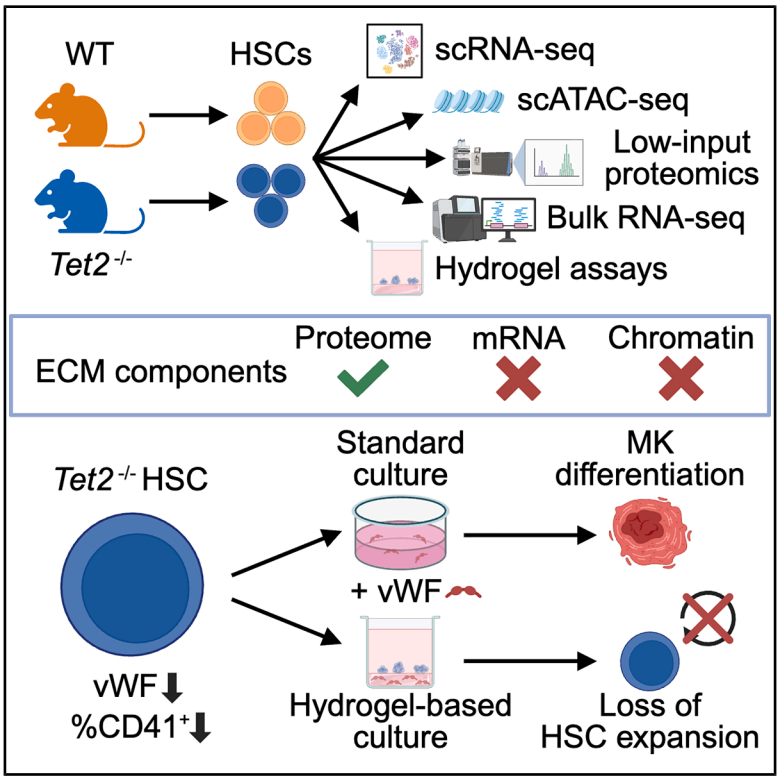
Cell Reports

Article

Low-input proteomics identifies vWF as a negative regulator of *Tet2* mutant hematopoietic stem cell expansion

Graphical abstract



Highlights

- A comprehensive molecular map of TET2 deficiency in hematopoietic stem cells (HSCs)
- Global mass spectrometry-based proteomics approach for 10,000 primary HSCs
- Application of 2D functionalized hydrogel to test ECM molecules in primary HSCs
- Physical and mechanical modulators impact *Tet2*-mutant HSC fate

Authors

Maria Jassinskaja, Daniel Bode, Monika Gonka, ..., Berthold Göttgens, Jyoti Choudhary, David G. Kent

Correspondence

david.kent@york.ac.uk

In brief

Jassinskaja, Bode et al. provide a multi-omics characterization of *Tet2*-mutated cells, including global proteomics revealing novel roles for extracellular matrix (ECM) molecules in selectively modifying self-renewal divisions. These findings point more broadly to physical and mechanical mediators of self-renewal, implicating integrins and cytokine signaling as extracellular drivers of clonal expansion.

Jassinskaja et al., 2026, Cell Reports *45*, 116770
January 27, 2026
https://doi.org/10.1016/j.celrep.2025.116770

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Low-input proteomics identifies vWF as a negative regulator of *Tet2* mutant hematopoietic stem cell expansion

Maria Jassinskaja,[1,2,9] Daniel Bode,[1,9] Monika Gonka,[1] Theodoros I. Roumeliotis,[3] Alexander J. Hogg,[1] Juan A. Rubio Lara,[1,4,5] Ellie Bennett,[1] Joanna Milek,[1] Samuel Elberfeld,[1] Bart Theeuwes,[4,5] M.S. Vijayabaskar,[4,5] Lilia Cabrera Cosme,[1] James Lok Chi Che,[1,4,5] Sandy MacDonald,[6] Sophia Ahmed,[6] Benjamin A. Hall,[7] Grace Vasey,[1] Helena Kooi,[1] Miriam Belmonte,[4,5] Mairi S. Shepherd,[4,5] William J. Brackenbury,[1] Iwo Kucinski,[4,5] Satoshi Yamazaki,[8] Andrew N. Holding,[6] Alyssa H. Cull,[1] Nicola K. Wilson,[4,5] Berthold Göttgens,[4,5] Jyoti Choudhary,[3] and David G. Kent[1,10,*]

[1]Centre for Blood Research, York Biomedical Research Institute, Department of Biology, University of York, York YO10 5DD, UK
[2]Lund Stem Cell Center, Division of Molecular Medicine and Gene Therapy, Lund University, Lund 221 84, Sweden
[3]Functional Proteomics Team, Chester Beatty Laboratories, The Institute of Cancer Research, London SW3 6JB, UK
[4]Wellcome MRC Cambridge Stem Cell Institute, University of Cambridge, Puddicombe Way, Cambridge CB2 0AW, UK
[5]Department of Haematology, University of Cambridge, Cambridge CB2 0AW, UK
[6]Department of Biology, University of York, York YO10 5DD, UK
[7]Department of Medical Physics and Biomedical Engineering, Malet Place Engineering Building, University College London, Gower Street, London WC1E 6BT, UK
[8]Division of Stem Cell Biology, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan
[9]These authors contributed equally
[10]Lead contact
*Correspondence: david.kent@york.ac.uk
https://doi.org/10.1016/j.celrep.2025.116770

SUMMARY

Despite rapid advances in mapping genetic drivers and gene expression changes in hematopoietic stem cells (HSCs), few studies exist at the protein level. We perform a deep, multi-omics characterization (epigenome, transcriptome, and proteome) of HSCs in a mouse model carrying a loss-of-function mutation in *Tet2*, a driver of increased self-renewal in blood cancers. Using state-of-the-art, multiplexed, low-input mass spectrometry (MS)-based proteomics, we profile TET2-deficient ($Tet2^{-/-}$) HSCs, revealing previously unrecognized molecular processes that define the pre-leukemic HSC molecular landscape. Specifically, we obtain more accurate stratification of wild-type and $Tet2^{-/-}$ HSCs than transcriptomic approaches and identify extracellular matrix (ECM) molecules as being dysregulated upon TET2 loss. HSC expansion assays using ECM-functionalized hydrogels confirm a selective effect on the expansion of *Tet2*-mutant HSCs. Taken together, our study represents a comprehensive molecular characterization of *Tet2*-mutant HSCs and identifies a previously unanticipated role of ECM molecules in regulating self-renewal of disease-driving HSCs.

INTRODUCTION

Hematological malignancies are commonly initiated by single hematopoietic stem cells (HSCs) that have acquired mutations, which confer a clonal advantage relative to non-mutated hematopoietic cells.[1] Loss-of-function (LoF) mutations in the gene encoding the DNA-demethylating enzyme Tet methylcytosine deoxygenase 2 (TET2) are commonly found in hematological malignancies, and evidence points toward its loss driving an increase in HSC self-renewal.[1–5] At the transcript level, mutations in *Tet2* are associated with altered gene expression in mouse HSCs; however, relatively few of these potential partner genes have been implicated in directly driving disease initiation and progression,[2–4] thus highlighting an urgent need to further explore the molecular landscape of mutated HSCs.

Historically, comprehensive profiling of HSCs beyond the transcriptome has been impeded due to their low numbers. To address the prohibitively large amount of material typically required for global proteomic characterization, multiple strategies for facilitating low cell number and single-cell mass spectrometry (MS)-based proteomics have begun to emerge,[6–12] including a recent effort which profiles thousands of hematopoietic stem and progenitor cells (HSPCs) at the single cell level.[13] These studies have revealed a generally poor correlation between proteome and transcriptome, especially in non-homeostatic contexts such as inflammation and disease,[10,11,14,15] further highlighting the need to develop low-cell-number approaches to facilitate the study of the global proteome in HSCs and assess the functional unit of molecular activity.

In this study, we use an MS-based method for global proteomic profiling of low numbers (10,000–20,000) of primary HSPCs as a part of a comprehensive multi-omics profiling of TET2-deficient (*Tet2*$^{-/-}$) long-term (LT)-HSCs and uncover previously undescribed regulators of *Tet2*-mutant HSC biology. Analysis of the proteome exclusively identifies extracellular matrix (ECM) interactions as a point of dysregulation upon loss of TET2 in HSCs, providing evidence for ECM molecules altering the differentiation and self-renewal of mutant HSCs relative to their non-mutant counterparts, and thus identifies previously unexplored pathways for therapeutic intervention. These data highlight the importance of assessing the proteome of pre-leukemic and leukemic HSCs in order to reveal novel biology that is typically hidden from genomic and transcriptomic studies.

RESULTS

Integrative single-cell ATAC-seq and RNA-seq analysis of TET2-deficient HSCs

To understand molecular changes in HSCs induced by TET2 LoF, we first assessed chromatin accessibility and gene expression at the single-cell level. Using a genetic knockout mouse model with targeted disruption of the TET2 catalytic domain,[5] we applied single-cell assay for transposase-accessible chromatin by sequencing (scATAC-seq; Figures 1A–1C) and single-cell RNA sequencing[16] (scRNA-seq; Figures 1D and 1E) to fluorescence-activated cell sorting (FACS)-isolated CD45^{+} CD48^{-} CD150^{+} EPCR^{+} Sca-1^{+} (ESLAM Sca-1^{+}) HSCs, a cell population highly enriched (>60%) for HSCs with LT serial reconstitution capacity[17] from *Tet2*$^{-/-}$ and wild-type (WT) mice. scATAC-seq analysis revealed changes in chromatin accessibility induced by TET2 LoF, and differential analysis revealed more accessible genomic regions in *Tet2*$^{-/-}$ HSCs compared to HSCs isolated from WT littermate controls (Figure 1A; Table S1). Nearly half (47% of peaks) of the regions deemed to be more accessible in the *Tet2*$^{-/-}$ HSCs were intronic (Figure 1B) and potentially related to specific gene regulation. In line with this, recent studies of chromatin accessibility revealed that HSCs with mutated *Tet2* have hypermethylation of enhancer sites.[18] Analysis of transcription factor (TF) binding sites identified specific motifs enriched in *Tet2*$^{-/-}$ HSCs, including binding motifs for known self-renewal regulator *Smarcc1*,[19] tumor suppressor *Runx3*,[20] and cell cycle and oxidative stress regulator *Bach1*,[21] among others (Figure 1C).

Next, we undertook plate-based (scRNA-seq) to determine changes in gene expression between WT and *Tet2*$^{-/-}$ HSCs. As in our scATAC-seq data, scRNA-seq indicated significant molecular changes upon *Tet2* loss (Figure 1D; Table S2). We identified 54 differentially expressed genes (log2 fold change [FC] > 1 and adjusted p value < 0.05), including 18 upregulated and 36 downregulated in *Tet2*$^{-/-}$ HSCs compared to WT (Figure 1E). In accordance with previous studies,[24] Gene Ontology (GO) analysis identified enrichment in pathways regulating transcription and cellular response to calcium ion (Figure 1F).

To further study the mechanisms underpinning the self-renewal differences in *Tet2*$^{-/-}$ compared to WT HSCs, scATAC-seq and scRNA-seq datasets were integrated. The closest genes to peaks identified in the scATAC-seq analysis were determined, and the expression of these genes was assessed in the scRNA-seq data (distance to transcription start site < 100,000 bp[22]). Out of 698 identified closest genes in the scATAC-seq analysis, only 16 had significantly altered expression between *Tet2*$^{-/-}$ and WT HSCs in the scRNA-seq data (Figures 1G and 1H), with 14 of these appearing among the top 40 differentially expressed genes in the scRNA-seq dataset. Analysis of predicted protein associations using the Genemania database[25] showed high connectivity, strongly suggesting biological relatedness (Figure 1I). For example, among the genes with both lower chromatin accessibility and gene expression in *Tet2*$^{-/-}$ HSCs were three members of the Krüppel-like factor (*Klf*) family (*Klf2*, *Klf4*, and *Klf6*) that regulate self-renewal[26] and *Fosb*, a member of the AP1 complex, recently shown by us to be downregulated in hibernating HSCs.[27] To further integrate the information from both analyses, we built a TF regulatory network using the DoRothEA database.[26] We selected all TFs whose binding motifs were significantly more accessible in *Tet2*$^{-/-}$ HSCs (as shown in our scATAC-seq analysis) and curated a list of all differentially expressed genes (as determined by our scRNA-seq analysis) that the selected TFs regulate. The constructed interaction network containing all defined regulons (TF regulatory gene pairs) identified a set of TFs potentially regulating the HSC fate (Figure 1J), among which we found well-known HSC regulators, such as *Gata2*, *Gata3*, *Fli1*, and *Runx1*.[28–30] Together, these data both identify factors important for normal HSC function and identify additional candidate regulators of the increased self-renewal observed in *Tet2*$^{-/-}$ HSCs.

Optimization of a low-input proteomic workflow

Our analyses of the epigenome and transcriptome of cells largely reflected current knowledge surrounding *Tet2* LoF mutations, whereby expression and chromatin accessibility of a number of genes' hematopoietic TFs with known roles in cellular differentiation are altered. These data further accord with the phenotypes observed in mouse models and patients with hematological malignancies (i.e., an accumulation of myeloid progenitor cells[1,5]). In order to explore changes in populations of primary HSCs that might occur downstream of the transcriptome, a global proteomic method for low cell numbers was required. To develop this method, we used the hematopoietic progenitor cell line HoxB8-FL[31] for protocol optimization. In an initial single-vessel sample preparation and protein-level tandem mass tag (TMT) 10plex isobaric labeling approach applied to a sample set of 3 × 10,000 cells and 7 × 20,000 cells, just 2,218 and 1,345 proteins were identified and quantified, respectively, across the entire multiplexed set (Figure 2A and 10K direct) with only modest benefit gained by increasing the smaller samples from 10,000 to 15,000 cells and carrier samples from 20,000 to 30,000 cells (Figure 2A and 15K direct). Next, through a multi-stage optimization process, we devised a method using single-vessel sample preparation, peptide-level isobaric labeling with a carrier approach, and high-resolution offline fractionation, which resulted in robust quantification of over 3,500 proteins from as few as 10,000 cells (Figures 2A–2C). These results suggest that the higher labeling efficiency achieved by peptide-level isobaric labeling outweighs the benefits of combining samples

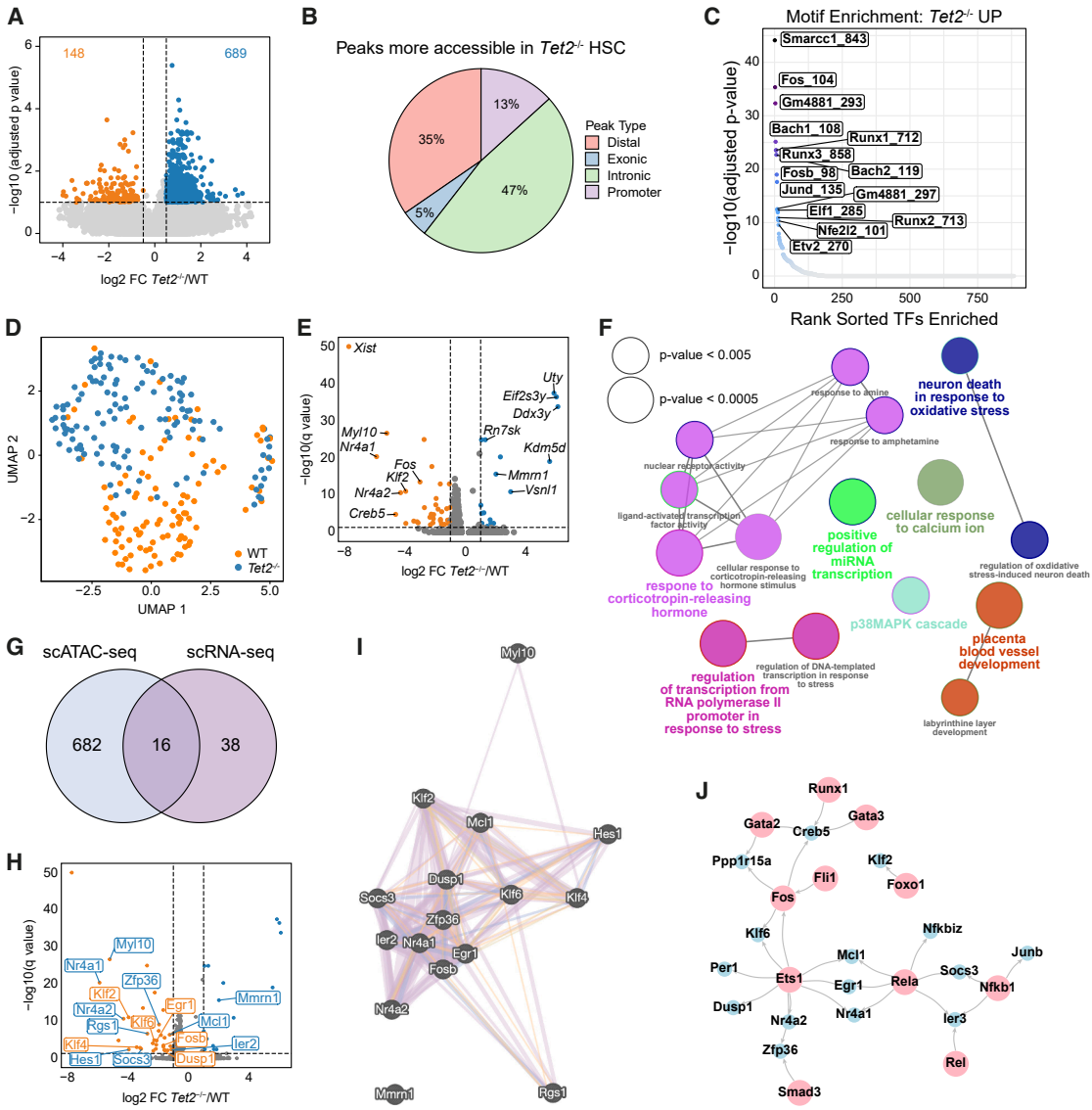


Figure 1. Single-cell analyses reveal increased chromatin accessibility and decreased transcriptional activity in *Tet2*$^{-/-}$ HSCs

(A) Volcano plot showing differentially accessible sites between WT and *Tet2*$^{-/-}$ HSCs in scATAC-seq analysis (log2 FC > 0.5, adjusted *p* value < 0.1).

(B) Proportion of significantly more accessible peak regions categorized by genomic feature in *Tet2*$^{-/-}$ HSCs.

(C) Motif enrichment for TF motifs more accessible in *Tet2*$^{-/-}$ HSCs.

(D) Uniform Manifold Approximation and Projection visualization of scRNA-seq data. Each dot represents one cell.

(E) Volcano plot showing differentially expressed genes between WT and *Tet2*$^{-/-}$ HSCs in scRNA-seq analysis (log2 FC > 1, adjusted *p* value < 0.05).

(F) GO biological processes enrichment analysis of differentially expressed genes in scRNA-seq. Node sizes reflect the statistical significance of the terms. Force-directed layout presented by the kappa score. The layout was adjusted to minimize label overlap.

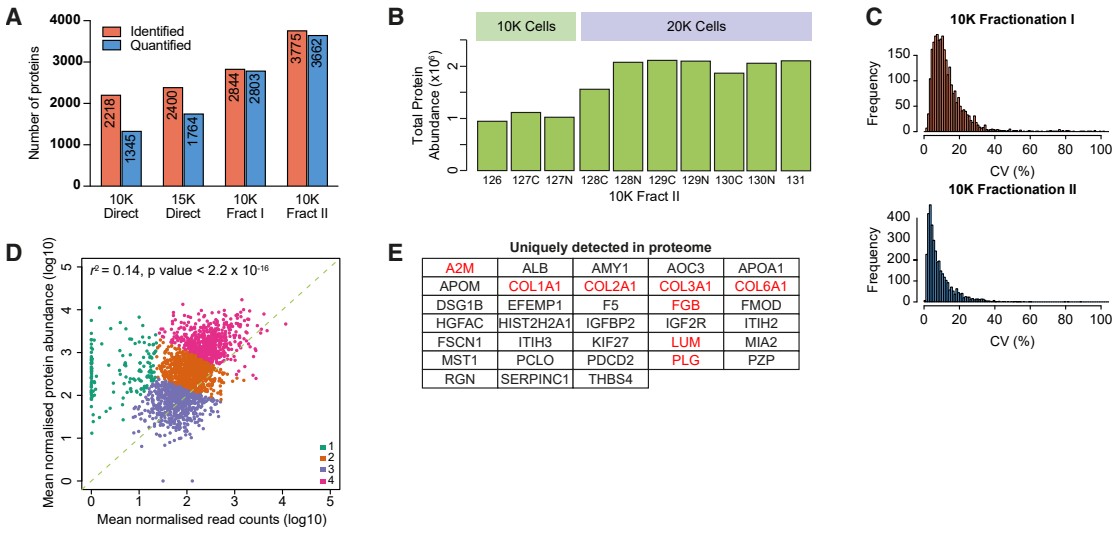
(G) Venn diagram showing the overlap of targets identified in both scATAC-seq and scRNA-seq modalities (*Tet2*$^{-/-}$ versus WT HSC pairwise testing).

(H) Volcano plot of scRNA-seq data with labeled selected genes identified in the scATAC-seq closest gene analysis (*Tet2*$^{-/-}$ versus WT HSC pairwise testing). Blue/orange dots indicate genes down/upregulated in scRNA-seq data analysis; blue/orange labels indicate less/more accessible markers in the scATAC-seq data analysis.

(I) Interaction graph for targets identified in the integrative scATAC-seq/scRNA-seq data analysis constructed using the Genemania database.[22] Purple, co-expression; blue, co-localization; orange, predicted; green, shared protein domains; and red, physical interaction.

(J) Interaction graph of identified regulons: TF (scATAC-seq analysis) and genes they regulate (scRNA-seq analysis). Network built using information from the DoRothEA database.[23] Pink, TF and blue, regulated genes.

See also Figure S1.



A2M	ALB	AMY1	AOC3	APOA1
APOM	COL1A1	COL2A1	COL3A1	COL6A1
DSG1B	EFEMP1	F5	FGB	FMOD
HGFAC	HIST2H2A1	IGFBP2	IGF2R	ITIH2
FSCN1	ITIH3	KIF27	LUM	MIA2
MST1	PCLO	PDCD2	PLG	PZP
RGN	SERPINC1	THBS4		

Figure 2. An optimized workflow enables deep proteomic profiling of low numbers of hematopoietic cells

(A) Number of identified and quantified proteins in HoxB8-FL cells using direct injection or offline fractionation into 6 fractions (Fract I) or 5 fractions × 2 runs each with different upper intensity precursor selection limit in the two runs (Fract II). K: 1000.

(B) Total protein abundance across TMT channels in the 10K Fract II experiment from (A). K: 1000.

(C) Coefficient of variance (CV) values for the 10K Fract I and 10K Fract II experiments. K:1000.

(D) Correlation between protein and mRNA abundance in HoxB8-FL cells. Points are colored according to the *k*-means cluster they belong to. Correlation assessed using Pearson's correlation coefficient r^2.

(E) Identifications unique to proteome data. Proteins in red are associated with ECM organization.

earlier in the protocol with protein-level labeling, and that reduction of sample complexity by offline fractionation prior to liquid chromatography-MS analysis is the most important factor in achieving high proteome coverage in low-input samples.

To assess the type of information gained by proteome-level characterization, we mapped protein abundance from the Fract II experiment (Table S3) against gene expression[32] (Figure 2D) and observed a positive correlation overall (r^2 = 0.13, $p < 2.2 \times 10^{-18}$). Clustering analysis identified a set of genes/proteins showing particularly strong enrichment at the protein level (Figure 2D, cluster 1, shown in green), including 33 that were uniquely detected in the proteome data (Figure 2E). Intriguingly, 8 of these 33 proteins (24%) were associated with ECM organization (Figure 2E, highlighted in red), suggesting that ECM components might be more readily captured by proteomic relative to transcriptomic analysis.

Global proteomics identifies distinct molecular changes in *Tet2*$^{-/-}$ HSPCs not captured by transcriptomics

To generate a comprehensive molecular map of WT and *Tet2*$^{-/-}$ HSPCs, we applied our low-input proteomic workflow (Figure 2) to WT and *Tet2*$^{-/-}$ HSPCs with primary and secondary transplantation capacity (lineage [Lin]$^-$ cKit$^+$ CD45$^+$ CD48$^-$ CD150$^+$; collectively called "CD150$^+$") or those limited to finite reconstitution capacity in a primary transplantation (Lin$^-$ cKit$^+$ CD45$^+$ CD48$^-$ CD150$^-$; collectively called "CD150$^-$"; Figure 3A).[33] We additionally analyzed the proteome of WT ESLAM HSCs (CD45$^+$ CD48$^-$ CD150$^+$ EPCR$^+$) as a reference LT-HSC population[33] and WT Lin$^-$ cKit$^+$ cells as a carrier proteome population to increase mapping efficiency.[9] To assess the degree of post-transcriptional regulation and the overall correlation between transcriptome and proteome in *Tet2*$^{-/-}$ HSPCs, we performed bulk RNA-seq on WT and *Tet2*$^{-/-}$ CD150$^+$ and CD150$^-$ cells. Using just 10,000–30,000 cells per sample, the MS analysis identified 4,133 unique proteins, out of which 3,989 (~97%) were reliably quantified across all cell populations (Table S4). Notably, our low cell number multiplex captured 55% of proteins previously quantified in HSCs[34] using only 2.5%–7.5% of total cell input per sample, and protein coverage was similar to recent studies using 40,000–100,000 hematopoietic progenitor cells per sample and a similar methodology for sample preparation and MS analysis.[11,12] Normalization against sample loading successfully corrected for the difference in cell number between samples (Figure S2A).

Following the generation of global proteomic datasets for WT and *Tet2*$^{-/-}$ HSPCs, we first compared proteome and transcriptome datasets. Principal component analysis (PCA) of proteomic data clearly separated cell populations according to genetic background in principle component (PC) 1 (Figure 3B), whereas separation in PC1 and PC2 for RNA-seq data was driven by cell type rather than *Tet2* mutational status (Figure 3C), indicating that the transcriptome and proteome have substantial global differences in regulation. Interestingly, in Figure 3B, the *Tet2*$^{-/-}$ CD150$^-$ samples are most distinct compared to WT populations and *Tet2*$^{-/-}$ CD150$^+$ cells, suggesting that proteome dysregulation in *Tet2*$^{-/-}$ hematopoietic cells is exacerbated as cells mature past the stem cell state, a finding potentially associated with the myeloid skewing observed with loss of TET2.[1,5] To a lesser extent, this was also reflected in PC2 of the transcriptome data (Figure 3C). To investigate which proteins best separated

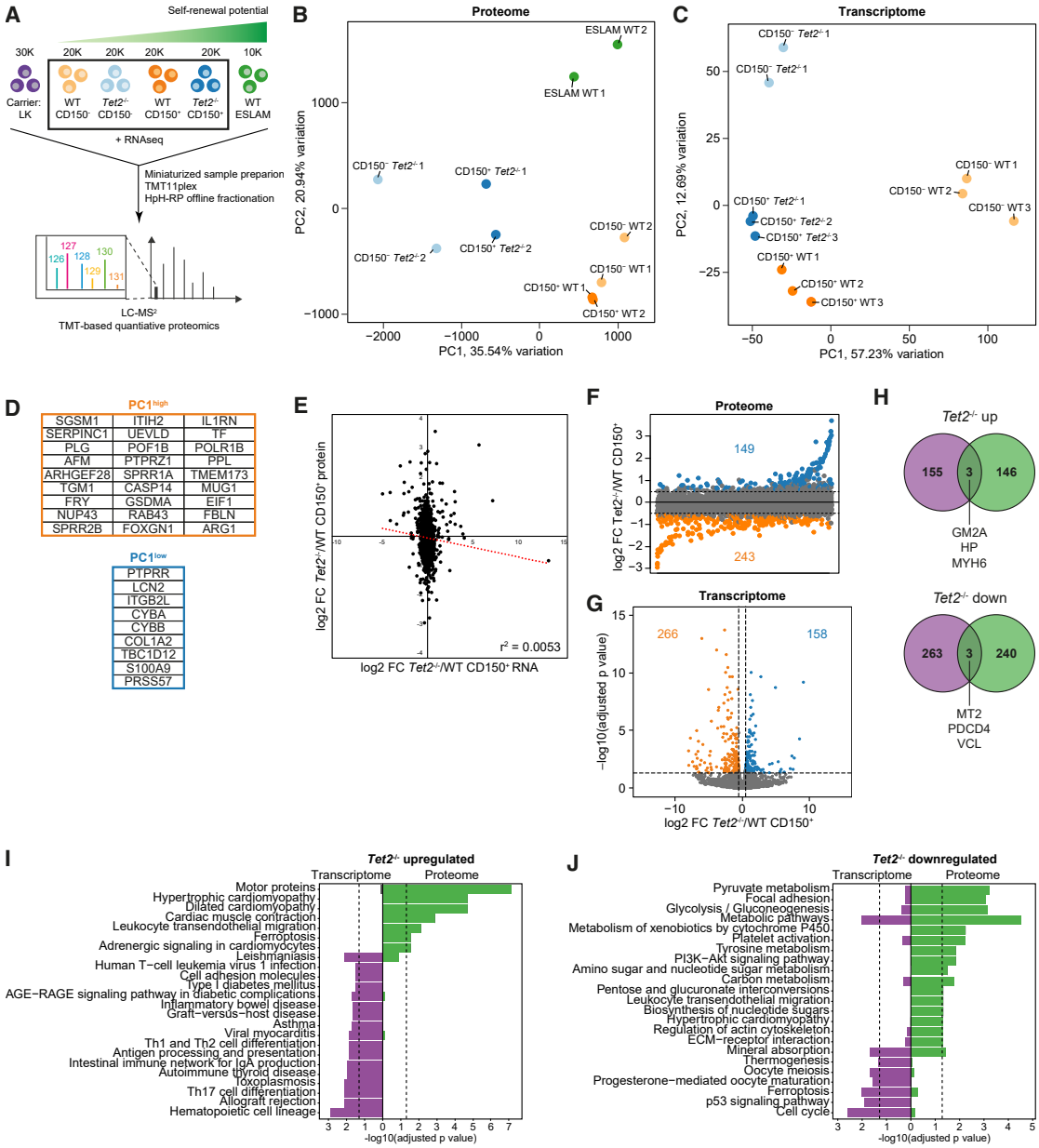


Figure 3. The proteome and transcriptome highlight alterations in disparate cellular processes upon loss of *Tet2* in HSPCs

(A) Workflow for liquid chromatography-tandem mass spectrometry analysis of WT and *Tet2*$^{-/-}$ HSPCs.

(B and C) PCA of proteomic (B) and transcriptomic (C) data.

(D) Top 15% loadings from PC1 in (B).

(E) Correlation between protein and gene expression differences (log2 FC) between *Tet2*$^{-/-}$ and WT CD150^{+} cells. The dotted red line indicates the linear trendline. r^2 represents the Pearson correlation coefficient.

(F) Protein expression difference between *Tet2*$^{-/-}$ and WT CD150^{+} cells. Candidate target proteins (log2 FC > 0.5 across all comparisons between *Tet2*$^{-/-}$ and WT CD150^{+} cells) enriched and depleted in *Tet2*$^{-/-}$ relative to WT CD150^{+} cells are shown in blue and orange, respectively.

(G) Gene expression difference between *Tet2*$^{-/-}$ and WT CD150^{+} cells. Candidate target genes (log2 FC > 0.5 and adjusted *p* value < 0.05) enriched and depleted in *Tet2*$^{-/-}$ relative to WT CD150^{+} cells are shown in blue and orange, respectively.

(H) Overlap between candidate target genes and proteins enriched or depleted in *Tet2*$^{-/-}$ relative to WT CD150^{+} cells.

(I and J) KEGG pathway analysis of candidate target proteins and genes enriched (I) or depleted (J) in *Tet2*$^{-/-}$ relative to WT CD150^{+} cells. The dotted lines mark adjusted *p* value = 0.05.

See also Figure S2.

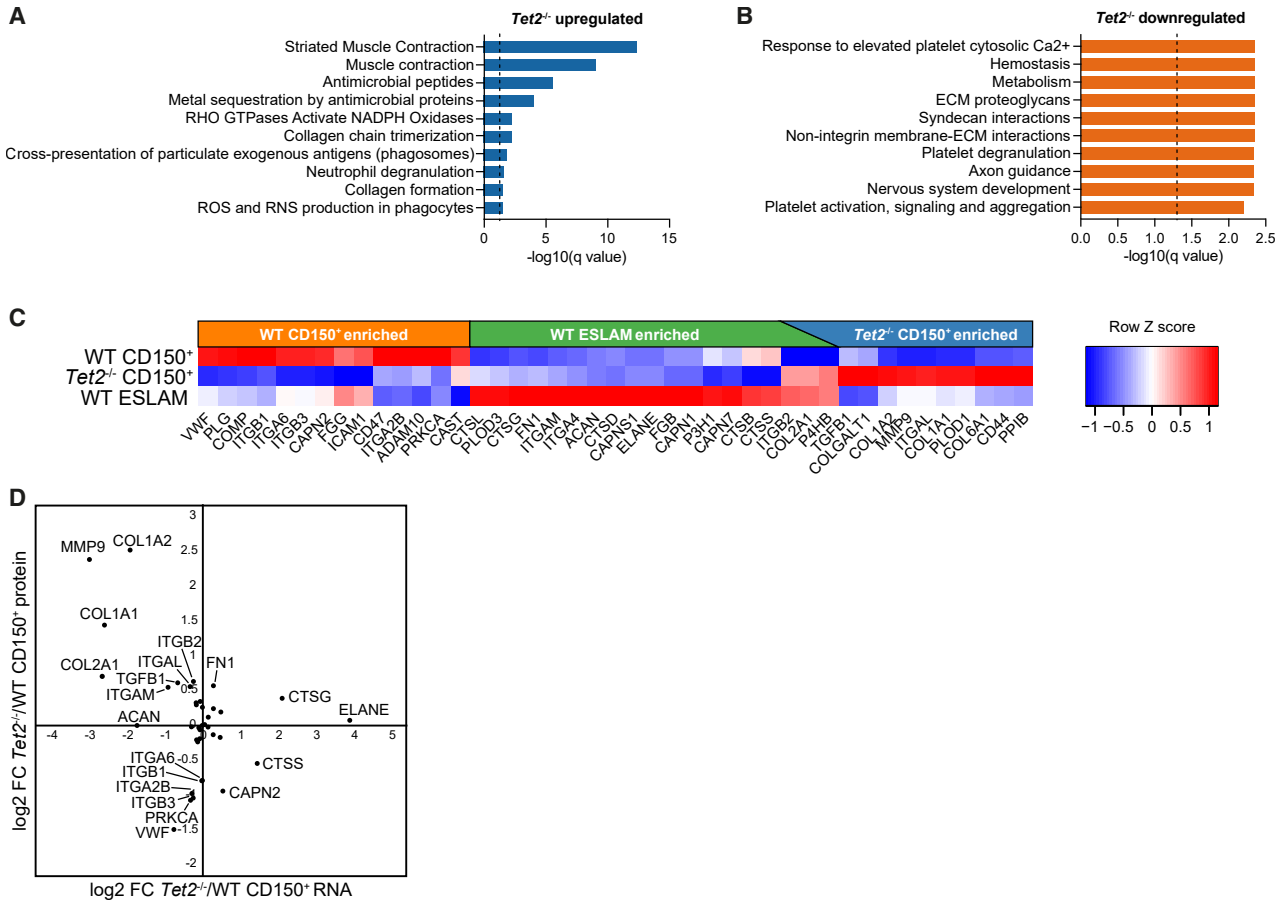


Figure 4. Expression of ECM proteins is altered upon loss of *Tet2* and correlates with self-renewal potential

(A and B) Reactome pathway analysis of candidate target proteins enriched (A) or depleted (B) in *Tet2*$^{-/-}$ relative to WT CD150^{+} cells. The dotted lines mark *q* value = 0.05.

(C) Relative abundance of proteins contained within the reactome pathway "ECM organization" in WT CD150^{+}, *Tet2*$^{-/-}$ CD150^{+}, and WT ESLAM cells.

(D) Correlation between protein and gene expression differences (log2 FC) of proteins/genes contained within the reactome pathway "ECM organization" between *Tet2*$^{-/-}$ and WT CD150^{+} cells.

WT and *Tet2*$^{-/-}$ cells, we extracted the top and bottom 15% of loadings for PC1 (27 WT-enriched and 9 *Tet2*$^{-/-}$-enriched; Figure 3D). Proteins associated with the TET2-deficient cell populations included inflammatory proteins S100A9 and LCN2, which have been reported to be involved in the pathogenesis of myelodysplastic syndrome and myelofibrosis, respectively.[35,36] WT-enriched loadings included interleukin-1 (IL-1) receptor antagonist protein IL-1RA, a protein important for dampening IL-1-driven inflammation, which was recently shown to contribute to the clonal outgrowth of *Tet2*$^{+/-}$ HSPCs during aging.[37] Two of the top loadings for WT cells in PC1 were anticoagulant proteins antithrombin III (SERPINC1) and plasminogen (PLG), suggesting potential dysregulation of clotting mechanisms in *Tet2*$^{-/-}$ HSPCs. Intriguingly, LoF mutations in *TET2* have recently been linked to an increased risk for myeloproliferative neoplasm (MPN)-associated thrombosis, and patients carrying *TET2* mutations have significantly lower levels of antithrombin III than those with intact *TET2* expression.[38] Our data suggest that the increased production of inflammatory proteins and impaired clotting functions associated with mutations in *Tet2* are evident already at the level of HSPCs.

Next, we specifically interrogated the immature CD150^{+} population. There was no correlation between the protein and RNA datasets for the CD150^{+} population (r^2 = 0.0053; Figure 3E), which accords with previous studies that reported a high degree of post-transcriptional regulation in early HSPCs,[10,14,15] and poor correlation between proteome and transcriptome in HSCs compared to downstream progenitors.[14] The correlation between transcriptome and proteome in the CD150^{-} population was equally poor (r^2 = 0.0056; Figure S2B). To further assess concordance between the two datasets, we next generated shortlists of candidate proteins and genes and compared the lists (Figures 3F and 3G; Table S4). In the CD150^{+} LT-HSC-enriched population, only 6 candidate targets overlapped between the two datasets; these included inflammatory biomarker haptoglobin, which showed elevated expression in *Tet2*$^{-/-}$ cells, and anti-inflammatory methallothionein 2, which was enriched in WT relative to knockout cells (Figure 4H). The overlap in the

$CD150^-$ population was substantially higher, with 33 and 97 candidate proteins (representing 15% and 21% of all candidate proteins, respectively) significantly enriched at the transcript level in *Tet2*$^{-/-}$ and WT cells, respectively (Figure S2C).

Despite the low overlap of specific targets between the MS and RNA-seq datasets, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of candidate proteins/genes in the $CD150^+$ LT-HSC-enriched population highlighted some commonalities (Figures 3I and 3J), suggesting significant dysregulation in HSPCs upon loss of *Tet2*. The most prominent of these categories were "metabolic pathways" and "mineral absorption," both strongly enriched among proteins and genes downregulated in *Tet2*$^{-/-}$ relative to WT $CD150^+$ cells. Surprisingly, the top enriched pathway among proteins upregulated in *Tet2*$^{-/-}$ relative to WT cells was "motor proteins," and the proteomic datasets highlighted changes in additional pathways related to the actomyosin motor and the ECM upon loss of TET2 in immature HSPCs, such as "leukocyte transendothelial migration" and "ECM-receptor interaction" (Figure 3J). Metabolism and the actomyosin motor were also altered in $CD150^-$ cells, pointing to their dysregulation as a general feature of TET2 loss in HSPCs (Figures S2D and S2E). In line with the deregulated expression of thrombosis-related proteins in *Tet2*$^{-/-}$ $CD150^+$ cells (Figure 3D), "platelet activation" was one of the pathways enriched among proteins with lower expression in mutant compared to WT cells (Figure 4J).

ECM interactions regulate self-renewal of *Tet2*$^{-/-}$ HSCs

Reactome pathway enrichment analysis (Figures 4A and 4B) identified several pathways related to the actomyosin motor and to ECM organization among proteins more highly expressed in either *Tet2*$^{-/-}$ (Figure 4A) or WT (Figure 4B) cells. Analysis of individual proteins contained within the reactome pathway "ECM organization" revealed distinct expression patterns of ECM proteins in the three most HSC-enriched cell populations in our proteomic data (Figure 4C), with proteins either being depleted or enriched with increasing self-renewal potential. A subset of these proteins, including four collagen family members, displayed particularly high expression in *Tet2*$^{-/-}$ $CD150^+$ cells, and three of these (ITGB2, COL2A2, and P4HB) overlapped with the more highly HSC-enriched WT ESLAM population. Notably, bulk RNA-seq data did not capture these changes, instead showing that ECM-associated genes were either unchanged between WT and TET2-deficient cells or had opposing expression patterns compared to the proteome data (Figure 4D).

All ECM proteins identified in the MS data are known interaction partners, and, intriguingly, proteins enriched in *Tet2*$^{-/-}$ and WT $CD150^+$ cells were positioned in distinct areas of the interaction network (Figure 5A). Several findings in our proteome analysis pointed toward a deregulation of platelet-associated pathways in *Tet2*-mutant HSCs (Figures 3D and 3J). We therefore next decided to focus on ECM proteins with known roles in megakaryopoiesis—ITGA2B/CD41, ITGB3/CD61, and von Willebrand factor (vWF). Intracellular flow cytometry confirmed that a significantly smaller proportion of *Tet2*$^{-/-}$ $CD150^+$ cells express vWF and ITGA2B/CD41 (Figures 5B and 5C). The proportion of ITGB3/$CD61^+$ $CD150^+$ cells was also reduced in 2/3 assayed animals (Figure S3A). Immunofluorescence measured by confocal microscopy provided further proof of a lower abundance of vWF in mutant relative to WT $CD150^+$ HSPCs (Figures 5D, S3B, and S3C). As both ITGA2B/CD41 and vWF mark HSCs primed toward platelet production,[15,39] our findings here suggest that the low expression of proteins associated with platelet activation, *Tet2*$^{-/-}$ $CD150^+$, relative to WT cells, is due to a loss of megakaryocyte-biased HSCs within the stem cell pool.

In order to test the functional impact of HSC-vWF interaction, we cultured ESLAM HSCs derived from WT and *Tet2*$^{-/-}$ animals using recently published HSC culture conditions[41] in tissue culture plates functionalized with vWF (Figure 5E). In these conditions, vWF did not affect the expansion of primitive Lin^- $Sca\text{-}1^+$ $cKit^+$ (LSK) HSPCs and $EPCR^+$ LSK (ELSK)[40] HSCs (Figures S3D and S3E). However, in line with our findings regarding deregulated expression of platelet-associated proteins in *Tet2*$^{-/-}$ HSCs (Figures 3D and 3J) and a smaller proportion of platelet-biased cells within the *Tet2*-deficient HSC pool (Figures 5B and 5C), we observed a lower output of megakaryocytic (MK; $CD41^+$ $CD42d^+$) cells from *Tet2*$^{-/-}$ relative to WT HSCs at steady-state, which was partially rescued by vWF (Figure 5F). WT HSCs were unaffected by vWF in this regard, suggesting that TET2-deficient HSCs are more sensitive to the pro-thrombotic effects of vWF signaling.

We hypothesized that presenting ECM proteins to cells in a setting more closely resembling the bone marrow (BM) niche may elicit different effects on HSC expansion and differentiation. To this end, we utilized STEMBOND hydrogels functionalized with vWF or hyaluronan (HA) as HA receptors CD44 and ITGB1/CD29 were among proteins showing self-renewal and *Tet2*-status-associated differences in expression (Figure 4C). As expected, loss of TET2 confers a self-renewal advantage in HSCs[1,5] that presents itself as an increase in the frequency of LSK HSPCs and ELSK HSCs compared to WT controls (Figures 5G, 5H, S3F, and S3G). Strikingly, this mutant self-renewal advantage was abolished when cells were cultured in the presence of hydrogel-anchored vWF (Figures 5G and 5H), and, to a lesser extent, HA (Figures S3F and S3G). The reduced fraction of progenitors (LSK) and HSCs (ELSK) was only observed in the *Tet2*$^{-/-}$ cultures, suggesting that vWF is selectively (and negatively) influencing TET2-mediated HSC expansion. Importantly, transplantation of cells cultured in the presence of niche-anchored vWF revealed that vWF-exposed *Tet2*$^{-/-}$ HSCs have an impaired capacity to engraft in the BM (Figure 5I). Together these data show that exposure to different ECMs can regulate both the differentiation and self-renewal capacity of *Tet2*$^{-/-}$ HSCs and firmly establish that features captured exclusively by proteomic analysis provide novel insight into HSC biology.

DISCUSSION

TET2 has been widely studied in the context of clonal hematopoiesis and myeloid malignancies due to its recurring LOF mutations and its functional role in HSC self-renewal. The exact molecular mechanism of the increased self-renewal initiated by loss of TET2 function remains unclear, and transcriptomic studies to date have been unable to identify clear drivers of increased HSC self-renewal. Our integrated ATAC-seq and

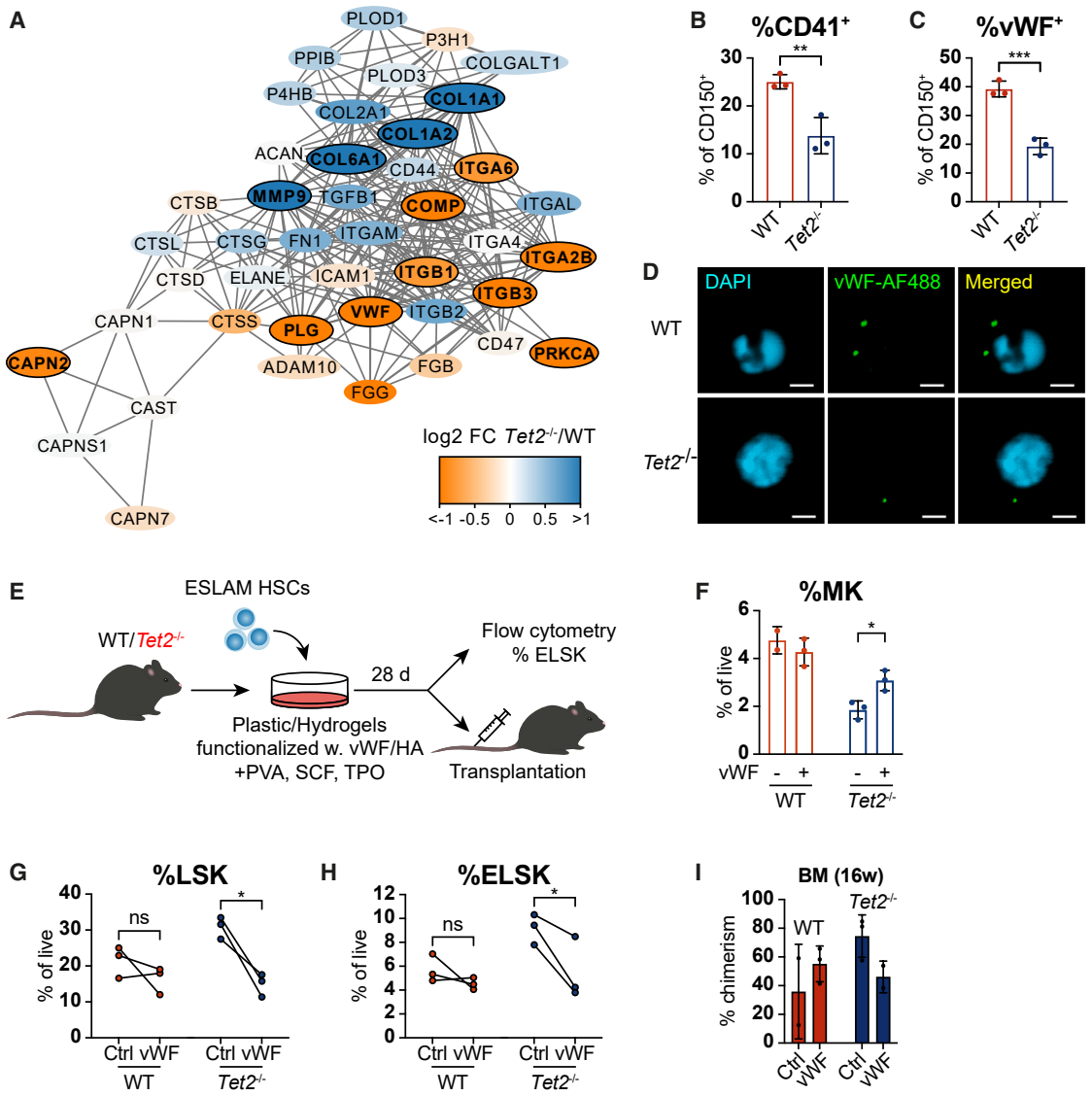


Figure 5. Interaction with niche-anchored vWF inhibits expansion of *Tet2*$^{-/-}$ HSCs

(A) Protein interaction network of proteins contained within the reactome pathway "ECM organization." Outlined nodes represent candidate target proteins.

(B and C) Proportion of cells expressing CD41 (B) and vWF (C) of WT and *Tet2*$^{-/-}$ CD150^{+} cells. $n = 3$ individual mice per genotype.

(D) Representative immunofluorescence images of vWF in WT and *Tet2*$^{-/-}$ CD150^{+} HSPCs. Scale bars, 2 μM.

(E) Experimental workflow for assessing the effect of ECM-HSC interaction *ex vivo*.

(F) Frequency of megakaryocytic (MK) cells (CD41^{+} CD42d^{+}) following 28-day culture in tissue culture plates functionalized with vWF. $n = 2$ and $n = 3$ individual mice for WT-vWF and all other conditions, respectively.

(G and H) Frequency of HSPCs (LSK; G) and HSCs (ELSK[40]; H) following 28-day culture of WT and *Tet2*$^{-/-}$ ESLAM HSCs on hydrogels functionalized with vWF. $n = 3$ individual mice per genotype.

(I) Bone marrow chimerism at 16 weeks post-transplantation of WT and *Tet2*$^{-/-}$ ESLAM HSCs cultured for 28 days on hydrogels functionalized with vWF. $n = 2$ and $n = 3$ individual mice for WT-vWF and all other conditions, respectively. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; and ns, non-significant. Error bars = SD.

See also Figure S3.

scRNA-seq analyses implicated KLF family members and key hematopoietic TFs *Gata2*, *Gata3*, *Fli1*, and *Runx1* in the regulation of *Tet2*-mutant HSCs, confirming that loss of TET2-driven demethylase activity causes widespread and functionally relevant dysregulation of the epigenome and transcriptome of HSCs. Specifically interrogating changes in methylation status upon loss of TET2 using, for example, bisulfite-seq, represents an important future direction to fully understand the effects of *TET2* mutations on the epigenome.

Moving beyond transcriptomics to capture global proteome level data, we identify a distinct set of ECM molecules with specific roles in altering the function of *Tet2*-mutated HSCs.

Intriguingly, ECM protein abundance does not correlate with the expression of associated transcripts as determined by RNA-seq, suggesting a high degree of post-transcriptional regulation in this group of proteins. Alternatively, ECM proteins may be transcribed by other cell types and bind to or be taken up by HSCs; however, our similar findings in hematopoietic cell line HoxB8-FL, which is a system devoid of other cell types, speak against this theory.

Techniques to undertake global proteomics in limited numbers of cells are rapidly evolving, such that even over the course of this study, single-cell proteomic methodologies have been developed that obtain thousands of unique proteins in individual human HSCs.[13] While our method still captures more proteins in the HSC population, it remains completely blind to cell-cell heterogeneity. Application of single-cell proteomic approaches to larger numbers of cells in normal and diseased states will permit dissection of more complete pathways in addition to understanding the functional and molecular heterogeneity of different cell types.

Functionalizable hydrogels represent a novel tool to study components of the ECM and their impact on cell function. In this study, we utilize STEMBOND hydrogels,[42] which permit robust matrix tethering and have tunable stiffness, to test the ECM components that emerged from our proteomic studies. This permits the investigation of biophysical properties of cells with LoF *Tet2* mutations, and we demonstrate a clear role for vWF in specifically restricting TET2-mediated HSC expansion *in vitro* and BM engraftment *in vivo*. While our study shows a clear selective effect of exogenous vWF in *Tet2*-mutant HSCs, whether mutant and healthy cells differ in their interaction with vWF and other ECM *in vivo* remains to be determined.

Furthermore, our proteomic data point to actomyosin motor control as being dysregulated when *Tet2* is mutated. While TET2 has been previously implicated in cytoskeleton organization in ovarian cells[43] and in smooth muscle cell plasticity,[44] its role in the actomyosin motor of HSPCs has not yet been described. In this vein, myosins are upregulated during inflammatory stress in HSCs,[15] and their inhibition impairs growth and survival of acute myeloid leukemia cells,[45] implying that their high expression in *Tet2*-mutant HSPCs may be linked to the aberrant phenotype of the cells.

There were also a number of proteins more highly expressed in WT cells, in particular a set of proteins related to platelet function. Interestingly, the WT-enriched proteins representing these processes include those with anti-coagulant (e.g., PLG, SERPINC1, and ANXA5) as well as pro-thrombotic (e.g., FGG, VWF, and ITGB3) functions, suggesting that TET2 loss results in a general decrease in expression of proteins related to platelet production and coagulation in HSPCs and perhaps even related to a shift in HSC subtypes away from MK-biased HSCs. In line with this, we found that the HSC pool of *Tet2*-deficient animals contains fewer vWF^+ and $CD41^+$ MK-biased cells and that $Tet2^{-/-}$ HSCs produce fewer MK cells *in vitro*, a defect that can be partially reversed by the addition of vWF.

Clonal hematopoiesis and myeloid malignancies driven by mutations in *TET2* predominantly affect individuals over 70 years of age[46] (corresponding to 18–24 months of age in mice), and *TET2* loss in these patients is often hetero- rather than homozygous.[47] Our proteome analysis was performed in 30-week-old $Tet2^{-/-}$ animals and as such provides insight into molecular changes occurring in fully TET2-deficient HSCs in middle age. Conducting the same analyses in older and $Tet2^{+/-}$ animals will be an important future direction in order to fully understand *TET2* LoF-driven pathology in a clinically relevant setting. Of further note, our current data are in a transplantation setting, and selection in people typically operates in the absence of transplantation and occurs over many decades, and the relative role of ECM molecules in mediating these selection pressures *in vivo* is completely unknown. Further work to modulate ECM over sustained periods *in vivo* would therefore be of great future interest.

Overall, our study emphasizes the importance of moving beyond transcriptomic studies to reveal new aspects of mutant cell biology during processes of HSC self-renewal and leukemogenesis. In particular, proteomic studies have triggered the investigation of the mechanisms by which the ECM alters HSC self-renewal and influences clonal advantage competition during aging and disease. How changes in ECM composition throughout aging might contribute to the clinical observations of clonal hematopoiesis and pre-leukemic cell expansion is an intriguing concept that accords with recent studies showing that integrins and their molecular regulators underpin healthy aging.[48–50] This in turn, opens up new lines of thinking regarding potential therapies, and new tools such as functionalizable hydrogels will accelerate discoveries that reach well beyond the HSC system for applications in numerous other stem cell systems as has already been pioneered for oligodendrocyte precursors[51] and pluripotent stem cells.[42]

Limitations of the study

While the ECM-functionalized hydrogels provide evidence that *Tet2*-mutant cells respond differently than WT cells to distinct components of the microenvironment (e.g., vWF), these data do not formally demonstrate that increased vWF protein *in vivo* would drive a functional decline of TET2-mutant HSCs. Similarly, while many components of the ECM are dysregulated in *Tet2*-mutant HSCs, their relevance to steering clonal selection or clonal hematopoiesis more broadly remains unclear and will require future work in spatial-omics, mechanobiology, and new methods to dissect HSC competition *in vitro* and *in vivo*. Finally, our data are not yet sufficient to understand the heterogeneity of ECM composition in individual cells, which will require new technological advances in single-cell proteomics and/or *in vivo* HSC-niche reporters coupled with spatial-omics tools.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, David G. Kent (david.kent@york.ac.uk).

Materials availability

This study did not generate new, unique reagents.

Data and code availability

- The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[52] partner repository with the dataset identifier PXD059814. scRNA-seq and scATAC-seq data have been

deposited to the Sequence Read Archive with dataset identifiers PRJNA1210137 and PRJNA1210127, respectively.
- Code used to analyze raw data in this manuscript is available from the lead contact upon request.
- Any additional information is available from the lead contact upon request.

ACKNOWLEDGMENTS

The work in Prof. D.G.K.'s laboratory was supported by the European Research Council Starting Grant (ERC-2016-STG-715371), the Cancer Research UK Programme Foundation Award (DCRPGF\100008), the MRC-AMED joint award (MR/V005502/1), the UK Medical Research Council (MC_PC_21043; MR/Y011945/1), and the Blood Cancer UK. The Kent lab is also supported by the National Institute for Care and Health Research Leeds Biomedical Research Centre (NIHR203331). Dr. M.J. was supported by a Swedish Research Council International Postdoc grant (2021-00185) and project grants from The Royal Physiographic Society of Lund and Siv-Inger and Per-Erik Anderssons Memorial Fund. M.G. was the recipient of a Biotechnology and Biological Sciences Research Council White Rose DTP PhD Studentship (BB/T007222/1). Dr. M.S.S. was the recipient of a Biotechnology and Biological Sciences Research Council Industrial Collaborative Award in Science and Engineering (iCase) PhD Studentship. Dr. J.L.C.C. was supported by an MRC PhD Studentship under the University of Cambridge Doctoral Training Program. Dr D.B. was supported by a Wellcome PhD Studentship. Work in Cambridge was further supported by core support grants from the Wellcome and Medical Research Council (MRC) to the Wellcome - MRC Cambridge Stem Cell Institute (203151/Z/16/Z). The work in Prof. B.G.'s laboratory was supported by Wellcome (206328/Z/17/Z and 203151/Z/16/Z), Blood Cancer UK, Cancer Research UK (C1163/A21762), and UKRI Medical Research Council (MC_PC_17230). The authors would like to thank Professor Anjana Rao for originally providing the Tet2$^{-/-}$ mouse; and Drs. Veronique Voisin, Andrew Zeng, Matthew Care, and Alastair Droop for their valuable input on scRNA-seq and scATAC-seq analyses. The authors also greatly acknowledge the expert technical assistance provided by the Genomics, Imaging & Cytometry, and Metabolomics and Proteomics Facilities within the Biosciences Technology Facility at the University of York; and the CIMR Flow Cytometry core (Reiner Schulte, Chiara Cossetti, and Gabriela Grondys-Kotarba). We also thank Tina Hamilton and Dean Pask for technical assistance and the staff of the animal facilities at the University of York and the University of Cambridge. The Viking cluster was used during this project, which is a high-performance computing facility provided by the University of York. We are grateful for additional computational support from the University of York, IT Services, and the Research IT team.

AUTHOR CONTRIBUTIONS

D.G.K., J.C., and B.G. designed the study together with D.B. and M.J.; D.B., I.K., and M.G. performed scRNA-seq and scATAC-seq experiments under the supervision of N.K.W., B.G., and D.G.K.; D.B. and T.I.R. performed proteome analysis under the supervision of J.C.; D.B. and M.G. analyzed scRNA-seq data with S.M.; M.G. analyzed scATAC-seq data with B.T., M.S.V., A.N.H., and S.A.; M.G. performed integrative analysis of scRNA-seq and scATAC-seq data with B.T. and B.A.H.; M.J. and D.B. analyzed the proteome data; M.B., M.J., J.L.C.C., M.S.S., A.J.H., E.B., J.M., S.E., M.G., J.A.R.L., and L.C.C. performed HSC isolation, flow cytometry, and cell culture experiments with experimental design input from S.Y. and W.J.B.; M.J. analyzed all flow cytometry data; G.V., H.K., A.C.H., E.B., and J.M. performed animal work; M.J. prepared the figures; D.G.K. supervised the study; and M.J., M.G., and D.G.K. wrote the paper together with D.B., A.J.H., E.B., J.M., and input from other authors.

DECLARATION OF INTERESTS

The D.G.K. lab has received research funding from STEMBOND Inc. (Cambridge, UK) to conduct experiments using hematopoietic cells that were unrelated to this manuscript.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODELS AND STUDY PARTICIPANTS DETAILS
  - Mice
  - Cell lines
- METHOD DETAILS
  - Flow cytometry and FACS
  - Immunofluorescence
  - HSC gel culture
  - Bulk RNA-seq
  - scATAC-seq
  - Plate-based scRNA-seq
  - Sample preparation for proteome analysis
  - LC-MS/MS analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - scATAC-seq
  - Plate-based scRNA-seq
  - Integrative scATAC-seq and scRNA-seq data analysis
  - Protein identification and quantification
  - Bioinformatic analysis of proteomic and bulk RNA-seq data
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.116770.

Received: March 8, 2025
Revised: August 19, 2025
Accepted: November 28, 2025

REFERENCES

1. Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell *20*, 11–24. https://doi.org/10.1016/j.ccr.2011.06.001.
2. Shepherd, M.S., Li, J., Wilson, N.K., Oedekoven, C.A., Li, J., Belmonte, M., Fink, J., Prick, J.C.M., Pask, D.C., Hamilton, T.L., et al. (2018). Single-cell approaches identify the molecular network driving malignant hematopoietic stem cell self-renewal. Blood *132*, 791–803. https://doi.org/10.1182/blood-2017-12-821066.
3. Kameda, T., Shide, K., Yamaji, T., Kamiunten, A., Sekine, M., Taniguchi, Y., Hidaka, T., Kubuki, Y., Shimoda, H., Marutsuka, K., et al. (2015). Loss of TET2 has dual roles in murine myeloproliferative neoplasms: disease sustainer and disease accelerator. Blood *125*, 304–315. https://doi.org/10.1182/blood-2014-04-555508.
4. Chen, E., Schneider, R.K., Breyfogle, L.J., Rosen, E.A., Poveromo, L., Elf, S., Ko, A., Brumme, K., Levine, R., Ebert, B.L., and Mullally, A. (2015). Distinct effects of concomitant Jak2V617F expression and Tet2 loss in mice promote disease progression in myeloproliferative neoplasms. Blood *125*, 327–335. https://doi.org/10.1182/blood-2014-04-567024.
5. Ko, M., Bandukwala, H.S., An, J., Lamperti, E.D., Thompson, E.C., Hastie, R., Tsangaratou, A., Rajewsky, K., Koralov, S.B., and Rao, A. (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc. Natl. Acad. Sci. USA *108*, 14566–14571. https://doi.org/10.1073/pnas.1112317108.
6. Dimayacyac-Esleta, B.R.T., Tsai, C.-F., Kitata, R.B., Lin, P.-Y., Choong, W.-K., Lin, T.-D., Wang, Y.-T., Weng, S.-H., Yang, P.-C., Arco, S.D., et al. (2015). Rapid High-pH Reverse Phase StageTip for Sensitive

Small-Scale Membrane Proteomic Profiling. Anal. Chem. *87*, 12016–12023. https://doi.org/10.1021/acs.analchem.5b03639.

7. Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat. Methods *11*, 319–324. https://doi.org/10.1038/nmeth.2834.
8. Schoof, E.M., Furtwängler, B., Üresin, N., Rapin, N., Savickas, S., Gentil, C., Lechman, E., Keller, U.A.d., Dick, J.E., and Porse, B.T. (2021). Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. Nat. Commun. *12*, 3341. https://doi.org/10.1038/s41467-021-23667-y.
9. Slavov, N. (2021). Single-cell protein analysis by mass spectrometry. Curr. Opin. Chem. Biol. *60*, 1–9. https://doi.org/10.1016/j.cbpa.2020.04.018.
10. Jassinskaja, M., and Hansson, J. (2022). The Opportunity of Proteomics to Advance the Understanding of Intra- and Extracellular Regulation of Malignant Hematopoiesis. Front. Cell Dev. Biol. *10*, 824098.
11. Jassinskaja, M., Pimková, K., Arh, N., Johansson, E., Davoudi, M., Pereira, C.-F., Sitnicka, E., and Hansson, J. (2021). Ontogenic shifts in cellular fate are linked to proteotype changes in lineage-biased hematopoietic progenitor cells. Cell Rep. *34*, 108894. https://doi.org/10.1016/j.celrep.2021.108894.
12. Jassinskaja, M., Ghosh, S., Watral, J., Davoudi, M., Claesson Stern, M., Daher, U., Eldeeb, M., Zhang, Q., Bryder, D., and Hansson, J. (2024). A complex interplay of intra- and extracellular factors regulates the outcome of fetal- and adult-derived MLL-rearranged leukemia. Leukemia *38*, 1115–1130. https://doi.org/10.1038/s41375-024-02235-5.
13. Furtwängler, B., Üresin, N., Richter, S., Schuster, M.B., Barmpouri, D., Holze, H., Wenzel, A., Grønbæk, K., Theilgaard-Mönch, K., Theis, F.J., et al. (2025). Mapping early human blood cell differentiation using single-cell proteomics and transcriptomics. Science *390*, eadr8785. https://doi.org/10.1126/science.adr8785.
14. Zaro, B.W., Noh, J.J., Mascetti, V.L., Demeter, J., George, B., Zukowska, M., Gulati, G.S., Sinha, R., Flynn, R.A., Banuelos, A., et al. (2020). Proteomic analysis of young and old mouse hematopoietic stem cells and their progenitors reveals post-transcriptional regulation in stem cells. eLife *9*, e62210. https://doi.org/10.7554/eLife.62210.
15. Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, Á.M., Schnell, A., Hexel, K., et al. (2015). Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. Cell Stem Cell *17*, 422–434. https://doi.org/10.1016/j.stem.2015.07.007.
16. Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181. https://doi.org/10.1038/nprot.2014.006.
17. Wilson, N.K., Kent, D.G., Buettner, F., Shehata, M., Macaulay, I.C., Calero-Nieto, F.J., Sánchez Castillo, M., Oedekoven, C.A., Diamanti, E., Schulte, R., et al. (2015). Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. Cell Stem Cell *16*, 712–724. https://doi.org/10.1016/j.stem.2015.04.004.
18. Rasmussen, K.D., Jia, G., Johansen, J.V., Pedersen, M.T., Rapin, N., Bagger, F.O., Porse, B.T., Bernard, O.A., Christensen, J., and Helin, K. (2015). Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev. *29*, 910–922. https://doi.org/10.1101/gad.260174.115.
19. Schaniel, C., Ang, Y.-S., Ratnakumar, K., Cormier, C., James, T., Bernstein, E., Lemischka, I.R., and Paddison, P.J. (2009). Smarcc1/Baf155 couples self-renewal gene repression with changes in chromatin structure in mouse embryonic stem cells. Stem Cell. *27*, 2979–2991. https://doi.org/10.1002/stem.223.
20. Balogh, P., Adelman, E.R., Pluvinage, J.V., Capaldo, B.J., Freeman, K.C., Singh, S., Elagib, K.E., Nakamura, Y., Kurita, R., Sashida, G., et al. (2020). RUNX3 levels in human hematopoietic progenitors are regulated by aging and dictate erythroid-myeloid balance. Haematologica *105*, 905–913. https://doi.org/10.3324/haematol.2018.208918.
21. NandyMazumdar, M., Paranjapye, A., Browne, J., Yin, S., Leir, S.-H., and Harris, A. (2021). BACH1, the master regulator of oxidative stress, has a dual effect on CFTR expression. Biochem. J. *478*, 3741–3756. https://doi.org/10.1042/BCJ20210252.
22. Ton, M.-L.N., Keitley, D., Theeuwes, B., Guibentif, C., Ahnfelt-Rønne, J., Andreassen, T.K., Calero-Nieto, F.J., Imaz-Rosshandler, I., Pijuan-Sala, B., Nichols, J., et al. (2023). An atlas of rabbit development as a model for single-cell comparative genomics. Nat. Cell Biol. *25*, 1061–1072. https://doi.org/10.1038/s41556-023-01174-0.
23. Garcia-Alonso, L., Holland, C.H., Ibrahim, M.M., Turei, D., and Saez-Rodriguez, J. (2019). Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. *29*, 1363–1375. https://doi.org/10.1101/gr.240663.118.
24. Luchsinger, L.L., Strikoudis, A., Danzl, N.M., Bush, E.C., Finlayson, M.O., Satwani, P., Sykes, M., Yazawa, M., and Snoeck, H.-W. (2019). Harnessing Hematopoietic Stem Cell Low Intracellular Calcium Improves Their Maintenance In Vitro. Cell Stem Cell *25*, 225–240.e7. https://doi.org/10.1016/j.stem.2019.05.002.
25. Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. *38*, W214–W220. https://doi.org/10.1093/nar/gkq537.
26. Park, C.S., Lewis, A., Chen, T., and Lacorazza, D. (2019). Concise Review: Regulation of Self-Renewal in Normal and Malignant Hematopoietic Stem Cells by Krüppel-Like Factor 4. Stem Cells Transl. Med. *8*, 568–574. https://doi.org/10.1002/sctm.18-0249.
27. Oedekoven, C.A., Belmonte, M., Bode, D., Hamey, F.K., Shepherd, M.S., Che, J.L.C., Boyd, G., McDonald, C., Belluschi, S., Diamanti, E., et al. (2021). Hematopoietic stem cells retain functional potential and molecular identity in hibernation cultures. Stem Cell Rep. *16*, 1614–1628. https://doi.org/10.1016/j.stemcr.2021.04.002.
28. Ross, K., Sedello, A.K., Todd, G.P., Paszkowski-Rogacz, M., Bird, A.W., Ding, L., Grinenko, T., Behrens, K., Hubner, N., Mann, M., et al. (2012). Polycomb group ring finger 1 cooperates with Runx1 in regulating differentiation and self-renewal of hematopoietic cells. Blood *119*, 4152–4161. https://doi.org/10.1182/blood-2011-09-382390.
29. de Pater, E., Kaimakis, P., Vink, C.S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P.S., Camper, S.A., Speck, N., and Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. J. Exp. Med. *210*, 2843–2850. https://doi.org/10.1084/jem.20130751.
30. Frelin, C., Herrington, R., Janmohamed, S., Barbara, M., Tran, G., Paige, C.J., Benveniste, P., Zuñiga-Pflücker, J.-C., Souabni, A., Busslinger, M., and Iscove, N.N. (2013). GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. Nat. Immunol. *14*, 1037–1044. https://doi.org/10.1038/ni.2692.
31. Redecke, V., Wu, R., Zhou, J., Finkelstein, D., Chaturvedi, V., High, A.A., and Häcker, H. (2013). Hematopoietic progenitor cell lines with myeloid and lymphoid potential. Nat. Methods *10*, 795–803. https://doi.org/10.1038/nmeth.2510.
32. Kucinski, I., Wilson, N.K., Hannah, R., Kinston, S.J., Cauchy, P., Lenaerts, A., Grosschedl, R., and Göttgens, B. (2020). Interactions between lineage-associated transcription factors govern haematopoietic progenitor states. EMBO J. *39*, e104983. https://doi.org/10.15252/embj.2020104983.
33. Kent, D.G., Copley, M.R., Benz, C., Wöhrer, S., Dykstra, B.J., Ma, E., Cheyne, J., Zhao, Y., Bowie, M.B., Zhao, Y., et al. (2009). Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. Blood *113*, 6342–6350. https://doi.org/10.1182/blood-2008-12-192054.
34. Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D.B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S., et al. (2014). Identification of Regulatory Networks in HSCs and Their Immediate Progeny via Integrated Proteome, Transcriptome, and DNA

Methylome Analysis. Cell Stem Cell *15*, 507–522. https://doi.org/10.1016/j.stem.2014.07.005.

35. Cheng, P., Eksioglu, E.A., Chen, X., Kandell, W., Le Trinh, T., Cen, L., Qi, J., Sallman, D.A., Zhang, Y., Tu, N., et al. (2019). S100A9-induced overexpression of PD-1/PD-L1 contributes to ineffective hematopoiesis in myelodysplastic syndromes. Leukemia *33*, 2034–2046. https://doi.org/10.1038/s41375-019-0397-9.

36. Lu, M., Xia, L., Liu, Y.-C., Hochman, T., Bizzari, L., Aruch, D., Lew, J., Weinberg, R., Goldberg, J.D., and Hoffman, R. (2015). Lipocalin produced by myelofibrosis cells affects the fate of both hematopoietic and marrow microenvironmental cells. Blood *126*, 972–982. https://doi.org/10.1182/blood-2014-12-618595.

37. Caiado, F., Kovtonyuk, L.V., Gonullu, N.G., Fullin, J., Boettcher, S., and Manz, M.G. (2023). Aging drives Tet2+/− clonal hematopoiesis via IL-1 signaling. Blood *141*, 886–903. https://doi.org/10.1182/blood.2022016835.

38. Wang, Z., Liu, W., Wang, D., Yang, E., Li, Y., Li, Y., Sun, Y., Wang, M., Lv, Y., and Hu, X. (2022). TET2 Mutation May Be More Valuable in Predicting Thrombosis in ET Patients Compared to PV Patients: A Preliminary Report. J. Clin. Med. *11*, 6615. https://doi.org/10.3390/jcm11226615.

39. Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232–236. https://doi.org/10.1038/nature12495.

40. Che, J.L.C., Bode, D., Kucinski, I., Cull, A.H., Bain, F., Becker, H.J., Jassinskaja, M., Barile, M., Boyd, G., Belmonte, M., et al. (2022). Identification and characterization of in vitro expanded hematopoietic stem cells. EMBO Rep. *23*, e55502. https://doi.org/10.15252/embr.202255502.

41. Wilkinson, A.C., Ishida, R., Kikuchi, M., Sudo, K., Morita, M., Crisostomo, R.V., Yamamoto, R., Loh, K.M., Nakamura, Y., Watanabe, M., et al. (2019). Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. Nature *571*, 117–121. https://doi.org/10.1038/s41586-019-1244-x.

42. Labouesse, C., Tan, B.X., Agley, C.C., Hofer, M., Winkel, A.K., Stirparo, G.G., Stuart, H.T., Verstreken, C.M., Mulas, C., Mansfield, W., et al. (2021). StemBond hydrogels control the mechanical microenvironment for pluripotent stem cells. Nat. Commun. *12*, 6132. https://doi.org/10.1038/s41467-021-26236-5.

43. Gou, M., Li, J., Yi, L., Li, H., Ye, X., Wang, H., Liu, L., Sun, B., Zhang, S., Zhu, Z., et al. (2023). Reprogramming of ovarian aging epigenome by resveratrol. PNAS Nexus *2*, pgac310. https://doi.org/10.1093/pnasnexus/pgac310.

44. Liu, R., Jin, Y., Tang, W.H., Qin, L., Zhang, X., Tellides, G., Hwa, J., Yu, J., and Martin, K.A. (2013). Ten-Eleven Translocation-2 (TET2) Is a Master Regulator of Smooth Muscle Cell Plasticity. Circulation *128*, 2047–2057. https://doi.org/10.1161/CIRCULATIONAHA.113.002887.

45. Mali, R.S., Ramdas, B., Ma, P., Shi, J., Munugalavadla, V., Sims, E., Wei, L., Vemula, S., Nabinger, S.C., Goodwin, C.B., et al. (2011). Rho kinase regulates the survival and transformation of cells bearing oncogenic forms of KIT, FLT3, and BCR-ABL. Cancer Cell *20*, 357–369. https://doi.org/10.1016/j.ccr.2011.07.016.

46. Jaiswal, S., and Ebert, B.L. (2019). Clonal hematopoiesis in human aging and disease. Science *366*, eaan4673. https://doi.org/10.1126/science.aan4673.

47. Delhommeau, F., Dupont, S., Valle, V.D., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.P., Robert, F., Alberdi, A., et al. (2009). Mutation in TET2 in Myeloid Cancers. N. Engl. J. Med. Overseas. Ed. *360*, 2289–2301. https://doi.org/10.1056/NEJMoa0810069.

48. Skinder, N., Sanz Fernández, I., Dethmers-Ausema, A., Weersing, E., and de Haan, G. (2023). CD61 identifies a superior population of aged murine HSCs and is required to preserve quiescence and self-renewal. Blood Adv. *8*, 99–111. https://doi.org/10.1182/bloodadvances.2023011585.

49. Su, T.-Y., Hauenstein, J., Somuncular, E., Dumral, Ö., Leonard, E., Gustafsson, C., Tzortzis, E., Forlani, A., Johansson, A.-S., Qian, H., et al. (2024). Aging is associated with functional and molecular changes in distinct hematopoietic stem cell subsets. Nat. Commun. *15*, 7966. https://doi.org/10.1038/s41467-024-52318-1.

50. Widjaja, A.A., Lim, W.-W., Viswanathan, S., Chothani, S., Corden, B., Dasan, C.M., Goh, J.W.T., Lim, R., Singh, B.K., Tan, J., et al. (2024). Inhibition of IL-11 signalling extends mammalian healthspan and lifespan. Nature *632*, 157–165. https://doi.org/10.1038/s41586-024-07701-9.

51. Segel, M., Neumann, B., Hill, M.F.E., Weber, I.P., Viscomi, C., Zhao, C., Young, A., Agley, C.C., Thompson, A.J., Gonzalez, G.A., et al. (2019). Niche stiffness underlies the ageing of central nervous system progenitor cells. Nature *573*, 130–134. https://doi.org/10.1038/s41586-019-1484-9.

52. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. *50*, D543–D552. https://doi.org/10.1093/nar/gkab1038.

53. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

54. Granja, J.M., Corces, M.R., Pierce, S.E., Bagdatli, S.T., Choudhry, H., Chang, H.Y., and Greenleaf, W.J. (2021). ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat. Genet. *53*, 403–411. https://doi.org/10.1038/s41588-021-00790-6.

55. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978. https://doi.org/10.1038/nmeth.4401.

56. Wickham, H. (2016). ggplot2 (Springer International Publishing). https://doi.org/10.1007/978-3-319-24277-4.

57. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

58. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304. https://doi.org/10.1038/s41587-023-01767-y.

59. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. *13*, 2498–2504. https://doi.org/10.1101/gr.1239303.

60. Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–1093. https://doi.org/10.1093/bioinformatics/btp101.

61. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

62. Yu, G. (2020). In Methods in Molecular Biology, *2117*, B. Kidder, ed. (New York, NY: Humana), pp. 207–215. https://doi.org/10.1007/978-1-0716-0301-7_11.

63. Chen, H., and Boutros, P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinf. *12*, 35. https://doi.org/10.1186/1471-2105-12-35.

64. Csárdi, G., Nepusz, T., Müller, K., Horvát, S., Traag, V., Zanini, F., and Noom, D. (2024). igraph for R: R interface of the igraph library for graph

theory and network analysis. Version v2.1.1 (Zenodo). https://doi.org/10.5281/zenodo.13964143.

65. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47. https://doi.org/10.1093/nar/gkv007.

66. Yu, G., and He, Q.-Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol. Biosyst. *12*, 477–479. https://doi.org/10.1039/C5MB00663E.

67. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A.L., Fang, T., Doncheva, N.T., Pyysalo, S., et al. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. *51*, D638–D646. https://doi.org/10.1093/nar/gkac1000.

68. Wilkinson, A.C., Ishida, R., Nakauchi, H., and Yamazaki, S. (2020). Long-term ex vivo expansion of mouse hematopoietic stem cells. Nat. Protoc. *15*, 628–648. https://doi.org/10.1038/s41596-019-0263-2.

69. McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297. https://doi.org/10.1093/nar/gks042.

70. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. *11*, R25. https://doi.org/10.1186/gb-2010-11-3-r25.

71. Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914. https://doi.org/10.1126/science.aab1601.

72. McInnes, L., Healy, J., Saul, N., and Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. J. Open Source Softw. *3*, 861. https://doi.org/10.21105/joss.00861.

73. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

74. Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431–1443. https://doi.org/10.1016/j.cell.2014.08.009.

75. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

76. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

77. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

78. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. *2008*, P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
FITC anti-mouse CD45	BioLegend	Clone 30-F11, Cat#103107
Brilliant Violet 785™ anti-mouse CD45	BioLegend	Clone 30-F11, Cat#103149
PE/Cyanine7 anti-mouse CD150 (SLAM)	BioLegend	Clone TC15-12F12.2, Cat#115913
APC anti-mouse CD48	BioLegend	Clone HM48-1, Cat#103412
Brilliant Violet 421™ anti-mouse CD48	BioLegend	Clone HM48-1, Cat#103427
PE anti-mouse CD201 (EPCR)	eBioscience	Clone eBio1560, Cat#12-2012-82
APC/Cyanine7 anti-mouse CD117 (c-kit)	BioLegend	Clone 2B8, Cat#105826
Brilliant Violet 421™ anti-mouse Ly-6A/E (Sca-1)	BioLegend	Clone D7, Cat#108128
Brilliant Violet 510™ anti-mouse Ly-6A/E (Sca-1)	BioLegend	Clone D7, Cat#108129
Brilliant Violet 605™ anti-mouse Ly-6A/E (Sca-1)	BioLegend	Clone D7, Cat#108134
PE anti-mouse/rat CD61	BioLegend	Clone 2C9.G2, Cat#104307
Biotin anti-mouse CD201 (EPCR)	Stem Cell Technologies	Clone 1560, Cat#60038BT
Alexa Fluor® 647 Streptavidin	BioLegend	Cat#405237
Alexa Fluor® 488 anti-vWF	Abcam	Cat#AB307389
Chemicals, peptides, and recombinant proteins		
Animal-free recombinant mouse SCF	Peprotech	AF-250-03
Animal-free recombinant mouse TPO	Peprotech	AF-315-14
POLY(VINYL ALCOHOL), 87–90% HYDROLYZED	Sigma-Aldrich	P8136
Ham's F-12 Nutrient Mix	Gibco	11510586
Insulin-Transferrin-Selenium-Ethanolamine (ITS -X) (100X)	Gibco	10524233
Penicillin-Streptomycin-Glutamine (PSG) (100X)	Gibco	12090216
HEPES, 1M Buffer Solution	Gibco	11550496
Recombinant Mouse Von Willebrand Factor	antibodies.com	A317514
Hyaluronan	R&D Systems	GLR002
Fibronectin human plasma	Sigma Aldrich	F0895
Critical commercial assays		
TMT10plex™ Isobaric Label Reagents	Thermo Fisher Scientific	90110
PicoPure™ RNA Isolation Kit	Thermo Fisher Scientific	KIT0204
EasySep™ Mouse Hematopoietic Progenitor Cell Isolation Kit	Stem Cell Technologies	19856
BD Cytofix/Cytoperm™ Fixation/Permeabilization Kit	BD	554714
Chromium Next GEM Single Cell ATAC Library & Gel Bead Kit	10X Genomics	PN-1000176
ERCC RNA Spike-In Mix	Invitrogen	4456740
Deposited data		
Raw scRNA-seq data	This paper	Sequence Read Archive PRJNA1210137
Raw scATAC-seq data	This paper	Sequence Read Archive PRJNA1210127
Raw proteome data	This paper	ProteomeXchange Consortium PXD059814
Raw HoxB8-FL RNA-seq data	Kucinski et al.[32]	Gene Expression Omnibus GSE146128
Experimental models: Cell lines		
Human: HoxB8-FL cell line	Laboratory of Dr Hans Häcker	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
C57BL/6$^{W41/W41}$-Ly5.1 (W41) mouse	In-house breeding	N/A
C57BL/6 mouse	In-house breeding	N/A
Tet2$^{-/-}$ mouse (derived from B6(Cg)-*Tet2*$^{tm1.2Rao}$/J)	In-house breeding	N/A
Software and algorithms		
FlowJo	BD	https://www.flowjo.com/
edgeR package	Robinson et al.[53]	https://bioconductor.org/packages/release/bioc/html/edgeR.html
ArchR package	Granja et al.[54]	https://www.archrproject.com/
chromVAR package	Schep et al.[55]	https://bioconductor.org/packages/release/bioc/html/chromVAR.html
ggplot2 package	Wickham[56]	https://cran.r-project.org/web/packages/ggplot2/index.html
Seurat toolkit	Hao et al.,[57] Hao et al.[58]	https://satijalab.org/seurat/
Cytoscape	Shannon et al.[59]	https://cytoscape.org/
ClueGO	Bindea et al.[60]	https://apps.cytoscape.org/apps/cluego
clusterProfiler package	Wu et al.[61]	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
enrichPlot package	Yu[62]	http://bioconductor.org/packages/release/bioc/html/enrichplot.html
DoRothEA	Garcia-Alonso et al.[23]	https://saezlab.github.io/dorothea/
VennDiagram package	Chen et al.[63]	https://cran.r-project.org/web/packages/VennDiagram/index.html
igraph package	Csárdi et al.[64]	https://r.igraph.org/
Proteome Discoverer (version 2.2)	Thermo Fisher Scientific	https://www.thermofisher.com/se/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html
PCAtools package	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/PCAtools.html
limma package	Ritchie et al.[65]	https://bioconductor.org/packages/release/bioc/html/limma.html
ReactomePA package	Yu and He[66]	https://bioconductor.org/packages/release/bioc/html/ReactomePA.html
STRING	Szklarczyk et al.[67]	https://string-db.org/
Prism software	GraphPad	https://www.graphpad.com/
Other		
96-well CELLview^{TM} plates	Greiner	N/A

EXPERIMENTAL MODELS AND STUDY PARTICIPANTS DETAILS

Mice

Wild-type C57BL/6N and *Tet2*$^{-/-}$ mice bred in-house were used for all experiments. Adult animals aged 12–16 weeks were used for all experiments except for MS analysis where used animals were 30 weeks old. A mix of male and female animals was used. Animals were housed in individually ventilated cages (IVC) and provided with sterile food and water *ad libitum*. All mice were kept in specified pathogen-free conditions, and all procedures performed according to the United Kingdom Home Office regulations, in accordance with the Animal Scientific Procedure Act.

Cell lines

Hoxb8-FL cells were cultured in RPMI 1640 media (Sigma), supplemented with 10% FBS (Gibco), 0.1% mercaptoethanol (Invitrogen), 1% penicillin-streptomycin (Sigma), 1% glutamine (Sigma), 1 μM estradiol and 5% FLT3L conditioned media from the B16-FL cell line. Cells were maintained in culture at concentrations of 10^5-10^6 cells/ml.

METHOD DETAILS

Flow cytometry and FACS

For sorting of primary HSPCs, bone marrow was extracted from hind limbs, hips, sternum, and spine collected in ice-cold phosphate buffered saline (PBS). Bones were crushed using a mortar and pestle and cell suspension was mechanically dissociated using a pipette and passed through a 40 μM filter. Red blood cells were lysed by incubation with ammonium chloride (Stem Cell Technologies). Mature cells were magnetically depleted from the cell suspension using the EasySep Mouse Hematopoietic Stem and Progenitor Cell Isolation Kit (Stem Cell Technologies). For sorting for MS analysis, cell suspensions were stained with fluorophore-conjugated antibodies against CD45, CD150, CD48, EPCR, and cKit by incubation for 30 min on ice protected from light. For all other experiments, anti-cKit antibody was omitted and anti-Sca-1 antibody was included. Where indicated, antibodies against CD61 and CD41 were included in the panel. For flow cytometric analysis of cultured HSCs, cell suspensions were stained with fluorophore-conjugated antibodies against CD45, CD11b, Gr-1, cKit, Sca-1, and EPCR. In all flow cytometry and FACS experiments, 7-aminoactinomycin D (7-AAD) staining was used to exclude dead cells. For intracellular flow experiments, bone marrow cells were stained with Fc block for 15 min on ice prior to cell surface staining with antibodies described above. Following cell surface staining, cells were fixed and permeabilized using BD Cytofix/Cytoperm kit in accordance with manufacturer's protocol. Staining with intracellular antibodies was carried out overnight at 4°C. In cases where a biotinylated primary antibody was used, secondary staining with Streptavidin-Alexa Fluor 647 was performed. FACS experiments were performed on a BD Influx at the Cambridge Institute for Medical Research, or a Beckamn Coulter MoFlo Astrios or BD FACS Discoverer S8 at the Imaging & Cytometry Technology Facility at the University of York. All flow cytometric analyses were performed on a Beckman Coulter CytoFlex LX or BD LSRFortessa X20 at the Imaging & Cytometry Technology Facility at the University of York. All flow cytometry data were analyzed using FlowJo software (BD).

Immunofluorescence

SLAM cells were isolated by FACS as described above directly into fibronectin-coated (10 μg/cm^2) 8-well ibidi chamber slides containing HSC expansion media without cytokines. Cells were incubated overnight at 37°C and 5% CO_2 to allow adherence to the slide. Following removal of media, cells were fixed with 2% PFA for 10 min at RT. Cells were washed and blocked for 1 h at RT in PBS containing 5% FBS and 0.01% tween. Blocking buffer was removed and cells were stained with fluorophore conjugated antibodies against ECM proteins for 1 h at RT. Following staining for ECM proteins, cells were stained with 0.5 μg/mL DAPI for 15 min at RT. Cells were imaged on a Zeiss LSM 880 confocal microscope at the Imaging & Cytometry Technology Facility at the University of York.

HSC gel culture

96-well CELLview plates (Greiner) were activated to allow the binding of StemBond. Plates were treated inside a plasma system (Henniker HPT-200) and functionalized using 5% Bind Silane solution (GE Healthcare). Plates were washed thoroughly with 100% ethanol. 3 mL soft hydrogel solutions were prepared using 40% acrylamide (210 μL), 2% Bis-acrylamide (120 μL), TEMED (15 μL), 10% Ammonium Persulfate (APS, 30uL), and water (2461.8 μL) and transferred to the CELLview plates. Following polymerisation, gels were rinsed twice in methanol, followed by a PBS rinse. Prior to activation with EDAC/NHS solution (Sigma Aldrich), gels were rinsed with pH 6.1 MES buffer. Once activated, gels were rinsed with chilled 60% methanol in PBS, followed by a 50 mM pH 8.5 HEPES buffer rinse. Gels and plastic control wells were coated 100–200 μg/mL of ECM protein diluted in HEPES buffer and incubated overnight at 4°C. Following incubation, the protein solution was removed and gels were rinsed with HEPES buffer. Ethanolamine solution (0.5 M; ChemCruz) in HEPES buffer was used to block the gels for 30 min at room temperature. Gels were rinsed for a final time with pH 7.4 HEPES buffer and PBS to equilibrate the pH. Gels were stored at 4°C until use. 50 ESLAM HSCs per well were sorted directly onto gels into HSC expansion media[41,68] and maintained in culture at 37 C and 5% CO2 for 28 days with media changes every 2–3 days.

Bulk RNA-seq

To match the cell populations extracted for proteomic profiling, Lin^- $CD45^+$ $CD48^-$ $CD150^+$ $cKit^+$ (collectively called $CD150^+$) and Lin^- $CD45^+$ $CD48^-$ $CD150^-$ $cKit^+$ (collectively called $CD150^-$) were isolated using FACS (as described above). A total of 1000 cells were collected per biological sample from homogenized cell extracts (femurs, tibiae, hips and spines) of *Tet2*$^{-/-}$ and WT mice. RNA extraction was performed using the Picopure RNA Isolation Kit (Thermo Scientific) according to manufacturer's protocol. Library preparation and sequencing was performed at the Cancer Research UK Cambridge Institute Genomics Core as previously described.[40] Data processing was conducted as previously described.[40] In brief, adapter trimming was performed using trim_galore (parameters: –paired –quality 30 –clip_R2 3). Reads alignment against the Mus musculus genome build (mm10) was conducted using STAR (default parameters). Gene counts were computed using HTSeq (parameters: –format = bam –stranded = reverse –type = exon –mode = intersection-nonempty

–additional-attr = gene_name). Downstream processing and quality control was conducted using EdgeR[53,69] (version 3.28.1), with read counts being transformed to counts per million (cpm), genes with fewer than 2 samples expressing >1 cpm being excluded and read count normalization being performed using the trimmed mean of M values (TMM) method.[70]

scATAC-seq

scATAC-seq data were generated from *Tet2*$^{-/-}$ mice and WT littermate controls. ESLAM HSCs were isolated as described above. Cells were isolated from 5 WT mice and 5 *Tet2*$^{-/-}$ mice (4,000 cells per genotype). Libraries were prepared using the 10x Genomics Chromium Next GEM Single Cell ATAC Reagent Kits v1.1. Sequencing was run at Leeds University Next Generation Sequencing facility using a NextSeq 2000.

Plate-based scRNA-seq

scRNA-seq data were generated from *Tet2*$^{-/-}$ mice and WT littermate controls. ESLAM HSCs were FACS-purified as previously described. Freshly isolated HSCs were subjected to single-cell RNA SmartSeq2 sequencing (4 x 96-well plates). RNA was extracted using the Picopure RNA isolation kit (Thermo Fisher). Libraries were prepared using a protocol adapted from the msSCRB-seq workflow and quality control was performed using the Bioanalyzer system (Agilent). ERCC external RNA Spike-In controls were used (ERCC RNA Spike-In Mix; ThermoFisher). Constructed libraries were sequenced using the Illumina NovaSeq X and Novogene systems using a paired end 150 bp run.

Sample preparation for proteome analysis

Prior to proteomic analysis, Hoxb8-FL cells were resuspended in phosphate buffered saline (PBS). For experiments 'Direct 10K' and 'Direct 15K', cell lysis was performed using 2% sodium dodecyl sulfate (SDS) with subsequent boiling at 95°C. Cell lysates were sonicated and dried using vacuum centrifugation. Samples were re-suspended in 100 mM TEAB. Reduction and alkylation of cysteine residues was performed by incubation with a final concentration of 5 mM tris-2-carboxyethyl phosphine (TCEP) at 60°C for 30 min followed by final concentration 10 mM iodoacetamide (IAA) for 30 min at RT protected from light. Protein-level isobaric labeling was performed using TMT 10plex reagents (Thermo Scientific) in accordance with manufacturer's protocol. 100% (w/v) trichloroacetic acid (TCA) was added to the sample mixture at a ratio of 1–4, followed by incubation for 10 min. The sample was centrifuged at 14,000 rpm and the resulting protein pellet was resuspended in 100 mM TEAB buffer. Trypsin was added and proteins were digested overnight at 37°C. For experiments '10K Fract I' and '10K Fract II', cells in a volume of 20 μL of PBS were thawed on ice and 2 μL 1 M TEAB, 1 μL 2% SDS, and 1 μL Halt Protease & Phosphatase inhibitor cocktail (pre-diluted 1:5 in water) was added. Cells were lysed by bath sonication for 5 min followed by 3 min incubation at 90°C. Reduction and alkylation of cysteine residues were performed by incubation with 2 μL 50 mM TCEP at 40°C for 30 min followed by 1 μL 200 mM IAA for 30 min at RT protected from light. 0.5 μg trypsin was added, and proteins were digested overnight at RT. Peptide-level Isobaric labeling was performed using TMT 10plex reagents (Thermo Scientific) in accordance with manufacturer's protocol. Following quenching of the reaction with 5% hydroxylamine, samples were combined and dried completely by vacuum centrifugation. High pH Reversed-Phase (RP) fractionation was performed with the Waters XBridge C18 column (2.1 × 150 mm, 3.5 μm, 120 Å) on a Dionex UltiMate 3000 HPLC system. Ammonium hydroxide at 0.1% v/v was used as mobile phase A and mobile phase B was set as 100% acetonitrile/0.1% v/v ammonium hydroxide. The peptide mixture was reconstituted in 100 μL mobile phase A and subjected to gradient elution at 200 μL/min as follows: 5 min isocratic at 5% B, for 15 min gradient to 35% B, for 5 min gradient to 80% B, isocratic for 5 min and re-equilibration to 5% (B). The chromatogram was recorded at 215 and 280 nm and fractions were collected every minute. Fractions were dried completely by vacuum centrifugation and stored at −20°C until further use.

For primary mouse samples, 10,000–30,000 cells from *Tet2*$^{-/-}$ and WT mice were FACS-sorted into 0.1 mL PCR tubes containing 20 μL ice-cold PBS and processed as described above for the '10K Fract I' and '10K Fract II' experiments. 6, 5 and 8 fractions were finally subjected to LC-MS analysis for the '10K Fract I', '10K Fract II' and primary mouse samples, respectively.

LC-MS/MS analysis

LC-MS/MS analysis was performed on a Dionex UltiMate 3000 UHPLC system coupled with an Orbitrap Lumos Mass Spectrometer (Thermo Scientific). Each peptide fraction was reconstituted in 10 μL 0.1% formic acid and 7 μL were loaded on the Acclaim PepMap 100, 100 μm × 2 cm C18, 5 μm, trapping column with the μlPickUp method at a flow rate of 10 μL/min. The samples were subjected to a multi-step gradient elution on an EASY-Spray (75 μm × 50 cm, 2 μm) C18 capillary column (Thermo Scientific) at 45°C. Mobile phase A was 0.1% formic acid and mobile phase B was 80% acetonitrile/0.1% formic acid. The gradient separation method at flow rate 300 nL/min was as follows: for 90 min gradient 5%–38% B, for 10 min up to 95% B, for 5 min isocratic at 95% B, re-equilibration to 5% B in 5 min, for 10 min isocratic at 5% B. Precursor ions were selected with mass resolution of 120k, AGC 4×10^5 and max IT 50 ms in the top speed mode within 3 s. Peptides were isolated for HCD fragmentation with quadrupole isolation width 0.7 Th and 50k resolution. Collision energy was set at 38% with AGC 1×10^5 and max IT 105 ms. Targeted precursors were dynamically excluded from further isolation and activation for 45 s with 7 ppm mass tolerance. For the 'Direct 10K' and 'Direct 15K' runs a 150 min 5%–38% B gradient was used. For the '10K Fract II' experiment, the 5 fractions were injected twice by setting a maximum intensity threshold at 5×10^6 in the second run (from 5×10^{20}).

QUANTIFICATION AND STATISTICAL ANALYSIS

scATAC-seq

Read alignment to a reference genome (mm10) was performed using CellRanger pipeline (CellRanger-ATAC, 10x Genomics, cellranger-atac count). For downstream data analyzes, the ArchR workflow was used.[54] Due to the low intra-sample heterogeneity, the ArchR's simulation of synthesized *in silico* doublets over the data was not used to exclude potential doublet cells and doublets were removed by filtering cells containing less than 40,000 fragments instead. The term frequency-inverse document frequency (TF-IDF) normalization and the singular value decomposition (SVD) were performed (latent semantic indexing, LSI[71]) using ArchR's addIterativeLSI. Uniform Manifold Approximation and Projection (UMAP) dimension reduction[72] was run with ArchR's addUMAP. Pseudo-bulk replicates were created, and peaks were called using MACS2.[73] Marker peaks unique to individual groups were identified with ArchR's getMarkerFeatures. Paired samples Wilcoxon test was used to compare WT and *Tet2*$^{-/-}$ samples (FDR 0.1 & absolute log2 FC > 0.5)). Transcription factor binding motifs were annotated using ArchR's addMotifAnnotations function, and motif set from the cisbp database[74] was used (chromVAR package[55]). Differentially accessible peaks were tested for motif enrichment with ArchR's peakAnnoEnrichment function. Closest genes to the accessible regions were identified if the distance to the transcription start site was <100k base pairs (FDR $\leq$ 0.1 and log2 FC $\geq$ 0.5). Computational analysis was performed using the University of York Research High Performance Computing Cluster (Rocky 8.8, Viking2). Plots were made with ArchR,[54] ggplot2,[56] Seurat,[57,58] Cytoscape[59] and ClueGO.[60]

Plate-based scRNA-seq

Sequenced reads were aligned to the GRCm39 (Genecode version M33) reference mouse genome using STAR aligner[75] and gene counts were computed using featureCounts.[76] For downstream scRNA-seq data analysis, the Seurat workflow[57,58] was used. Sequencing data from the scRNAseq experiments were integrated using Harmony.[77] Data were normalized using regularized negative binomial regression (Seurat's SCTransform). Principal component analysis (PCA) reduction analysis was performed with Seurat's RunPCA (default parameters), and the top 5 principal components were selected. UMAP[72] was run with Seurat's RunUMAP (default parameters). Local neighbourhoods were defined with Seurat's FindNeighbors, and cells were clustered using the Louvain algorithm[78] (Seurat's FindClusters). Differentially expressed genes for *Tet2*$^{-/-}$ and WT cell groups were found with Seurat's FindMarkers. Significantly up/down regulated genes were defined as q value <0.05 and absolute average log2 FC > 1. To compare functional profiles for identified genes, clusterProfiler[61] and enrichplot[62] were used. Gene Ontology analysis was performed with Cytoscape[59] and ClueGO.[60] Computational analysis was performed using the University of York Research High Performance Computing Cluster (Rocky 8.8 and Viking2). Plots were made with ArchR,[54] ggplot2,[56] Seurat,[57,58] Cytoscape[59] and ClueGO.[60]

Integrative scATAC-seq and scRNA-seq data analysis

The lists of more/less accessible regions in the scATAC-seq closest gene analysis determined by the *Tet2*$^{-/-}$ HSC versus WT HSC pairwise testing were intersected with the lists of genes showing higher/lower expression in the *Tet2*$^{-/-}$ HSC versus WT HSC scRNA-seq analysis. To find genes related to identified scATAC-seq/scRNA-seq targets, the Genemania database[25] was used. Network of TFs (scATAC-seq analysis) and genes they regulate (scRNA-seq analysis) was built using information from the DoRothEA database.[23] Plots were made with with VennDiagram,[63] ggplot2,[56] Genemania,[25] DoRothEA[23] and igraph.[64]

Protein identification and quantification

MS raw data was searched against the SwissProt human or mouse database using the SequestHT node in Proteome Discoverer 2.2. Precursor mass tolerance was 20 ppm and fragment ion mass tolerance was 0.02 Da. Spectra were searched for fully tryptic peptides with no more than 2 missed cleavages and a minimum length of 6 amino acids. TMT6plex at N-termini and lysine residues and carbamidomethyl at cysteine residues were set as fixed modifications. Methionine oxidation and glutamine and asparagine deamidation were set as dynamic modifications. Peptide FDR was set to 0.01 and validation was based on q-value and target-decoy database search using the Percolator node. The Reporter Ion Quantifier node included a custom TMT-10plex quantification method with an integration window tolerance of 15 ppm. At least one unique peptide was required for identification and only unique peptides were used for quantification.

Bioinformatic analysis of proteomic and bulk RNA-seq data

Scaled quantitative values were obtained by dividing each TMT signal-to-noise (S/N) ratio by the mean TMT S/N across samples per protein. For bulk RNA-seq data, the gene list was filtered for genes with a minimum of 2 libraries with a minimum count per million (CPM) of 1. Remaining CPM values were normalized using the trimmed mean of M values (TMM) method in the edgeR R package[53] (version 3.40.2). Principal component analysis (PCA) was performed using the R package PCAtools (version 2.10.0). The bottom 10% least variable genes/proteins were not included in PCA. Correlation between proteome and bulk transcriptome data was assessed using the Pearson correlation coefficient. For shortlisting targets for follow-up analysis, proteins with an absolute log2 FC of >0.5 across all comparisons between the two *Tet2*$^{-/-}$ and WT replicates were considered potential targets. For the bulk RNAseq dataset, a Students' *t* test was performed and genes with an absolute log2 FC > 0.5 and adjusted *p*-value <0.05 were considered potential targets. KEGG and Reactome pathway analysis were performed using the limma[65] (version 3.54.2) and ReactomePA[66] (version

1.42.0) package, respectively. Interaction network analysis was performed using the STRING database[67] with a combined score cut-off of 0.4. Networks were visualized in Cytoscape[59] (version 3.10.0).

Statistical analysis

For all other experiments, differences between groups were assessed by one or two-tailed Students' *t* test (two groups) or one-way ANOVA with Tukey's post hoc test (three or more groups) using Prism software (GraphPad). Details about number of replicates used for experiments can be found in the respective figure legends. Error bars represent SD. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$ and ns = non-significant.