-PE-Cy5, Gr-1-PE, CD11b-PE-Cy7, CD117-APC, Sca-1-AlexaFluor700, Ter119-eFluor450 (1:100 in 1× PBS/2% FBS), and fluorescence signals were analyzed on a BD FACSAria I system (Becton Dickinson). Data were analyzed with FlowJo V10.7.2 (RRID:SCR_008520).

In Vitro *Colony Formation and Serial Replating Assay*

Clonogenic capacity of cells was characterized by plating 50,000 TBM cells in mouse methylcellulose complete media containing IL3, IL6, SCF, and EPO, according to the manufacturer's recommendation. Colonies were scored 7 to 9 days after plating using an inverted microscope (Olympus CKX53), an EVOS FL Auto (Life Technologies), or Cytation5 cell imaging multimode reader (BioTek). For morphology analysis, colony assay-derived cells were used to generate cytospins that were subjected to May-Grünwald-Giemsa stain and analyzed using an Axiovert 200M microscope (Zeiss). For FACS analysis, cells were stained with the following antibody cocktail: CD4-PE-Cy5, CD8a-PE-Cy5, CD19-PE-Cy5, B220-PE-Cy5, Gr-1-PE, CD11b-AlexaFluor488, CD117-APC-eF780, Sca-1-APC, FcγRII/II-PE-Cy7, CD34-Alexa Fluor700, CD115-BrilliantViolet421, Ter119-Biotin (all 1:100 in 1× PBS/2% FBS). Then cells were washed and stained with Streptavidin Brilliant Violet 510 (1:100 in 1× PBS/2% FBS). Fluorescence signals were analyzed on a BD FACSAria I system (Becton Dickinson). Data were analyzed with FlowJo V10.7.2 (RRID:SCR_008520). After the first plating, cells (5,000–40,000 cells/mL) were serially replated in mouse methylcellulose complete media until colony formation ceased, and colonies were scored after 7 to 9 days and stained for morphology and FACS analyses as above.

RNA Purification, Real-time PCR, and Gene-Expression Analysis by RNA Sequencing

RNA was extracted from FACS-purified hematopoietic cells using the RNeasy Micro kit (Qiagen) or TRIzol Reagent (Ambion). RNA quantity and quality were assessed using a NanoDrop 1000 full-spectrum spectrophotometer (Thermo Fisher) and 2100 Bioanalyzer (Agilent) device, respectively. For real-time PCR, RNA was reverse transcribed using Superscript II (Invitrogen) or iScript cDNA synthesis kit (Bio-Rad). Amplification of target genes was performed using the PrimeTime Gene-Expression Master Mix (Integrated DNA Technologies). cDNA was amplified in a final volume of 20 μL in 96-well or 384-well microtiter plates according to the manufacturer's recommendation. Triplicate samples and six serial dilutions of standards were prepared for each target gene. Primers and probes used for real-time PCR can be found in Supplementary Table S9. We performed real-time PCR using an ViiA7 instrument (Life Technologies) with one cycle of 50°C (for 2 minutes) and 95°C (for 3 minutes) followed by 40 cycles of amplification, with each cycle comprising the steps: 95°C (for 15 seconds) and 60°C (for 1 minute). Target gene-expression quantification was calculated using the Pfaffl model and normalized to *Gapdh* expression levels. For global gene-expression analysis using RNA-sequencing, high-quality RNA (RNA integrity number ≥8) was prepared and sent for library construction and 100 bp, paired-end, Illumina-HiSeq 2500/4000 sequencing to the Beijing Genomics Institute (BGI) Group.

Improved Assay for Transposase-Accessible Chromatin by Sequencing

Improved Assay for Transposase-Accessible Chromatin by Sequencing (Omni-ATAC-seq) was performed as previously described (36), with a few alterations: FACS-purified, high viability (>90%) cKit$^+$ cells were used. At the nuclei centrifugation step, nuclei were pelleted at 13,000 rcf, for 1 minute, at 4°C, and all supernatants were aspirated. Transposition reactions were incubated at 37°C for 30 minutes in a thermomixer with 1200 RPM mixing. Transposed fragments were preamplified for 3 cycles (72°C for 5 minutes, 98°C for 30 seconds followed by 3 cycles of 98°C for 10 seconds, 63°C for 30 seconds and

65°C for 45 seconds) using the following reaction: 10 μL transposed DNA, 10 μL nuclease-free H_2O, 2.5 μL 25 μmol/L Custom Nextera PCR Primer 1 (Ad1 noMX), 2.5 μL 25 μmol/L Custom Nextera Barcoded PCR primer 2 (Ad2.X) and 25 μL 2× NEBNext Ultra II Q5 Master Mix. Then samples were immediately used in qPCR amplification to determine additional cycles. Amplified libraries were double-sided purified (removal of primer dimers and large >1,000 bp fragments) using AMPure XP beads: 0.5× volume (22.5 μL) of AMPure XP beads was added to each library, mixed thoroughly, incubated at room temperature for 10 minutes and tubes were placed on a magnetic rack for 5 minutes. Supernatants were transferred to new Eppendorf tubes, and 1.3× original volume (58.5 μL) AMPure XP beads was added (1.8× bead buffer:sample ratio), mixed thoroughly, incubated at room temperature for 10 minutes, and tubes were placed on a magnetic rack for 5 minutes. Supernatants were discarded, and beads were washed with 200 μL 80% ethanol (pipetted 10 times over beads), which were then discarded. Tubes were left on a magnetic rack with the cap open for 6 minutes, until all ethanol was removed, without overdrying. Beads were resuspended in 20 μL nuclease-free H_2O and thoroughly mixed by pipetting up and down at least 10 times. Tubes were then placed on the magnetic rack for 1 to 5 minutes, and supernatants were transferred to new Eppendorf tubes. Purified libraries were stored at –20°C. Library quantity was measured using the dsDNA HS Assay kit, and quality was assessed by running 5 ng of each library on a High-Sensitivity DNA Bioanalysis chip. Libraries were sent for 150 bp paired-end Illumina HiSeqXTen sequencing to the BGI group.

Bisulfite Sequencing

High-quality genomic DNA from FACS-purified cKit$^+$Lymph$^-$Ery$^-$ cells was prepared using the Puregene Cell and Tissue kit (Qiagen) and was subsequently used for BS-sequencing library preparation with the Ovation Ultralow Methyl-Seq with TrueMethyl oxBS kit (Nugen/Tecan), according to the manufacturer's guidelines (Ovation Ultralow Methyl-Seq Library Systems, Nugen/Tecan, M01320 v10). Libraries were sent for 150 bp paired-end Illumina HiSeqXTen sequencing to the BGI group.

Analysis of RNA-seq Data

For differential RNA expression, a set of enhancer RNA (eRNA) single-exon transcripts using the coordinates of cCREs was generated, and Gencode transcript gtf files containing this information were used to count reads overlapping these transcripts. For differential gene-expression analysis, the count matrix was subsequently split into eRNAs and non-eRNAs, and differential expression analysis was performed separately for eRNAs and regular genes. For differential gene-expression analysis, counts from the combined gene and cCRE alignment were used, filtered on genes only, and DEGs were determined using the DESeq2 package (RRID:SCR_015687; ref. 67) under R/Bioconductor. For PCA, counts from the combined gene and cCRE alignment were used, filtered on genes only, prefiltered expressed genes and took the top 90% genes with highest variance to perform PCA using vst-transformed counts. GSEA was performed using GSEA software (68). Pathway analyses were performed using Ingenuity Pathway Analysis (IPA, Qiagen; https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/; RRID:SCR_008653).

Analysis of Omni-ATAC-seq Data

ATAC-seq reads were cleaned (adapters removed and quality trimmed with bbduk.sh from the BBTools suite), aligned to mm10 with bowtie2 (v. 2.4.1; ref. 69) with the parameters: no-mixed, no-discordant, very-sensitive -k 10 -X 2000. Duplicates were removed with Samblaster v.0.1.26 (RRID:SCR_000468; ref. 70). Mitochondrial reads were removed with Sambamba v0.7.1 (RRID:SCR_002105; ref. 71), and the remaining reads were sorted by name using sambamba and then filtered to keep uniquely aligned read pairs using samtools (1.10.2) with the parameters -F 0 × 900 -f 0 × 0002. More precisely, we defined the peakset to be tested as the cCRE regions, excluded the mm10 blacklisted regions from ENCODE (RRID:SCR_015482), and restricted the peak calling to the standard chromosomes (autosomes and chrX and chrY). Individual peaks were called in the leukemic and nonleukemic samples, and the union of the peaks from both groups was used as potential peaks to perform differential testing within the *csaw* framework (peaks called by local enrichment), as previously described (72). In the csaw pipeline, we used *peak.counts.loess* to define the working.windows, merged nearby windows up to 500 bp apart with a max merged window width of 50,000 bp, and used the most significant window as statistical representation for *P* value and FDR for merged windows. We then defined significant merged peaks as those with FDR <0.001. Closed chromatin regions were defined as the loci that were present in the nonleukemic samples, but not in the leukemic animals, whereas open chromatin loci were present in leukemic mice.

Analysis of Bisulfite-seq Data

Adapters were trimmed (trim galore v0.14), Illumina phiX mapped reads were filtered away, and then bisulfite-treated sequencing reads were mapped to the GRCm38 reference genome with Bismark v0.14.5 (RRID:SCR_005604) and bowtie2 v2.2.3 (RRID:SCR_016368; refs, 69, 73). Duplicate reads were marked (picard-tools v1.92) before methylation calls. The degree of methylation in any given region was determined by the count of converted/unconverted bases in the pileup of the bisulfite alignment. Though the depth of sequencing was not suitable for analysis at single-base resolution, a composite analysis was achieved by summing counts over ranges of chromosomal positions, specifically PU.1 bound, Tet2 bound, and PU.1 motif loci. Regions with significant differential methylation across comparison groups were determined via the Fisher exact test on the numbers of methylated/unmethylated bases and corrected for multiple comparisons with the Benjamini–Hochberg procedure.

Integrative Analysis with Published Data Sets

For direct target gene identification, we intersected binding peaks from published data sets of whole-genome chromatin immunoprecipitation (ChIP-seq) for PU.1 (GSE50762) and Tet2 (GSE115965). We next identified genes with PU.1 or Tet2 occupancy within enhancer (enhancer-like, nuclease-accessible, active; E, N, A cCREs) and promoter regions (enhancers were defined as cCRE present within 1 Mb of TSS of genes, excluding cCRE at ±1 kb of the TSS that we termed corresponding to promoters), which were also differentially expressed in leukemic mice. HOMER (RRID:SCR_010881) motif analysis software (74) was used to retrieve information about closed chromatin regions with 80+ % overlap with cCRE-regulatory regions and for the identification of known DNA-binding motifs. Comparative pathway analyses between DEG in mouse AML and human preleukemia, as well as mouse and human AML, were performed using IPA. Briefly, to identify different or common pathways with differential activation in mouse and human preleukemia or myeloid malignancies, we used the comparative analysis module under IPA. For human preleukemic stem cells, samples were used either all combined or separated by CH (Tet2, Dnmt3a, and IDH1/2)- or other AML (Npm1, Flt3)-associated mutation for DEG analysis. For the identification of cCRE-associated epigenetic state and gene expression, we used the VISION data to contrast our RNA-seq or ATAC-seq data to cCRE regions that overlapped by 80+ % with the eRNA regions or differential chromatin peaks were considered functional enhancer regulatory elements. The subsetByOverlaps function was used to identify closed/open chromatin regions as cCREs (enhancers were defined as present within 1 Mb of TSS of genes, excluding ±1 kb of the TSS that we termed corresponding to promoters). Neutrophil-, monocyte-specific enhancers (transcribed, enhancer-like; T, E and E,

N, A epigenetic states combined) and promoters (±1 kb of the TSS of genes) were defined as the ones present in these cells but not in B and T cells; B- and T-cell–specific enhancers (T, E and E, N, A epigenetic states combined) and promoters (±1 kb of the TSS of genes) were defined as the ones present in the latter two but not in neutrophils and monocytes. Wilcoxon ranked-sum tests were performed using the averaged normalized coverage for the metagene regions (cell type–specific cCREs) to calculate significance. Multiple-testing–adjusted *P* values are reported.

Statistical Analysis

Statistical analysis of group comparisons was performed using the Student unpaired *t* test, ordinary one-way ANOVA, log-rank test, Fisher exact (corrected with the Benjamini–Hochberg), or Wilcoxon ranked-sum test with correction for multiple testing, in GraphPad Prism (RRID:SCR_002798) or R, as indicated. Statistical significance was set at $P < 5\%$. Statistical evaluation of NGS data was performed using the built-in functions of IPA (RRID:SCR_008653), GSEA, HOMER, or the respective R/Bioconductor packages. Sample sizes chosen are indicated in the individual figure legends and were not based on formal power calculations to detect prespecified effect sizes.

Data Availability

The data discussed in this publication have been deposited in NCBI's Gene-Expression Omnibus (RRID:SCR_005012; ref. 75) and are accessible through GEO Series accession number GSE191053.

The data analyzed in this study were obtained from GSE50762, GSE115965, and GSE143271. All scripts for data analyses and data visualization from GSE143271 were accessed through https://github.com/rosshardison/VISION_mouseHem_code (37).

Details on experimental models, materials, reagents, software, and algorithms can be found in Supplementary Table S10.

Authors' Disclosures

B. Will has received funds for research projects and serving on advisory boards from Novartis Pharmaceuticals.

Authors' Contributions

M.M. Aivalioti: Data curation, formal analysis, investigation, visualization, methodology, writing–original draft, project administration. **B.A. Bartholdy:** Conceptualization, resources, data curation, formal analysis. **K. Pradhan:** Software, investigation. **T.D. Bhagat:** Software, investigation, methodology. **A. Zintiridou:** Investigation, methodology. **J.J. Jeong:** Investigation, methodology. **V.J. Thiruthuvanathan:** Investigation, methodology. **M. Pujato:** Formal analysis, investigation, visualization, methodology. **A. Paranjpe:** Software, formal analysis, visualization. **C. Zhang:** Formal analysis, visualization. **R.L. Levine:** Software, formal analysis, investigation, methodology. **A.D. Viny:** Investigation, methodology. **A. Wickrema:** Data curation, formal analysis, investigation. **A. Verma:** Conceptualization, resources, formal analysis, investigation, methodology. **B. Will:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, validation, investigation, visualization, methodology, project administration, writing–review and editing.

Acknowledgments

We thank the members of the Will lab for feedback on and discussions of the study. We further thank Dr. Daqian Sun and Swathi-Rao Narayanagari from the Stem Cell Isolation and Xenotransplantation Core facility for technical assistance with FACS analysis and cell sorting; David Reynolds from the Einstein Genomics Core facility for feedback and advice on the use of the Bioanalyzer or Fragment analyzer in library quality assessments, the Einstein Histology & Comparative Pathology and Biomarker Analytic Research Cores, the Einstein Institute for Animal Studies, and Research Facilities Coordinator Peter Schultes for expert support and technical assistance. This study was supported by grants from the NIH (CA230756 and DK105134 to B. Will), Cancer Center Support Grant (P30CA013330, pilot project award to B. Will), the Feldstein Medical Foundation (to B. Will), and the Leukemia Research Foundation (to B. Will). M.M. Aivalioti was the recipient of a training grant in Stem Cell Research by the NYSTEM (C30292GG; PI: P. Frenette).

Note

Supplementary data for this article are available at Blood Cancer Discovery Online (https://bloodcancerdiscov.aacrjournals.org/).

Received December 12, 2021; revised April 5, 2022; accepted July 7, 2022; published first August 12, 2022.

REFERENCES

1. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med 2014;371:2488–98.
2. Fuster JJ, MacLauchlan S, Zuriaga MA, Polackal MN, Ostriker AC, Chakraborty R, et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. Science 2017;355:842–7.
3. Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. N Engl J Med 2017;377:111–21.
4. Bonnefond A, Skrobek B, Lobbens S, Eury E, Thuillier D, Cauchi S, et al. Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. Nat Genet 2013;45:1040–3.
5. Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P, Weissman IL, Quake SR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med 2012;4:149ra18.
6. Shlush LI, Zandi S, Mitchell A, Chen WC, Brandwein JM, Gupta V, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature 2014;506:328–33.
7. Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med 2014;371:2477–87.
8. McKerrell T, Park N, Moreno T, Grove CS, Ponstingl H, Stephens J, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. Cell Rep 2015;10:1239–45.
9. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat Commun 2016;7:12484.
10. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood 2016;127:2391–405.
11. Hwang SM. Classification of acute myeloid leukemia. Blood Res 2020;55:S1–S4.
12. Juliusson G, Antunovic P, Derolf A, Lehmann S, Mollgard L, Stockelberg D, et al. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. Blood 2009;113:4179–87.
13. Shih AH, Jiang Y, Meydan C, Shank K, Pandey S, Barreyro L, et al. Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. Cancer Cell 2015;27:502–15.
14. Meyer SE, Qin T, Muench DE, Masuda K, Venkatasubramanian M, Orr E, et al. DNMT3A haploinsufficiency transforms FLT3ITD myeloproliferative disease into a rapid, spontaneous, and fully penetrant acute myeloid leukemia. Cancer Discov 2016;6:501–15.
15. Poitras JL, Heiser D, Li L, Nguyen B, Nagai K, Duffield AS, et al. Dnmt3a deletion cooperates with the Flt3/ITD mutation to drive leukemogenesis in a murine model. Oncotarget 2016;7:69124–35.

16. Loberg MA, Bell RK, Goodwin LO, Eudy E, Miles LA, SanMiguel JM, et al. Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. Leukemia 2019;33:1635–49.
17. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell 2011;20:11–24.
18. Zhang X, Su J, Jeong M, Ko M, Huang Y, Park HJ, et al. DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. Nat Genet 2016;48:1014–23.
19. Vangala RK, Heiss-Neumann MS, Rangatia JS, Singh SM, Schoch C, Tenen DG, et al. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. Blood 2003;101:270–7.
20. Mizuki M, Schwable J, Steur C, Choudhary C, Agrawal S, Sargin B, et al. Suppression of myeloid transcription factors and induction of STAT response genes by AML-specific Flt3 mutations. Blood 2003;101:3164–73.
21. Steidl U, Steidl C, Ebralidze A, Chapuy B, Han HJ, Will B, et al. A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. J Clin Invest 2007;117:2611–20.
22. Mueller BU, Pabst T, Fos J, Petkovic V, Fey MF, Asou N, et al. ATRA resolves the differentiation block in t(15;17) acute myeloid leukemia by restoring PU.1 expression. Blood 2006;107:3330–8.
23. Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno S, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood 2005;106:1590–600.
24. Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science 1994;265:1573–7.
25. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. Nat Genet 2004;36:624–30.
26. Staber PB, Zhang P, Ye M, Welner RS, Nombela-Arrieta C, Bach C, et al. Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. Mol Cell 2013;49:934–46.
27. Will B, Vogler TO, Narayanagari S, Bartholdy B, Todorova TI, da Silva Ferreira M, et al. Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. Nat Med 2015;21:1172–81.
28. Tulstrup M, Soerensen M, Hansen JW, Gillberg L, Needhamsen M, Kaastrup K, et al. TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. Nat Commun 2021;12:6061.
29. Corces MR, Buenrostro JD, Wu B, Greenside PG, Chan SM, Koenig JL, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet 2016;48:1193–203.
30. Pan F, Wingo TS, Zhao Z, Gao R, Makishima H, Qu G, et al. Tet2 loss leads to hypermutagenicity in haematopoietic stem/progenitor cells. Nat Commun 2017;8:15102.
31. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, et al. Memorial sloan kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J Mol Diagn 2015;17:251–64.
32. Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev 2015;29:910–22.
33. Dahl R, Walsh JC, Lancki D, Laslo P, Iyer SR, Singh H, et al. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor. Nat Immunol 2003;4:1029–36.
34. Lio CW, Zhang J, Gonzalez-Avalos E, Hogan PG, Chang X, Rao A. Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. eLife 2016;5:e18290.
35. Gupta P, Gurudutta GU, Saluja D, Tripathi RP. PU.1 and partners: regulation of haematopoietic stem cell fate in normal and malignant haematopoiesis. J Cell Mol Med 2009;13:4349–63.
36. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 2017;14:959–62.
37. Xiang G, Keller CA, Heuston E, Giardine BM, An L, Wixom AQ, et al. An integrative view of the regulatory and transcriptional landscapes in mouse hematopoiesis. Genome Res 2020;30:472–84.
38. Ungerback J, Hosokawa H, Wang X, Strid T, Williams BA, Sigvardsson M, et al. Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). Genome Res 2018;28:1508–19.
39. Ye M, Ermakova O, Graf T. PU.1 is not strictly required for B cell development and its absence induces a B-2 to B-1 cell switch. J Exp Med 2005;202:1411–22.
40. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell 2010;143:46–58.
41. Minderjahn J, Schmidt A, Fuchs A, Schill R, Raithel J, Babina M, et al. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. Nat Commun 2020;11:402.
42. Hiebert SW, Sun W, Davis JN, Golub T, Shurtleff S, Buijs A, et al. The t(12;21) translocation converts AML-1B from an activator to a repressor of transcription. Mol Cell Biol 1996;16:1349–55.
43. Cancer Genome Atlas Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 2013;368:2059–74.
44. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science 2017;356:eaaj2239.
45. Kaasinen E, Kuismin O, Rajamaki K, Ristolainen H, Aavikko M, Kondelin J, et al. Impact of constitutional TET2 haploinsufficiency on molecular and clinical phenotype in humans. Nat Commun 2019;10:1252.
46. Mueller BU, Pabst T, Osato M, Asou N, Johansen LM, Minden MD, et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. Blood 2002;100:998–1007.
47. Gerloff D, Grundler R, Wurm AA, Brauer-Hartmann D, Katzerke C, Hartmann JU, et al. NF-kappaB/STAT5/miR-155 network targets PU.1 in FLT3-ITD-driven acute myeloid leukemia. Leukemia 2015;29:535–47.
48. Bonadies N, Pabst T, Mueller BU. Heterozygous deletion of the PU.1 locus in human AML. Blood 2010;115:331–4.
49. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell 2012;150:264–78.
50. Busque L, Patel JP, Figueroa ME, Vasanthakumar A, Provost S, Hamilou Z, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. Nat Genet 2012;44:1179–81.
51. Suzuki T, Shimizu Y, Furuhata E, Maeda S, Kishima M, Nishimura H, et al. RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. Blood Adv 2017;1:1699–711.
52. Rasmussen KD, Berest I, Kebetaler S, Nishimura K, Simon-Carrasco L, Vassiliou GS, et al. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. Genome Res 2019;29:564–75.
53. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 2010;18:553–67.
54. Li Y, Okuno Y, Zhang P, Radomska HS, Chen H, Iwasaki H, et al. Regulation of the PU.1 gene by distal elements. Blood 2001;98:2958–65.
55. Walter MJ, Park JS, Ries RE, Lau SK, McLellan M, Jaeger S, et al. Reduced PU.1 expression causes myeloid progenitor expansion and increased leukemia penetrance in mice expressing PML-RARalpha. Proc Natl Acad Sci U S A 2005;102:12513–8.
56. An N, Khan S, Imgruet MK, Gurbuxani SK, Konecki SN, Burgess MR, et al. Gene dosage effect of CUX1 in a murine model disrupts HSC

homeostasis and controls the severity and mortality of MDS. Blood 2018;131:2682–97.
57. Starkova J, Madzo J, Cario G, Kalina T, Ford A, Zaliova M, et al. The identification of (ETV6)/RUNX1-regulated genes in lymphopoiesis using histone deacetylase inhibitors in ETV6/RUNX1-positive lymphoid leukemic cells. Clin Cancer Res 2007;13:1726–35.
58. Gu X, Hu Z, Ebrahem Q, Crabb JS, Mahfouz RZ, Radivoyevitch T, et al. Runx1 regulation of Pu.1 corepressor/coactivator exchange identifies specific molecular targets for leukemia differentiation therapy. J Biol Chem 2014;289:14881–95.
59. Hosokawa H, Ungerback J, Wang X, Matsumoto M, Nakayama KI, Cohen SM, et al. Transcription factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. Immunity 2018;48:1119–34.
60. Yun H, Narayan N, Vohra S, Giotopoulos G, Mupo A, Madrigal P, et al. Mutational synergy during leukemia induction remodels chromatin accessibility, histone modifications and three-dimensional DNA topology to alter gene expression. Nat Genet 2021;53:1443–55.
61. Coughlan AM, Harmon C, Whelan S, O'Brien EC, O'Reilly VP, Crotty P, et al. Myeloid engraftment in humanized mice: impact of granulocyte-colony stimulating factor treatment and transgenic mouse strain. Stem Cells Dev 2016;25:530–41.
62. Shultz LD, Lyons BL, Burzenski LM, Gott B, Chen X, Chaleff S, et al. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. J Immunol 2005;174:6477–89.
63. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012;9:671–5.
64. Piaton E, Fabre M, Goubin-Versini I, Bretz-Grenier MF, Courtade-Saidi M, Vincent S, et al. Technical recommendations and best practice guidelines for May-Grünwald-Giemsa staining: literature review and insights from the quality assurance. Ann Pathol 2015;35:294–305.
65. Jeong JJ, Gu X, Nie J, Sundaravel S, Liu H, Kuo WL, et al. Cytokine-regulated phosphorylation and activation of TET2 by JAK2 in hematopoiesis. Cancer Discov 2019;9:778–95.
66. Uhrig S, Ellermann J, Walther T, Burkhardt P, Frohlich M, Hutter B, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res 2021;31:448–60.
67. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
68. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
69. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.
70. Faust GG, Hall IM. SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics 2014;30:2503–5.
71. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics 2015;31:2032–4.
72. Reske JJ, Wilson MR, Chandler RL. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. Epigenetics Chromatin 2020;13:22.
73. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. Bioinformatics 2011;27:1571–2.
74. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38:576–89.
75. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.