







# Inducible disruption of *Tet* genes results in myeloid malignancy, readthrough transcription, and a heterochromatin-to-euchromatin switch

Hiroshi Yuita<sup>a,1,2</sup>, Isaac F. López-Moyado<sup>a,b,1</sup>, Hyeongmin Jeong<sup>c,1</sup> , Arthur Xiuyuan Cheng<sup>a,b</sup>, James Scott-Browne<sup>d,e</sup>, Jungeun An<sup>f</sup> , Toshinori Nakayama<sup>g,h</sup> , Atsushi Onodera<sup>a,g,i,3</sup>, Myunggon Ko<sup>c,j,3</sup>, and Anjana Rao<sup>a,b,3</sup> 

Contributed by Anjana Rao; received August 29, 2022; accepted December 29, 2022; reviewed by Grant Challen and Golnaz Vahedi

The three mammalian TET dioxygenases oxidize the methyl group of 5-methylcytosine in DNA, and the oxidized methylcytosines are essential intermediates in all known pathways of DNA demethylation. To define the *in vivo* consequences of complete TET deficiency, we inducibly deleted all three *Tet* genes in the mouse genome. *Tet1/2/3-inducible TKO (iTKO)* mice succumbed to acute myeloid leukemia (AML) by 4 to 5 wk. Single-cell RNA sequencing of *Tet iTKO* bone marrow cells revealed the appearance of new myeloid cell populations characterized by a striking increase in expression of all members of the *stefin/cystatin* gene cluster on mouse chromosome 16. In patients with AML, high *stefin/cystatin* gene expression correlates with poor clinical outcomes. Increased expression of the clustered *stefin/cystatin* genes was associated with a heterochromatin-to-euchromatin compartment switch with readthrough transcription downstream of the clustered *stefin/cystatin* genes as well as other highly expressed genes, but only minor changes in DNA methylation. Our data highlight roles for TET enzymes that are distinct from their established function in DNA demethylation and instead involve increased transcriptional readthrough and changes in three-dimensional genome organization.

TET proteins | myeloid expansion | readthrough transcription | heterochromatin-to-euchromatin transition | Stefins

The biochemical function of Ten-eleven translocation (TET) methylcytosine dioxygenases is to convert 5-methylcytosine to 5-hydroxymethylcytosine (5hmC) and other oxidized methylcytosines in DNA, culminating in DNA demethylation (1–3). TET enzymes are recruited to active enhancers by transcription factors, as shown for TET2 in several hematopoietic cell types (4–7), and TET2 has been shown to bind enhancer regions in mouse embryonic stem cells (8). Moreover, TET2 has been shown to associate with the SET/COMPASS complex, an H3K4 methyltransferase complex that travels with RNA polymerase II (9). In consequence, 5hmC is primarily found in highly transcribed regions in euchromatin, including the gene bodies of highly transcribed genes and the most active enhancers marked by high levels of H3K27 acetylation (H3K27Ac) (10).

TET enzymes have repeatedly been shown to function as tumor suppressors. *TET2* loss-of-function mutations are frequent in diverse human hematopoietic malignancies, including myelodysplastic syndrome (MDS), myeloproliferative neoplasms, chronic myelomonocytic leukemia (CMML), diffuse large B cell lymphoma, peripheral T cell lymphoma, angioimmunoblastic T cell lymphoma, and acute myeloid leukemia (AML), among others (2, 11). In mouse models, loss of *Tet1* or *Tet2* skews the differentiation of hematopoietic stem cells (HSC) toward lymphoid or myeloid commitment and results in hematologic malignancies with features of B lymphoid and myeloid lineages, respectively (12–15). *Tet1/2* double-knockout (DKO) mice develop B cell tumors and display a median survival of 20 mo (16), whereas mice with acute, inducible deletion of *Tet2* and *Tet3* (*Tet2/3 DKO*) develop myeloid leukemias within 3 to 7 wk (17). We extended these studies to triple deletion of TET genes for two reasons: to ask whether inducible deletion of all three TET proteins in adult mice would also give rise to cancers and whether the complete absence of TET proteins would skew hematopoietic differentiation and leukemogenesis in the B cell and/or myeloid direction.

To assess these and other outcomes, we generated *Tet1<sup>fl/fl</sup> Tet2<sup>fl/fl</sup> Tet3<sup>fl/fl</sup>* (*Tet triple-floxed, Tet Tfl*) mice that also harbored *UBC-Cre-ERT2*, a fusion protein of the Cre recombinase with ERT2, an engineered estrogen receptor that is normally cytoplasmic but moves to the nucleus after tamoxifen injection. The mice also possessed a *Rosa26-YFP<sup>LSL</sup>* allele in which the Enhanced Yellow Fluorescent Protein (EYFP, hereafter referred to as YFP) gene is preceded by a *floxed* transcriptional STOP cassette; thus, YFP expression serves as a

## Significance

TET loss of function is strongly linked to cancer. We generated a mouse model of inducible deletion of all three *Tet* genes (*Tet iTKO* mice) to study the *in vivo* effects of loss of TET enzymes. *Tet iTKO* mice develop rapid and fatal myeloid expansion within 4 to 5 wk. The expanded myeloid cells include new cell populations characterized by upregulation of the *Stefin* gene cluster on mouse chromosome 16, accompanied by a heterochromatin-to-euchromatin switch encompassing the same region. *Stefin/cystatin* genes encode cysteine protease inhibitors that inhibit cathepsin proteases among others. We show that high *stefin/cystatin* gene expression correlates with poor clinical outcomes in acute myeloid leukemia patients and discuss the connection of TET deficiency with changes in genome organization.

Reviewers: G.C., Washington University in St. Louis; and G.V., Raymond and Ruth Perelman School of Medicine at the University of Pennsylvania.

Competing interest statement: The authors have organizational affiliations to disclose. A.R. is on the scientific advisory board of Cambridge Epigenetics (Cambridge, UK). The other authors declare no competing interests.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

<sup>1</sup>H.Y., I.F.L.-M., and H.J. contributed equally to this work.

<sup>2</sup>Present address: Shinagawa R&D Center, Daiichi Sankyo Co., Ltd, Tokyo 140-8710, Japan.

<sup>3</sup>To whom correspondence may be addressed. Email: onodera@lji.org, mgko@unist.ac.kr, or arao@lji.org.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2214824120/-DCSupplemental>.

Published February 1, 2023.

reporter for nuclear translocation and activation of Cre-ERT2. After tamoxifen injection, 100% of injected *Tet1/2/3 inducible TKO (iTKO)* mice succumbed to AML within 4 to 5 wk. Loss of all TET family enzymes in hematopoietic stem/precursor cells (HSPC) led to the appearance of a new population of myeloid cells bearing neutrophil markers that displayed a striking increase in expression of the *stefin/cystatin* gene cluster on mouse chromosome 16, coincident with a compartment switch of this region from heterochromatin to euchromatin as judged by Hi-C analysis. Additionally, we document the occurrence of transcriptional read-through downstream of the *stefin/cystatin* gene cluster and other highly expressed genes in *Tet iTKO* cells, consistent with a previous report that readthrough transcription causes heterochromatin-to-euchromatin transitions. Finally, we showed that in patients with AML, there was a clear association of *stefin/cystatin* gene expression with clinical outcome: high CSTA, CSTB, CST3, and CST7 expression was associated with increased percentages of monocytes and neutrophils in peripheral blood (PB), and genomic amplification or increased expression of CSTB was associated with decreased patient survival.

## Results

**Inducible Deletion of All Three *Tet* Genes in Mouse HSPC Results in Myeloid, Not Lymphoid, Leukemia.** To determine whether complete elimination of all TET proteins and activity would skew hematopoietic differentiation and leukemogenesis in the B cell or myeloid direction, we treated *Tet Tfl Rosa26-YFP<sup>LSL</sup> UBC-Cre-ERT2* mice with tamoxifen for 5 d. The experiments were performed in two separate locations, the United States and South Korea, respectively, with identical results. After tamoxifen injection, there were no notable abnormalities in the hematopoietic systems of control *Tet Tfl* or *Tet Tfl Rosa26-YFP<sup>LSL</sup>* mice that lacked Cre-ERT2 expression, defined here as wild type (WT) for convenience. In contrast, between 20 and 40 d after the final tamoxifen injection, 100% of injected *Tet1/2/3 iTKO* mice became very sick and were sacrificed, whereas all WT mice survived (Fig. 1*A*). Five weeks after tamoxifen treatment, cells from bone marrow (BM) and spleen of the injected mice showed complete loss of all three *Tet* mRNAs (SI Appendix, Fig. S1*A*) as well as almost complete loss of 5hmC by anti-5hmC dot blot (18) (SI Appendix, Fig. S1*B*).

In BM, spleen, and PB, the percentages of white-blood cells were hugely increased relative to control mice, and the vast majority were Gr-1<sup>+</sup> Mac-1<sup>+</sup> myeloid cells (Fig. 1*B* and *C* and SI Appendix, Fig. S1*C* and *D*). The frequencies of c-kit<sup>+</sup> Mac-1<sup>+</sup> myeloid progenitors were also substantially increased (Fig. 1*B* and *C*). In all *Tet TKO* mice, myeloid expansion was accompanied by progressive and massive splenomegaly, enlargement of the liver and lungs, and lymph node hypertrophy (SI Appendix, Fig. S1*E*); the bones were pale, and the mice developed profound anemia and lymphopenia (SI Appendix, Fig. S1*F*). Injection of control and *Tet iTKO* mice with BrdU, a thymidine analog that is incorporated into DNA during replication, followed by organ harvest 16 h later, resulted in a higher fraction of BrdU-positive CD11b<sup>+</sup> (Mac-1<sup>+</sup>) and granulocyte–macrophage progenitors (GMP) cells in *Tet iTKO* compared to WT mice (SI Appendix, Fig. S1*G*), indicating that *Tet iTKO* myeloid cells proliferate to a greater extent than WT cells. Since CD11b is one component of the Mac-1 heterodimer, we use the designations Mac-1<sup>+</sup> and CD11b<sup>+</sup> interchangeably in the figures and text hereafter. Analysis of cellular populations by flow cytometry showed that the frequencies of Gr-1<sup>+</sup> Mac-1<sup>+</sup> myeloid cells were increased in BM and spleen (Fig. 1*B* and *C*, *Top*). The BM also showed a decreased frequency of Ter-119<sup>+</sup> CD71<sup>+</sup> erythroid precursors (Fig. 1*B*, *2nd panel*) and CD19<sup>+</sup> B220<sup>+</sup> B cells

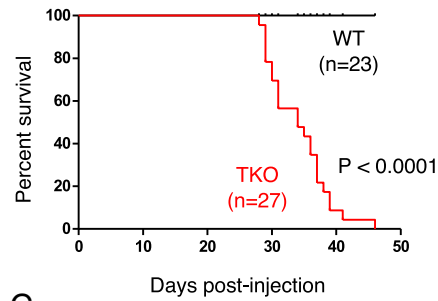
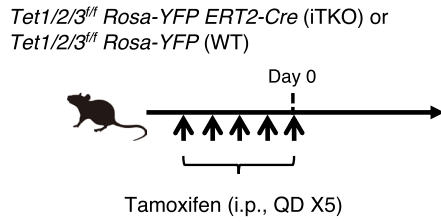
(Fig. 1*B*, *3rd panel*), while the spleen showed decreased frequencies of CD19<sup>+</sup> B220<sup>+</sup> B cells and CD8<sup>+</sup> T cells (Fig. 1*C*, SI Appendix, Fig. S1*G*). The Mac-1<sup>+</sup> population in BM and spleen (Fig. 1*B* and *C*, *Top*) expressed more c-Kit, a hematopoietic progenitor cell marker (*Bottom*). Overall, these results indicate that *Tet iTKO* mice develop a primarily myeloid malignancy, thus resembling *Tet2/3 iDKO* mice in which *Tet2/3* deficiency is induced with poly-I:C or tamoxifen (17).

In BM of *Tet iTKO* mice, Lin<sup>−</sup> Sca-1<sup>+</sup> c-Kit<sup>+</sup> (LSK) cells, which are enriched for HSCs and multipotent progenitors (MPPs), were decreased in both frequency and number (SI Appendix, Fig. S2*A* and *B*), as were megakaryocyte–erythrocyte progenitors (MEPs) (SI Appendix, Fig. S2*A* and *B*), long-term and short-term HSC (LT- and ST-HSC, SI Appendix, Fig. S2*C* and *D*) and common lymphoid progenitors (CLP) (SI Appendix, Fig. S2*E* and *F*), whereas the frequencies and numbers of GMP, a myeloid-committed progenitor cell, were increased (SI Appendix, Fig. S2*A* and *B*). These features are very similar to those noted previously in *Tet2/3 iDKO* mice (17). Subdivision of MPPs into MPP2, MPP3, and MPP4 subpopulations (19) showed decreased frequencies and numbers of MPP2 and MPP4 cells, which yield MEPs and lymphoid precursors, respectively (SI Appendix, Fig. S2*C–F*), leading to a relative increase in the frequency of MPP3 cells, myeloid-biased MPPs similar to pregranulocyte/macrophage (pre-GM) progenitors, within the LSK population (19). These results are consistent with our previous findings that genes with a pre-GM gene signature are up-regulated in *Tet2/3 iDKO* LSK cells, compared to WT LSK, whereas those with lymphoid and premegakaryocyte/erythrocyte gene signatures are down-regulated (17). A schematic showing the skewed differentiation is shown in SI Appendix, Fig. S2*G*. Overall, we conclude that mice with triple *TET* deletion are distinct from *Tet1/2 DKO* mice but resemble *Tet2/3 DKO* mice in displaying primarily myeloid rather than lymphocyte (B or T cell) expansion.

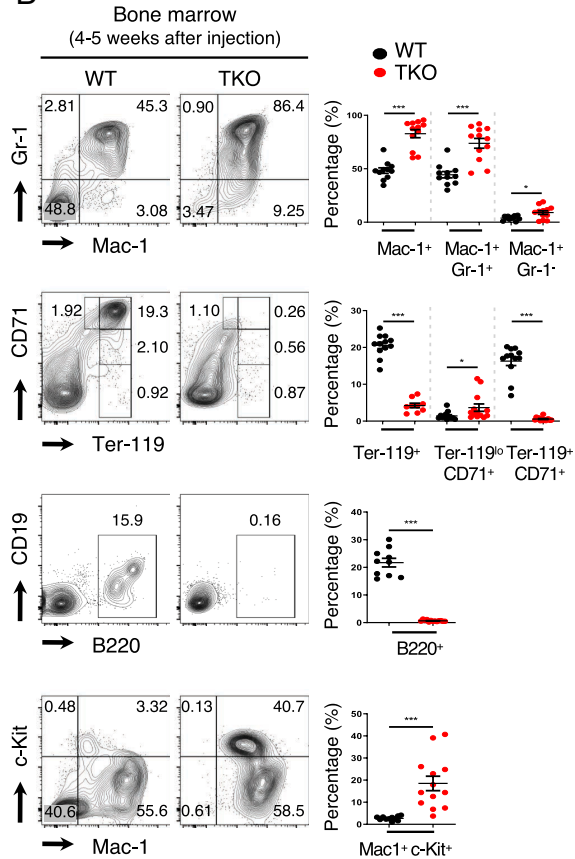
To identify regions of altered chromatin accessibility as well as transcription factors in LSK cells potentially involved in myeloid expansion observed after *TET* deletion, we performed ATAC-seq on LSK cells from *Tet2<sup>fl/fl</sup> Tet3<sup>fl/fl</sup> Mx1-Cre (Tet2/3 Dfl, Tet2/3 iDKO)* mice 3 wk after poly-I:C treatment (17) (SI Appendix, Fig. S3*A*). The 661 regions that showed increased accessibility in *Tet2/3 iDKO* LSK cells compared to WT (SI Appendix, Fig. S3*B*) displayed enrichment of consensus binding motifs for CCAAT/enhancer binding protein (C/EBP) and PU.1, two transcription factors that are essential during the period when cells commit to the myeloid lineage from HSPC (20, 21) (SI Appendix, Fig. S3*C*). These regions corresponded to myeloid enhancers that displayed high levels of H3K27Ac in WT GMP, macrophages (MF), granulocytes (GN), and monocytes (Mono) (SI Appendix, Fig. S3*D*). In contrast, the 183 regions showing decreased accessibility in *Tet2/3 iDKO* LSK cells compared to WT corresponded to enhancers with high H3K27Ac in LT-HSC, ST-HSC, MPP, and CLP (SI Appendix, Fig. S3*D*), consistent with the strong myeloid skewing and decreased lymphoid cell populations observed in *Tet2/3 iDKO* mice (17).

Considering the similar phenotypes between our previously published (17), *Mx1-Cre*-driven *Tet2/3 iDKO* model and the presently described, *Cre-ERT2*-driven *Tet1/2/3 iTKO* models, we directly compared differences at the gene expression level between *Tet2/3* deletion and *Tet1/2/3* triple deletion (SI Appendix, Fig. S4). We lentivirally transduced BM cells from *Tet2<sup>fl/fl</sup> Tet3<sup>fl/fl</sup>* or *Tet1<sup>fl/fl</sup> Tet2<sup>fl/fl</sup> Tet3<sup>fl/fl</sup>* animals with Cre recombinase, transferred them to irradiated recipients, and performed total (ribodepleted) RNA-seq on the expanded CD11b<sup>+</sup> cells (SI Appendix, Fig. S4*A*). Consistent with the myeloid expansion that occurs in both cases, the gene expression profiles of

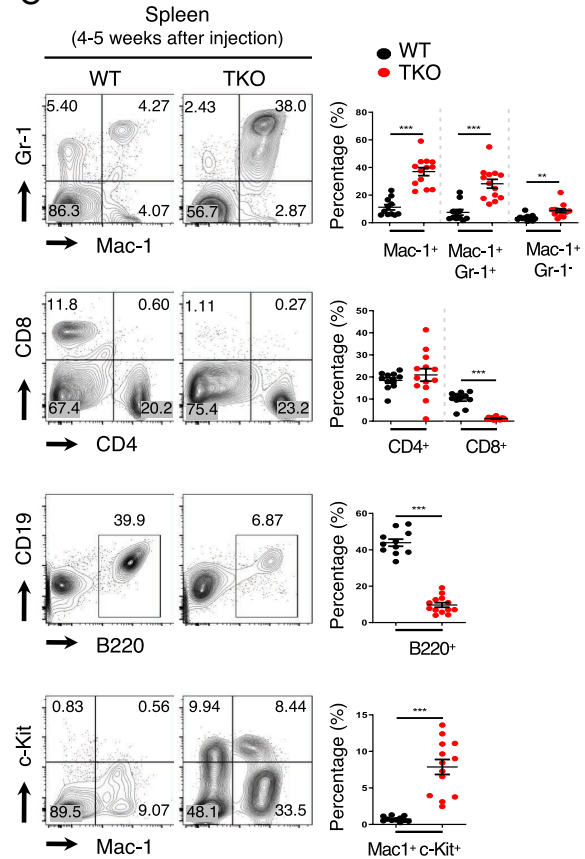
A



B



C

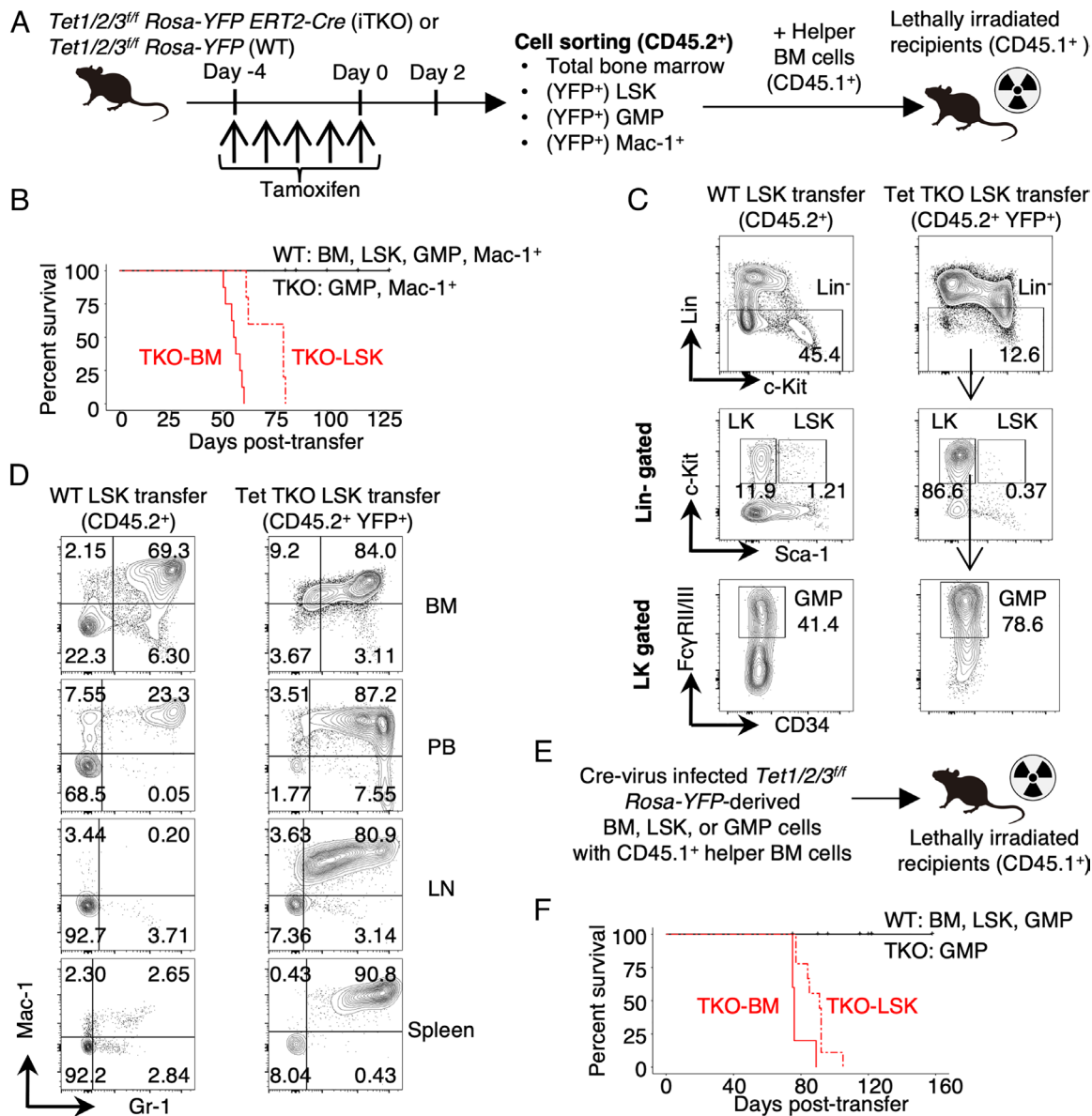


**Fig. 1.** Acute deletion of *Tet1/2/3* genes results in myeloid leukemia. (A) *Left*, flowchart of experiments. Mice of the indicated genotypes were injected daily with tamoxifen intraperitoneally, over five consecutive days, and monitored thereafter. *Right*, Kaplan-Meier curve representing percent survival of WT (n = 23) and *Tet iTKO* (n = 27) mice over time after tamoxifen injection. (B and C) Analysis of hematopoietic cell populations following simultaneous deletion of *Tet1*, *Tet2*, and *Tet3*. Flow cytometry was performed to assess myeloid (Gr-1/Mac-1), erythroid (CD71/Ter-119), B (B220/CD19), T (CD4/CD8), and myeloid progenitor (c-Kit/Mac-1) cell populations in the BM (B, n = 9 ~ 13 per each genotype) and spleen (C, n = 11 ~ 13 per each genotype) of WT or *Tet iTKO* mice. Representative flow cytometry plots (*Left*) and summary graphs (*Right*) at 4 ~ 5 wk after tamoxifen injection with means ± SEM are shown. \*\*\*P < 0.005, \*\*\*\*P < 0.0005 (Student's t test).

*Tet2/3* DKO and *Tet1/2/3* TKO expanded myeloid cells were also very similar (*SI Appendix, Fig. S4 B and C*), and the gene expression changes between *Tet2/3* DKO and *Tet1/2/3* TKO CD11b<sup>+</sup> cells largely occurred in the same direction relative to WT—*Tet1/2/3* TKO and *Tet2/3* DKO myeloid cells showed 6,110 and 6,745 differentially expressed genes (DEGs) respectively compared to WT, of which 5,265 genes were similarly differentially expressed in both comparisons (*SI Appendix, Fig. S4C*). Only 23 DEGs displayed opposite changes in gene expression in *Tet1/2/3* TKO and *Tet2/3* DKO myeloid cells with respect to WT, e.g., only four DEGs were significantly up-regulated in *Tet2/3* DKO and simultaneously down-regulated in *Tet1/2/3* TKO compared to WT myeloid cells (*SI Appendix, Fig. S4C*).

***Tet iTKO*-Related Myeloid Leukemia Arises from Early Progenitor HSPCs.** To ask what cell populations could transfer the myeloid

leukemia early after *TET* deletion, we isolated LSK, GMP, and Mac-1<sup>+</sup> cells from mice 2 d after final tamoxifen administration and transferred them to lethally irradiated CD45.1<sup>+</sup> mice together with WT CD45.1<sup>+</sup> BM cells (Fig. 2A). At this early time after tamoxifen treatment, the cell populations derived from BM of *Tet iTKO* mice before transplantation were phenotypically identical to those derived from BM of WT mice (*SI Appendix, Fig. S2H*). *TET*-deficient BM or LSK cells transferred leukemia, and the mice succumbed within 60 to 90 d, but *TET*-deficient GMP and Mac-1<sup>+</sup> cells failed to induce leukemia and the recipient mice survived beyond 125 d after cell transfer (Fig. 2B). In mice receiving WT LSK cells, CD45.2<sup>+</sup> donor cells were observed in multiple organs—BM, PB, lymph nodes (LN), and spleen (Fig. 2C and D), whereas in mice transplanted with *Tet iTKO* LSK cells, immature GMP-like and more mature Mac-1<sup>+</sup> cells were dominantly observed in several organs (Fig. 2C and D).



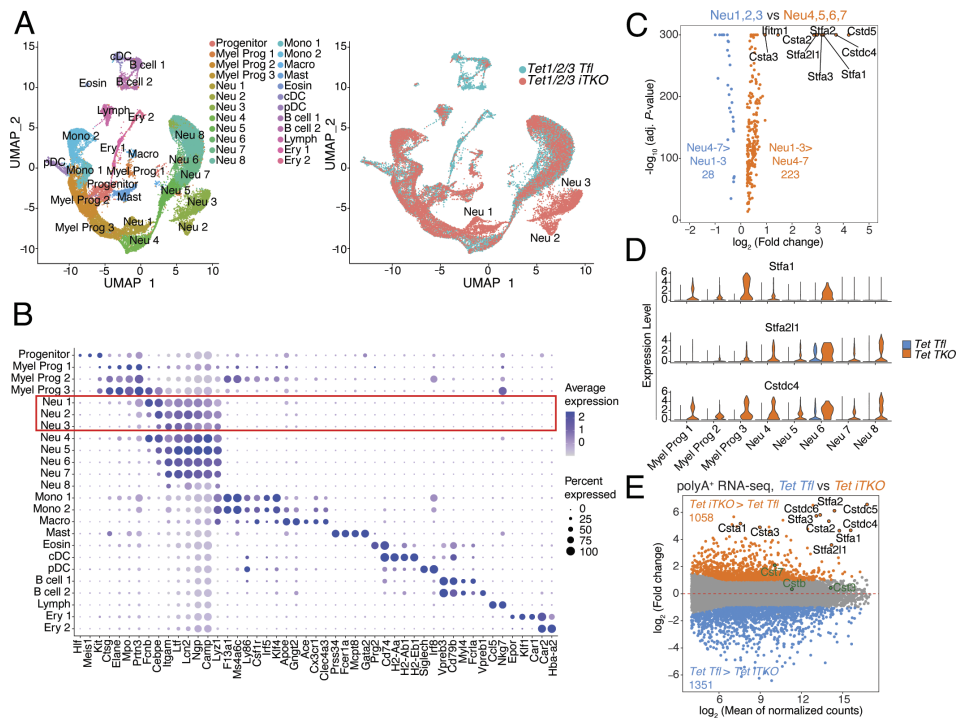
**Fig. 2.** Only early HSPCs of *Tet iTKO* mice can transfer myeloid leukemia. (A) Flowchart of experiments in which whole BM cells or LSK, GMP, or Mac-1<sup>+</sup> cells isolated from WT and *Tet iTKO* mice were transferred into irradiated recipient mice. Transfer of *Tet iTKO* BM or LSK cells induces myeloid malignancy within 100 d, whereas transfer of *Tet iTKO* GMP or Mac-1<sup>+</sup> cells does not. (B) Kaplan–Meier survival plot of mice transplanted with WT or *Tet iTKO* cells;  $2 \times 10^6$  total BM cells,  $1 \times 10^6$  Mac-1<sup>+</sup>, 5,000 GMP, or 2,000 LSK cells from control (WT) and *Tet iTKO* mice were transferred into lethally irradiated CD45.1<sup>+</sup> recipient mice together with  $2 \times 10^6$  helper BM cells from CD45.1<sup>+</sup> mice. All recipient mice given total BM (8/8) or LSK (5/5) cells from *Tet iTKO* mice, but none of the mice that received GMP or Mac-1<sup>+</sup> cells from *Tet iTKO* mice, developed myeloid leukemia. (C and D) Representative flow cytometric analysis of BM HSPCs (C) and myeloid lineage cells (D) in BM, PB, LN, or spleen of recipient mice 7 to 8 wk after adoptive transfer. (E) Flowchart of experiments in which whole BM cells, LSK, or GMP cells from WT or *Tet Tfl Rosa26-YFP<sup>LSL</sup>* mice were lentivirally transduced with Cre before adoptive transfer into irradiated recipient mice. (F) Kaplan–Meier survival plot of mice transplanted with Cre- or control-transduced cells from WT or *Tet Tfl Rosa26-YFP<sup>LSL</sup>* mice.  $2 \times 10^6$  BM cells, 5,000 GMP, or 2,000 LSK from *Tet Tfl Rosa26-YFP<sup>LSL</sup>* mice were isolated by FACS and infected with Cre lentivirus and then transferred into lethally irradiated CD45.1<sup>+</sup> recipient mice together with  $2 \times 10^6$  helper BM cells from CD45.1<sup>+</sup> mice. All recipient mice given total BM (11/11) or LSK (9/9) cells, but none of the mice that received GMP cells from *Tet iTKO* mice or BM, LSK, or GMP cells from WT control cells, developed myeloid leukemia.

These phenotypes were nearly identical to those of the parental (donor) mice (Fig. 1 and *SI Appendix, Fig. S1*), indicating that early progenitor HSPCs rather than GMP-like cells are the cell of origin in the myeloid leukemias developing in *Tet iTKO* mice.

Because the Cre-ERT2 system has some drawbacks (for instance, the rapid decline in erythropoiesis), we established a second model system to confirm our observations. *Tet Tfl Rosa26-YFP<sup>LSL</sup>* cells were lentivirally transduced with Cre recombinase, yielding a mixed population of untransduced *Tet Tfl* and transduced *Tet TKO* progenitor cells; this is more relevant as a cancer model since only a subset of progenitor cells develop somatic *Tet* gene deletion. Whole BM from *Tet Tfl Rosa26-YFP<sup>LSL</sup>* mice, or

LSK or GMP cells purified from BM of these mice by fluorescence-activated cell sorting (FACS), was infected with a Cre-expressing lentivirus, and the transduced cells were transferred into lethally irradiated CD45.1<sup>+</sup> mice together with WT CD45.1<sup>+</sup> BM cells (Fig. 2E). Again, TET-deficient LSK cells, but not GMP cells, were capable of transferring leukemia, and the mice succumbed within 120 d (Fig. 2F).

**Cell-Autonomous Leukemia Development in *Tet iTKO* Mice.** To demonstrate cell-autonomous leukemia development in *Tet iTKO* mice, we isolated red blood cell-depleted leukemic splenocytes from WT or *Tet iTKO* mice at 4 wk following tamoxifen administration



**Fig. 3.** Appearance of new myeloid cell populations in *Tet iTKO* BM. (A) Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction for single-cell RNA-seq data, displaying the different identified populations (*Left*) and the new cell populations present in *Tet iTKO* BM (*Right*). Myel. Prog., myeloid progenitor; Neu, neutrophil; Mono, monocyte; Macro, macrophage; Mast, mast cell; Eosin, eosinophil; cDC, conventional dendritic cell; pDC, plasmacytoid dendritic cell; Lymph, lymphoid progenitor; Ery, erythroid progenitor. (B) Expression of key marker genes across the different cell populations identified in (A). The dot size encodes the percentage of cells within each cell population that show expression of the indicated gene, and the color intensity encodes the average expression level of the indicated gene across all cells within each cell population. (C) Volcano plot comparing gene expression in cell populations Neu1 to Neu3 (specific to *Tet iTKO*) compared to Neu4 to Neu7, using single-cell RNA-seq data. (D) Expression levels of *stefin/cystatin* genes *Stfa1*, *Stfa211*, and *Cstcd4* in WT and *Tet iTKO* cells, across different myeloid progenitor and neutrophil cell populations. (E) Mean average plots of pairwise comparison of gene expression determined by polyA<sup>+</sup> RNA-seq in *Tet iTKO* vs *Tet Tfl* CD11b<sup>+</sup> cells. P-values were calculated using the Wald test (as implemented in DESeq2), and adjusted using the Benjamini-Hochberg method. DEGs (adjusted  $P < 0.05$ , fold change ( $\log_2$ , scale)  $\geq 1$  or  $\leq -1$ ) are highlighted in orange (*Tet iTKO* > WT) or blue (WT > *Tet iTKO*). Differentially expressed *Stefin/Cystatin* genes found on mouse chromosome 16 are labeled.

and transferred them to sublethally irradiated CD45.1<sup>+</sup> recipient mice, respectively (*SI Appendix, Fig. S5*). None of the recipient mice transplanted with *Tet iTKO* splenocytes survived longer than 25 d after transplantation (*SI Appendix, Fig. S5A*). Similar to acute *TET* deletion in primary mice, mice receiving *Tet TKO* cells showed leukocytosis and enlarged spleen and liver (*SI Appendix, Fig. S5 B and C*). The leukocytosis was associated with significant expansion of myeloid lineage cells and with severe anemia (*SI Appendix, Fig. S5 D and E*). Furthermore, mice receiving *Tet iTKO* cells showed expansion of Mac-1<sup>+</sup> cells in the BM, spleen, and PB, whereas cells of lymphoid and erythroid lineages were reduced (*SI Appendix, Fig. S5 F–H*). Again, *Tet* deficiency led to a substantial increase in the proportion of cells expressing the progenitor marker, c-Kit (*SI Appendix, Fig. S5 F and G, Bottom*) in BM and spleen, and to an increase in the percentage of GMP cells within Lin<sup>−</sup>c-Kit<sup>+</sup> (LK) cells (*SI Appendix, Fig. S5I*). Similar results were obtained when leukemic BM cells from WT or *Tet iTKO* mice were transferred into recipient mice (*SI Appendix, Fig. S6*). Together, these results show that inducible deletion of all three *TET* genes results in the rapid, cell-autonomous development of leukemia in *Tet iTKO* mice.

**Increased Expression of the *Stefin/Cystatin* Gene Cluster in *TET iTKO* BM Cells.** We used single-cell RNA-sequencing (scRNA-seq) to determine whether cell populations in BM were altered in *Tet iTKO* compared to WT mice (Fig. 3 *A and B* and *SI Appendix, Fig. S7A*). We profiled a total of 40,151 single cells from total BM of two WT and two *Tet iTKO* mice using the 10x genomics

platform (*SI Appendix, Fig. S7A* and *Methods*) and partitioned them into 24 different cell populations (Fig. 3 *A and B*) based on lineage markers known to be highly expressed in specific hematopoietic cell populations (22) (Fig. 3*B* and *SI Appendix, Fig. S7 B and C*). For instance, the erythroid Ery 1 population was defined based on expression of the known erythroid genes *Epor*, *Klf1*, *Car1*, and *Car2* (carbonic anhydrase 1 and 2) and *Hba-a2* (hemoglobin A) (Fig. 3*B* and *SI Appendix, Fig. S7 B and C*). In agreement with the data from flow cytometry (Fig. 1*B*), *Tet iTKO* BM showed a striking decrease in erythroid (Ery1 and Ery2), lymphoid (lymph, B cell 1, and B cell 2), and certain myeloid (cDC) populations (*SI Appendix, Fig. S7C*).

Notably, in *Tet iTKO* BM cells, we observed three new populations of myeloid cells largely absent in WT cells, which we labeled Neu1, Neu2, and Neu3, based on the fact that they expressed at least nine neutrophil markers (Fig. 3 *A and B*). We also identified five additional cell populations, Neu4–8, in WT cells expressing the same nine neutrophil markers. We then directly compared gene expression in these two sets of neutrophil populations, the novel cell subpopulations Neu1 to Neu3 present only in *Tet iTKO* BM cells and the subpopulations Neu4 to Neu7 represented predominantly in WT cells. Of nine genes strongly up-regulated ( $\log_2$  fold change > 1) in the novel Neu1 to Neu3 myeloid populations observed upon *Tet* triple deletion, eight were members of the *stefin/cystatin* gene cluster located on mouse chromosome 16 (Fig. 3*C* and *SI Appendix, Figs. S7D and S8A*), a family of cysteine protease inhibitors whose expression is often dysregulated in human cancers (23–25). The new myeloid populations Neu1 to

Neu3 showed the highest levels of expression of the *stefin/cystatin* gene cluster on mouse chromosome 16 (SI Appendix, Fig. S7D); however, upregulation of these genes was not limited to the Neu1 to Neu3 populations since we also observed increased expression of these genes in *Tet iTKO* cells compared to WT cells in other myeloid populations (Fig. 3D). We confirmed this observation arising from scRNA-seq data analysis by performing bulk polyA<sup>+</sup> RNA-sequencing of expanded CD11b<sup>+</sup> cells from *Tet iTKO* and WT mice, isolated ~4 wk after tamoxifen injection, which revealed that among the 1,058 differentially up-regulated genes were all 10 members of the *stefin/cystatin* gene cluster on mouse chromosome 16 (Fig. 3E); similar results were observed in bulk total (ribodepleted) RNA-seq from Cre-transduced *Tet1<sup>fl/fl</sup> Tet2<sup>fl/fl</sup> Tet3<sup>fl/fl</sup>* cells (SI Appendix, Fig. S4F). In the mouse genome, in addition to the chromosome 16 cluster mentioned above, there is a second *stefin/cystatin* gene cluster on mouse chromosome 2 (SI Appendix, Fig. S8B) and two isolated *stefin/cystatin* genes, *Cstb* and *Cst6*, located on chromosomes 10 and 19, respectively. Of these additional genes, we were able to detect expression of only *Cst3*, *Cstb*, and *Cst7*, and *Cst7* showed perceptible upregulation only in the polyA<sup>+</sup> RNA-seq dataset (Fig. 3E and SI Appendix, Fig. S4F).

*Stefin/cystatin* genes were previously reported to be up-regulated in other mouse models of aggressive AMLs (26). To define the stage of differentiation at which the *stefin/cystatin* genes were up-regulated, we analyzed RNA-seq data from LSK cells isolated from WT and *Tet iDKO* mice (17) 2.5 wk after *Tet2/3* deletion (SI Appendix, Fig. S7E). Three genes in the *stefin/cystatin* gene cluster—*Stfa1*, *Cstbc5*, and *Cstbc6*—were significantly up-regulated even in *Tet iDKO* LSK cells, which are enriched in HSPC, indicating that the change in expression of this gene cluster was initiated at the early precursor stage in HSPC, prior to myeloid differentiation.

**Role of the *Stefin/Cystatin* Gene Cluster in Human AML.** *Stefin/cystatin* gene products are cysteine protease inhibitors whose functions have been primarily studied in the context of inhibition of the cathepsin family of cysteine proteases, which have critical roles in protein turnover within the lysosome as well as in other intracellular and extracellular locations (27). *Stefins/cystatins* have been categorized into two classes based on their structure (25): Type I cystatins (also known as *Stefins*) are unglycosylated, primarily intracellular proteins that are ~100 amino acids in length and lack disulfide bonds, whereas type II cystatins are extracellular inhibitors that are ~120 amino acids in length and have two intrachain disulfide bonds. SI Appendix, Fig. S8C shows a genome browser view and table of the human cystatin genes located in a cluster on human chromosome 20 and also lists three isolated cystatin genes, *CSTA*, *CSTB*, and *CST6* located on chromosomes 3, 21, and 11, respectively. New members of the *stefin/cystatin* gene family are still emerging with each new genome annotation, and in many cases, the subcellular localizations and functions of the newly annotated gene products remain obscure.

To explore whether *stefin/cystatin* expression could play a role in human myeloid malignancies, we examined the association between *stefin/cystatin* mRNA expression and the clinical features observed in patients with AML (28). There was a positive correlation of *CSTA* mRNA expression with the percentage of circulating monocytes and neutrophils present in the PB of AML patients (Fig. 4A and B). A similar positive correlation was observed between *CSTB* and *CST3* expression and the percentage of circulating monocytes (Fig. 4A) and between *CST7* and neutrophil percentage (Fig. 4B). In a different AML dataset (29), we observed that *CSTB* genomic gain/amplification was associated with shorter overall survival (Fig. 4C, median survival of 4.96 mo in patients with *CSTB* amplification, vs. 12 mo in patients with two copies of the *CSTB* locus).

Additionally, when stratifying patients by *CSTB* expression (1st quartile vs. 4th quartile), high *CSTB* expression (4th quartile) was associated with shorter survival (median survival of 7.96 mo, compared to 16.08 mo in the 1st quartile) (Fig. 4D).

Given these data, we attempted a functional analysis of the role of *stefin/cystatin* genes in *Tet iTKO* cells. We generated *Tet Tfl Rosa26-YFP<sup>SL</sup> Cas9-IRES-GFP Cre-ERT2* mice, treated them with tamoxifen for 5 d, isolated LK cells 2 d later, transduced them with a mixture of lentiviral vectors encoding different sgRNAs targeting individual *Stefin* genes in the chromosome 16 cluster as well as several positive and negative control sgRNAs, and transferred the transduced cells to lethally irradiated recipient mice. Approximately 7 wk after transfer, we extracted DNA from BM cells and amplified and sequenced the region containing the different sgRNAs (SI Appendix, Fig. S7F). By measuring sgRNA abundance in expanded *Tet iTKO* cells, we found that sgRNAs targeting maintenance DNA methyltransferase *Dnmt1* dropped in abundance (SI Appendix, Fig. S7G and H). However, gRNAs targeting individual *Stefin* genes were not significantly enriched or depleted relative to nontargeting sgRNAs (SI Appendix, Fig. S7G), most likely due to redundancy among genes in this family (30). Because of this redundancy, it may be challenging to determine whether *stefin/cystatin* depletion affects oncogenic progression in mouse or human systems.

**The Increased Expression of *Stefin/Cystatin* Genes Is Not Related to Changes in DNA Methylation.** We previously showed that *Tet* gene disruption was associated with a striking genome-wide alteration of DNA methylation patterns in all cell types analyzed (31). Briefly, *Tet*-deficient cells showed focal increases of DNA methylation in euchromatin (Hi-C A compartment), coupled with widespread losses of DNA methylation in heterochromatin (Hi-C B compartment) (31). To ask whether increased expression of the *stefin/cystatin* gene cluster correlated with changes in DNA CpG methylation and HiC compartments, we performed whole-genome bisulfite sequencing (WGBS) and Hi-C on WT and *Tet TKO* CD11b<sup>+</sup> cells (SI Appendix, Fig. S4A and Methods).

From the WGBS analysis, we identified 85,635 high-confidence regions that were differentially methylated (DMRs) genome-wide. Of these, 68,927 DMRs were hypermethylated, with a median size of 1,182 bp, and 16,709 DMRs were classified as hypomethylated, with a median size of 2,400 bp, consistent with the focal increases and widespread decreases in DNA methylation that we previously reported for TET-deficient cells (31). DMRs that were hypermethylated in *Tet TKO* (Cre-transduced *Tet Tfl Rosa26-YFP<sup>SL</sup>*) compared to WT control (Cre-transduced *Rosa26-YFP<sup>SL</sup>*) CD11b<sup>+</sup> cells were mostly intragenic and 20% of them overlapped a promoter region; in contrast, 63% of hypomethylated DMRs in *Tet TKO* compared to control Cd11b<sup>+</sup> cells were intergenic and distal to transcription start sites (TSS) (SI Appendix, Fig. S9A). Of the 8,549 genes with hypermethylated DMRs within 1 kb of their TSS, most (6,050; 71%) showed no changes in gene expression, and a similar number showed down-regulated or up-regulated expression compared to WT (~1,250; ~15% in each case) (SI Appendix, Fig. S9B). Moreover, there were only very minor changes in DNA methylation around the *stefin/cystatin* gene cluster in *Tet TKO* cells (Fig. 5A); data for a few hypermethylated DMRs surrounding *Stfa1*, including one directly upstream of the *Stfa* gene, are shown in Fig. 5A and SI Appendix, Fig. S9C). Thus, as we have noted previously for other models of TET deficiency (17, 32), there is not a straightforward relation between gene expression and DNA methylation at gene promoters (SI Appendix, Fig. S9B).

**The *Stefin/Cystatin* Gene Cluster Undergoes a Compartment Switch from Heterochromatin to Euchromatin in *Tet iTKO* Cells.** From the Hi-C analysis (at 50kb resolution), we found

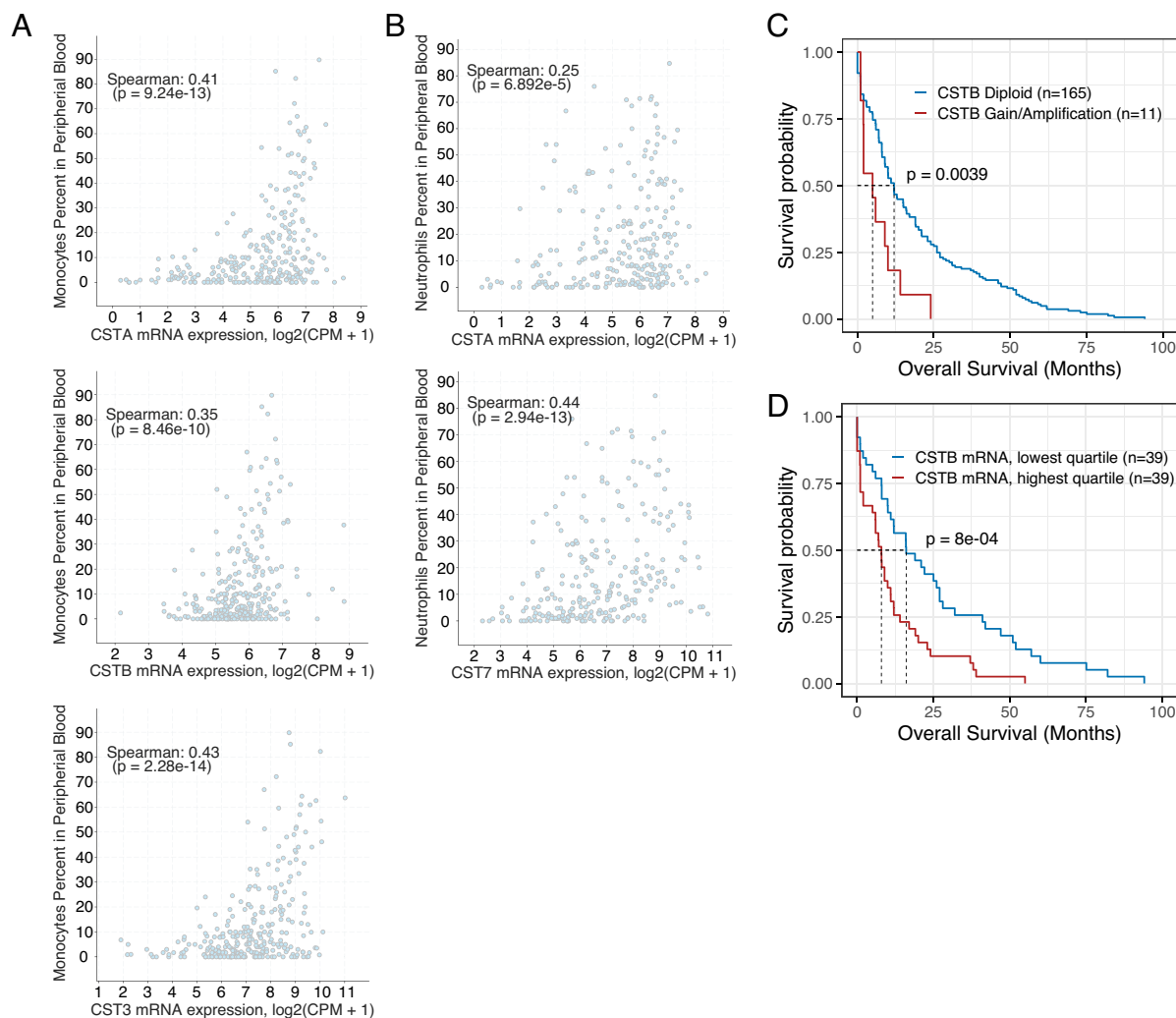
that the whole region of 300 kb encompassing the *stefin/cystatin* gene cluster underwent a change from the Hi-C B compartment in WT (*Tet Tfl* CD11b<sup>+</sup>) cells to the Hi-C A compartment in *Tet iTKO* CD11b<sup>+</sup> cells (Fig. 5 A and B). Triple *Tet* gene deletion resulted in a heterochromatin-to-euchromatin switch of this region (Fig. 5A, tracks 5 and 6), associated with a striking upregulation of the expression of all genes in the *stefin/cystatin* gene cluster (Fig. 5A, tracks 1-4; Fig. 5C). This compartment switch was observed only in the *stefin/cystatin* gene cluster and few other genomic regions, however, since at the whole-genome level, the vast majority (>98%) of 50kb windows remained in the same compartment (heterochromatin or euchromatin) following *Tet* gene deletion (Fig. 5B and *SI Appendix*, Fig. S10A); only a few (641/43,844; ~1.5%) underwent compartment switching in *Tet*-deficient compared to *Tet Tfl* (WT) CD11b<sup>+</sup> cells. Of these, most windows (537/641; 83.8%) were located in heterochromatin in *Tet Tfl* (WT) cells and moved into euchromatin upon *Tet* gene deletion (*SI Appendix*, Fig. S10A).

As expected from our previous analysis (31), the majority of genomic regions (50kb windows) that remained in euchromatin after *Tet* gene deletion (A-to-A) gained DNA methylation on average (*SI Appendix*, Fig. S10B, purple violin plots; *SI Appendix*,

Fig. S10C); this is expected since TET proteins and 5hmC are predominantly localized to euchromatin where they mediate DNA demethylation. In contrast, consistent with our previous findings (31), the majority of genomic regions that remained in heterochromatin after *Tet* gene deletion (B-to-B) showed a decrease in average DNA methylation (*SI Appendix*, Fig. S10B, blue violin plots; *SI Appendix*, Fig. S10C). At a global level, DNA methylation did not appear strikingly altered in the 50kb windows that changed genomic compartments (*SI Appendix*, Fig. S10B, green and orange violin plots; *SI Appendix*, Fig. S10C); however, at the level of individual CpGs, DNA methylation was increased at the majority of euchromatic CpGs in *Tet iTKO* cells, regardless of whether these CpGs were in euchromatin or heterochromatin in the parental *Tet Tfl* cells (see *Stefin* locus in Fig. 5A).

#### ***Tet iTKO* Cells Display ReadThrough Transcription in Highly Expressed Genes.**

In human monocyte-derived macrophages infected with influenza A, transcription into heterochromatin, particularly readthrough transcription past gene ends, was shown to be causal for compartment switching from heterochromatin to euchromatin (33). Along the same lines, by analyzing either total transcripts or nascent (unspliced) transcripts in total (ribodepleted) RNA-seq data, we observed RNA signal downstream of *stefin* genes throughout

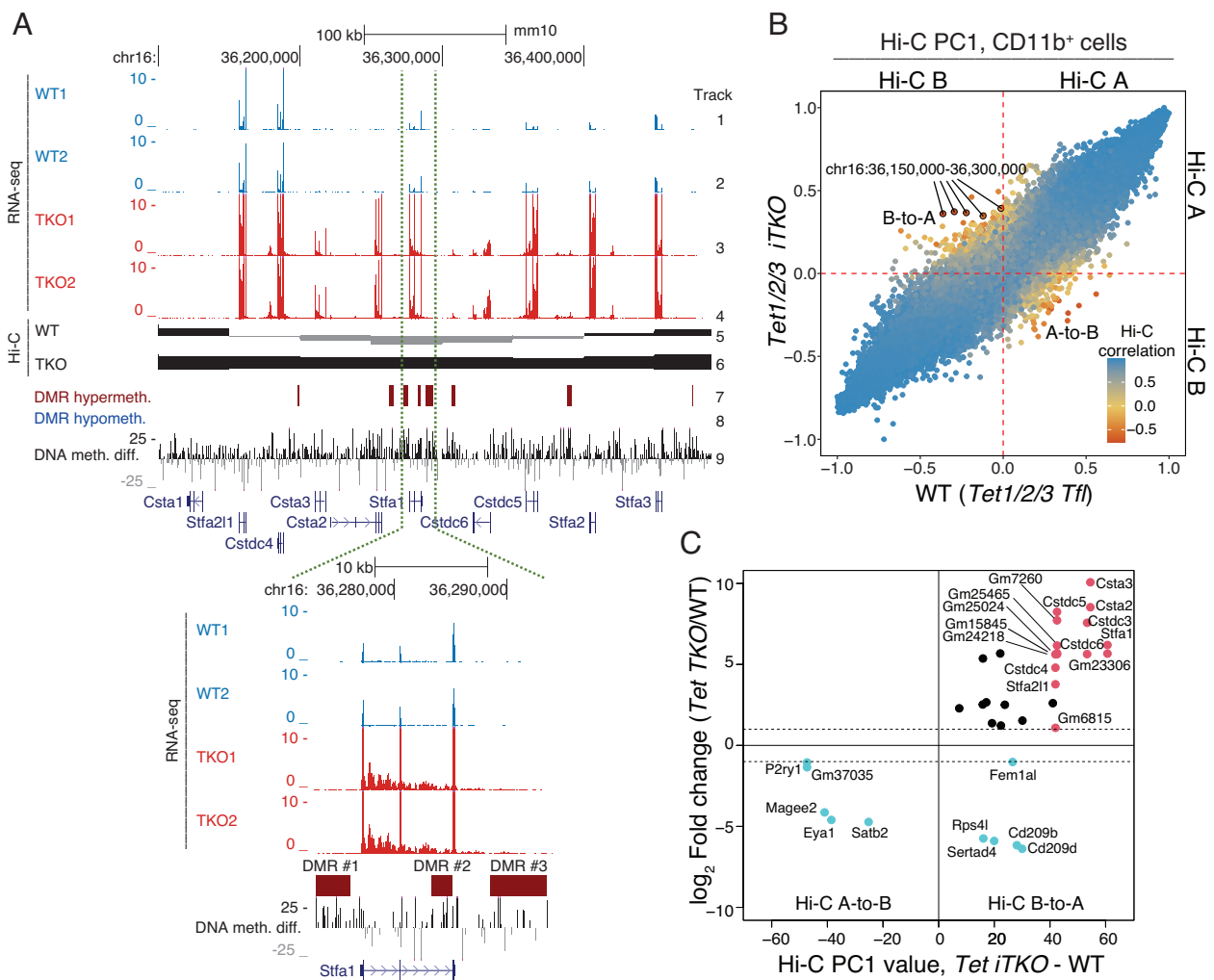


**Fig. 4.** Association between cystatin expression and clinical features in human AML. (A) Spearman correlation between CSTA, CSTB, and CST3 gene expression with percentage of monocytes in peripheral blood (PB) in AML patients. Each dot represents a different patient. (B) Spearman correlation between CSTA and CST7 gene expression with percentage of neutrophils in PB in AML patients. Each dot represents a different patient. (C) Kaplan–Meier survival curves of human AML patients, stratifying by CSTB genomic gain/amplification. *P* values were calculated using the log-rank test. (D) Kaplan–Meier survival curves of human AML patients, stratifying by highest quartile vs. lowest quartile in terms of CSTB gene expression. *P* values were calculated using the log-rank test.

the *stefin/cystatin* cluster on chromosome 16 in *Tet TKO* but not in control myeloid cells, and the RNA signal extended for kilobases past the gene ends (Fig. 6 A and B). Thus, the heterochromatin-to-euchromatin transition of the *stefin/cystatin* gene cluster in *Tet TKO* cells was accompanied by increased intronic and readthrough transcription and coordinated upregulation of *stefin/cystatin* genes.

The readthrough transcription observed in influenza-infected human macrophages was most apparent in highly transcribed genes (33, 34). Since as noted above, genes in the *stefin/cystatin* cluster were among the most differentially up-regulated genes in *Tet TKO* myeloid cells (Fig. 3 C–E and *SI Appendix*, Fig. S7D), we asked whether the readthrough transcription observed in the *stefin/cystatin* gene cluster might be a general phenomenon occurring primarily in highly transcribed genes in *Tet TKO* cells. By quantifying the ratio of transcription downstream of gene ends to that within the gene body using RNA-seq data (34) (Fig. 6C), we calculated gene expression and readthrough transcription across two biological replicates from control or *Tet TKO* cells and quantified

transcriptional readthrough in genes categorized by their expression level in each sample. The distributions of readthrough ratios were shifted toward higher values in *Tet TKO* cells compared to WT control CD11b<sup>+</sup> cells, particularly among the most highly expressed genes (Fig. 6D and *SI Appendix*, Fig. S11A). While we observed a clear shift in the distribution among the top 100 most highly expressed genes, this shift disappeared as one went down the ranking by gene expression levels (*SI Appendix*, Fig. S11A). Readthrough transcription in *Tet TKO* cells occurred in other differentially up-regulated genes, such as downstream of the *myeloperoxidase* gene (*Mpo1*) (*SI Appendix*, Fig. S11B), but it was not limited to this category, as genes equally highly expressed in *Tet TKO* and WT CD11b<sup>+</sup> cells—such as the *histone* gene cluster in chromosome 13 (*SI Appendix*, Fig. S11C), the *glutathione peroxidase-1* (*Gpx1*) gene (Fig. 6C), and the *metalloproteinase-8* (*Mmp8*) gene, in which readthrough transcription continued into the downstream gene *Mmp27* (*SI Appendix*, Fig. S11D)—displayed higher levels of transcriptional readthrough in *Tet TKO* cells.



**Fig. 5.** Compartment switching from heterochromatin to euchromatin in *Tet iTKO* cells is strongly apparent at the *stefin/cystatin* gene cluster. (A) *Top*, Genome browser view of the *stefin/cystatin* gene cluster comparing WT and *Tet iTKO* CD11b<sup>+</sup> cells. RNA expression profiles (tracks 1–4), Hi-C PC1 values (track 5–6), hypermethylated (track 7) and hypomethylated (track 8) differentially methylated regions, and DNA methylation changes (track 9) are shown. *Bottom*, Zoomed-in view of the *Stfa1* gene, comparing gene expression and DNA methylation in control (WT) and *Tet iTKO* CD11b<sup>+</sup> cells. RNA expression profiles (tracks 1–4), differentially hypermethylated regions (DMR #1–3, track 5), and DNA methylation differences (%*Tet TKO*–%WT) (track 6) are shown. (B) Scatterplot of pairwise comparison of Hi-C PC1 values between WT (x-axis) and *Tet iTKO* (y-axis) CD11b<sup>+</sup> cells; color scale indicates the correlation of the Hi-C interaction profile between both conditions (lower correlation values indicate that the same locus interacts with different regions in WT and *Tet iTKO* cells). Highlighted are the windows overlapping the *stefin/cystatin* gene cluster (chr16: 36,150,000 to 36,300,000). (C) Relation between RNA expression changes ( $\log_2$  fold change) determined by total (ribodepleted) RNA-seq and euchromatin/heterochromatin compartment changes in WT vs *Tet iTKO* CD11b<sup>+</sup> cells. Only the genes that exhibited compartment switching (from B to A on the right and from A to B on the left) and significant upregulation (red and black) or downregulation (blue) are plotted. The significant gene expression change was defined as greater than twofold change with a *P*-value < 0.05. Red dots correspond to the *Stefin/cystatin* cluster genes on chromosome 16.



Together, these observations suggest a role for TET proteins in modulating transcription termination and emphasize that increased intronic and readthrough transcription of highly expressed genes, especially of long gene clusters, may underlie heterochromatin-to-euchromatin transitions in *Tet TKO* cells.

## Discussion

In this study, we show that acute, inducible depletion of all three TET proteins results in rapid myeloid expansion, culminating in the development of aggressive AML in mice. Only BM cells and LSK cells could transfer the disease to irradiated recipient mice, suggesting that triple *Tet1/2/3* deficiency resulted in myeloid skewing, increased proliferation, and oncogenic transformation of early HSPC. Early HSPC are also thought to be cancer-initiating cells in malignancies associated with *TET2* loss-of-function mutations in humans (11) as well as in myeloid expansion arising from double *Tet2/3* deficiency (17) or *Dnmt3a* deficiency (35) in mice. Notably, the expanded myeloid cells in *Tet iTKO* mice contained new myeloid populations whose predominant feature was a heterochromatin-to-euchromatin switch, associated with increased expression of a single gene cluster encoding a subset of *stefin/cystatin* genes.

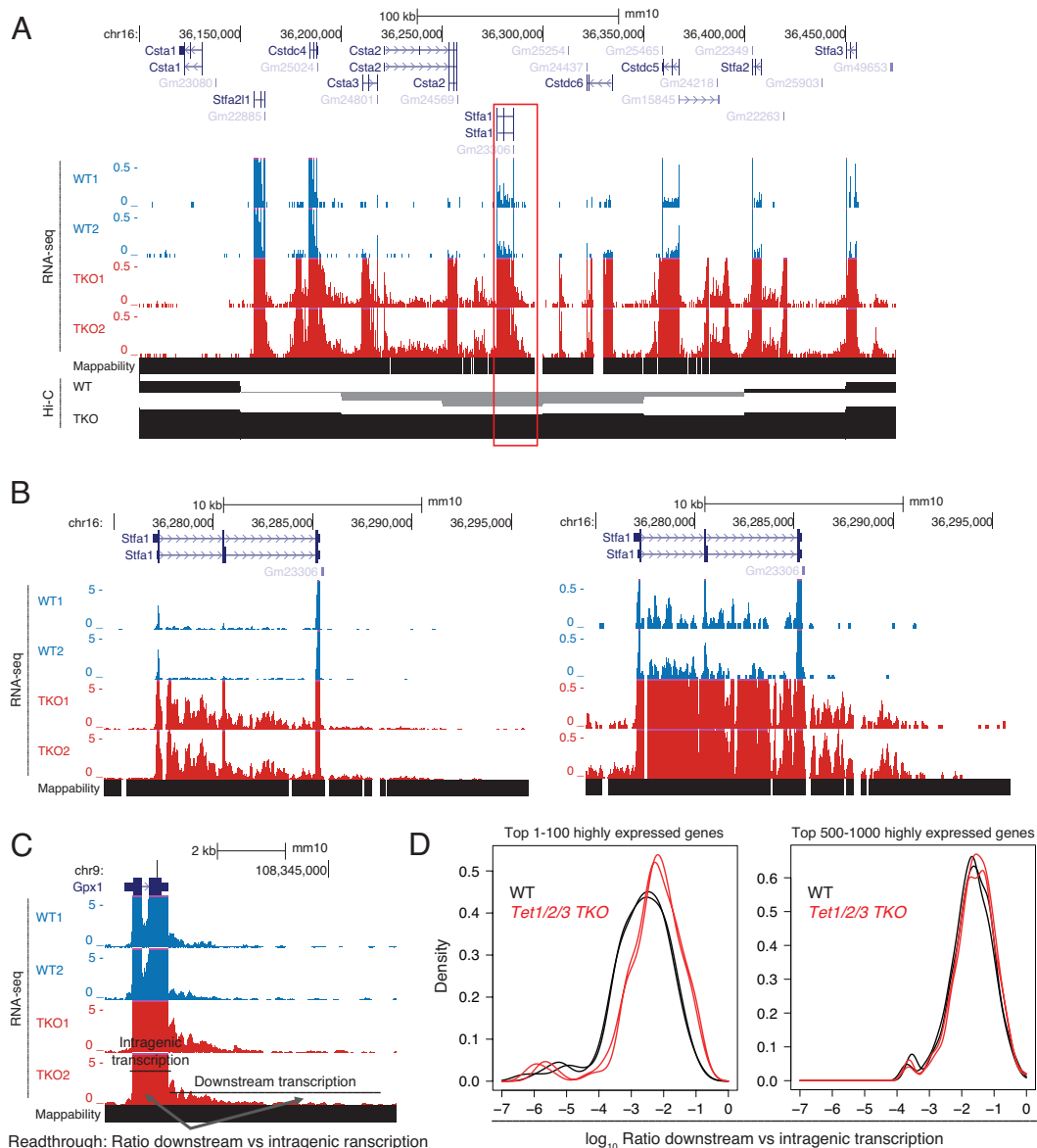
We previously showed that inducible double deletion of *Tet2* and *Tet3* in mice resulted, unexpectedly, in the rapid onset of AML that was fatal within 4 to 5 wk (17). This outcome was very different from that observed in mice with germline deletion of *Tet2* alone, which developed late-onset diseases resembling human MDS and CMML (11, 13, 14). The difference indicated that *Tet2* and *Tet3* have partially redundant activities in mouse HSPC [as later confirmed by another study (36)], prompting us to examine the physiological outcome of complete TET loss of function through acute, inducible deletion of all three *Tet* genes. Indeed, *Tet iTKO* mice developed aggressive AML with the same rapid time course as *Tet2/3 iDKO* mice and displayed the same skewing of hematopoietic differentiation to the myeloid lineage and subsequent myeloid transformation. Given that mice with germline disruption of both *Tet1* and *Tet2* exhibit predominantly B cell rather than myeloid expansion (16), it is possible that *Tet3* contributes more prominently than *Tet2* to maintaining the proper balance of HSPC differentiation to myeloid and lymphoid lineages in mice. This hypothesis is currently being tested.

The most striking feature of the expanded myeloid cells in *Tet iTKO* mice was the appearance of new myeloid populations that bore neutrophil markers and showed a consistent upregulation of the *stefin/cystatin* gene cluster on mouse chromosome 16. Stefins are cysteine protease inhibitors (23–25) whose physiological roles are discussed in more detail below. The mouse genome contains two *stefin/cystatin* gene clusters, on chromosomes 2 and 16 respectively, as well as two isolated *Cst* genes—*Cstb* on chromosome 10 and *Cst6* on chromosome 19. Of these, only genes in the chromosome 16 cluster were consistently up-regulated in expanded *Tet iTKO* myeloid cells. Our study reveals the importance of *cystatin/stefin* overexpression in myeloid expansion in human AML. Expression of mRNA encoding the type I cystatins CSTA and CSTB, 98-amino-acid proteins which display 54% amino-acid sequence identity (24), correlated positively with the percentage of circulating neutrophils and monocytes in the PB of AML patients. Furthermore, CSTB DNA amplification and mRNA overexpression were associated with shorter survival in these same AML patients (28, 29). Supporting the human data, the percentage of myeloid cells in PB of germline *Tet2*-deficient mice correlated with several markers of the preleukemic myeloproliferative disorder observed in these mice—myeloid cell expansion, extramedullary hematopoiesis and splenomegaly (37). Moreover, another aggressive mouse AML

initiated by the simultaneous expression of NUP98-HOXD13 and NUP98-PHF23 also displays a dramatic upregulation of *stefin* genes (26), pointing to an association between *stefin/cystatin* overexpression and the aggressiveness of AML. There appears to be considerable redundancy between *stefin/cystatin* superfamily members, however (30), making it challenging to determine in either mouse or human systems whether *stefin/cystatin* depletion or overexpression affects oncogenic progression or survival. Future studies will be needed to determine the effects of genetic manipulation of *stefin/cystatin* genes in HSPC in *Tet iTKO* cells and to determine the contribution of *stefin/cystatin* expression to human cancers.

The increased expression of a subset of chromosome 16 *stefin/cystatin* genes in *TET*-deleted cells was already apparent, albeit less pronounced, in hematopoietic progenitor (LSK) cells; this was accentuated in the expanded myeloid cells and was associated with large scale changes in genome organization that occurred after *TET* deletion. Specifically, increased expression of the chromosome 16 *stefin/cystatin* gene cluster correlated with compartment switching of the entire cluster from the transcriptionally inactive Hi-C B heterochromatic compartment to the transcriptionally permissive Hi-C A euchromatic compartment. The cause-and-effect relationship between chromatin compartment switching and transcriptional upregulation is not well understood. Loss of CCCTC-binding factor (CTCF) and cohesin did not result in a global change in the organization of chromatin compartments measured in bulk populations (38); in contrast, both replication timing (39, 40) and transcriptional elongation (33) have been reported to affect the 3-dimensional compartmentalization of the genome. The effects of replication timing appear to vary with cell type since depletion of a regulator of replication timing, RIF1, resulted in different types of compartment switching in a human embryonic stem cell line (H9) compared to the colon cancer cell line HCT116 (40). A potentially more generally applicable model invokes increased transcription in heterochromatin (33, 41). In an in vitro model of influenza A infection, transcriptional elongation, particularly readthrough transcription into heterochromatin, resulted in compartment switching from heterochromatin to euchromatin by disrupting cohesin-mediated loops (33). Noncoding transcription has been reported to affect chromatin compartmentalization (42) and topologically associating domains (43) and to regulate expression of nearby genes (44). For instance, transcription of the noncoding RNA ThymoD during T cell development results in heterochromatin-to-euchromatin compartment switches and affects local DNA methylation levels, leading to demethylation of specific CpG residues which impact CTCF binding (42).

In conclusion, our molecular analysis of the aggressive myeloid leukemias developing in *Tet iTKO* mice in vivo has identified several effects of profound TET deficiency that appear unrelated to the established role of TET proteins in reversing DNA demethylation. The overall pattern of gene expression changes in *Tet TKO* myeloid cells shows no clear relation to changes in DNA methylation; moreover, the heterochromatin-to-euchromatin compartment switch and the associated increase in expression of the *stefin/cystatin* gene cluster in *Tet iTKO* myeloid cells are more closely related to high-level expression and transcriptional readthrough than to changes in DNA methylation. A few previous studies have identified functions for TET proteins that are independent of DNA demethylation. Notably, all of these involve heterochromatin: among others, maintenance of DNA methylation in heterochromatin (31), transcriptional repression of repetitive elements through modulation of KAP1/TRIM28 binding (45), and establishment of the heterochromatic histone modifications H3K9me3 and H4K20me3 (46). Our results suggest two additional functions for TET proteins that appear unrelated to DNA demethylation: TET proteins preserve three-dimensional genome organization and euchromatin-heterochromatin compartmentalization, at least



**Fig. 6.** *Tet iTKO* cells display intronic and readthrough transcription downstream of highly expressed genes. (A) Genome browser tracks of total (ribodepleted) RNA-seq (tracks 1–4) and Hi-C PC1 compartment changes (tracks 6–7) at the *stefin/cystatin* gene cluster in WT and *Tet iTKO* CD11b<sup>+</sup> cells. A region containing the *Stfa1* gene is highlighted and shown in (B). (B) Total RNA levels in the *Stfa1* locus in WT and *Tet iTKO* CD11b<sup>+</sup> cells, at two different scales (Left: 0 to 5 RPM (reads per million of mapped reads); Right: 0 to 0.5 RPM). (C) Genome tracks of total RNA levels in WT and *Tet iTKO* CD11b<sup>+</sup> cells at the *Gpx1* locus. RNA levels within the gene body and downstream of the gene end are indicated in the figure. These regions are used to quantify transcriptional readthrough, following a previously published approach (34). *Gpx1* is not differentially expressed between WT and *Tet iTKO* cells and is among the top 100 most highly expressed genes in both conditions. (D) Distribution of readthrough transcription levels ( $\log_{10}$  ratio of reads downstream of gene ends vs. read within gene body) for top 100 (Left) or top 501 to 1000 (Right) most highly expressed genes, across two biological replicates from WT or *Tet iTKO* CD11b<sup>+</sup> cells.

at specific genomic loci, and facilitate proper transcriptional termination, especially of highly expressed genes. The latter observation may be related to the finding that the most highly expressed genes also have the highest levels of 5hmC (10, 47). Future studies will establish whether these two functions are interrelated, whether they are conserved among different cell types, and whether they require TET catalytic activity.

## Materials and Methods

Detailed materials and methods are described in *SI Appendix, Materials and Methods*.

**Acute Deletion of Tet1, Tet2, and Tet3 Genes.** *Tet1<sup>fl/fl</sup>*, *Tet2<sup>fl/fl</sup>*, and *Tet3<sup>fl/fl</sup>* mice have been described previously (48–50). *Rosa26-stop-EYFP* (*Rosa26-YFP<sup>LSL</sup>*) (strain #006148) and *UBC-Cre-ERT2* (strain #008085) mice were purchased from Jackson

Laboratories. To delete floxed alleles using the *Cre-ERT2* recombinase, tamoxifen (Sigma) was solubilized at 10 mg mL<sup>-1</sup> in corn oil (Sigma) and delivered into mice by intraperitoneal injection of 2 mg tamoxifen per mouse every day for five consecutive days. The day of the last tamoxifen injection was designated Day 0 (Fig. 1A). All animal procedures were approved by the La Jolla Institute (LJI) or Ulsan National Institute of Science and Technology (UNIST) Institutional Animal Care and Use Committee and were conducted in accordance with institutional guidelines.

**Data, Materials, and Software Availability.** Next-generation sequencing data have been deposited in the Gene Expression Omnibus (GEO) repository, accession number *GSE222726* (51). Previously published RNA-seq data of *Tet2/3 iDKO* LSK cells are available under accession number *GSE72630*. All study data are included in the article and/or *SI Appendix*.

**ACKNOWLEDGMENTS.** We thank C. Kim, L. Nosworthy, R. Simmons, D. Hinz, and C. Dillingham of the LJI Flow Cytometry Core Facility for cell sorting; J. Day, S. Wlodychak, and C. Kim of the LJI Next Generation Sequencing Facility for

next-generation sequencing; the Department of Laboratory Animal Care and the animal facility for excellent support; and K. Jepsen and E. Ricciardelli of the UCSD IGM Genomics Center. This work was supported by NIH grant R35 CA210043 (to A.R.), equipment grants S10OD016262 and S10RR027366 to L.J.; JSPS KAKENHI Grant Numbers JP19H05650, AMED under Grant Numbers JP20gm1210003 (to T.N.); and National Research Foundation of Korea grants (2018R1A6A1A03025810, 2019R1F1A1063340 to M.K.). I.F.L.-M. was supported by a UC MEXUS-CONACYT Fellowship. M.K. is supported by UNIST (Ulsan National Institute of Science & Technology (1.180075.01, 1.220023.01) and Center for Genomic Integrity, Institute for Basic Science (IBS-R022-D1). A.O. is supported by Ministry of Education, Culture, Sports, Science and Technology (Japan) Grants-in-Aid for Fostering Joint International Research (A) JP20KK0351 and Scientific Research (B) JP22H02885.

Author affiliations: <sup>a</sup>Division of Signaling and Gene Expression, La Jolla Institute for Immunology, La Jolla, CA 92037; <sup>b</sup>Sanford Consortium for Regenerative Medicine, La Jolla, CA 92093; <sup>c</sup>Department of Biological Sciences, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea; <sup>d</sup>Department of Immunology and Genomic Medicine, National Jewish Health, Denver, CO 80206; <sup>e</sup>Department of Immunology and Microbiology, University of Colorado, Anschutz Medical Campus, Aurora, CO 80045; <sup>f</sup>Department of Life Sciences, Jeonbuk National University, Jeonju 54896, Republic of Korea; <sup>g</sup>Department of Immunology, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan; <sup>h</sup>Japan Agency for Medical Research and Development (AMED), Core Research for Evolutional Science and Technology (CREST), Chiba 260-8670, Japan; <sup>i</sup>Institute for Advanced Academic Research, Chiba University, Inage-ku, Chiba 263-8522, Japan; and <sup>j</sup>Center for Genomic Integrity, Institute for Basic Science, Ulsan 44919, Republic of Korea

Author contributions: H.Y. and H.J. performed mouse experiments; I.F.L.-M. and A.O. performed bioinformatic analyses; H.Y. and J.S.B. prepared libraries for next-generation sequencing; I.F.L.-M. and A.X.C. performed the Stefin CRISPR screen; H.Y., I.F.L.-M., H.J., J.S.B., J.A., T.N., A.O., M.K., and A.R. designed research and analyzed data; and H.Y., I.F.L.-M., A.O., M.K., and A.R. wrote the paper.

- M. Tahiliani *et al.*, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
- C. J. Lio, H. Yuita, A. Rao, Dysregulation of the TET family of epigenetic regulators in lymphoid and myeloid malignancies. *Blood* **134**, 1487–1497 (2019).
- Y. F. He *et al.*, Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
- C. W. Lio *et al.*, Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife* **5**, e18290 (2016).
- C. J. Lio *et al.*, TET enzymes augment activation-induced deaminase (AID) expression via 5-hydroxymethylcytosine modifications at the Aicda superenhancer. *Sci. Immunol.* **4**, 7523 (2019).
- Y. Wang *et al.*, WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol. Cell* **57**, 662–673 (2015).
- J. L. Sardina *et al.*, Transcription factors drive Tet2-mediated enhancer demethylation to reprogram cell fate. *Cell Stem Cell* **23**, 727–741.e9 (2018).
- G. C. Hon *et al.*, 5mC Oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell* **56**, 286–297 (2014).
- R. Deplus *et al.*, TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* **32**, 645–655 (2013).
- A. Tsagaratou *et al.*, Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3306–3315 (2014).
- M. Ko, J. An, A. Rao, DNA methylation and hydroxymethylation in hematologic differentiation and transformation. *Curr. Opin. Cell Biol.* **37**, 91–101 (2015).
- L. Cimmino *et al.*, TET1 is a tumor suppressor of hematopoietic malignancy. *Nat. Immunol.* **16**, 653–662 (2015).
- K. Moran-Crusio *et al.*, Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
- C. Quivoron *et al.*, TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* **20**, 25–38 (2011).
- R. Shrestha *et al.*, Molecular pathogenesis of progression to myeloid leukemia from TET-insufficient status. *Blood Adv.* **4**, 845–854 (2020).
- Z. Zhao *et al.*, Combined loss of Tet1 and Tet2 promotes B cell, but not myeloid malignancies, in Mice. *Cell Rep.* **13**, 1692–1704 (2015).
- J. An *et al.*, Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat. Commun.* **6**, 10071 (2015).
- Y. Huang *et al.*, The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5**, e8888 (2010).
- E. M. Pietras *et al.*, Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell* **17**, 35–46 (2015).
- S. Monticelli, G. Natoli, Transcriptional determination and functional specificity of myeloid cells: Making sense of diversity. *Nat. Rev. Immunol.* **17**, 595–607 (2017).
- D. Álvarez-Erico, R. Vento-Tormo, M. Sieweke, E. Ballestar, Epigenetic control of myeloid cell differentiation, identity and function. *Nat. Rev. Immunol.* **15**, 7–17 (2015).
- A. Giladi *et al.*, Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat. Cell Biol.* **20**, 836–846 (2018).
- V. Turk, W. Bode, The cystatins: Protein inhibitors of cysteine proteinases. *FEBS Lett.* **285**, 213–219 (1991).
- S. Magister, J. Kos, Cystatins in immune system. *J. Cancer* **4**, 45–56 (2013).
- B. Breznik, A. Mitrović, T. T. Lah, J. Kos, Cystatins in cancer progression: More than just cathepsin inhibitors. *Biochimie* **166**, 233–250 (2019).
- S. Kundu *et al.*, Thymic precursor cells generate acute myeloid leukemia in NUP98-PHF23/NUP98-HOXD13 double transgenic mice. *Sci. Rep.* **9**, 17213 (2019).
- T. Yadaï, T. Houben, A. Bitorina, R. Shiri-Sverdlov, The Ins and Outs of Cathepsins: Physiological function and role in disease management. *Cells* **9**, 1679 (2020).
- J. W. Tyner *et al.*, Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).
- T. J. Ley *et al.*, Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
- M. Bilodeau *et al.*, Analysis of blood stem cell activity and cystatin gene expression in a mouse model presenting a chromosomal deletion encompassing Csta and Stfa211. *PLoS One* **4**, e7500 (2009).
- I. F. López-Moyado *et al.*, Paradoxical association of TET loss of function with genome-wide DNA hypomethylation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16933–16942 (2019).
- A. Tsagaratou *et al.*, TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. *Nat. Immunol.* **18**, 45–53 (2017).
- S. Heinz *et al.*, Transcription elongation can affect genome 3D structure. *Cell* **174**, 1522–1536.e22 (2018).
- S. J. Roth, S. Heinz, C. Benner, ARTDeco: Automatic readthrough transcription detection. *BMC Bioinformatics* **21**, 214 (2020).
- L. Yang *et al.*, DNMT3A loss drives enhancer hypomethylation in FLT3-ITD-associated leukemias. *Cancer Cell* **29**, 922–934 (2016).
- L. Cimmino *et al.*, Restoration of TET2 function blocks aberrant self-renewal and leukemia progression. *Cell* **170**, 1079–1095.e20 (2017).
- M. Meisel *et al.*, Microbial signals drive pre-leukaemic myeloproliferation in a Tet2-deficient host. *Nature* **557**, 580–584 (2018).
- M. J. Rowley, V. G. Corces, Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
- J. Sima *et al.*, Identifying cis elements for spatiotemporal control of mammalian DNA replication. *Cell* **176**, 816–830.e18 (2019).
- K. N. Klein *et al.*, Replication timing maintains the global epigenetic state in human cells. *Science* **372**, 371–378 (2021).
- Q. Du *et al.*, DNA methylation is required to maintain both DNA replication timing precision and 3D genome organization integrity. *Cell Rep.* **36**, 109722 (2021).
- T. Isoda *et al.*, Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* **171**, 103–119.e18 (2017).
- Y. Zhang *et al.*, Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat. Genet.* **51**, 1380–1388 (2019).
- J. Joung *et al.*, Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**, 343–346 (2017).
- F. Lu, Y. Liu, L. Jiang, S. Yamaguchi, Y. Zhang, Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.* **28**, 2103–2119 (2014).
- P. Stolz *et al.*, TET1 regulates gene expression and repression of endogenous retroviruses independent of DNA demethylation. *Nucleic Acids Res.* **50**, 8491–8511 (2022).
- W. A. Pastor *et al.*, Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**, 394–397 (2011).
- M. Ko *et al.*, Ten-eleven-translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14566–14571 (2011).
- M. Ko *et al.*, TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol. Rev.* **263**, 6–21 (2015).
- J. Kang *et al.*, Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4236–E4245 (2015).
- Isaac F. López-Moyado, H. Yuita, J. Scott-Browne, M. Ko, Inducible disruption of Tet genes results in myeloid malignancy, readthrough transcription, and a heterochromatin-to-euchromatin switch. Gene Expression Omnibus (GEO) <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2227265> Deposited 11 January 2023.