THE UNIVERSITY OF CHICAGO

ESTABLISHING CUX1 AS A DETERMINANT OF HEMATOPOIETIC STEM CELL HETEROGENEITY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

INTERDISCIPLINARY SCIENTIST TRAINING PROGRAM:

IMMUNOLOGY

BY

TANNER CLARK MARTINEZ

CHICAGO, ILLINOIS

DECEMBER 2024

Copyright 2024 Tanner C. Martinez
Some material originally appeared in the Annual Review.
Reproduced with permission from the Annual Review of *Pathology: Mechanisms of Disease*,
Volume 19 © 2024 by Annual Reviews

Table of Contents

List of Figures ..... *vi*

List of Tables ..... *vii*

Acknowledgments ..... *viii*

Abstract ..... *x*

Introduction *1*

Abstract **1**

Introduction **1**

The paradigm of haploinsufficiency **2**

Hematopoietic stem cell heterogeneity **6**

Dose in Hematopoiesis **10**

Mechanisms of Transcription Factor Haploinsufficiency **14**

Transcription Factor Dynamics 14

Clustering the Characters 17

Models of Transcription Factor Action at Target Genes 19

Compensation 20

How to lose a partner in seven ways **21**

Aneuploidy 21

Focal Deletions 27

Balanced Translocations 28

Mutations 30

Mutation of Another Gene 33

Change in Isoform Balance 35

Epigenetic Silencing 37

Therapeutic Options **37**

Conclusion **38**

Summary Points, Future Issues, Definitions **39**

Summary Points 39

Future Issues 40

Definitions 40

Results: CUX1 regulates hematopoietic stem cell fate through regulation of innate immune activation *43*

Abstract **43**

Introduction **43**

Results **45**

CUX1 protein levels predict HSC activity 45

CUX1 levels correlate with the gradual loss of the HSC gene expression program 55

CUX1 knockdown restores and extends HSC activity 61
CUX1 knockdown alters differentiation trajectories 69
Loss of CUX1 results in Endogenous Retroelement de-repression 79
Innate immune activation is required for CUX1-mediated self-renewal 83

Discussion 87
Engraftment and quiescence 90
HSCs count divisions 92
Intrinsic and extrinsic viral response 93
Dose-dependent control of HSCs 94
Somatic stem cells 95
Limitations of the study 95

Methods 97

Experimental Models and Subject Details 97
Generation of CUX1 Reporter Mice 97
Mouse models 97
Experimental Diets 98
Cell isolation 98
Primary cell culture 99

Method Details 99
Flow Cytometry and Cell Sorting 99
Complete Blood Count 103
Cell cycle analysis 103
Serial colony-forming-unit assays (CFU) 103
In vitro differentiation assay 103
Intracellular flow cytometry 104
Ex vivo OPP-incorporation assay 104
HSC competitive transplantation 105
Homing Assay 105
Western Blot 105
Cycloheximide treatment 106
Immunofluorescence 106
qPCR 106

Chemicals and Reagents 107
Bulk RNA-Seq 107
ATAC-Seq 108
scRNA-Seq 110
Quantification and Statistical Analysis 113

Discussion 114

Endogenous Retroelements and Chromatin 114
Expression 114
Repression 119
Revisiting the Quality Control Hypothesis 123
Limitations 125

Self-Nonself Discrimination **126**
Interferon and homeostasis **127**
Trajectories of LT-HSC differentiation **129**
Asymmetric division **133**
A paradigm for studying CUX1 ..... **136**
Conclusion **138**
Remaining Questions/Future Directions **139**
References ***140***

List of Figures

FIGURE 1. PRINCIPLES OF HITFS 5
FIGURE 2. STEM CELL PROPERTIES ARE CONTINUOUS CHARACTERISTICS. 13
FIGURE 3. LOSS OF A HITF RESULTS IN A MYELOID DIFFERENTIATION BLOCK AND GREATER TRANSCRIPTIONAL PLASTICITY. 15
FIGURE 4. MULTIPLE WAYS TO LOSE A HITF. 22
FIGURE 5. *CUX*1^{mCherry} MICE EXHIBIT NO GROSS OR HEMATOPOIETIC ABNORMALITIES 46
FIGURE 6. CUX1 LEVELS INCREASE DURING HEMATOPOIETIC DIFFERENTIATION 48
FIGURE 7. CUX1 LEVELS EXHIBIT HIGH VARIANCE IN LT-HSCS 49
FIGURE 8. CUX1 LEVELS PREDICT HSC ACTIVITY 50
FIGURE 9. CUX1^{DIM} LT-HSCS EXHIBIT INITIAL PLATELET BIAS 53
FIGURE 10. CUX1^{INT} LT-HSCS REPRESENT A DISTINCT CELL STATE 54
FIGURE 11. HIGH INTERFERON SIGNALING IN CUX1^{DIM} HSCS 56
FIGURE 12. CUX1 LEVELS CORRELATE WITH THE GRADUAL LOSS OF THE HSC GENE EXPRESSION PROGRAM 58
FIGURE 13. KNOCKDOWN OF CUX1 RESTORES AND EXTENDS HSC ACTIVITY 62
FIGURE 14. HSC HETEROGENEITY OBSCURES EFFECTS OF CUX1 KNOCKDOWN 65
FIGURE 15. KNOCKDOWN OF CUX1 POTENTIATES INTERFERON SIGNALING 66
FIGURE 16. CUX1 KNOCKDOWN RESTORES CUX1^{DIM} TRANSCRIPTOME IN CUX1^{BRIGHT} HSCS 67
FIGURE 17. CUX1 KNOCKDOWN DOES NOT INDUCE AN ALTERNATIVE CELL STATE 70
FIGURE 18. CUX1 KNOCKDOWN PROMOTES RETENTION IN LESS DIFFERENTIATED CELL STATES 72
FIGURE 19. CUX1 KNOCKDOWN ALTERS DIFFERENTIATION TRAJECTORIES 73
FIGURE 20. RNA VELOCITY ANALYSIS REVEALS KNOCKDOWN OF CUX1 REVERSES EARLIEST HSC DIFFERENTIATION TRAJECTORY 74
FIGURE 21. KNOCKDOWN OF CUX1 INDUCES SIMILAR CHROMATIN ACCESSIBILITY LANDSCAPES IN CUX1^{DIM} AND CUX1^{BRIGHT} HSCS 76
FIGURE 22. KNOCKDOWN OF CUX1 IN CUX1^{BRIGHT} HSCS REINSTATES THE CHROMATIN ACCESSIBILITY LANDSCAPE OF CUX1^{DIM} HSCS 77
FIGURE 23. LOW LEVELS OF CUX1 PERMIT ENDOGENOUS RETROELEMENT EXPRESSION 80
FIGURE 24. CUX1 MOTIF ENRICHED IN OPEN LTR-CONTAINING ENDOGENOUS RETROELEMENTS . 84
FIGURE 25. INNATE IMMUNE RESPONSE RESULTING FROM ENDOGENOUS RETROELEMENT REVERSE TRANSCRIPTION IS REQUIRED FOR CUX1-MEDIATED HSC SELF-RENEWAL 85
FIGURE 26. IRF1 IS REQUIRED FOR HSC SELF-RENEWAL 88
FIGURE 27. CUX1 CONTROLS HSC SELF-RENEWAL THROUGH THE REGULATION OF INNATE IMMUNE ACTIVATION 91
FIGURE 28. BONE MARROW, SPLEEN, AND PERIPHERAL BLOOD GATING SCHEMES 101

List of Tables

TABLE 1. FREQUENCY OF HITF GENETIC ABERRATIONS AND GENETIC ABERRATIONS AFFECTING THE DOSE OF HITFS ACROSS THE SPECTRUM OF MYELOID DISEASES. 12
TABLE 2. GENOTYPING PRIMERS FOR *CUX1* EXON 24 97
TABLE 3. PRIMERS FROM DE ALMEIDA ET AL. 2017[313] 107

Acknowledgments

The McNerney Lab has provided a supportive environment for completing my thesis. I am thankful for all the current and former members of the lab. Everyone contributed to a positive atmosphere for science and helped shape me into the scientist that I am today.

Dr. McNerney has been an encouraging and supportive mentor. Her prompt and thorough feedback on my writing and presentations has significantly enhanced my scientific communication skills. I particularly appreciate her willingness to provide constructive critiques, which have helped me grow as a writer. Dr. McNerney's insightful approach to science initially captivated me, and she has consistently encouraged me to broaden my ideas and consider the larger implications of my work. Most importantly, she has given me the independence to explore my scientific interests while offering necessary support along the way.

Dr. Angela Stoddart has also been an invaluable mentor. She has guided this project down exciting avenues and helped me focus on the most critical aspects of my work. Her vast scientific expertise has been an invaluable resource for both me and the rest of the lab.

I am grateful to Dr. Ningfei An for her patience in teaching me everything about working with mice as a model organism. Saira Khan’s positive attitude and expertise have significantly contributed to the lab, particularly through her preparation of the ATAC-seq libraries.

Lastly, I am incredibly thankful for Dr. Matthew Jotte's extensive contributions to this project. He has done a tremendous amount of work and has also been a great sounding board for ideas, providing valuable insights throughout the research process. He has been great to work with and has made this process much more fun.

My thesis committee of Drs. Marcus Clark, Eric Pietras, and Samantha Riesenfeld have given valuable feedback on this project over the past four years. I appreciate the interdisciplinary nature of our meetings and the helpful assessments of my work. I am indeed a much better

scientist because of it. Dr. Riesenfeld and Hope Anderson have been particularly helpful for the RNA velocity analysis and presentation of scRNA-seq data.

This work is dependent on the help of the core facilities of The University of Chicago. In particular, the Animal Resources Center, the Flow Cytometry Core Facilty, the Functional Genomics Core, and the Transgenic Animal Core have all been essential for this work.

My parents have always been supportive of me and my ambitions, and I thank them for their encouragement during the last few years. Finally, I met Grant Hayes while rotating in the McNerney Lab. His support throughout this process has been remarkable, making the journey more enjoyable and manageable. Without his help, this process would have been much more challenging.

Abstract

CUX1 is a dose-dependent transcription factor frequently mutated or deleted in high-risk myeloid neoplasms. We sought to understand the function of CUX1 in hematopoiesis. Here, utilizing a novel CUX1 reporter mouse, we established that CUX1 levels are dynamically regulated throughout the differentiation of hematopoietic progenitors. Initially, we established a strong correlation between CUX1 levels and the functional capacity of hematopoietic stem cells (HSCs). As CUX1 levels increase, HSCs lose repopulating capacity. The progressive loss of functional capacity can be reversed by the knockdown of CUX1 in HSCs. Accompanying the reversal of CUX1 levels includes the re-establishment of the transcriptome and chromatin accessibility landscape of HSCs with naturally low levels of CUX1. These findings establish CUX1 as a determinant of HSC fate and heterogeneity. Low levels of CUX1 also permit the high expression of interferon-stimulated genes (ISGs), which is a feature of somatic stem cells, including HSCs with low levels of CUX1. We establish that this interferon response and innate immune system activation results from the de-repression, expression, and reverse transcription of endogenous retroelements. The current paradigm posits that high ISG expression is imperative for protection from viral infection. We extend this paradigm to establish that high ISG expression and ERE reverse transcription are also essential for HSC self-renewal. This mechanism of ISG-driven self-renewal could be exploited by myeloid malignancies that have lost CUX1. Our findings advance our understanding of HSC fate and self-renewal, demonstrating that CUX1 finely controls the earliest steps of hematopoiesis by regulating innate immune system activation.

Introduction

Abstract

Many transcription factors (TFs) function as tumor suppressor genes with heterozygous phenotypes, yet haploinsufficiency generally has an underappreciated role in neoplasia. This is no less true in myeloid cells, which are normally regulated by a delicately balanced and interconnected transcriptional network. A detailed understanding of TF dose in this circuitry sheds light on the leukemic transcriptome. In this review, we discuss the emerging features of haploinsufficient transcription factors (HITFs). We posit that: (*a*) monoallelic and biallelic losses can have distinct cellular outcomes; (*b*) the activity of a TF exists in a greater range than the traditional Mendelian genetic doses; and (*c*) how a TF is deleted or mutated impacts the cellular phenotype. The net effect of a HITF is a myeloid differentiation block and increased intercellular heterogeneity in the course of myeloid neoplasia.

Introduction

By the classical genetic definition, malignancy was initially thought to be a dominant phenotype. Malignant cells fused with nonmalignant counterparts produced malignant hybrids[1]. Subsequent work demonstrated that in some combinations, the reverse is true, i.e., fusions result in nonmalignant hybrids, creating the paradigm of a recessive mechanism of tumor suppression[2]. This paradigm was formalized with the discovery of the first tumor suppressor gene, *RB1*. As described in Knudson's two-hit hypothesis, both copies of a given tumor suppressor gene must be inactivated to cause cancer[3]. Nearly 30 years later, this hypothesis was challenged by the finding that loss of only one copy of $P27^{Kip}$ (*CDKN1B*) is sufficient to promote the development

of tumors[4]. $P27^{Kip}$ thus founded the growing class of genes that act as haploinsufficient tumor suppressor genes and implicated the dose of tumor suppressors in oncogenesis. As discussed below, extensive tumor sequencing suggests that Knudson's hypothesis may in fact apply only to a minority of tumor suppressor genes.

The paradigm of haploinsufficiency

The classical definition of haploinsufficiency is the production of a dominant phenotype in a diploid organism after one allele of a gene undergoes a loss-of-function event[5,6]. The context of the phenotype measured is critical and nuanced, as demonstrated in yeast. In genome-wide proliferation screens, the fraction of haploinsufficient yeast genes increases from 3% in nutrient-replete media to 20% in nutrient-deplete media. When morphological traits are also assessed, 33–59% of genes produce haploinsufficient phenotypes[7-9]. The extent of haploinsufficiency is harder to quantify in mammals; however, efforts to systematically knock out and phenotype all mouse genes reveal that a striking 42% of genes have significant phenotypes in heterozygous mice[10].

Surveys of healthy human genomes have revealed which human genes are among the most evolutionarily constrained, defined as genes that are the most sensitive to deleterious mutations. Almost half of coding genes have fewer loss-of-function variants than would be expected by chance, indicating a high degree of genetic constraint genome wide. Among constrained genes, haploinsufficient genes are the most conserved, as a deleterious variant in one allele is selected against[5,11-14]. Transcription factors (TFs) appear more likely to produce haploinsufficient phenotypes when compared with other genes, and homeodomain-containing TFs are overrepresented among haploinsufficient transcription factors (HITFs)[15]. Indicating the

importance of finely tuned protein dose, genes predicted to produce haploinsufficient phenotypes are also likely to be triplosensitive, in that gain of an additional copy of the gene produces a dominant phenotype[12]. However, measures characterizing the extent of haploinsufficiency genome wide in healthy individuals do not capture genes with deleterious alleles that do not come into play until after reproduction (e.g., *BRCA1*). Overall, these studies demonstrate the profound dose-sensitive nature of the human genome.

Just as unbiased sequencing studies have demonstrated greater haploinsufficiency across the genome than previously appreciated, similar sequencing studies of tumors highlight the inadequacy of the two-hit model of tumor suppressor genes. To measure haploinsufficient genes in cancer, Elledge and colleagues[16] performed a sophisticated pan-cancer analysis for patterns of loss-of-function mutations and gene deletions. They estimate that at least 30% of all genes exhibit haploinsufficiency[16,17]. Notably, almost all sporadic tumor suppressor genes are predicted to produce haploinsufficient phenotypes. The correlation between haploinsufficiency and triplosensitivity is also observed in tumor suppressor genes. Overall, haploinsufficiency in cancer is much more prevalent than previously appreciated. Furthermore, there is genetic pressure, even in cancer, to maintain dose-sensitive genes within a minimum and maximum range.

To highlight the cardinal features of haploinsufficient genes in cancer, we turn to PTEN. While not a TF, PTEN ranks among the most well-studied proteins in this class. In prostate cancer, *Pten* heterozygous loss promotes proliferation; *Pten* homozygous loss leads to greater proliferation, but only after *Trp53* inactivation to avoid cellular senescence[18,19]. Similarly, in hematopoietic malignancies, inactivation of both copies of *PTEN* results in hematopoietic stem cell depletion, and additional gene mutations are necessary to avoid this loss[20]. Even a small 20% reduction in PTEN protein levels can promote tumor development[21]. Likewise, subtle, subdiploid

reduction in the expression of the TF PU.1 (*SPI1*) is sufficient to promote a preleukemic state[22,23]. Furthermore, in prostate cancer, mutation of *PTEN* is less common than copy number alteration of the locus containing *PTEN*, 10q23. This latter phenomenon may be related to the fact that in addition to PTEN, there are five other tumor-suppressive genes at this locus that contribute to the phenotype in prostate cancer[24].

The above examples allow us to underscore some principles of HITFs, which we discuss further (Figure 1):

- Principle 1: One is enough, but sometimes two is better—HITFs may exhibit haploinsufficiency in that loss of one allele produces a phenotype, but in some instances loss of the second allele produces a greater or distinct effect.
- Corollary to principle 1: Myeloid neoplasms require a minimal myeloid cell identity—TFs involved in the terminal differentiation of myeloid cells (e.g., RUNX1, CEPBα, PU.1) are rarely mutated to zero remaining activity. This is not true of TFs that are not important in myeloid cell identity (e.g., TP53).
- Principle 2: There are more than three doses—phenotypes associated with the dose of a TF can occur on a continuum of TF activity and are not limited to the Mendelian diploid, haploid, and null levels.
- Principle 3: The monoallelic loss of multiple genes is cumulative—the concurrent losses of genes neighboring a HITF in the setting of aneuploidies and focal deletions can modulate the phenotype.

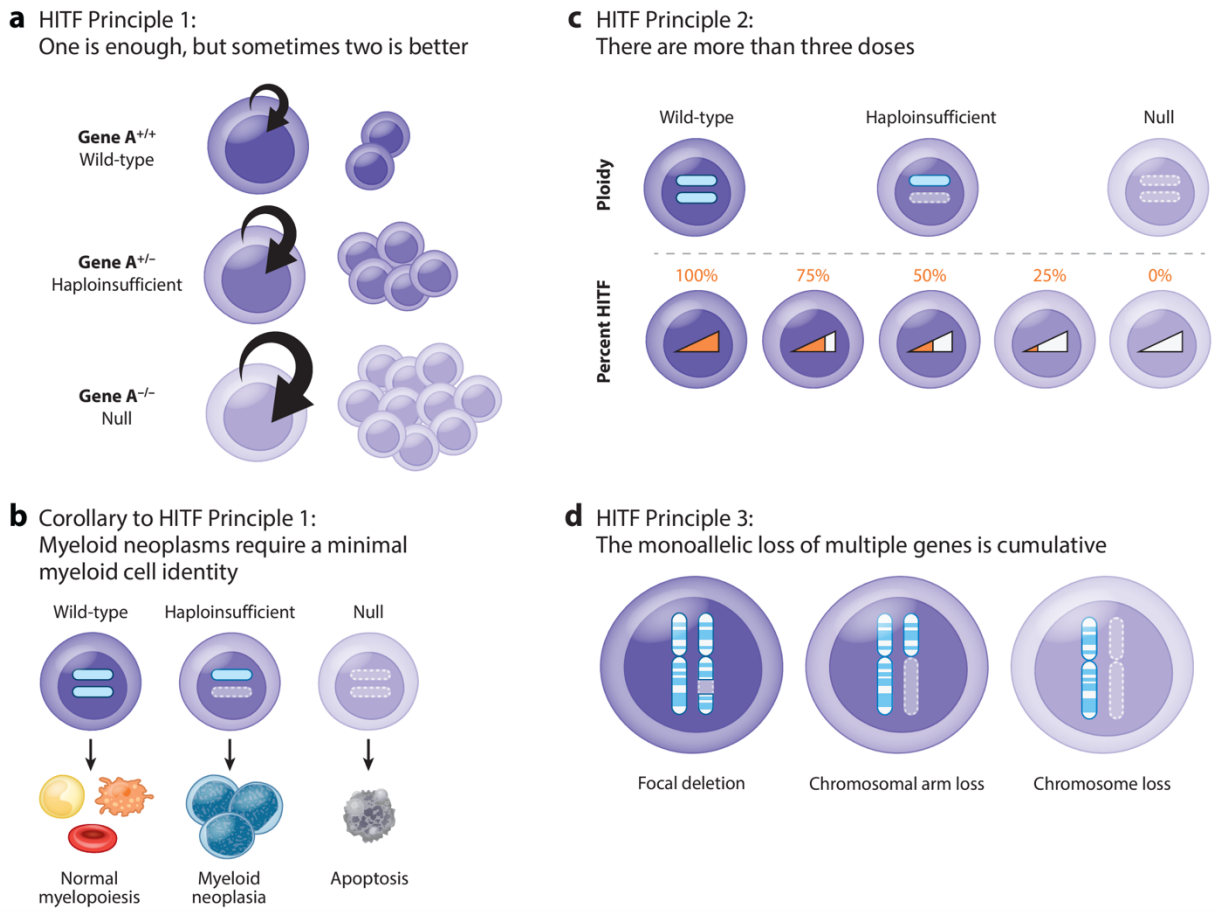

Figure 1. Principles of HITFs.
(a). Some genes are haploinsufficient in that monoallelic loss in a diploid organism produces a phenotype. However, this does not preclude an increased effect or distinct phenotype upon a further reduction in activity or loss of the second allele. (b) Differentiation into the full panoply of myeloid cells is possible for progenitors with a fully intact myeloid transcriptional network (left). A block in myeloid differentiation is imposed with the loss of a HITF involved in the differentiation of myeloid cells, and leukemogenesis can proceed with the acquisition of additional genetic aberrations (center). However, in the absence of myeloid cell identity, for example, through the total ablation of CEPBα activity, myeloid leukemogenesis cannot proceed (right). (c) Dose has traditionally been conceptualized as the number of intact alleles of a gene. However, further evidence has given rise to a model of TF activity that extends beyond diploid, haploid, and null. Specifically, preleukemic states can arise in cells with slight, sub-diploid alterations in the dose of a HITF, and sub-haploid but not null doses of a HIFT efficiently drive tumorigenesis. (d) HITFs are only part of the story when it comes to the phenotype of deletion events that result in the loss of multiple genes. The ultimate cellular phenotypes of deletion events depend on the other co-deleted genes. Abbreviations: HITF, haploinsufficient transcription factor; TF, transcription factor.
Sourced from Martinez and McNerney 2024[25]

Hematopoietic stem cell heterogeneity

Hematopoietic stem cells (HSCs) are responsible for generating the full complement of blood cells for the longevity of the organism. From the earliest conceptions of the hematopoietic stem cell, investigators pondered the potential heterogeneity of this cell population. Unitarians like Maximow and Neuman, who had previously attributed the bone marrow to the site of blood production[26], believed the "*stammzelle*" gave rise to all blood lineages[27]. Earlier work from Haeckel used the term *stammzelle* to refer to the stem cell that gave rise to the primordial germ layers, envisioning one cell as the root of a tree that would ultimately give rise to all cell types of an organism; however, Maximow is credited with applying the term to hematopoiesis[28,29]. The unitarian or monophyletic view was not without detractors. The dualist doctrine supported by Ehrlich conceived of separate stem cells for granulocytic and lymphocytic lineages, even going as far as to suggest that lymphocytes might develop in different organs than granulocytes[30,31]. These contrasting conceptions of the cell of origin for all hematopoietic lineages established the foundational principle of multipotency, which is essential in defining an HSC. Inherent in the characterization of an HSC is its capacity to differentiate into all mature blood lineages.

Early conceptions of HSCs were supported by peering into the microscope. The inability to isolate and manipulate these cells hampered refining the definition and understanding of HSCs. But even before HSCs could be isolated, parabiosis experiments demonstrated that a healthy parabiont could haematopoietically rescue an irradiated parabiont[32]. Transplanting bone marrow cells to irradiated recipients demonstrated one cell or combinations of cells were sufficient to restore hematopoiesis after bone marrow ablation[33]. Finally, reintroducing populations derived from spleen colony-forming units to irradiated recipients established that hematopoietic reconstitution depends on the ability of a single cell to generate a clonal panoply

of descendants[34,35]. These experiments established self-renewal as the second definitional principle of HSCs. From the beginning, McCulloch and Till noticed the heterogeneity in colony-forming potential and self-renewal. They proposed that the differences they observed could result from heritable heterogeneity in the population or the loss of the ability to self-renew by taking on the specialized task of differentiation[34,35]. This question of progenitor heterogeneity in self-renewal stands the test of technological advancement because, 60 years on, a clear answer is still lacking.

Fluorescence-activated cell sorting (FACS) was the first technological advancement to herald a refined understanding of hematopoiesis[36]. This technique allows for the characterization of cells based on immunophenotype and the isolation of specific cell populations to test their functional properties. From here, the first murine HSCs were FACS-purified and characterized[37]. With an understanding of the two definitional properties of HSCs—multipotency and self-renewal—it was apparent that $Thy1^{low}$ Lin^{-} $Sca1^{+}$ cells were potentially a 1000-fold enrichment of HSCs over whole bone marrow but still constituted a heterogeneous population[37]. Thus, the quest for ever-more pure markers to isolate was started, motivated by the potential secrets of self-renewal that a pure population of HSCs would unlock[38]. CD34 and then SLAM family markers further refined the immunophenotype of the HSC compartment[39-41]. Even still, the expression of more surface markers, such as cKit, CD49b, CD41, and CD201, have shown the ability to divide the HSC compartment based on the potential for self-renewal. Setting aside the immunophenotype of HSCs, an alternative method of isolating pure HSCs by exploiting the rates of dye efflux even reveals heterogeneity in this population[42-44]. These studies have demonstrated that heterogeneity still dominates even in relatively pure populations of HSCs[45].

Functional studies of HSCs have revealed four factors by which HSCs exhibit heterogeneity: hierarchy, lineage bias, cell cycle participation, and repopulating capacity[46]. Flow cytometry enables the analysis of single cells, but cell surface markers provide a limited set of parameters to dissect these properties. The second technological advancement that enabled further refinement of the heterogeneity in the HSC compartment, single-cell RNA sequencing (scRNA-seq), vastly expands the parameters to compare cells[47]. Bulk techniques have the asset of greater depth of information, but they always carry the risk of obfuscating meaningful differences between single cells. scRNA-seq, on the other hand, largely mitigates this risk of missing biological heterogeneity but creates the new risk of potentially amplifying noise in the population from nonsynchronous cellular processes like cell cycle and probabilistic gene expression[48,49]. Nevertheless, the first use of scRNA-seq began to link the gene expression program of single HSCs with high repopulation capacity to the gene expression programs[50]. Importantly, linking transcriptome with functional potential is one of the most challenging tasks, but it is required to establish biologically meaningful connections. The ability to sequence more cells helped further refine the gene expression program of HSCs[51,52].

Another fundamental insight from McCulloch and Till was that the environment created in an irradiated recipient was not akin to that of a healthy bone marrow[35]. Bone marrow and HSC transplant are necessary methods to assert the cell-intrinsic properties of progenitors, but the behavior of a cell in the transplant setting might not recapitulate its behavior in the native setting. Specifically, transplant reveals the differentiation trajectories of progenitors when forced to repopulate an empty mouse. The third technological advancement, sequencing-based lineage tracing, allows the study of hematopoiesis in the absence of a transplant[53]. Clonal lineage tracing relies on the ability to discern the descendants of a single HSC unambiguously. Early uses of

retroviral-[54] or lentiviral-mediated[55] lineage tracing revealed heterogeneity in HSC behaviors, resulting in differences in HSC output. However, viral lineage tracing necessitates the isolation and transplantation of cells after transduction. Thus, it is similar to the prospective isolation of HSCs by FACS but still allows for the clonal study of HSCs[53]. PolyLox- and CRISPR-based murine lineage tracing strategies rely on the ability to induce the generation of a diverse set of barcodes from arrays of DNA sequences and then read out the barcodes in single cells, even in tandem with the cellular transcriptome. These models have made multiple major contributions to our understanding of hematopoiesis. First, these models controvert the idea that HSCs play a major role in steady-state hematopoiesis. These models suggest that downstream progenitors play a major role in adult hematopoiesis[56]. Second, transcriptomic HSCs can produce cells of only a single lineage or a small subset of lineages over a long period of time[57-61]. An important caveat of this finding is that it is unknown if these cells are truly restricted in their ability to contribute to other lineages or have a strong bias toward one lineage over another. Third, there is a distinct transcriptome of quiescent HSCs that do not participate in steady-state hematopoiesis[59,62]. These Childless HSCs can retain quiescence in the face of stress cues to differentiate and are thus tasked with preserving stem cell potential in the event that the organism survives the temporary stress[62]. Finally, these models reveal that the transcriptome of progenitors is not sufficient to predict their behavior entirely[57,63]. The DNA methylation epigenome is among the hidden variables in predicting HSC behavior, highlighting the important role of transcription factors and transcription co-regulators in HSC fate decisions[57]. Linking the lineage output of an HSC to its transcriptome in the native setting has provided important insights into hematopoiesis, highlighted persistent gaps in understanding, and questioned important assumptions about the definition of HSCs.

These advancements in understanding hematopoiesis have led to the current paradigm in hematopoiesis, which posits that differentiation is a continuous process where HSCs gradually adopt the transcriptome of terminally differentiated cells and slowly lose the ability to self-renew over time[51,52,64,65]. This model accepts that the HSC compartment is functionally and transcriptionally heterogeneous. There are important implications of this model. First, heterogeneity could arise from sampling HSCs at different points along the continuum of differentiation[55,58,59,62,63,66-70]. Second, the existence of lineage-biased HSCs controverts the assumption that self-renewal and differentiation are mutually antagonistic processes[42,57-59,68,69,71]. The platelet-biased stem cell may be a special case as the phenotype can be observed, even in single HSC transplantation[66]. This could be an evolutionary remnant. Multicellular organisms of a certain size require a circulatory system, and the primary requirement of a circulatory system is the maintenance of its integrity. The first blood-producing cells likely started by producing platelets. Thus, this bias is retained in apex HSCs. Third, the starting cell state can influence differentiation trajectories. Not only does lineage-biased transcription result in lineage-biased outcomes, but HSCs closer to the approximation of the HSC transcriptome exhibit greater repopulation capacity[50]. Finally, the multiple differentiation trajectories likely contribute to the plasticity of hematopoiesis and accommodations for demand adaptation[53].

Dose in Hematopoiesis

The classical, branching tree–based model of hematopoiesis presupposes discrete steps at which developing cells make unforeseen and binary fate choices. While a helpful tool in conceptualizing the hierarchy of development, this model fails to capture the heterogeneous and continuous nature of hematopoiesis[72]. Prospective fractioning of hematopoietic stem cells (HSCs) on the basis of cell surface or cytoplasmic protein expression reveals the cell-intrinsic

heterogeneity in key HSC functions, namely, cell cycle participation, lineage bias, and self-renewal capacity[68]. HSC heterogeneity is further confirmed by retrospective lineage tracing models that demonstrate functional and transcriptional differences present in HSCs[59,73]. In other words, lineage biases are already present at the apex of hematopoiesis, are continuously accrued, and sometimes entail more than one route to arrive at a terminal differentiated cell[51,58]. TFs and coregulators are the likely origins of this cell-intrinsic HSC heterogeneity, as differentiation results from the proper expression and timing of TFs[74].

Imagining the three properties of HSC heterogeneity—self-renewal capacity, cell cycle participation, and lineage bias—as continuous spectra upon which HSCs can reside rather than as discrete phenotypes that cells either possess or do not possess is more consistent with the continuous model of hematopoiesis and the dose-dependent nature of TFs (Figure 2). Loss of one copy of a TF can reposition cells along these axes and potentially endow a cell with increased capacity for self-renewal and/or a block in lineage-specific differentiation. For example, monoallelic loss or mutation of the nonclustered, homeobox domain-containing TF CUX1 is found in approximately 10–15% of myeloid neoplasms[75,76]. Loss of CUX1 in mouse models promotes an expansion of hematopoietic stem and progenitor cells (HSPCs), progressive anemia, and an expansion of mature myeloid cells, leading to a lethal myelodysplastic syndrome (MDS)[77,78]. In other words, the change in CUX1 dose influences HSPC decisions and moves the HSPCs along the axes. Similar changes can be conceptualized for the other TFs discussed in this review.

Gene	Genetic aberration	CH	CCUS	CMML	MPN	MDS	AML	t-MN
CEPBA	Mutation	0%	0%	1%	0%	2%	9%	4%
CUX1	Mutation	0.03–0.4%	3.4%	10%	0.2–3%	2–6%	1%	2%
	−7/del(7q)	0.0002%	NR	1–10%	0%	2–16%	9–14%	33–50%
EGR1	Mutation	NR	NR	0%	NR	NR	0.5%	NR
	del(5q)	0.0003%	NR	1.5%	0.5%	4–15%	7–16%	49%
ETV6	Mutation	0.1%	0%	0%	1.6%	3–6%	1%	1%
	−12/del(12p)	0.0001%	NR	1%	0.2%	3–5%	2–7%	12%
GATA1	Mutation	0–0.1%	0%	0%	0%	<1%	0.2%	0%
GATA2	Mutation	0–0.1%	1.1%	0%	0.2–2.6%	2–5%	2–3.5%	1%
	inv(3)	NR	NR	NR	NR	Rare	1–1.5%	2%
MYBL2	Mutation	NR	NR	0%	0.7%	NR	1–5%	NR
	−20/del(20q)	0.001%	NR	1%	0%	4–5%	2–6%	9%
RUNX1	Mutation	0–0.2%	3.4%	13.2%	0.5–2%	9–14%	6–12%	10%
	t(8;21)	NR	NR	ED	ED	ED	1–4%	1%
	t(3;21)	NR	NR	NR	NR	NR	0%	3.6%
	inv(16)	NR	NR	ED	ED	ED	3–6%	1–2%
	+21	0.00025%	NR	1%	NR	2–4%	2–5%	11%
TP53	Mutation	2.6–4.4%	1.1%	1%	2–8%	6–13%	6–9%	25–37%
	−17/del(17p)	0.0002%	NR	1%	2%	7%	7%	13%
PML-RARA	t(15;17)	NR	NR	ED	ED	ED	3–9%	2%
Selected mutations that alter the dose of HITFs								
MLL1	t(11q23;x)	NR	NR	NR	NR	NR	3%	3.3%
NPM1	Mutation	0–0.1%	0%	2.6%	0.7%	1.5–2%	22–27%	15%
FLT3	Mutation	0–0.1%	0%	2.6%	0.5%	1.5–5%	28–33%	20%
Reference(s)		79-82	83	84-88	84,89-91	84,92-97	47,98-104	95,105-109

Table 1. Frequency of HITF genetic aberrations and genetic aberrations affecting the dose of HITFs across the spectrum of myeloid diseases.

Abbreviations: AML, acute myeloid leukemia; CCUS, clonal cytopenia of undetermined significance; CH, clonal hematopoiesis; CMML, chronic myelomonocytic leukemia; ED, excluded on the basis of diagnostic criteria; HITF, haploinsufficient transcription factor; MDS, myelodysplastic syndrome; MPN, myeloproliferative neoplasm; NR, not reported; t-MN, therapy-related myeloid neoplasm.

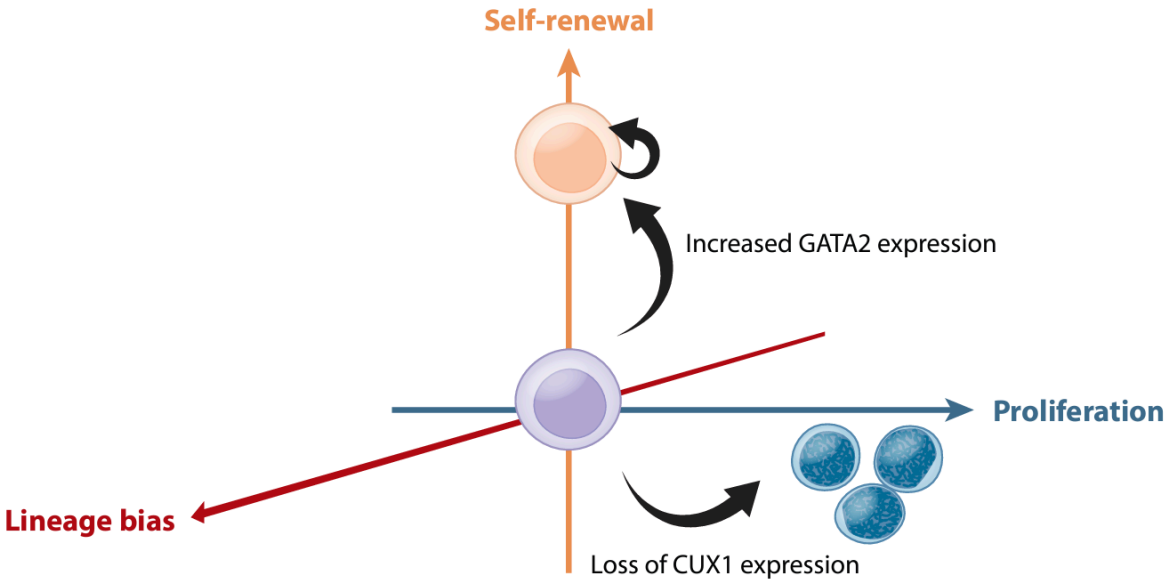

Figure 2. Stem cell properties are continuous characteristics.
Self-renewal, proliferation, and lineage bias are properties of hematopoietic stem cells that exist as continuous variables. These properties are dictated by the dose of HITFs. Therefore, alteration of the dose of a HITF can have profound effects on these properties. For example, GATA2 overexpression, which is found in primitive AMLs, results in excess self-renewal. In addition, loss of CUX1 results in increased proliferation and a myeloid bias at the expense of erythroid development. Abbreviations: AML, acute myeloid leukemia; HITF, haploinsufficient transcription factor.

Sourced from Martinez and McNerney 2024[25]

Mechanisms of Transcription Factor Haploinsufficiency

Transcription Factor Dynamics

TFs actualize the final cell state by reciprocally reading and influencing the enhancer code to promote the expression of effector genes. Transcription is a dynamic and noisy process that depends on the genomic colocalization of many factors. Thus, fundamentally, transcription is a concentration-dependent process. The noise of transcription refers to the phenomenon where rather than stable, consistent RNA polymerase II (Pol II) activity at a given gene, Pol II transcribes the gene in short, probabilistic bursts, which are influenced by enhancers and TFs[110,111]. *PU.1*, *GATA1*, and *GATA2* mRNA all exhibit noisy transcription with all stem cells exhibiting some expression of these genes, although not necessarily at the same time. Perhaps counterintuitively, these dynamic properties may increase the stability of stemness in the HSC pool as whole by maintaining HSCs that are sampling every transcriptional state[112,113].

Transcriptional noise impacts HITFs on two ends: the transcription of the TF itself and transcription of the downstream target genes. TFs generally have low expression levels at baseline[114]. Further loss of one TF allele portends particularly noisy expression, and the amount of the TF may even intermittently approach the null state as a result[115]. While transcription is inherently noisy, monoallelic loss of a HITF amplifies this noise, resulting in an inconsistent transcriptome and increased intercellular heterogeneity[116]. This variation is the precondition for selection, allowing a rare cell transcriptional state to be reinforced at the chromatin level, stabilizing an increase in fitness of a preleukemic cell[117-119].

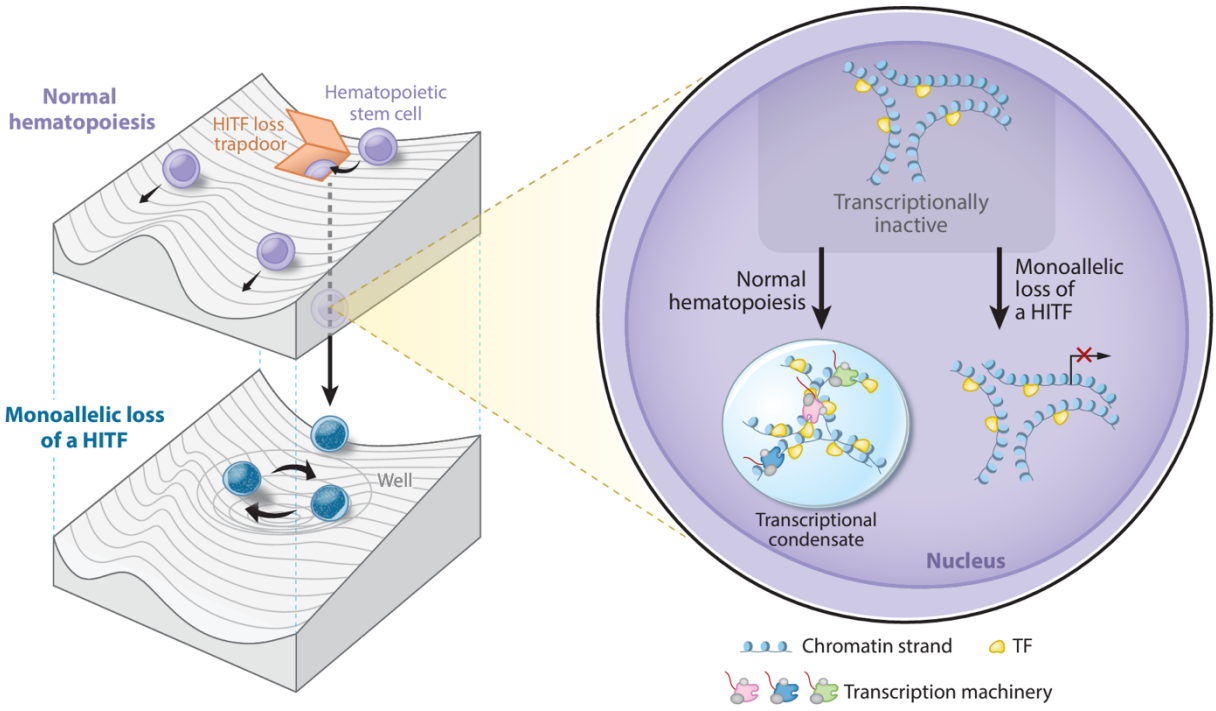

Figure 3. Loss of a HITF results in a myeloid differentiation block and greater transcriptional plasticity.

HSCs can develop normally by following one of the differentiation paths out of the bone marrow. Or, in a variation on Waddington's landscape, HSCs can undergo monoallelic loss of a HIFT, represented as falling through the trapdoor. The loss of a HIFT has profound effects on the nuclear environment. Normally, TFs bridge DNA sequences at enhancers with the general transcription machinery to hold all of these molecules in close proximity to transcribe a gene. However, with decreased TF molecules present in the nucleus, clustering of all the requisite DNA sequences and transcriptional machinery happens less frequently. It is difficult to achieve a sufficiently high local concentration of a TF to move enhancers and genes into an environment permissive for transcription, which can result in genes failing to be relocated and thus transcribed. Consequently, increased transcriptional plasticity ensues as cells fail to maintain a consistent epigenome and transcriptome. Furthermore, such cells have an inability to properly differentiate to myeloid cells due to the dysregulated enhancer landscape. These two ramifications of HITF loss—increased transcriptional plasticity and a myeloid differentiation block—are depicted as the cell rolling around in the well, sampling many different transcriptomes, and being unable to escape down the path of normal myeloid differentiation. Abbreviations: HITF, haploinsufficient transcription factor; HSC, hematopoietic stem cell; TF, transcription factor.

Sourced from Martinez and McNerney 2024[25]

In summary, the loss of a HITF has two major ramifications for leukemogenesis (Figure 3): First, it provides the requisite myeloid differentiation block. Second, it increases the intercell heterogeneity by increasing the number of sampled transcriptomes, giving the cell an advantage in adapting to the environment.

In addition to increasing transcriptional noise, suboptimal TF concentrations impede the efficiency of TFs binding to cognate DNA sequences. TF dynamics are determined by search pattern, search time, and on/off rates[74]. The search pattern of a TF refers to how the protein moves about the nucleus to find target sequences. MYC, for instance, explores the entire nucleus, while other TFs, such as CTCF, linger in smaller areas containing a high concentration of binding sites[120,121]. TF search time refers to the interval between DNA binding events and can be influenced by mitigating factors such as the chromatin milieu[122-124]. TFs spend the most time unbound and in search mode [74]. Per the law of mass action, a reduction in TF molecules will lower the overall likelihood that the DNA targets are occupied by the HITF. While this conclusion is intuitive, the downstream consequences can be complex. For instance, TFs regularly interact with other TFs and/or cooperatively bind DNA. SOX2 is one such example and guides OCT4 to cobind DNA. Decreased SOX2 concentration not only curtails SOX2 DNA targeting but also increases the search time of OCT4[125]. Numerous TFs physically interact with one another and bind DNA in concert, suggesting that they could have linked search patterns as well[126]. Prolonged searching for DNA binding sites could be further exacerbated in situations where the HITF is a pioneer factor, wherein many TFs depend on the HITF for access to DNA[127]. These scenarios highlight the wide-ranging detriments associated with inappropriate concentrations of a HITF.

Even in normal conditions, cells temporally manipulate the nuclear concentration of TFs to

bring about distinct cellular fates[128]. This phenomenon is exemplified by nuclear hormone receptors, which shuttle into the nucleus upon ligand binding, but also extends to other classes of TFs[129]. For instance, the effects of the dosage-sensitive TF TP53 are conveyed by nuclear localization. Pulses of TP53 into the nucleus promotes the repair of DNA damage, whereas sustained TP53 nuclear localization causes senescence[130,131]. In this example, the extent of DNA damage is communicated through the nuclear concentration of TP53, producing different outcomes. The developmental TF HES1 also exhibits distinct nuclear oscillation patterns and concentrations, producing different outcomes in neural progenitors[132,133]. Nuclear factor kappa B (NFκB) signaling patterns in murine HSPCs similarly result in differential gene expression of developmental or apoptosis genes in response to inflammation[134]. Given the finely tuned patterns necessary for these TFs to convey unique cell state information, it is unlikely that only one allele of a TF could achieve these same nuanced dynamics. These examples indicate not only the importance of TF concentration but also the ability of dose dynamics to convey different responses to environmental stimuli.

Clustering the Characters

Context-dependent gene expression is largely mediated by enhancers[135]. To influence transcription, an enhancer must have sustained interaction with a target promoter, and TFs bestow DNA sequence specificity on this interaction[125,136,137]. One enhancer can regulate multiple genes, and multiple genes can be coordinately regulated in turn, a process bridged by TFs[125,135,138-143]. These actors can be found in physical proximity to each other and have been described as transcriptional hubs or factories[143-146]. More recently, liquid-liquid phase separation has been considered an etiology of nuclear subcompartmentalization[147]. These biophysical structures may explain the observed clustering of the requisite transcriptional regulators and *cis*-

regulatory elements that regulate Pol II activity. Much more work is warranted to understand the role of phase separation in transcriptional regulation. Nonetheless, in both the hub and condensate models, TFs migrate into subcompartments of the nucleus, increasing local TF concentration[144,146,148,149].

KLF1 and EBF1 provide two examples of how TFs can be important for the relocation of target genes. In murine erythroid progenitors, KLF1 forms dynamic transcription factories. KLF1 shepherds together erythroid genes from across the genome, including hemoglobin genes, to coordinately drive transcription promoting erythropoiesis[144]. Reiterating the importance of the proper nuclear localization of TFs and target genes, monoallelic loss of the B cell HITF EBF1 results in less efficient relocation of a target transgene out of heterochromatin, leading to decreased transgenic expression. In this scenario, there is insufficient EBF1 to pull the transgene into a nuclear environment more conducive for transcription[150]. This finding supports a model of transcription where the TF must attain a certain local concentration at a given locus, which is associated with moving the gene into a permissive transcriptional environment. This model presupposes the ability to generate high local concentrations of a TF, a feat that is impaired in the setting of haploinsufficiency. The decreased frequency of activation of the transgene underscores the probabilistic nature of transcription and the buffering capacity of two alleles.

The biochemical properties mediating the clustering of *cis*-regulatory elements and the general transcription machinery is an active area of investigation. Consequently, intrinsically disordered regions (IDRs) have received renewed attention for the ability of these protein domains to mediate clustering, condensation, and liquid-liquid phase separation. The overwhelming majority of TFs have IDRs, which lack fixed three-dimensional structure but can still convey interaction specificity[151,152]. Mutations in IDRs are common in cancer and enriched

in known drivers of hematopoietic malignancy[152]. The transactivation domains of TFs have repeatedly been shown to be IDRs, which form condensates with general transcription machinery, including Pol II, Mediator, and BRD4, leading to high local concentrations of the important players for transcription[146,153-157]. These findings lead to a model wherein the IDRs and DNA binding domains of TFs serve to couple the transcriptional machinery with DNA. As a result, the TFs are held in increased local concentration to facilitate repetitive DNA binding and transcriptional bursting[145,147,158,159]. The loss of one copy of a HITF decreases the probability of target gene compartmental sequestration and activation. Overall, sequestration of transcriptional players and processes within the diffusion space of transcription condensates reduces transcriptional noise, as important molecular players are more likely to bump into each other[160].

Models of Transcription Factor Action at Target Genes

Regardless of the physical relationship between TFs, enhancers, and transcription initiation, some genes are more sensitive than others to diminished TF levels, begging the question of how a TF segregates in the nucleus under limiting conditions. For example, in murine HSPCs, heterozygosity of *RUNX1* leads to a loss of binding to and expression of the *Cepbε* locus but not the *Cepbα* locus, suggesting an underlying mechanism governing the actions of dose-limited RUNX1[161]. The differential sensitivity of some genes to the loss of a TF has led some investigators to conceptualize TFs as having either analog or digital outputs. RUNX1 in this case is functioning in a digital or binary manner; at a certain threshold of expression, RUNX1 either activates a gene target or does not. The TF ETV2 also exhibits digital effects. ETV2 can promote the development of both endothelial cells and erythroid cells, and this dichotomy is conveyed through different thresholds of binding and activation for genes of each lineage. Specifically, low levels of ETV2 are sufficient to promote expression of endothelial genes, but a much higher

threshold must be met to promote erythroid gene expression[162]. In contrast, NFκB TFs exhibit analog or gradient-like effects where the expression of target genes is directly proportional to the level of NFκB activation[163]. At a limited number of loci tested, PU.1 also appears to act in an analog manner at loci related to proliferation[164]. Interestingly, *SOX9*, which is mutated in devastating craniofacial developmental disorders, exhibits both analog and digital activity. Enhancers that exhibit analog activity are more likely to have a SOX9 palindrome motif and are associated with different pathways than digital enhancers, which are more likely to have a weak SOX9 motif and be bound by other TFs[165]. This later work represents the most advanced analysis of a dose-dependent TF, accomplished through the use of a degron tag to generate a spectrum of TF doses. As such tools are further developed to manipulate TFs at a single-cell level in an inducible manner, we may find that more TFs resist simplified classification into analog and digital regulators.

Compensation

In addition to the biochemical sequestration that condensates provide, cells have other mechanisms to safeguard the robustness of the transcriptome. In some cases, however, these mechanisms may be insufficient to compensate the monoallelic loss of a HITF. The long half-lives of mRNA molecules and posttranscriptional regulators are thought to buffer against the probabilistic nature of transcription. However, TFs are generally less stable on the mRNA and protein levels as compared with the rest of the genome[114]. Hence, compensatory mechanisms are less efficacious for TFs.

Some proteins have closely related family members that can minimally compensate for the loss of one of the members. With respect to RUNX1, family members RUNX2 and RUNX3 can partially compensate for the loss of RUNX1[166,167]. In the setting of CEPBα loss, CEPBβ

(*CEBPB*) can compensate in the setting of emergency granulopoiesis induced by inflammation[168-170]. Yet, as a class, HIFTs are less likely to have paralogs, contributing to the necessity of two alleles of these genes for proper development[171].

How to lose a partner in seven ways

Having considered the molecular mechanisms of TF action and the implications for haploinsufficiency and dose dependence, we now focus on the common means by which the abundance of a gene product can be altered and how the mechanism of gene inactivation can have a profound impact on the phenotype manifested (Figure 4). For example, the chromatin landscape of RUNX1-translocated acute myeloid leukemia (AML) is counterintuitively more akin to that of double-mutant CEPBα AML ($CEPB\alpha^{DM}$) than AMLs with RUNX1 point mutations[172]. Below, we consider the different types of HITF loss and how the context impacts the phenotype.

Aneuploidy

Losing one copy of part or all of a chromosome is seemingly the most straightforward method to lose one allele of a TF; however, the implications of losing potentially hundreds of neighboring genes in tandem as well have not been sufficiently appreciated[173]. In most cases, aneuploidy is a detriment to the fitness of a cell—this is evinced in the finding that the more genes on a chromosome, the less likely cancer cells are to tolerate loss or gain of that chromosome[174]. It is only under specific circumstances that aneuploidy confers an advantage. And while aneuploidy has been thought to be a late event of advanced-stage neoplasms, aneuploidy can be an early and founding event, as evidenced by aneuploidy in clonal hematopoiesis[79,175].

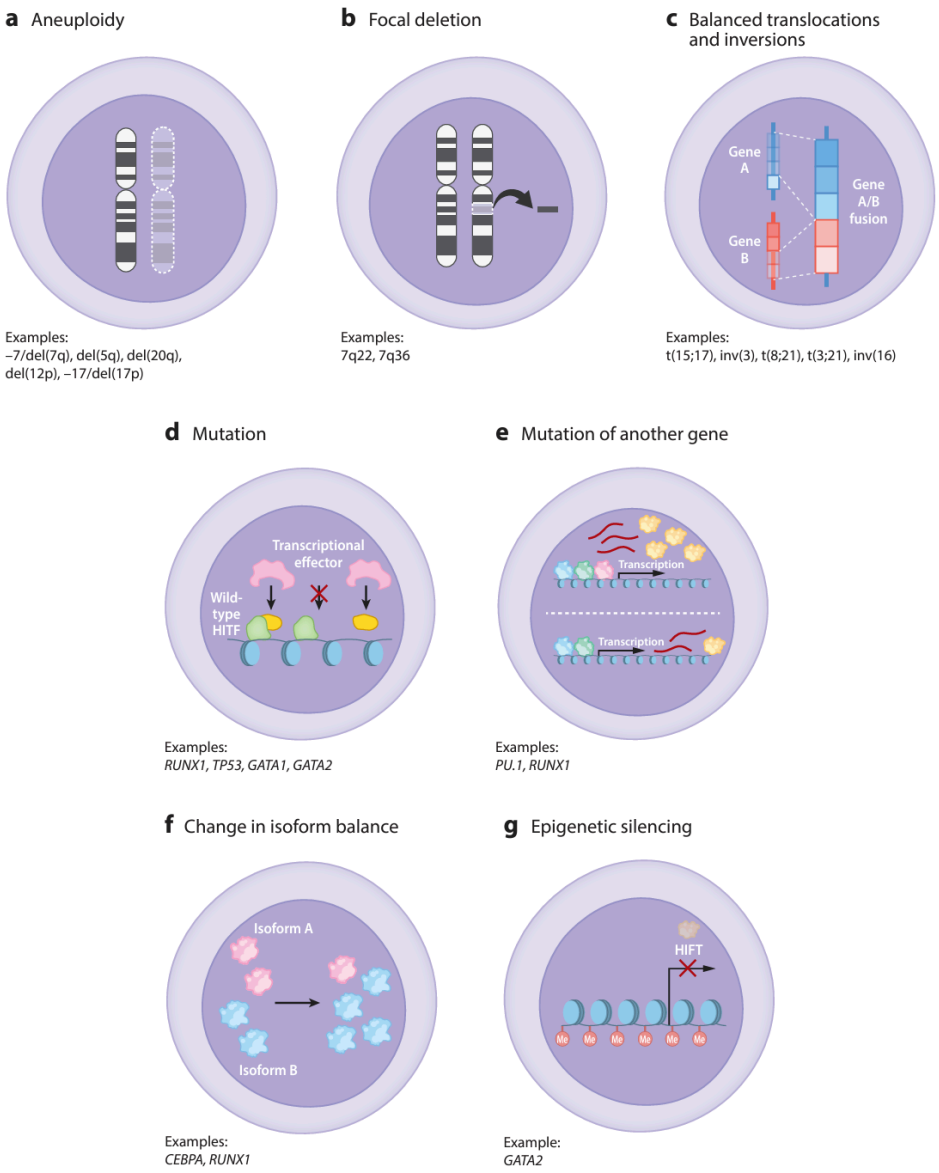

Figure 4. Multiple ways to lose a HITF.

(a) Aneuploid events result in the loss of a HITF in the context of the loss of an entire chromosome or chromosomal arm. (b) Focal deletions result in the loss of a HITF without the ensuing cellular stress of aneuploidy, but the phenotype is still the consequence of the loss of multiple co-deleted genes. (c) Balanced translocations and inversions create a potent fusion oncogene or lead to the overexpression of an oncogene, but the genetic disruption also creates loss-of-function events in the rearranged genes. (d) TFs have two basic parts: the domain that binds DNA and the part(s) of the protein that assemble with cofactors, other TFs, and/or general transcription machinery. Mutations in either of these parts can have diverse effects on the mutant protein. The TF can lose the ability to bind DNA and thus wild-type function. A TFs can also lose the capacity to associate with necessary cofactors and thus forfeit the ability to serve as a bridge between specific DNA sequences and other proteins to regulate transcription. Dominant

Figure 4., Continued
negative proteins arise on the basis of how much the mutant protein interferes with the functions of the remaining wild-type allele. (e) The complexity of the myeloid transcription network necessitates that dose alterations in one protein can have cascading disturbances on the effective dose of another gene. Thus, mutation of a HITF or other critical myeloid gene can have profound effects on the transcription or stability of other HITFs. (f) TFs have multiple isoforms with different functions. One way that the effective dose of a HITF can be disrupted is a change in the balance of these isoforms away from the wild-type ratios. This departure from the normal ratio changes the function of the TF pool given that the isoforms have different capabilities and functions. (g) The final way that HIFT dose can be affected is the loss of expression of a HIFT through aberrant epigenetic silencing through DNA methylation, repressive chromatin methylation, or a loss of accessibility at important sites regulating the expression of the gene. Abbreviations: HITF, haploinsufficient transcription factor; Me, methylation; TF, transcription factor.

Sourced from Martinez and McNerney 2024[25]

In general, any monosomy will have profound effects on the cell, including inducing ribosomal stress and gene expression changes that are not limited to the affected chromosome. The ensuing proteotoxic stress activates a TP53 response that can promote cell death[176]. Consequently, aneuploidy is largely incongruous with intact TP53 signaling, likely due to proteotoxic stress, and this may be why AML and MDS with complex karyotypes often also present with *TP53* mutations[177,178]. That said, while there are common features of aneuploidies, there is variability to the cellular response to aneuploidy as well. In one report, cells harboring an identical aneuploid event can diverge with respect to cell cycle alterations, response to extracellular and intracellular signals, and gene expression[179]. This aneuploidy-induced transcriptional plasticity likely also fosters increased drug resistance[180,181]. Others have argued that there is extensive posttranscriptional compensation for aneuploidy in cancer. As a result, protein changes are not as dramatic as one might surmise, including for oncogenes and tumor suppressors[182]. Nonetheless, an aneuploid event results in not only the loss of a HIFT within a greater genetic deletion but also the concomitant cellular stress of aneuploidy.

A second distinction of aneuploidy-induced HITF loss is that several pathogenic genes may be simultaneously deleted en bloc. A contiguous gene syndrome (CGS) can arise in this event, which refers to the phenotypic consequences of the loss of multiple neighboring genes. It is technically challenging to experimentally model CGS; however, there is mounting evidence that multiple genes encoded on a chromosome arm can cooperatively drive cancer when lost[183]. In some cases, CGS genes may converge on similar pathways; in others, genes may impact complementary pathways[184]. It would be remiss to neglect co-deleted genes when considering HITFs impacted by aneuploidy (HITF principle 3). In the remainder of this section, we highlight HITFs within recurrent aneuploid events in myeloid neoplasms. However, we acknowledge that

aneuploidies and focal deletions have more complex phenotypes than solely the loss of the HITF alone.

EGR1 encodes an ETS family TF and is frequently deleted in the context of del(5q). Consistent with del(5q) as a recurrent event in myeloid neoplasms arising in patients with a history of cytotoxic therapy [therapy-related myeloid neoplasms (t-MNs)], monoallelic loss of *EGR1* is sufficient for myeloid neoplasia upon treatment with an alkylating agent[185]. The loss of *EGR1* conveys a fitness advantage by blunting the DNA damage and inflammation response in HSPCs[186]. Highlighting principle 3, EGR1 cooperates with another 5q resident, *APC*, encoding a Wnt signaling component, as well as *Trp53* to promote de novo AML. In contrast, heterozygosity for only one or two of the three genes does not result in AML[184]. This example highlights the aggregate phenotype that can arise due to cumulative haploinsufficiency and emphasizes the need to consider the impact of genes deleted concomitantly with the HITF.

CUX1, a homeodomain-containing TF, has conserved dose sensitivity from *Drosophila melanogaster* to humans. In flies, the homologue of CUX1, Cut, has dose-dependent roles in the development of at least two organ systems[187,188]. Interestingly, up to four different doses of Cut are associated with different developmental fate decisions, suggesting that manipulation of the level of this TF is one way that developmental decisions are made. Different fates associated with different doses of CUX1 may be realized by the analog nature of CUX1[189]. The graded effects of CUX1 dose are true at the level of not only transcription but also the elicited phenotype. In mice, knockdown of CUX1 to an intermediate level results in a mild form of MDS with anemia, while further knockdown results in a more severe anemia, an expansion of mature myeloid cells, and a lethal form of MDS[78]. Certain features unify CUX1 and its orthologues across model systems and human disease: Knockdown of CUX1 results changes in proliferation

and impaired differentiation[76,77]. In humans, *CUX1* is deleted in the context of –7/del(7q) without silencing or mutation of the remaining allele[76,77,98]. In de novo myeloid malignancies, *CUX1* mutations or –7/del(7q) are more associated with MDS. In fact, in some cases of pediatric MDS, –7/del(7q) appears sufficient for the disease, as it is the sole somatic genetic change identified [190]. Also, –7/del(7q) is strongly associated with t-MNs[109]. One explanation for this association is that CUX1 has an early role in the DNA damage response, hence CUX1 deficiency aberrantly allows continued HSPC proliferation in the setting of DNA damage stemming from cytotoxic therapy[191]. The resultant clonal advantage of CUX1-deficient cells ultimately leads to the development of an overt myeloid neoplasm. This mechanistic work highlights the role of the HITF levels in normal differentiation and how deviant levels promote transformation.

Del(20q) results in the loss of *MYBL2* (bMyb), a TF that regulates the differentiation and self-renewal of myeloid progenitors[192]. Homozygous deletion of *MYBL2* is embryonic lethal; however, haploinsufficiency drives a range of myeloid disorders, including MDS and myeloproliferative neoplasms (MPNs)[193]. Downregulation of *MYBL2* is sufficient for murine HSPC expansion, which is consistent with *MYBL2* downregulation as a common feature of MDS, regardless of 20q status[194]. This finding indicates that del(20q) is not the only route to downregulate *MYBL2*.

Del(12p) is a small interstitial deletion found in MDS with a complex karyotype. Due to its small size, del(12p) is underreported in standard cytogenetic analyses (99). The ETS-family TF ETV6 is lost in del(12p) events, and the remaining allele is generally not methylated or mutated, cementing this TF as a HITF[195,196]. ETV6 is a homodimeric transcriptional repressor required for the establishment of hematopoiesis and survival of adult HSCs[197]. Like many of the other HITFs discussed here, including *RUNX1*, *CEPBα*, and *GATA2*, *ETV6* also harbors germline

mutations[198]. Similar to individuals carrying a germline *RUNX1* mutation, patients with inherited *ETV6* mutations have a high likelihood of developing thrombocytopenia due to defects in megakaryocyte and platelet maturation, underscoring the necessity of both alleles of these two genes for proper platelet generation[199].

Unlike the genes we have considered thus far, *TP53* frequently harbors biallelic alterations. *TP53* alterations include mutations, loss of heterozygosity, and deletions in the form of −17/del(17p). As *TP53* status is usually assessed amid full-blown malignancy, when biallelic alterations are common, *TP53* can appear to abide by the two-hit hypothesis. Even so, there is ample evidence from mouse and human studies that monoallelic loss of *TP53* is sufficient to promote leukemogenesis[200]. *Trp53* heterozygosity is conducive to murine HSPC clonal outgrowth after DNA damage[201]. Likewise, heterozygosity for *Trp53* is permissive for tumor formation in mice, and the remaining allele can remain intact and functional[202]. In MDS, while many patients have monoallelic or biallelic *TP53* mutations, a small subset of patients present with loss of 17p without mutation or loss of heterozygosity of the remaining allele[200]. Notably, like other aneuploid events, del(17p) deletes additional tumor suppressor genes that independently drive tumorigenesis in mice[183]. Hence, our understanding of del(17p) pathogenesis remains incomplete. Notwithstanding, these studies indicate that TP53 produces haploinsufficient phenotypes that may become more pronounced with biallelic loss (HITF principle 1).

Focal Deletions

In contrast to aneuploidy, focal deletions are smaller than a chromosome arm[174]. Focal deletions can affect HITFs while potentially avoiding some of the fitness defects induced by aneuploidy. In fact, there is selection to delete tumor suppressor genes while preserving essential genes and

oncogenes in focal deletions[17]. As a result, focal copy number alterations are more potent in driving proliferation signatures as compared with chromosome arm level changes[203]. That said, focal copy number alterations still correlate with mutations in DNA damage response genes. This later finding implicates the DNA damage response pathway as a gatekeeper for genomic instability in general, as remarked upon above with *TP53* mutations and aneuploidy[203].

Also analogous to aneuploidy, HITF principle 3 remains true for focal copy number changes. As an example, we return to del(7q). In addition to monosomy 7 and del(7q), smaller cryptic deletions can occur that are below the sensitivity of standard cytogenetics analysis. Higher-resolution approaches have revealed minimally deleted regions, including one spanning *CUX1* on 7q22[76,204]. While *CUX1* loss is sufficient to cause myeloid disease, deletion of a 7q22 syntenic region in mice that does not include *CUX1* also disrupted hematopoiesis[78,205,206]. These data implicate more than one haploinsufficient 7q22 gene. Furthermore, these findings of a CGS in a focal deletion extend to other copy number variations as well[207].

Balanced Translocations

Balanced translocations typically garner attention for the ensuing fusion product or overexpressed oncogene. For instance, t(15;17)(q22;q12) produces the PML-RARA fusion protein, pathognomonic for acute promyelocytic leukemia (APL). Less appreciated is the fact that such translocations inherently generate loss of function in the rearranged genes, in this case, PML and the TF RARA. Translocations can thereby result in three hits for the price of one. While the fusion product alone initiates leukemogenesis, haploinsufficiency of PML and RARA also further promotes disease[208-210].

Similarly, inv(3)(q21q26.2) or t(3;3)(q21;q26.2) repositions the *GATA2* distal enhancer away from *GATA2* to instead drive oncogenic overexpression of the EVI1 TF[211,212]. Inherently, this

inversion or translocation also attenuates *GATA2* expression. Indeed, GATA2 heterozygosity accelerates EVI1-mediated oncogenesis in mice[212]. A minority (15%) of patients also have mutations of the non-rearranged *GATA2* allele (HIFT principle 1)[213]. Thus, like t(15;17), inv(3) simultaneously produces both gain-of-function oncogenic events and loss-of-function events in HITFs.

RUNX1 is another HITF involved in balanced translocations. RUNX1 is a pleiotropic TF required for the initiation of definitive hematopoiesis and multilineage differentiation in adult hematopoiesis. However, RUNX1 does not work alone; it requires the cofactor CBFβ (*CBFB*) to achieve proper regulation of target genes[214,215]. There are three ways that this critical interaction can be interrupted through translocation to promote the development of myeloid malignancy: Two are translocations involving *RUNX1* and one is a translocation involving *CBFB*.

The first translocation is the t(8;21)(q22:q22.1) translocation, which fuses the RUNX1 DNA-binding Runt homology domain (RHD) to the Nervy-homology region (NHR) domains of ETO (RUNX1T1). The NHR recruits the repressive nuclear receptor corepressor (NCoR) complex[216-218]. Consequently, this translocation results in a neomorphic RUNX1 fusion protein that represses myeloid differentiation genes it normally activates, such as *PU.1*, *GATA1*, and *CEPBα* [219-222]. The second translocation is t(3;21), which is most commonly found in t-MNs; t(3;21) fuses the RHD of RUNX1 to the TF EVI1. RUNX1-EVI1 also results in the loss of *CEPBA* expression, and thus the promyeloid differentiation program, but uniquely upregulates *GATA2,* producing a more primitive AML[95,223]. Both translocations create a potent oncoprotein and simultaneously disrupt the involved *RUNX1* allele. The derivative RUNX1 cannot interact with CFBβ to drive proper gene expression and is effectively dysfunctional. The remaining wild-type RUNX1 is in competition with the oncogenic RUNX1 fusion for DNA binding sites, resulting in

a finely tuned transcriptional network optimized for cellular fitness[222]. This is made apparent upon loss of the wild-type *RUNX1* allele, as knockdown of RUNX1 in cells expressing RUNX1-ETO causes apoptotic cell death due to the overactivity of RUNX1-ETO[224]. The remaining wild-type RUNX1 activity is seemingly subhaploid but absolutely critical, highlighting the dose-dependent actions of RUNX1 (HITF principle 2).

The final translocation disrupting RUNX1 dose is inv(16)(p13q22), which rearranges *CBFB*, the gene encoding the obligate RUNX family TF cofactor, CBFβ. Inv(16) fuses *CBFB* to *MYH11* (SMMHC); the resultant CBFβ-SMMHC oncoprotein retains the ability to bind RUNX1. CBFβ-SMMHC actually has greater affinity for RUNX1 than CBFβ alone and co-opts RUNX1 to drive the leukemogenic program[225]. Wild-type RUNX1 activity is diminished but not abolished as a result. In fact, some residual RUNX1 activity is essential for the survival of this AML[225-227]. Further suggesting reduced RUNX1 dose as important in this disease, monoallelic loss of *Runx1* in mice phenocopies inv(16) AML with respect to granulocyte maturation failure[226]. In addition to disrupting RUNX1 function, CBFβ-SMMHC also reduces TP53 dose by indirectly mediating the inappropriate deacetylation and inactivation of TP53[228]. Thus, the mechanism of pathogenesis of inv(16) AML is the result of the oncogenic fusion protein compounded by the interference with two dose-dependent TFs, TP53, and RUNX1.

Mutations

Mutations are not created equal and can have widely varying effects on protein function, so we must consider the details of individual mutations to understand how they affect the dose of a given TF.

RUNX1 is among the most frequently mutated TFs in AML and MDS[104]. Patients with familial platelet disorder have inherited loss-of-function mutations in *RUNX1*, which results in

cell-intrinsic impairment of megakaryocyte differentiation and an increased risk of myeloid malignancy[229-231]. Expectedly, the loss of one copy of *RUNX1* in induced pluripotent stem cells (iPSCs) results in compromised megakaryocyte differentiation[229]. *RUNX1* mutations exist on a functional spectrum. N-terminal mutations result in pure loss of function, as they result in truncations or mutations that disrupt the ability to bind DNA or CBFβ. Mutations in the DNA binding domain (RHD) are weakly dominant negative. Finally, mutations in the C-terminal transactivation domain are strongly dominant negative (80% reduction in RUNX1 activity)[232-235]. The precise mechanism of strong dominant negative mutations is unclear and could be multifactorial. Underscoring the exquisite dose sensitivity of RUNX1, loss of one copy of *RUNX1* or a weak dominant negative mutation (R174Q) are both associated with thrombocytopenia and decreased megakaryocyte differentiation in iPSC culture. However, only the strong dominant negative mutation results in an expansion of myeloid progenitors and increased genomic instability[236]. Further emphasizing the dose-dependent nature of RUNX1, myeloid malignancies with lower RUNX1 activity have a higher likelihood of secondary leukemic progression[237]. The spectrum of RUNX1 activity and associated phenotypes capture the essence of HITF principle 2.

GATA1 is an X-linked gene, so while it cannot be considered classically haploinsufficient, it still exhibits dose-dependent effects. This characteristic is featured in the GATA1^{knockdown} mouse, which displays impaired erythropoiesis and megakaryopoiesis[238]. Total GATA1 ablation causes erythroid progenitor apoptosis, while a partial decrease promotes proliferation and a leukemia-like disease, underscoring the dose-dependent nature of GATA1[239]. A second example of this is apparent in Down syndrome. Of infants with Down syndrome, characterized by trisomy 21, 10% develop a self-limiting leukemia-like disease termed transient abnormal myelopoiesis (TAM).

Twenty percent of TAM cases will evolve to acute megakaryoblastic leukemia (AMKL). Virtually every case of TAM and AMKL has a mutation in *GATA1*[240]. *GATA1* mutations in the context of trisomy 21 are sufficient for TAM but not for AMKL, in which other mutations are acquired. Generally, the mutations in *GATA1* are 5′ nonsense mutations, allowing protein synthesis to restart at a 3′ start codon to create a shorter isoform, GATA1s[241]. The GATA1s isoform lacks part or all of the N-terminal transactivation domain, reducing the binding of an important cofactor, FOG1, ultimately leading to a TF with less transactivation capability. In this way, incomplete loss GATA1 activity is pathogenic in TAM and AMKL.

A second GATA family TF member, GATA2, is also a HITF. Loss of one copy of murine *Gata2* results in reduced HSC numbers, increased HSC proliferation, and impaired self-renewal upon transplant[242,243]. Conversely, overexpression of GATA2 results in a differentiation block, suggesting that GATA2 is also triplosensitive[244]. The vast majority of mutations in *GATA2* are predicted to generate a loss-of-function allele, via either nonsense or missense mutations in one of the two zinc finger domains[245,246]. Mutations in the first GATA2 zinc finger domain create a loss-of-function allele due to the inability of the mutant protein to bind DNA. These mutations are associated with CEPBα^{DM} in myeloid neoplasms[246]. In a mouse model of CEPBα^{DM} and a first zinc finger domain mutation of *Gata2*, the *Gata2* mutation causes decreased accessibility of myeloid TF binding sites, while both TFs increase erythroid TF expression[247]. This mechanistic study provides an explanation for the bilineage leukemia that arises from synergistic and concurrent mutations in *CEPBA* and *GATA2*.

We return to consider how TP53 is different than the other TFs discussed in this review. *TP53* mutations are enriched in older patients, t-MNs, and complex karyotype malignancies[106]. In myeloid neoplasms, loss-of-function, dominant negative, and gain-of-function *TP53* mutations

have been described[201,248,249]. All of these mutations result in differing degrees of TP53 impairment. There is an inverse relationship between TP53 activity and the ensuing fitness advantage in malignancies. Cells with the least amount of TP53 activity have the greatest fitness advantage, and cells can accordingly gain a further advantage by removing residual TP53 activity. For instance, of MDS cases with *TP53* aberrations, one-third are monoallelic and two-thirds are biallelic[200]. Monoallelic *TP53* mutations do not connote the poor prognosis that biallelic mutations carry. Some MDS patients with monoallelic *TP53* mutations developed secondary AML with biallelic *TP53* mutations, a transformation that correlates additional *TP53* loss with disease progression. With regards to the clonal structure, patients with monoallelic *TP53* mutations have greater clonal diversity than patients with biallelic *TP53* mutations, and biallelic *TP53* mutations are always in the dominant clone, further indicating the relative advantage of leukemic cells with biallelic *TP53* mutations (HITF principle 1). These results buttress the notion that monoallelic *TP53* loss is sufficient to drive MDS but is not synonymous with biallelic *TP53* disease.

Mutation of Another Gene

While rarely mutated itself, the ETS family TF PU.1 is often a concomitant casualty of other genetic aberrations. The centrality of PU.1 in the myeloid transcriptional network means that many common genetic aberrations in AML affect PU.1 dose. Functionally, PU.1 is required for HSCs to generate myeloid progenitors and for myeloid differentiation beyond the granulocyte monocyte progenitor stage[250]. In addition to orchestrating differentiation, PU.1 is a regulator of HSC proliferation and protein synthesis. Consequently, loss of PU.1 increases HSC cycling and decreases self-renewal[164,251,252]. The mouse models of PU.1 illustrate the dose-dependent aspects of this TF. Heterozygous PU.1-deficient mice have various hematopoietic defects but do not

spontaneously develop AML[253]. However, an 80% reduction in PU.1 by homozygous deletion of an enhancer upstream of *PU.1* results in a transplantable AML[23]. Interestingly, heterozygous deletion of this enhancer produces a 35% reduction in PU.1 and a preleukemic state, which progresses to an overt AML on a DNA-repair-deficient background[22]. These mouse models demonstrate that varying PU.1 doses from slight to more severe reductions are associated with unique leukemic and preleukemic states (HITF principle 2). In fact, low levels of PU.1 activity are required to form competent leukemic stem cells, as mutation of the RUNX1 binding sites in a *Pu.1* enhancer results in a severe decrease in leukemic stem cell activity and delayed onset of RUNX1-ETO-driven leukemia[167].

Many paths to myeloid neoplasms pass through PU.1, given its requirement in myeloid differentiation and cell cycle control. RUNX1 normally regulates PU.1 expression in a dynamic and lineage-specific manner, thus alteration of RUNX1 dose has repercussions for PU.1 levels[254]. RUNX1 fusion proteins, for instance, not only disrupt the normal function of RUNX1, as discussed above, but also directly reduce PU.1 expression. There are two active ways that RUNX1 fusion proteins reduce PU.1 dose: exploitation of an antisense promoter normally used to downregulate PU.1 during T cell development, and inactivation of a long noncoding RNA that normally facilitates PU.1 expression[255,256]. In addition, some N-terminal *RUNX1* mutations abrogate RUNX1 interaction with and recruitment of the activating H3K4 methyltransferase MLL (*KMT2A*) to the PU.1 promoter[257]. Moving beyond RUNX1, PU.1 levels are low in APL and increase after ATRA-induced differentiation; in fact, overexpression of PU.1 is sufficient to promote differentiation in this model[258]. Mutations in *CPEBA* result in the downregulation of PU.1, and the FLT3 internal tandem duplication mutation results in the downregulation of PU.1 as well as the downregulation of CEPBα[167,259-261]. Although different groups have found

conflicting results, mutation of the nucleolar protein NPM1 results in a strong nuclear export signal and can drag PU.1 along with it into the cytoplasm[262,263]. Finally, a reduction in PU.1 expression accelerates the development of AML when combined with loss of the DNA demethylase TET2. The residual PU.1 is not sufficient to overcome the aberrant methylation at PU.1 binding sites at critical myeloid enhancers[264]. The acquisition of a mutation in a TET2-deficient cell that results in the loss of PU.1 expression could rapidly shift the transcriptional balance toward myeloid differentiation and cell cycle dysregulation[265,266]. Overall, reduction in PU.1 dose is a common consequence of mutations in myeloid genes, underscoring the fundamental role of a stable myeloid transcriptional network in preventing leukemogenesis.

Change in Isoform Balance

TFs often have multiple isoforms—products with different functions that arise from alternative splicing or posttranslational regulation. CEBPα is one such TF with multiple isoforms containing distinct activities. CEBPα is required for granulocyte differentiation, and, apropos of the corollary to HITF principle 1, homozygous *Cebpa* knockout mice do not develop AML, even in the presence of a strong oncogene[170,267,268]. The full-length CEBPα p42 protein contains a C-terminal basic leucine zipper DNA binding domain and an N-terminal transactivation domain, the latter of which is omitted in the shorter p30 isoform[269]. Both isoforms are normally present, and their abundance is regulated translationally[270]. As a result of this unique arrangement, different CEPBα mutations have distinct results. N-terminal mutations result in the production of p30 only, and C-terminal mutations disrupt the ability of CEBPα to bind DNA or homodimerize. Most tumors with *CEPBA* mutations have biallelic mutations of both types[271]. The combination of these two mutations shifts the balance of isoforms toward p30 homodimers and away from normally predominant p42 homodimers[269]. The shift to entirely p30 homodimers resulting from

two N-terminal mutations is sufficient for leukemogenesis[272]. However, the combination of an N-terminal mutation and a C-terminal mutation results in leukemia with faster onset, indicating C-terminal mutations might be more potent in suppressing the myeloid transcriptional network[273]. These results have several implications. Importantly, biallelic mutations do not recapitulate the CEPBα null state, suggesting that there is a minimal dose of CEPBα that is required to form leukemic cells. Secondly, the mechanism by which leukemic cells alter the dose of CEPBα is partly due to an N-terminal mutation shifting the isoform balance toward the production of p30 homodimers. This later mechanism is exacerbated by further impairing p30 CEPBα activity with a C-terminal mutation. Of note, CEPBα isoforms can be deregulated through less direct means as well, as occurs with del(5q)[274].

While *RUNX1* can be inactivated, it can also be amplified by somatic or germline gains of chromosome 21, on which *RUNX1* is encoded. In TAM, a neoplasm dependent on RUNX1, the apparent gain in *RUNX1* expression through trisomy 21 actually translates into a loss of RUNX1 activity through dysregulation of the three RUNX1 isoforms. RUNX1b and RUNX1c are similar, full-length isoforms; however, RUNX1a is a primate-specific, shorter isoform that lacks the C-terminal transactivation domain[215]. Compared with the other isoforms, RUNX1a has a higher affinity for DNA and is thought to act in a dominant negative manner, similar to the impact of mutations in the RUNX1 transactivation domain[275]. RUNX1a expression is elevated in TAM, likely due to the inability of the aberrant GATA1s isoform to repress RUNX1a, ultimately resulting in a decrease in wild-type RUNX1 activity[275]. In addition, RUNX1a recruits other transcriptional regulators to inhibit megakaryocyte terminal differentiation and promote the proliferation of megakaryocytic blasts[275]. Gain of chromosome 21 is not the only way that RUNX1 isoforms can become dysregulated. A specific mutation in a splicing factor, $SRSF2^{P95H}$,

also promotes the expression of RUNX1a[276]. Hence, like CEPBα, RUNX1 dose can also be altered by a change in the balance of isoforms to promote abnormal proliferation and differentiation.

Epigenetic Silencing

DNA methylation is one of many ways to inhibit the expression of genes, including HITFs. AML co-opts the DNA methylation machinery to augment tumorigenesis in concert with driver mutations, and, therefore, AML can be subdivided on the basis of DNA methylation signatures[277]. In an unbiased look at diverse AMLs, GATA2 showed allele-specific expression in 60% of patients, indicating that one allele had been epigenetically silenced[278]. This is particularly pronounced in the case of CEPBα^{DM} AML, where 95% of cases show allele-specific expression of GATA2. In leukemias with a concomitant *GATA2* mutation, the wild-type allele is more likely to be silenced (HITF principle 1). Given the synergistic effects of *CEPBα^{DM}* and *GATA2* mutation, epigenetic silencing is another route to decrease wild-type GATA2 and amplify the potency of a *GATA2* mutation. Epigenetic silencing as a mechanism of decreasing HITF dose is not limited to *GATA2* and applies to other HITFs such as *CPEBA*[279].

Therapeutic Options

Haploinsufficiency opens at least two therapeutic windows: decreasing the activity of the residual protein to kill leukemic cells reliant on that protein or boosting the activity of the remaining wild-type protein to restore differentiation. These options can potentially avoid the toxic effects of nonspecific chemotherapy by specifically targeting cells containing the mutant allele. In an example of the first approach, knocking down the wild-type RUNX1 allele or inhibiting CBFβ binding impairs AML growth in *RUNX1* mutant xenografted mice[280]. This example ought to inspire further investigation into molecules that can be given therapeutically to

accomplish this goal of removing residual HITF activity in mutant cells.

Using a tetracycline-inducible (and reversible) small hairpin RNA transgene, knockdown of CUX1 expression results in de novo MDS/MPN or an alkylating agent–induced t-MN in mice[78,191]. In diseased mice, tetracycline withdrawal restores CUX1 levels and normalizes blood parameters. This example provides a genetic proof of principle for augmenting the wild-type allele as a method to treat disease driven by a HITF. Further research is warranted to ascertain methods that will allow the restoration of normal HITF protein expression.

Concordant with HITF principle 3, focal deletions and aneuploidies provide an additional opportunity to exploit the essentiality of concurrently deleted genes. Also described as collateral lethality, this concept assumes a therapeutic window for targeting cells that are haploinsufficient for an essential gene. Lenalidomide is a successful example of this. Lenalidomide and its derivatives are highly efficacious in the treatment of MDS with isolated del(5q) because lenalidomide mediates the degradation of CK1α, an essential protein encoded on 5q. Reducing CK1α levels to null results in TP53-mediated apoptosis[281]. This same strategy could be applied to other recurrent chromosomal deletions in myeloid malignancies.

Conclusion

We hope to have conveyed the importance of dose in a broad sense. Traditionally, dose has been thought of as diploid, haploid, and null, but many examples exist showing the pathogenic importance of non-Mendelian gene levels. In addition, there are ways to affect the dose of a HITF beyond its deletion or mutation. The myeloid cell identity is dependent on the complex transcriptional network brought about by the TFs discussed in this review, so it follows that changing the dose of one TF has secondary effects on the dose of other TFs. This balance brings added complexity to considering the genetic aberrations of a particular tumor. For example, a

patient presenting with a *TET2* mutation in the form of clonal hematopoiesis is at risk of having their disease transform into a myeloid neoplasm if they acquire another mutation that changes the dose of PU.1[264]. A tumor with a *FLT3*^{ITD} mutation also has an altered dose of CEPBα, and this is likely a source of the myeloid differentiation block associated with the disease. While myeloid cancers typically have fewer mutations than solid tumors, the recurrent mutations have multiple effects on the transcriptional network. A better understanding of this complexity will clarify why some patients respond to a given therapy and will open up treatment strategies for those who do not.

Summary Points, Future Issues, Definitions

Summary Points

1. A large portion of the coding genome is under genetic constraint due to haploinsufficiency, which includes almost all transcription factors (TFs) and tumor suppressor genes.
2. TFs function to influence the contact time between enhancers and target promoters to regulate the expression of target genes.
3. Loss of one allele of a gene can produce distinct effects that can be augmented by the loss of the second allele.
4. Given that myeloid neoplasms arise in cells with myeloid identity, TFs involved in myeloid cell identity are rarely mutated to zero residual activity. This is in contrast to TP53, the function of which is often completely lost in malignant cells.
5. There are TF doses not defined by strict Mendelian genetics that are consequential for the development of the preleukemic and leukemic state.

6. When one haploinsufficient transcription factor is altered, additional TFs are impacted due to the interwoven nature of the myeloid transcriptional network.

Future Issues

1. What is the role of transcription factor (TF) dose in normal hematopoiesis?
2. How do TF dynamics change for hematopoietic TFs in the haploinsufficient state?
3. How do mutations in or loss of the transactivation domain and/or intrinsically disordered region (IDR) change TF dynamics and impact haploinsufficient phenotypes?
4. How do mutations in IDRs of TFs change their interaction partners and ability to interact with promoters and enhancers?
5. How do cancer cells optimize the levels of a given TF to best drive the oncogenic state, and how is this altered by environmental stresses such as inflammation and chemotherapeutic pressure?
6. How do genes concomitantly lost with haploinsufficient transcription factors (HITFs) modulate the disease phenotype in the context of a contiguous gene syndrome?
7. What mechanisms of normal TF regulation can be exploited to therapeutically manipulate the dose of a HITF to treat disease?

Definitions

Haploinsufficient: a dominant phenotype in a diploid organism resulting from a loss-of-function event of only one allele

Loss-of-function: a mutation in which the mutant allele lacks some function of the wild-type allele

Genetic constraint: the degree to which DNA sequence is evolutionarily conserved and variations are selected against

Triplosensitive: a dominant phenotype arising from the additional copy of a gene in a diploid organism

Dose-sensitive: a trait arising from the gain or loss of expression of a gene

Myelodysplastic syndrome (MDS): hematologic neoplasms with clonal hematopoiesis, cytopenia(s), and abnormal cellular morphology that can be defined by morphology or recurrent genetic abnormalities

Enhancer: a cis-regulatory genomic element that influences the transcription of a target gene

Transcriptional noise: variations in transcription of a given gene between cells of the same type arising from the bursting and probabilistic nature of transcription initiation

Transcription condensates: assemblies of high local concentrations of the machinery required to transcribe a gene (TFs, transcription coregulators, and RNA Pol II) and their subsequent dissolution

Acute myeloid leukemia (AML): the clonal expansion of malignant HSPCs defined by genetic aberration or extent of differentiation of blasts

Clonal hematopoiesis: the overrepresentation of mature blood progeny derived from at least one mutant hematopoietic stem cell clone and the absence of detectable hematological abnormality

Contiguous gene syndrome (CGS): a disorder arising from the loss of multiple neighboring genes in a single genetic event resulting in a phenotype that represents the combinatorial haploinsufficiency of those genes

Myeloproliferative neoplasms (MPNs): clonal disorders of HSPCs resulting in overproduction of terminally differentiated myeloid cells

Essential gene: a gene for which a loss-of-function event results in severely impaired viability or fitness for a cell or organism

Gain-of-function: a mutation in which the mutant allele possesses a de novo function absent from the wild-type allele

Dominant negative: a mutation in which the mutant allele negatively affects the wild-type allele, producing a consequence greater than a loss-of-function variant in one allele alone

Results: CUX1 regulates hematopoietic stem cell fate through regulation of innate immune activation

Abstract

Hematopoietic stem cells (HSCs) possess properties of both self-renewal and multipotent differentiation; however, the regulation of these opposing properties remains elusive. Using a novel reporter mouse, we demonstrate the dose-dependent role of the transcription factor, CUX1, in HSC fate. As CUX1 levels naturally rise, HSCs gradually lose repopulating capacity. CUX1 knockdown reverses this process and reverts HSCs to an earlier cell state—restoring the transcriptome, chromatin accessibility, repopulating capacity, and self-renewal ability of HSCs with normally low CUX1 levels. In the absence of CUX1, endogenous retroelements (EREs) are aberrantly expressed, leading to a cell-intrinsic inflammatory response, which contributes to CUX1-dependent HSC activity. Likewise, loss of the transcription factor IRF1 attenuates CUX1-mediated HSC reprogramming. This mechanism could be relevant to leukemogenesis as acute myeloid leukemias with CUX1 deletions upregulate EREs and the interferon response. Our findings attribute CUX1 and innate immune activation to the complex molecular underpinnings of HSC fate decisions.

Introduction

Hematopoietic stem cells (HSCs) are responsible for generating the full complement of blood cells for the longevity of the organism. One current paradigm in hematopoiesis posits that differentiation is a continuous process where HSCs gradually adopt the transcriptome of terminally differentiated cells and lose the ability to self-renew over time[51,52,64,65]. This model accepts that the HSC compartment is functionally and transcriptionally

heterogeneous[55,58,59,62,63,66-70]. Given that the success of bone marrow transplantation, the future of HSC-based cellular therapies, and the devastation of myeloid malignancies depend on understanding heterogeneity in HSC potential, great effort has been made to elucidate the factors associated with self-renewal[282]. A non-exhaustive list of these factors includes oxygen content[283], metabolites including retinoic acid[284,285] and ascorbate[286], mitochondrial dynamics[287], fatty acid oxidation[288], protein synthesis rates[289,290], and lysosomal activity[291,292], not to mention the transcription regulators such as DNMT3A[293], MLLT3[294], and TET2[295], as well as transcription factors including EVI1/MECOM[296,297], GFI1[298], and HLF[299]. Despite all that is known, a molecular understanding of HSC self-renewal incorporating the HSC heterogeneity has remained elusive.

The haploinsufficient transcription factor CUX1 plays a role in the somatic stem cells of multiple tissues regulating diverse processes[78,300-303]. CUX1 is mutated or deleted in the context of monosomy 7 (-7) or del(7q) in myeloid disorders[76,98]. Loss of CUX1 occurs in up to 15% of acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS), which is independently associated with poor prognosis[75,93,94,98,103,104,304]. Not only is this homeodomain-containing transcription factor haploinsufficient, but different levels of CUX1 can give rise to different cellular phenotypes. In the development of pyramidal neurons and the respiratory system of *D. melanogaster*, different doses of the CUX1 homologue, cut, are associated with different developmental fates[187,188]. CUX1 plays a dose-dependent role in murine hematopoiesis as well. Progressively lower doses of CUX1 result in trilineage dysplasia and a more severe MDS, ultimately resulting in a lethal disease at the lowest level of CUX1[78]. Furthermore, CUX1 conveys dose-dependent activity in an analog manner, meaning that the level of CUX1 and the resulting transcriptional response are directly related[189,305]. Our objective was to elucidate how

the CUX1 integrates into the continuous process of hematopoiesis. Here, we demonstrate that the dose-dependent activity of CUX1 creates heterogeneity within the HSC compartment by controlling the extent of innate immune activation and, ultimately, HSC fate and self-renewal.

Results

CUX1 protein levels predict HSC activity

While not currently well-studied, there are examples of transcription factor levels influencing cell-intrinsic heterogeneity in HSC function.[305] To examine the role of CUX1 protein levels in HSC function, we generated a reporter mouse in which mCherry is fused in-frame with the 3' end of the endogenous locus of *Cux1* (Figure 6A). The resulting fusion protein has a normal half-life and is appropriately located in the nucleus (Figure 5A-C). In contrast to CUX1 knockdown or hypomorph mice, which have numerous phenotypic abnormalities in multiple tissues and occasional postnatal mortality[78,300-302], CUX1^{mCherry} mice did not exhibit any gross or measurable differences, including in body weight, spleen mass, or liver mass compared to wildtype littermates (Figure 5D). Complete blood counts revealed no differences in peripheral blood parameters (Figure 5E-F). A survey of mature, stem, and progenitor compartments revealed no difference in cell number or relative cell abundance compared to wildtype littermates (Figure 5G-J). The absence of gross or hematopoietic abnormality at 6-8 weeks of age indicates this mouse model is suitable for studying the role of CUX1 in murine hematopoiesis.

First, we examined CUX1 protein levels in hematopoietic cells. Within hematopoietic stem and progenitor cells (HSPCs), CUX1 levels are lowest in LT-HSCs, generally increase as differentiation proceeds to multipotent progenitors (MPPs), and ultimately decrease again with the lowest expression observed in terminally differentiated splenic compartments (Figure 6B).

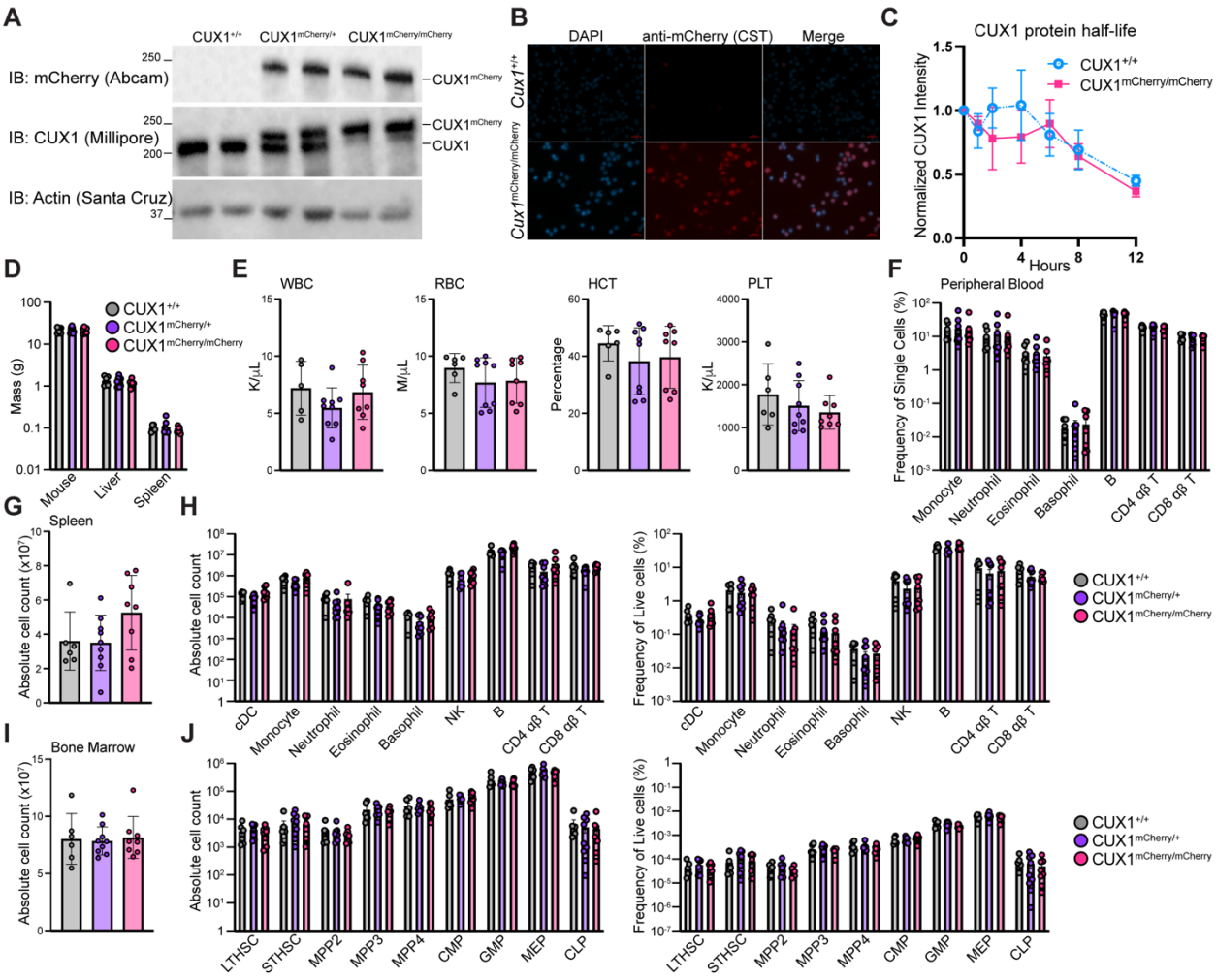

Figure 5. *Cux1*^{mCherry} mice exhibit no gross or hematopoietic abnormalities

(A) Western blot of thymocytes isolated from one male (left) and one female (right) *Cux1*$^{+/+}$, *Cux1*$^{mCherry/+}$, and *Cux1*$^{mCherry/mCherry}$ mouse. One representative blot of n=2 biological replicates.
(B) Immunofluorescence of mCherry indicating the CUX1-mCherry fusion protein produces an expected nuclear stain.
(C) Quantification of CUX1 protein from Western Blot after time course incubation with 50mg/mL cycloheximide, indicating the C-terminal mCherry tag does not affect the half-life of CUX1 (n=3 biological replicates).
(D-J) Comparison of *Cux1*$^{+/+}$, *Cux1*$^{mCherry/+}$, and *Cux1*$^{mCherry/mCherry}$ mice (n=5-9). One-way ANOVA with Tukey's test for multiple comparisons. Mean +/- SD. (D) Mass of the mouse, spleen, and liver. (E) WBC, RBC, and platelet (PLT) count and hematocrit by complete blood count. (F) Frequency of population measured by peripheral blood flow cytometry. (G) Total spleen cell count (H) spleen cell count and frequency (from G) by cell type. (I) Total bone marrow cell count (J) bone marrow hematopoietic stem and progenitor cells count and frequency (from I).

The dynamic expression of CUX1 suggests that different levels of CUX1 could be driving differentiation and transcriptomic heterogeneity.

To integrate CUX1 protein levels with transcriptomic data, we generated an additional reporter mouse in which the AviTag™ [306] is fused to the C-terminus of CUX1 along with the ectodomain of human NGFR (hNGFR) connected by the self-cleaving peptide P2A (Figure 6A). Thus, hNGFR protein is produced proportionally with CUX1 protein, enabling us to infer CUX1 levels via hNGFR. Given the dynamic expression of CUX1 protein, we performed single-cell RNA sequencing to capture transcriptomic heterogeneity from sorted murine HSPCs relative to CUX1 protein using an anti-hNGFR antibody-DNA conjugate (Figure 7A-E). CUX1 protein levels (Surface CUX1) exhibited similar dynamic expression as we observed by flow cytometry (Figure 6D). Monocle3 and Palantir[307] pseudotime trajectory inference corroborated that CUX1 protein levels broadly increase as cells differentiate (Figure 6E-F). We noted the greatest intra-cluster heterogeneity in CUX1 levels in the LT-HSC clusters, which was confirmed by flow cytometry as the variance in CUX1 expression is highest in the LT-HSC compartment (Figure 6D, 7F). Given the numerous characterizations of CUX1 as a dose-dependent transcription factor and the high variance within seemingly similar cells, we hypothesized that CUX1 plays a dose-dependent role in regulating HSC fate.

To hone our hypothesis further, we aimed to characterize CUX1 levels relative to the expression of immunophenotypic markers heterogeneously expressed in the LT-HSC compartment. Given the continuum of CUX1 expression, we segmented CUX1 expression in immunophenotypic LT-HSCs into $CUX1^{Dim}$ (lowest 15% of mCherry fluorescence), $CUX1^{Int}$ (middle 15% of mCherry fluorescence), and $CUX1^{Bright}$ (highest 15% of mCherry fluorescence) (Figure 8A). High expression of Sca1[50,308], CD150[42,67], and CD201[309] have been associated with

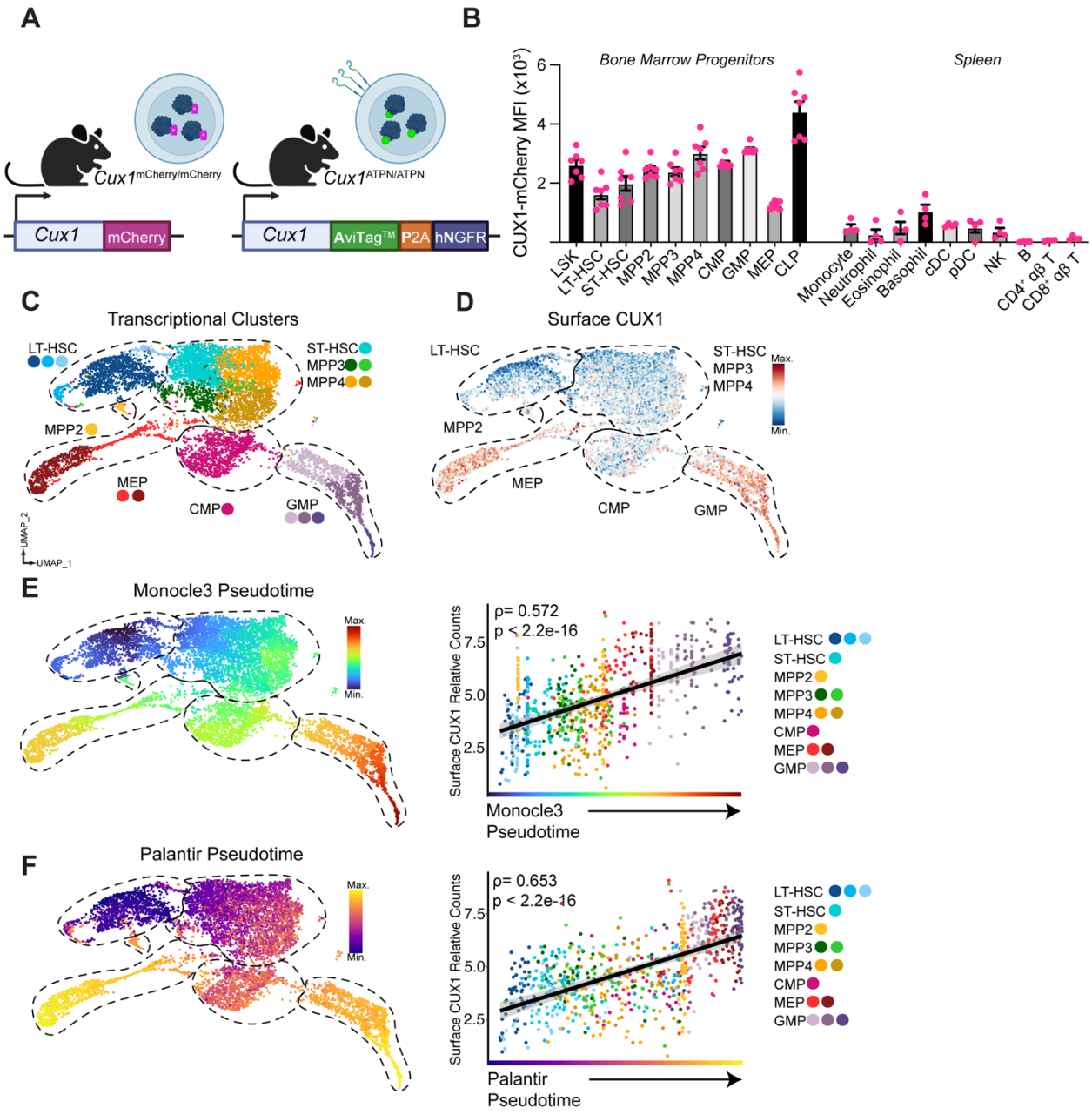

Figure 6. CUX1 levels increase during hematopoietic differentiation

(A) Description of CUX1 transgenic reporter mice.
(B) CUX1-mCherry MFI in hematopoietic progenitor compartments (left, n=7) and mature spleen compartments (right, n=4).
(C-F) scRNA-seq (10X Chromium) of sorted LT-HSCs and progenitors (n=8621) from *Cux1*^{ATPN} mouse (n=4). (D) Relative normalized counts of Surface CUX1 inferred from hNGRF antibody-DNA tag. (E) Monocle3 pseudotime and (F) Palantir[307] pseudotime rank and the Pearson correlations with Suface CUX1 expression. Error bars represent 99.99% confidence level.

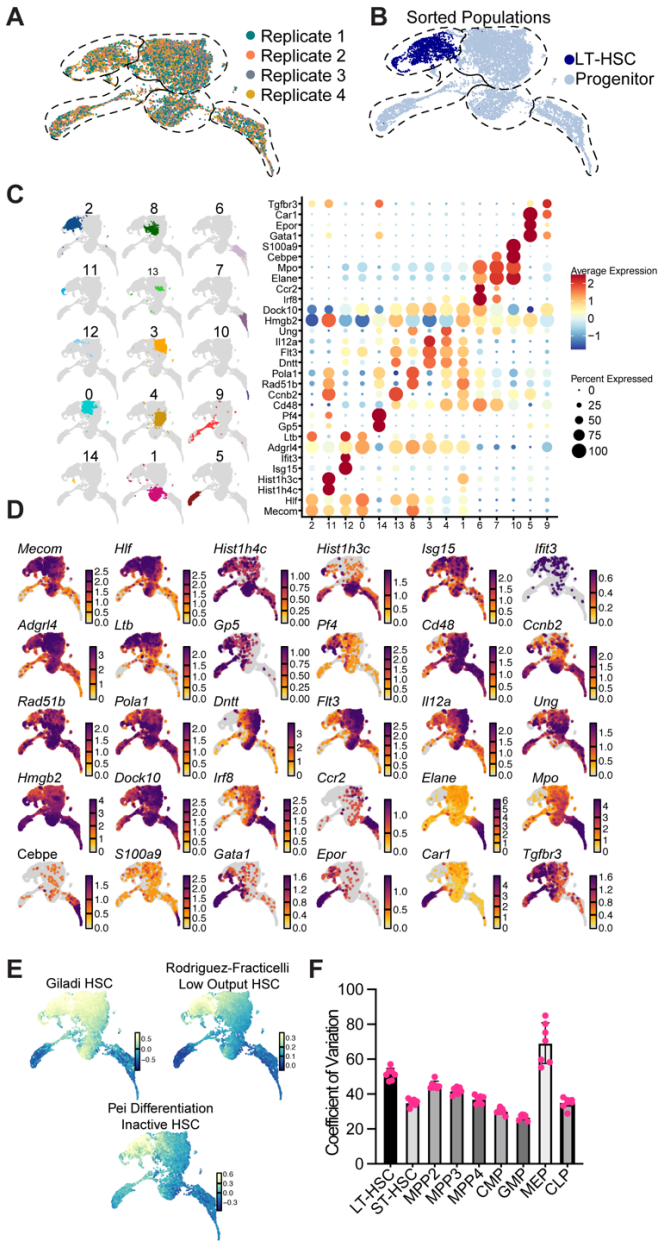

Figure 7. CUX1 levels exhibit high variance in LT-HSCs

(A-E) 10X Chromium single cell RNA-sequencing on sorted LT-HSCs and progenitor cells from four different mice. (A) UMAP of each replicate distinguished by cell multiplexing oligo. (B) UMAP of each sorted population. LT-HSCs were sorted as LSK CD135^{-}, CD48^{-}, CD150^{+}. Progenitors were sorted as a pooled population of ST-HSCs (LSK, CD135^{-}, CD48^{-}, CD150^{-}), MPPs (LSK, CD48^{+}), and myeloid progenitors (Lin^{-}, CD117^{High}, Sca1^{-}). Progenitors were pooled to overrepresent more low-frequency cell populations. (C-D) Dot plot and UMAP of two marker genes for each cluster. Marker genes were determined by pseudobulk analysis and the seurat feature FindMarkers. (O) UMAP of HSC gene signatures.
(P) Coefficient of variation of CUX1 levels in each HSPC compartment measured by flow cytometry (corresponding to 6B). Mean +/- SD.

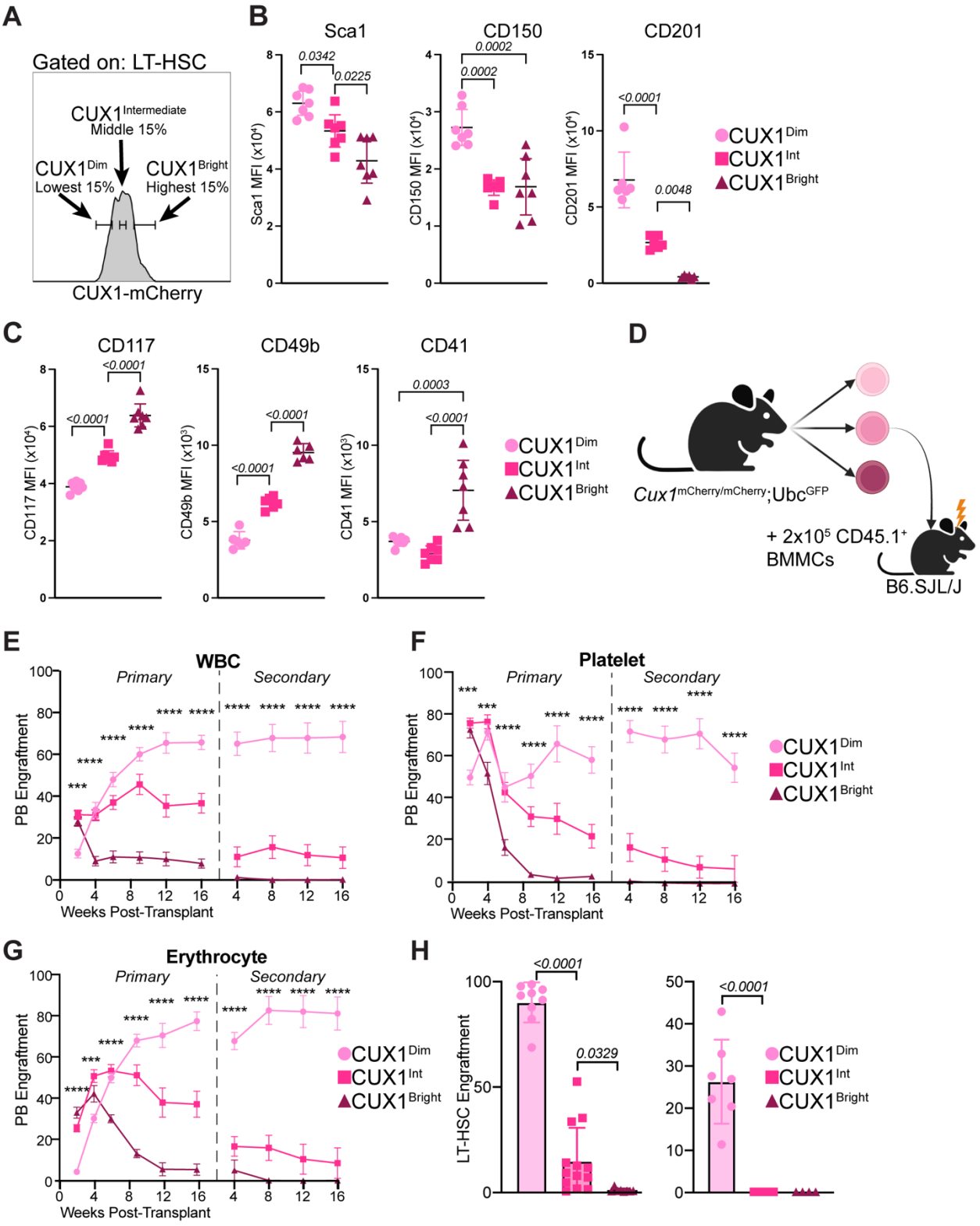

Figure 8. CUX1 levels predict HSC activity

Figure 8, continued
(A) Gating strategy for segmenting LT-HSCs CUX1^{mCherry} mice. CUX1^{Dim} LT-HSCs are defined as the lowest 15% of mCherry fluorescence, CUX1^{Int} LT-HSCs are defined as the middle 15% of mCherry fluorescence, and CUX1^{Bright} LT-HSCs are defined as the top 15% of mCherry fluorescence.
(B-C) Correlation of CUX1-mCherry with heterogeneously expressed surface markers (MFI) in LT-HSCs by flow cytometry (n=6-7). ANOVA and Tukey's test. Mean +/- 95% CI.
(D) Transplant strategy for 100 segmented LT-HSCs with 200,000 competitor bone marrow mononuclear cells (BMMCs). Secondary transplants were performed by pooling the bone marrow of two primary recipients and transplanting 1,000,000 cells into secondary recipients.
(E-G) Analysis of peripheral blood (PB) engraftment of donor cells (CD45.2^{+} or GFP^{+}) in the competitive primary and secondary transplant of CUX1^{Dim} (1°n=26, 2°n=14), CUX1^{Int,} (1°n=30, 2°n=14), or CUX1^{Bright} (1°n=26, 2°n=6) LT-HSCs. Two-way ANOVA with Tukey's test. Mean +/- SEM.
(N) Flow cytometry analysis of donor LT-HSC engraftment after primary (left, n=8-13) and secondary (right, n=4-7) transplant from (J). Frequency of donor LT-HSC shown. ANOVA and Tukey's. Mean +/- SD.

high stem cell activity, and these markers were most enriched on CUX1^{Dim} HSCs (Figure 8B). Conversely, high expression of cKit[310], CD49b[311], and CD41[69] have been associated with impaired stem cell activity, and these markers were most enriched in CUX1^{Bright} HSCs (Figure 8C).

We tested the refined hypothesis that CUX1 levels increase as stem activity decreases by transplanting 100 LT-HSCs segmented on CUX1 expression into lethally irradiated recipients (Figure 8D). To track the engraftment of erythrocytes and platelets, we crossed the CUX1^{mCherry} reporter mouse to a ubiquitous GFP reporter mouse[312]. CUX1^{Bright} HSCs gave rise to short-term, multilineage engraftment that extinguished upon secondary transplantation. In contrast, CUX1^{Dim} LT-HSCs gave rise to long-term engraftment and multilineage chimerism in both primary and secondary recipients. CUX1^{Int} LT-HSCs exhibited an intermediate phenotype relative to CUX1^{Dim} and CUX1^{Bright} HSCs (Figure 8D-G, 9A). The levels of peripheral blood engraftment were mirrored in the number of residual, donor-derived HSCs in primary and secondary recipients, confirming that HSCs with lower levels of CUX1 have greater repopulation capacity and potential for self-renewal (Figure 8H). CUX1^{Dim} HSCs embodied the paradigm of the platelet-biased HSC[58,68] because they were initially platelet-biased and sustained high platelet engraftment through two transplants (Figure 8F, 9B). In addition, we observed a previously described divergence between erythroid fates and self-renewal where CUX1^{Bright} HSCs are erythroid-biased and even upregulate GATA1 but lack long-term self-renewal (Figure 8G, 9B, 12A). All in all, for the first time, we attribute cell-intrinsic HSC heterogeneity to the levels of a transcription factor.

We next sought to understand the poor reconstitution capacity of CUX1^{Bright} HSCs. The lack of long-term reconstitution of CUX1^{Bright} LT-HSCs was not explained by a defect in bone

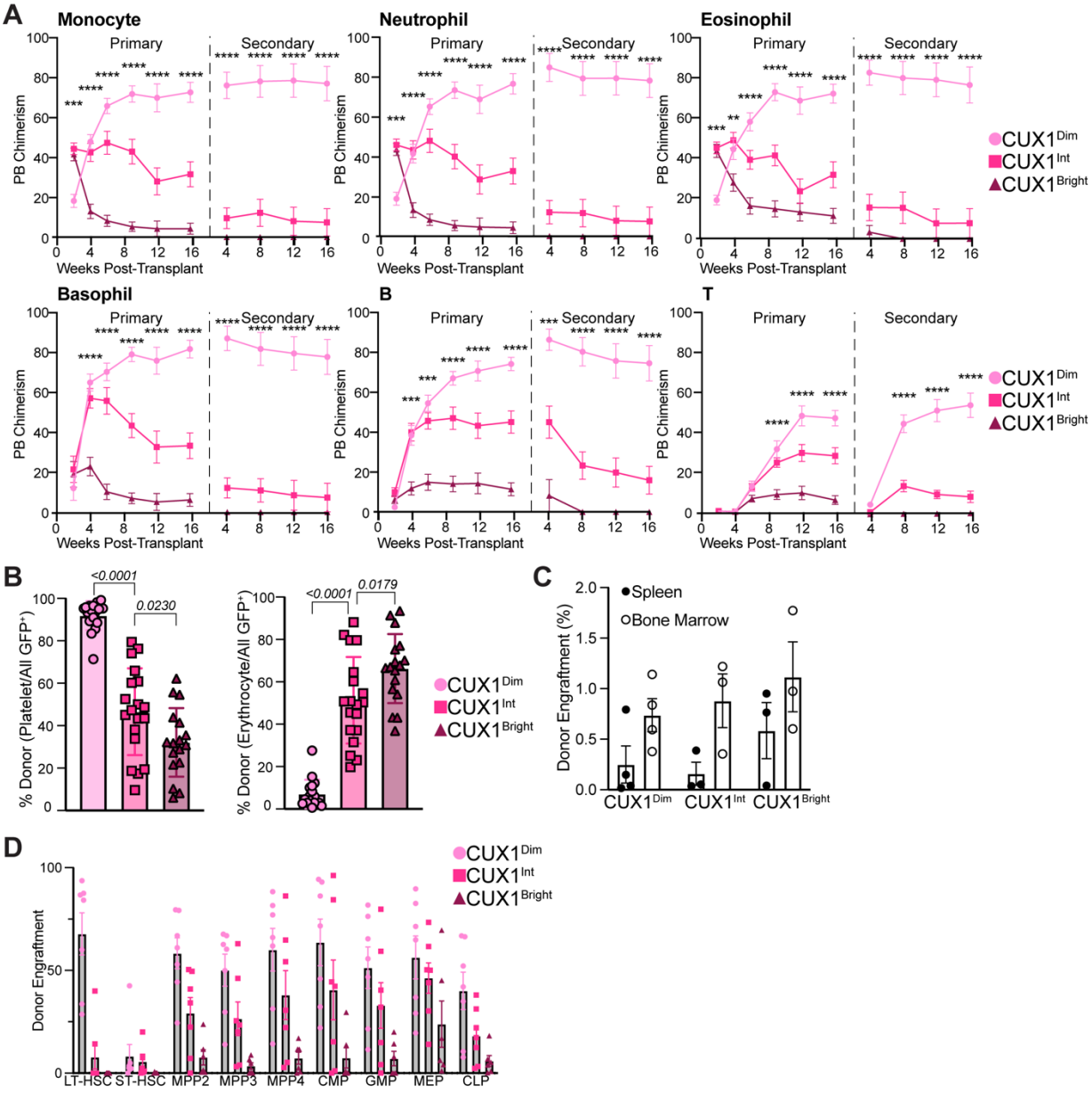

Figure 9. CUX1^{Dim} LT-HSCs exhibit initial platelet bias

(A) Corresponding to Figure 1J-K. Donor engraftment measured by flow cytometry. Two-way ANOVA with Tukey's test. Mean +/- SEM.
(B) Fraction of donor-derived cells (GFP$^+$) that were erythrocytes (left) and platelets (right) at two weeks post-transplant. ANOVA and Tukey's test. Mean +/- 95% CI.
(C) Transplant of 10,000 LSKs segmented by CUX1 level into lethally irradiated recipients. Donor engraftment (CD45.2$^+$) in the spleen and bone marrow was measured 16 hours after transplant by flow cytometry. Mean +/- 95% CI.
(D) Transplant of 100 segmented LT-HSCs with 200,000 BMMCs into lethally irradiated recipients. Donor engraftment (CD45.2$^+$) in each compartment was measured 3 weeks post-transplant by flow cytometry. Mean +/- 95% CI.

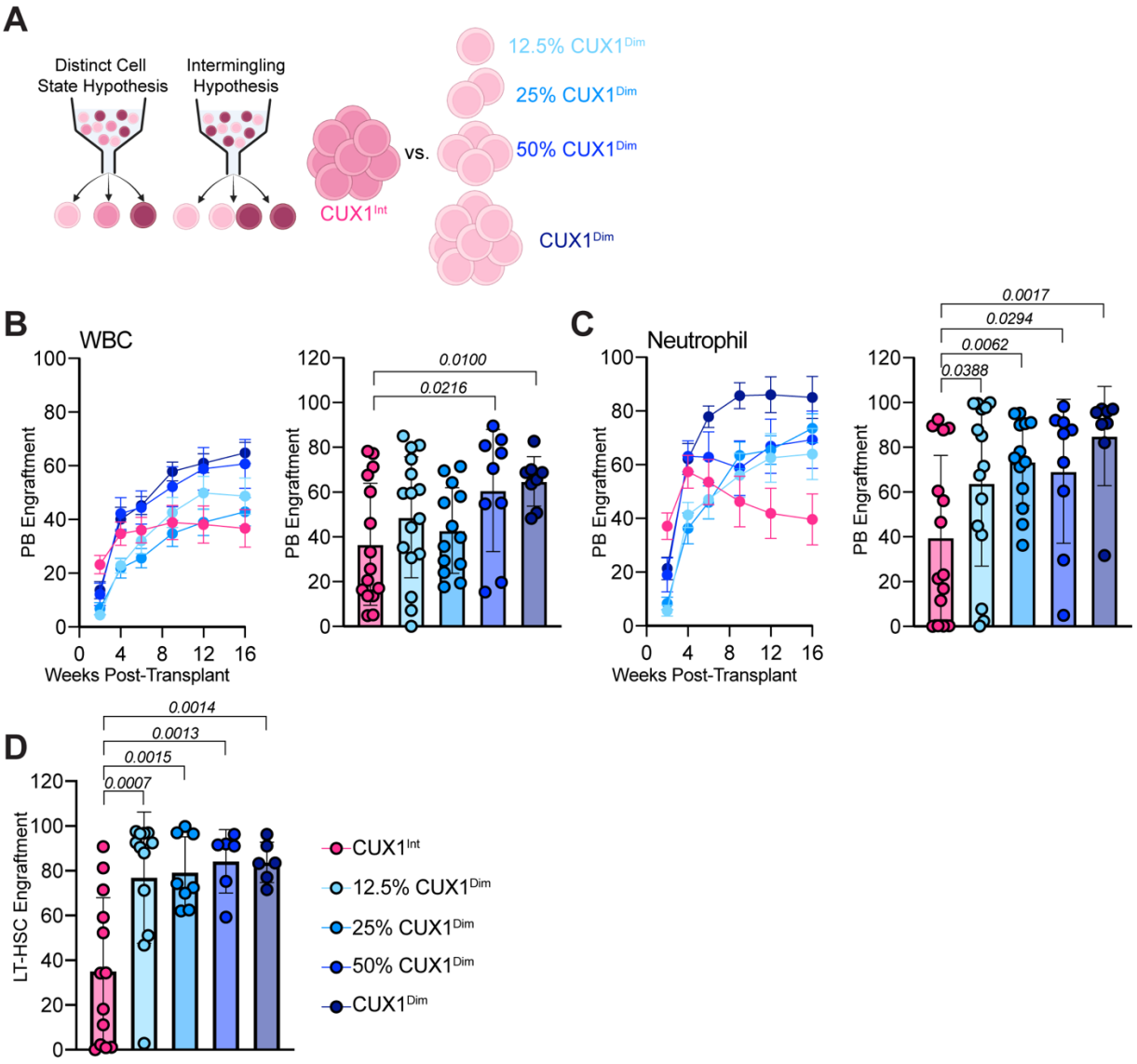

Figure 10. CUX1^{Int} LT-HSCs represent a distinct cell state

(A) This experiment was designed to test two hypotheses for the intermediate phenotype of CUX1^{Int} LT-HSCs. It was possible that CUX1^{Int} LT-HSCs represented a mix of CUX1^{Dim} and CUX1^{Bright} LT-HSCs and that poor resolution of these populations by flow cytometry and, ultimately, cell sorting resulted in a mixed population. What follows from the intermingling hypothesis is that CUX1^{Int} LT-HSCs could just be represented as a fraction of CUX1^{Dim} HSCs and that the decreased self-renewal capacity of CUX1^{Int} LT-HSCs was only observed because we had transplanted fewer CUX1^{Dim} LT-HSCs than our CUX1^{Dim} condition. To test this hypothesis, we transplanted varying numbers of CUX1^{Dim} LT-HSCs compared to CUX1^{Int} LT-HSCs. (F-H) Transplant of 80 CUX1^{Int} HSCs (CUX1^{Int}), 80 CUX1^{Dim} HSCs (CUX1^{Dim}), 40 CUX1^{Dim} HSCs (50% CUX1^{Dim}), 20 CUX1^{Dim} HSCs (25% (CUX1^{Dim}), or 10 CUX1^{Dim} HSCs (12.5% CUX1^{Dim}) plus 200,000 BMMCs into lethally irradiated recipient mice. Flow cytometry analysis of WBC (B), neutrophil engraftment (C), and LT-HSC engraftment (D) of donor-derived cells. ANOVA and Tukey's test. Mean +/- 95% CI.

marrow homing (Figure 9C). CUX1^{Bright} HSCs are not cells that coincidentally express stem cell markers because all immunophenotypic HSCs transcriptionally cluster together regardless of CUX1 level (Figure 7B). Also, CUX1^{Bright} HSCs gave rise to downstream progenitors three weeks post-transplant, suggesting these cells are truly HSCs at a more advanced differentiation state (Figure 9D).

We considered the possibility that the intermediate phenotype of CUX1^{Int} HSCs was due to the intermingling of CUX1^{Bright} and CUX1^{Dim} populations. To test this prospect, we compared decreasing proportions of CUX1^{Dim} HSCs to a fixed number of CUX1^{Int} HSCs. We reasoned that under the intermingling model, the phenotype CUX1^{Int} HSCs could be explained by a fraction of CUX1^{Dim} HSCs. This is not what we observed. Only 12.5% of CUX1^{Dim} HSCs relative to CUX1^{Int} resulted in higher myeloid and LT-HSC engraftment than CUX1^{Int} HSCs, suggesting that CUX1^{Int} HSCs are not a mixed population and represent a distinct cell state (Figure 10A-D). These findings create a model of LT-HSC differentiation where CUX1 levels gradually increase as HSCs lose reconstitution capacity.

CUX1 levels correlate with the gradual loss of the HSC gene expression program

To understand how CUX1 levels can influence stem cell potential, we performed bulk RNA-seq on HSCs segmented on CUX1 levels from *Cux1*^{mCherry} mice (Figure 11A). Using k-means clustering, we identified four signatures of differentially expressed genes. CUX1^{Bright} signatures represent genes that increase in expression from CUX1^{Dim} to CUX1^{Bright}, which are enriched terms associated with cell cycling and cellular respiration. CUX1^{Dim} signatures represent genes that decrease in expression from CUX1^{Dim} to CUX1^{Bright}. These signatures are enriched for quiescence, chromatin regulators, and innate immune activation (Figure 12A). Gene

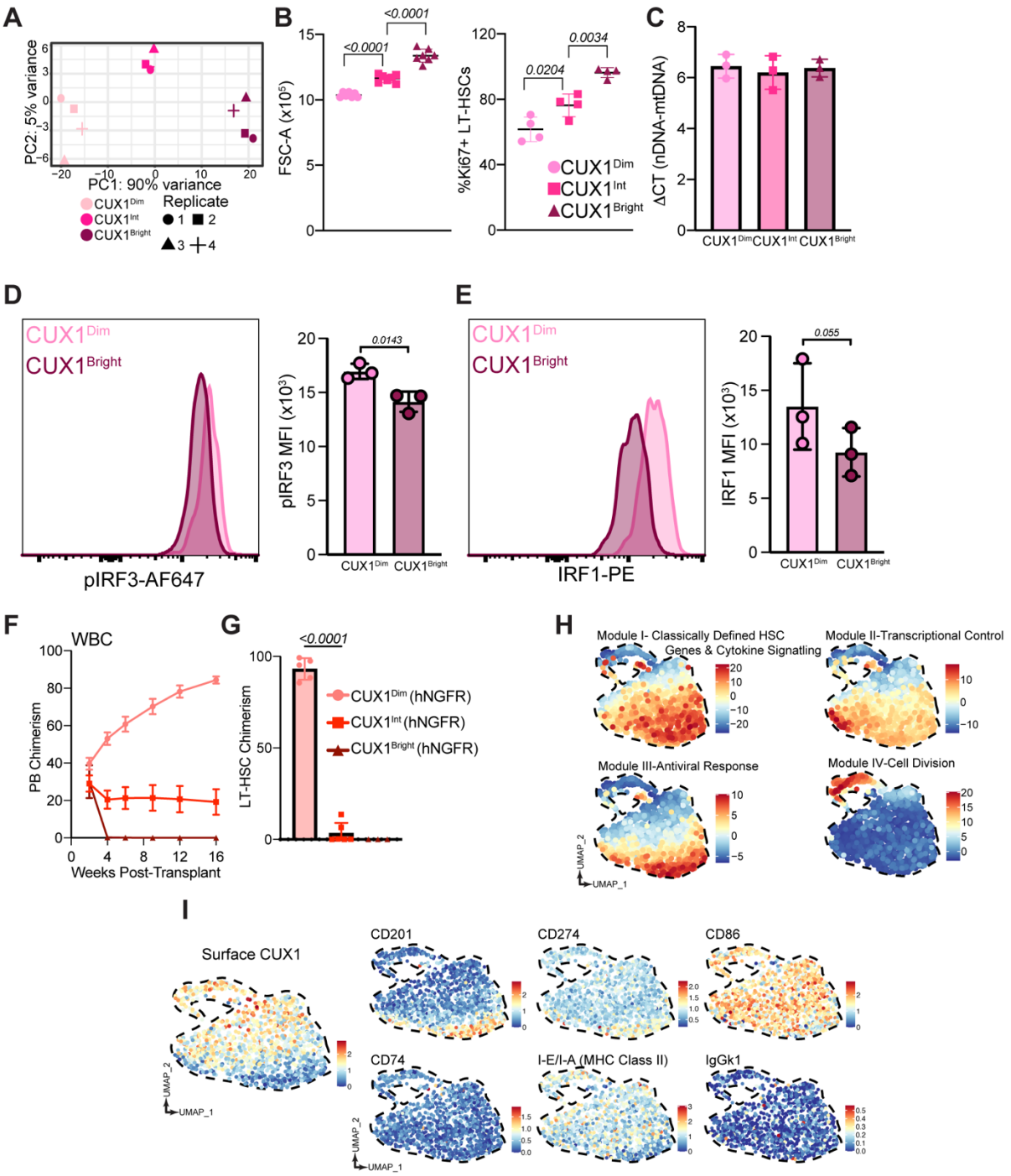

Figure 11. High interferon signaling in CUX1^{Dim} HSCs

(A) Batch-corrected PCA of bulk RNA-seq corresponding to Figure 12A (n=4).
(B) Flow Cytometry analysis of FSC-A and Ki-67+ in segmented LT-HSCs. ANOVA and Tukey's test. Mean +/- 95% CI.
(C) qPCR for nuclear and mitochondrial DNA on sorted LT-HSCs using protocol and primers from de Almeida et al. 2017[313]

Figure 11, continued
(D-E) Intracellular flow cytometry analysis of pIRF3 (D) and IRF1 (E) of sorted CUX1^{Bright} HSCs and CUX1^{Dim} HSCs cultured in Wilkinson et al. media for one week. Mean +/- 95% CI. Unpaired, two-tailed t-test.
(F-G) To compare our findings from the Cux1mCherry mouse model to our findings from the Cux1ATPN mouse model, we performed a similar experiment to Figure 8D-H. Transplant of 100 segmented LT-HSCs from *Cux1*^{ATPN} mouse with 200,000 BMMCs into lethally irradiated recipient mice. (F) Donor engraftment of WBC. Mean +/- SEM. (G) Donor Engraftment of LT-HSC. Mean +/- 95% CI. ANOVA and Tukey's test
(H) The expression of modules identified by Hotspot corresponding to 12I.
(O) Surface marker expression measured by antibody-DNA tag (ADT) corresponding to Figure 12F.

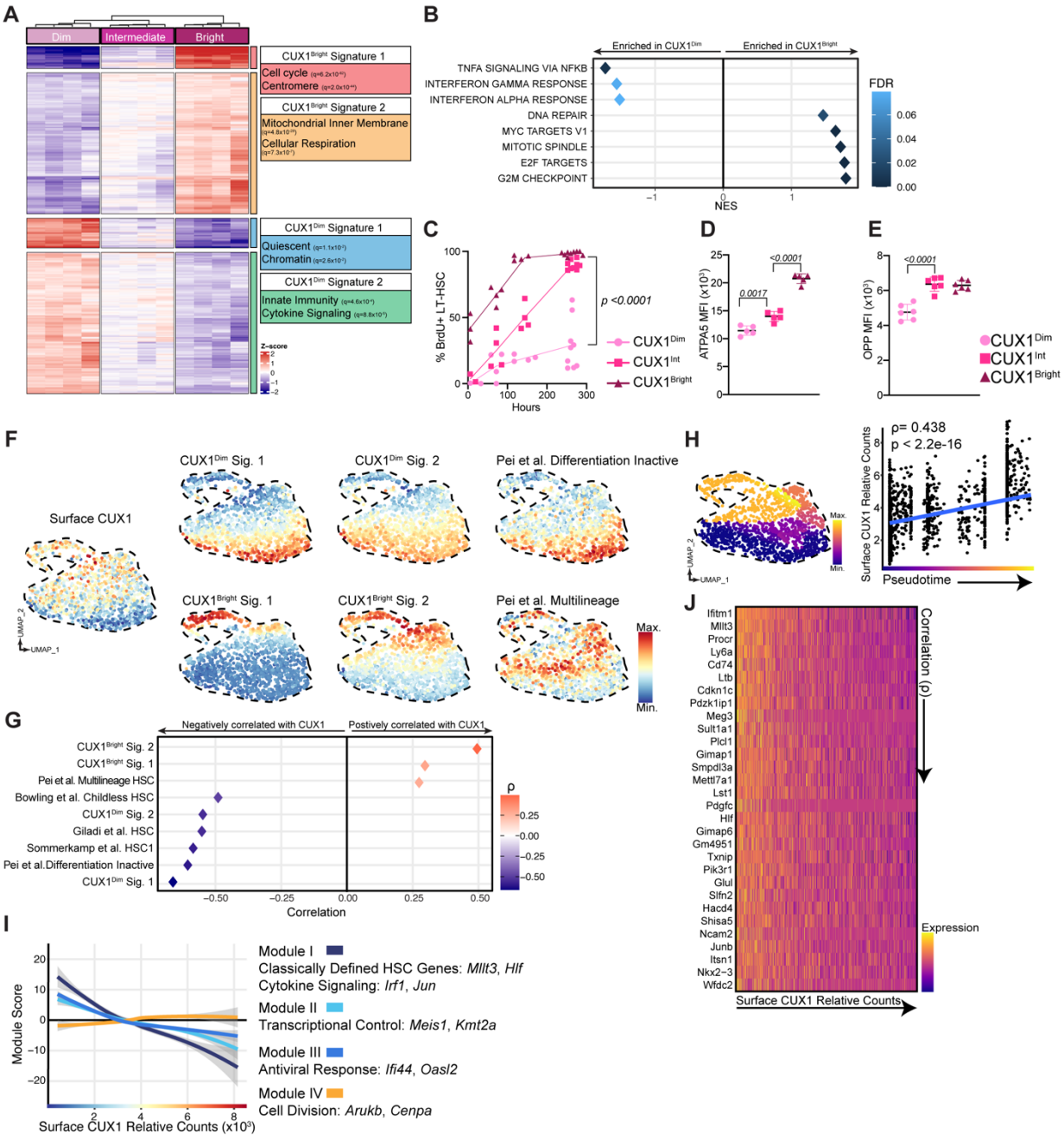

Figure 12. CUX1 levels correlate with the gradual loss of the HSC gene expression program

(A-B) Bulk RNA-Seq of sorted LT-HSCs sorted CUX1 levels (n=4). (A) Heatmap of differentially expressed genes (FDR < 0.01, n=3564) after k means clustering. GO terms generated using DAVID are shown for each cluster. (B) Hallmark GSEA of CUX1^{Dim} and CUX1^{Bright} LT-HSCs.

(C) Time course of in vivo BrdU labeling (n=3-9 per time point). The frequency of BrdU^{+} segmented LT-HSCs segmented on CUX1 is shown. One-way ANOVA and Tukey's test.

Figure 12, continued
(D) Expression of ATP5A (MFI) relative to CUX1 expression in LT-HSCs (n=5). Significance One-way ANOVA and Tukey's test. Mean +/- SD.
(E) Flow cytometry analysis of *ex-vivo* OP-Puro incorporation in segmented LT-HSCs (n=6). One-way ANOVA and Tukey's test. Mean +/- SD.
(F-J). scRNA-Seq (10X Chromium) of sorted LT-HSCs (n=1321). (F) UMAP of CUX1 expression (hNGFR) and HSC gene sets. (G) Pearson correlation of gene signatures with surface CUX1 expression. (H) Monocle3 pseudotime analysis and correlation with surface CUX1. (I) Hotspot modules in 1000 variable genes relative to surface CUX1. (J) Heatmap of top 30 genes ranked in order of Pearson correlation coefficient with surface CUX1 expression. Cells are ordered in increasing expression of CUX1 normalized counts.

set enrichment analysis (GSEA) comparing CUX1^{Dim} and CUX1^{Bright} HSCs confirmed the enrichment of innate immune gene signatures in CUX1^{Dim} HSCs (Figure 12B). Enrichment of interferon-stimulated genes (ISGs) in CUX1^{Dim} HSCs corresponds with the paradigm of high cell-intrinsic ISG expression in somatic murine and human stem cells[314]. In line with this finding, CUX1^{Dim} HSCs have higher levels of pIRF3 and IRF1 than CUX1^{Bright} HSCs (Figure 11D-E). GSEA also suggested that CUX1^{Bright} HSCs are more proliferative, which was confirmed by greater BrdU incorporation over 11 days of continuous labeling *in vivo*, as well as greater Ki-67 expression and larger size by FSC[315] (Figure 12C, 11B). In addition, CUX1^{Bright} HSCs showed greater protein translation rates as measured by *ex vivo* OP-Puro incorporation and greater expression of the mitochondrial inner membrane complex V component ATP5A with no difference in mitochondrial content (Figure 12D-E, 11C). Overall, these findings indicate that CUX1^{Dim} HSCs are dormant relative to the more metabolically active CUX1^{Bright} HSCs.

To achieve a finer resolution of CUX1 levels and the HSC transcriptome, we returned to the *Cux1*^{ATPN} mouse model. To validate this mouse model, we transplanted LT-HSCs segmented on CUX1 levels (Surface CUX1) as we did with mCherry and observed the same decline in HSC activity as CUX1 levels increase (Figure 11F-G). Single-cell RNA sequencing revealed that CUX1 protein levels (Surface CUX1) are anti-correlated with the previously mentioned CUX1^{Dim} signatures in single cells and previously published signatures of HSCs[52,59,62,316] (Figure 12F-G). Similarly, the CUX1^{Bright} signatures correlate with surface CUX1 levels (Figure 12F-G). These findings verify the *Cux1*^{ATPN} model and extend the gene expression signatures from bulk RNA-seq to the single-cell level. The four gene signatures defined by CUX1 expression create waves of gene expression that start at the apex of the HSC compartment and extend to the precipice of differentiation. Accordingly, Monocle3 pseudotime analysis supported our model

wherein CUX1 levels increase as HSCs approach the moment where they commit to terminal differentiation (Figure 12H).

These data support a model wherein CUX1 protein levels directly link HSC gene expression with functional capacity in single cells. We identified four gene modules in this data using Hotspot[317] (Figure 12I, 11H). Three modules decrease as HSCs exit the compartment and are associated with classically defined HSC genes (*Mllt3*[294], *Hlf*[318], *Son*[319], *Hoxa9*[320]) and cytokine signaling (*Irf1*, *Jun*, *Fos*), transcriptional control (*Mecom*, *Meis1*, *Kmt2a*), and the antiviral response (*Ifi44*, *Oasl2*, *Isg15*). One module associated with cell division (*Mki67*, *Cenpa*, *Aurkb*) increases with CUX1 protein levels. Digging deeper into genes most correlated with stem state, two surface genes, Procr and Cd74, emerged as two of the most anticorrelated genes with CUX1 (Figure 2J). However, only CD201 (*Procr*) protein is anti-correlated with surface CUX1, not CD74 protein (Figure 12J, 11I). Overall, the path out of the HSC compartment is marked by the reduction of antiviral response genes in addition to the classically defined stemness genes and the upregulation of cell division genes.

<u>CUX1 knockdown restores and extends HSC activity</u>

Moving beyond the correlation of CUX1 with HSC activity, we investigated if CUX1 plays a causal role in HSC fate and if we could reverse the progressive loss of HSC potential by reversing CUX1 levels. We crossed the *Cux1*^{mCherry} reporter mouse with our previously published doxycycline-inducible shRNA knockdown mice[78,321] (Figure 13A). This model allowed us to examine the effects of knocking down CUX1 at different starting CUX1 levels by sorting LT-HSCs segmented on CUX1 and then inducing shRNA expression. Given that HSCs start in the CUX1^{Dim} state and CUX1 levels increase with differentiation, we conceptualize shRNA induction in CUX1^{Dim} cells as limiting CUX1 expression rather than knocking down the

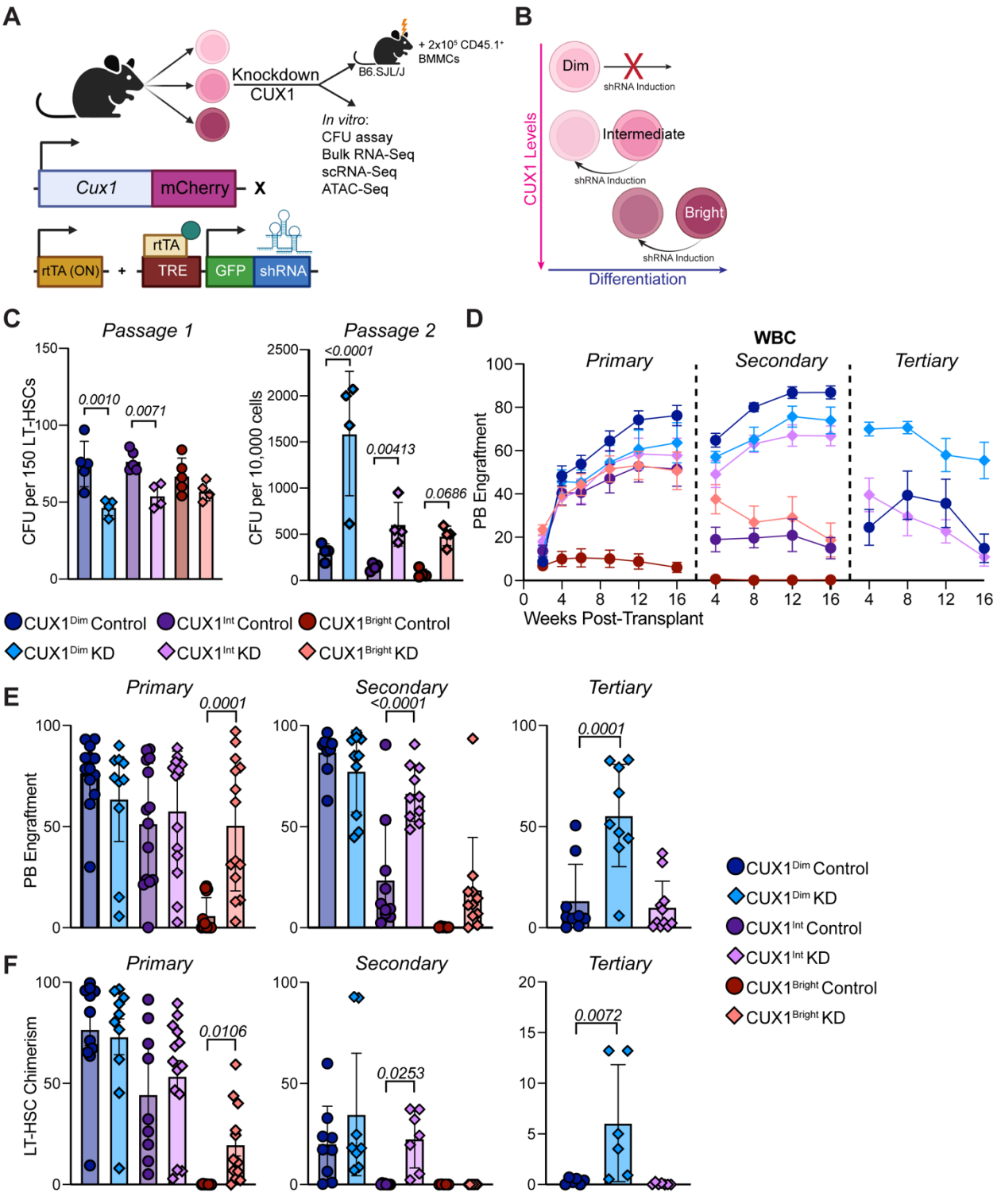

Figure 13. Knockdown of CUX1 restores and extends HSC activity

A) Description of reporter knockdown mouse model. The *Cux1*^{mCherry} mouse model was crossed to the previously published doxycycline-inducible shRNA TET-ON mouse model. Reverse

Figure 13, continued
tetracycline-controlled transactivator (rtTA) drives the expression of either a CUX1-targeting shRNA or a control shRNA targeting Renilla luciferase in the presence of doxycycline.
(B) Conceptualization of shRNA-mediated knockdown at three initial CUX1 levels. For $CUX1^{Bright}$ and $CUX1^{Int}$, shRNA induction reduces CUX1 levels; however, for $CUX1^{Dim}$ LT-HSCs, shRNA induction prevents CUX1 from turning on.
(C) CFU assay of sorted segmented LT-HSCs and knockdown initiated in vitro. One-way ANOVA and Fischer's LSD test. Mean +/- SD.
(D-F) Transplant of 100 segmented LT-HSCs into lethally irradiated recipients with 200,000 competitor BMMCs. Flow cytometry analysis of donor engraftment ($CD45.2^{+}$) for primary, secondary, and tertiary transplant ($CUX1^{Dim}$ control 1°n=13, 2°n=10, 3°n=9; $CUX1^{Dim}$ knockdown 1°n=10, 2°n=8, 3°n=9; $CUX1^{Int}$ control 1°n=8 2°n=8; $CUX1^{Int}$ knockdown 1°n=15, 2°n=7, 3°n=11; $CUX1^{Bright}$ control 1°n=8, 2°n=5; $CUX1^{Bright}$ knockdown 1°n=12, 2°n=7). Two-way ANOVA and Tukey's test. Mean +/- SEM. (E) Flow cytometry analysis of donor LT-HSC from (D) ($CUX1^{Dim}$ control n=11/9/3, $CUX1^{Dim}$ knockdown n=10/8/5, $CUX1^{Int}$ control n=8/8, $CUX1^{Int}$ knockdown n=15/7/4, $CUX1^{Bright}$ control n=8/5, $CUX1^{Bright}$ knockdown n=12/7). Mice were bled at the indicated intervals for 16 weeks before re-transplant of BMMCs. One-way ANOVA and Sidak's test. Mean +/- 95% CI.

gene (Figure 13B). CUX1 knockdown slightly decreased primary CFU formation but substantially increased CFU replating capacity at all three initial CUX1 doses, particularly for $CUX1^{Dim}$ HSCs (Figure 13C, 14A). Corroborating the *in vitro* increase in self-renewal, CUX1 knockdown in $CUX1^{Bright}$ HSCs restores HSC capacity such that $CUX1^{Bright}$ HSCs gave rise to multilineage engraftment throughout the primary transplant after CUX1 knockdown (Figure 13D-E 'Primary,' 14B). Knocking down CUX1 in $CUX1^{Bright}$ HSCs restores the capacity to self-renew, as $CUX1^{Bright}$ KD HSCs are recoverable 16 weeks post-transplant (Figure 13F 'Primary'). To control for potential shRNA off-target effects, we repeated this experiment with a second, less-potent shRNA, which also restores HSC activity to $CUX1^{Bright}$ HSCs (Figure 14C-E). The effect of knocking down CUX1 in a $CUX1^{Int}$ HSC became apparent in the secondary transplant, where $Cux1^{Int}$ HSCs gave rise to greater peripheral blood and LT-HSC engraftment (Figure 13D-F 'Secondary'). Limiting CUX1 expression in $CUX1^{Dim}$ HSCs resulted in increased engraftment in the tertiary transplant (Figure 13D-F 'Tertiary'). Ultimately, CUX1 knockdown results in an extension of HSC self-renewal (Figure 13F).

To understand the transcriptional effects of CUX1 knockdown and initial CUX1 level in the $CUX1^{Bright}$ to $CUX1^{Dim}$ reversion on the, we sorted segmented HSCs, knocked down CUX1 *in vitro*, and performed bulk RNA-seq (Figure 15A-B). Upon knockdown, Dim signature 1 and Bright signature 1 are positively and negatively enriched, respectively, at all three initial CUX1 levels (Figure 15C). In line with this finding, genes significantly associated with different CUX1 levels were strongly correlated with genes differentially expressed after CUX1 knockdown (Figure 16A). CUX1 knockdown tempered the expression of $CUX1^{Bright}$-associated genes (Figure 16A, upper left), which are enriched for processes associated with biosynthesis and cell

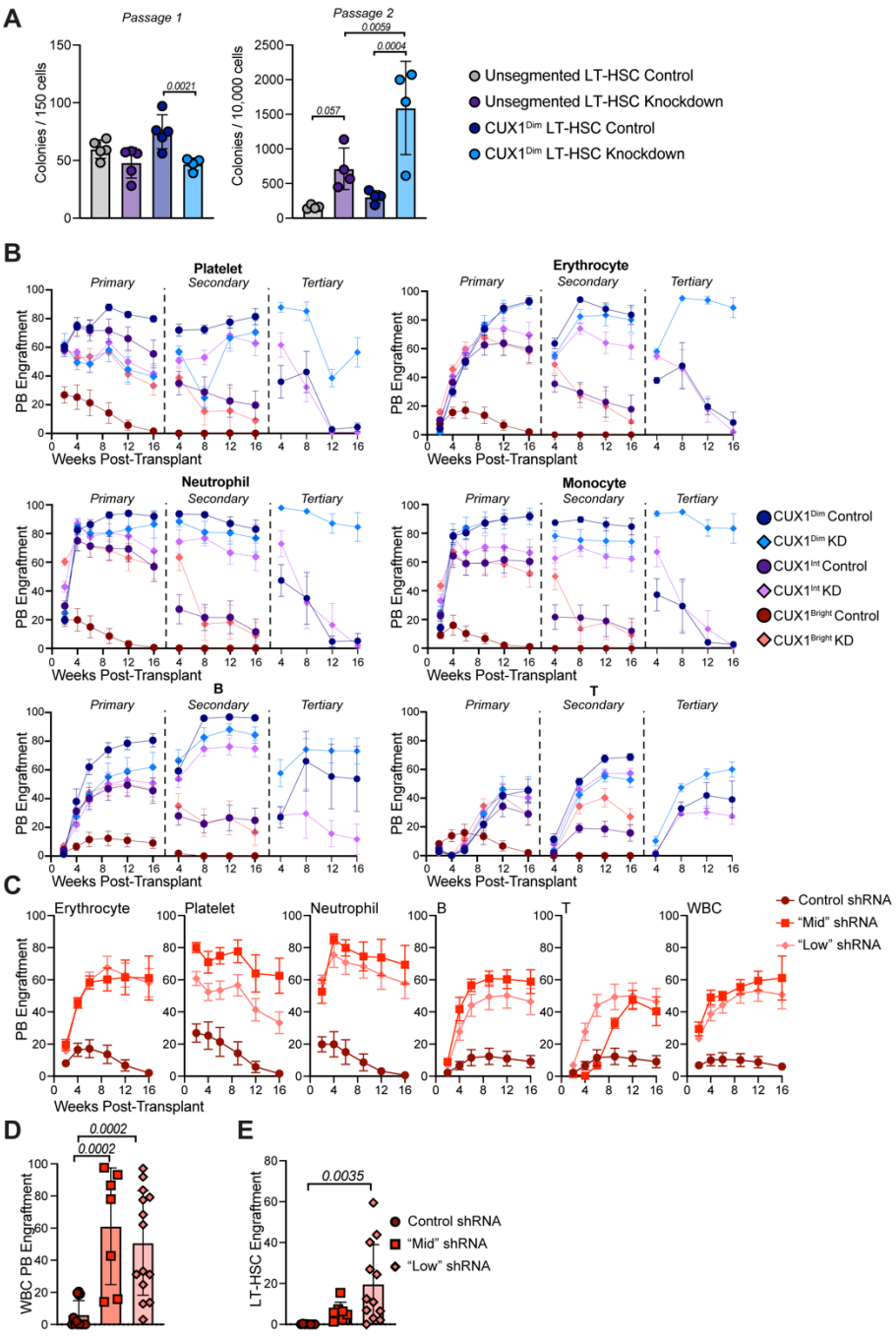

Figure 14. HSC heterogeneity obscures effects of CUX1 knockdown

(A) CFU assay of sorted CUX1Dim and unsegmented LT-HSCs and knockdown initiated in vitro. One-way ANOVA and Tukey's test. Mean +/- SD.
(B) Donor engraftment corresponding to Figure 13D. Mean +/- SEM
(C-E) PB WBC engraftment at 16 weeks post-transplant corresponding to Figure 3E-F (B) with second shRNA targeting CUX1 from An et al. 2018. (Control shRNA n= 8, "Mid" shRNA n=7, "Low" shRNA n=12). (C) LT-HSC engraftment at 16 weeks post-transplant. One-way ANOVA and Tukey's test. Mean +/- SD.

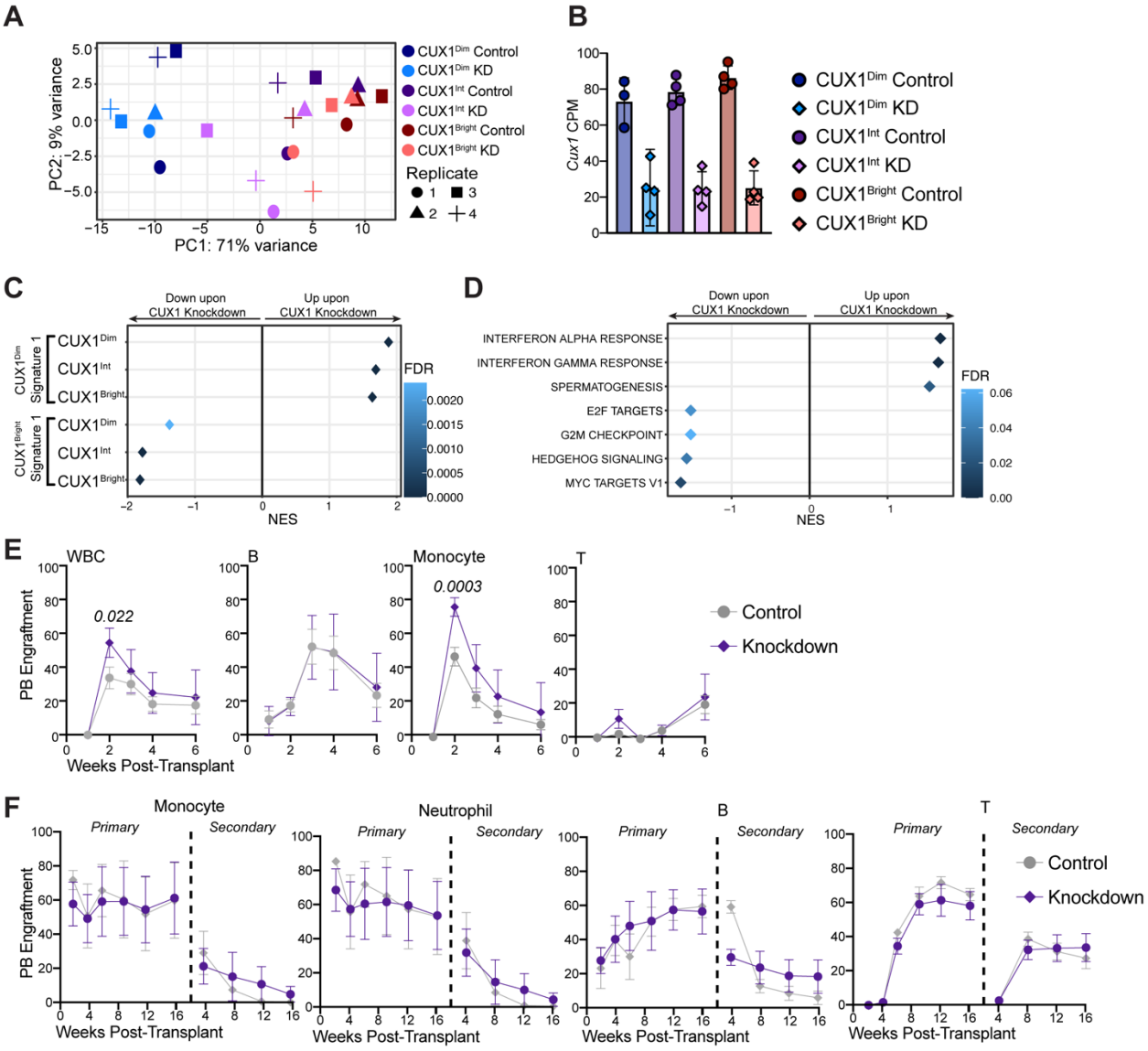

Figure 15. Knockdown of CUX1 potentiates interferon signaling

(A) Batch-corrected PCA of bulk RNA-seq corresponding to Figure 16A (n=3-4).
(B) Cux1 expression (counts per million) corresponding to Figure 16A (n=3-4).
(C) Corresponding to Figure 3F. GSEA for Dim signature 1 and Bright signature 1 upon CUX1 knockdown for each initial CUX1 level.
(C) GSEA of differentially expressed genes after CUX1 knockdown un CUX1^{Dim} LT-HSCs.
(E-F) Flow cytometry analysis of donor cell (CD45.2+) engraftment of 3000 transplanted MPPs (E) (n=9 per condition) and 300 ST-HSCs (F) (n=5 per condition) corresponding to Figure 16E-F. Two-way ANOVA and Sidak's test. Mean +/- SEM.

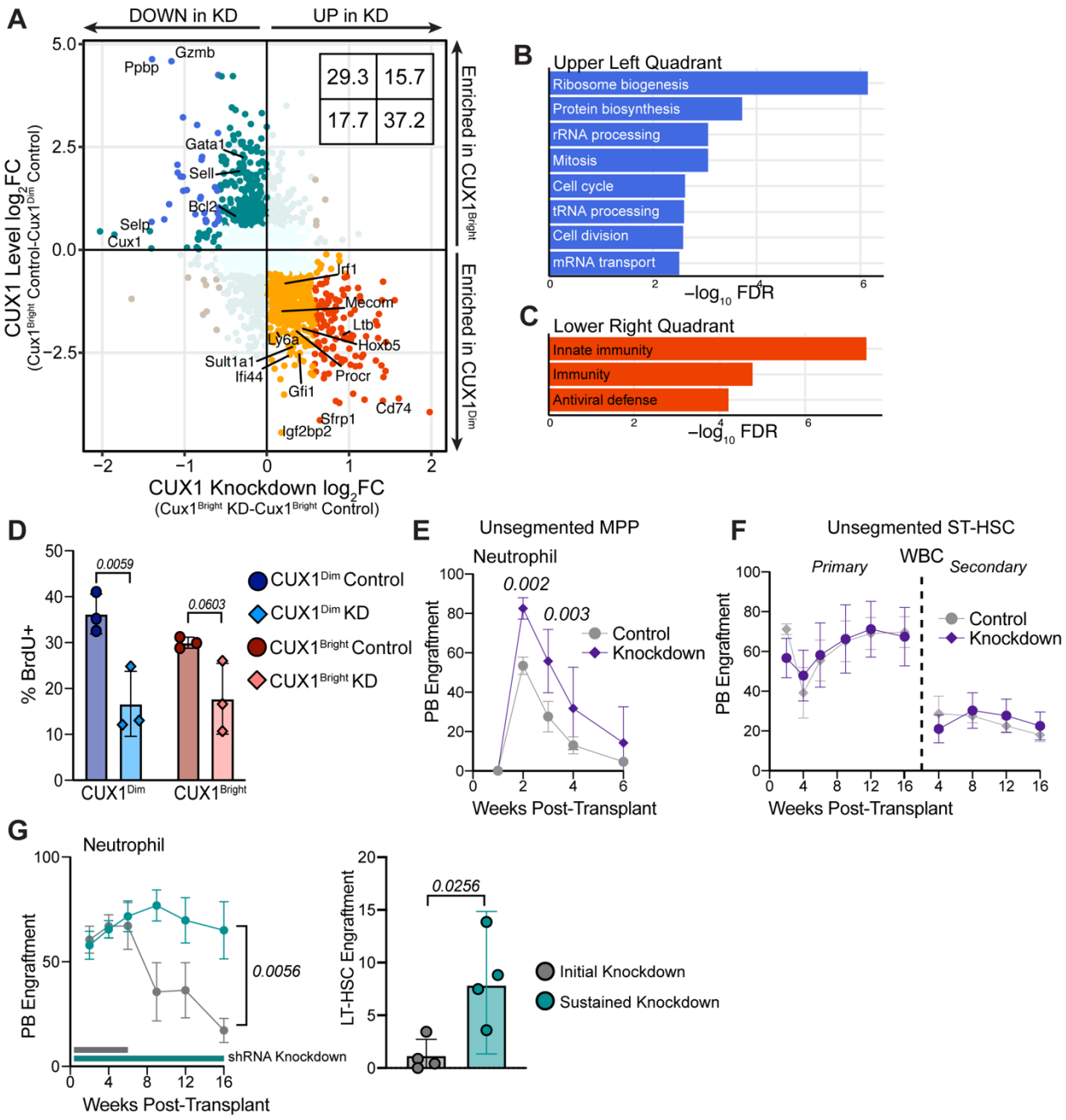

Figure 16. CUX1 knockdown restores CUX1^{Dim} transcriptome in CUX1^{Bright} HSCs

(A-C) Bulk RNA-Seq of sorted LT-HSCs segmented on CUX1 levels followed by 3 days of in vitro CUX1 knockdown in Wilkinson et al. media (n=3-4 per condition). Shown are the fold changes of differentially expressed genes (FDR < 0.01). The y-axis represents the fold change comparing CUX1 levels (CUX1^{Bright} Control – CUX1^{Dim} Control), and the x-axis represents the fold change comparing CUX1 knockdown in a CUX1^{Bright} cell (CUX1^{Bright} KD -CUX1^{Bright} Control). The proportion of genes in each quadrant is shown. (B-C) GO terms generated using DAVID for Quadrant II (B) and Quadrant IV (C).

Figure 16, continued
(D) Flow cytometry ex vivo BrdU incorporation of sorted LT-HSCs segmented on CUX1 levels with in vitro knockdown for 5 days in Wilkinson et al. media. Two-way ANOVA and Fischer's LSD test. Mean +/- SD.
(E-F) Flow cytometry analysis of donor cell (CD45.2+) engraftment of 3000 transplanted MPPs (E) (n=9 per condition) and 300 ST-HSCs (F) (n=5 per condition). Two-way ANOVA and Sidak's test. Mean +/- SEM.
(G-H) Flow cytometry analysis of donor cell engraftment (CD45.2+) of transplanted $CUX1^{Bright}$ LT-HSCs comparing initial CUX1 knockdown to sustained CUX1 knockdown (n=4-5 per condition). Two-way ANOVA and Sidak's test. Mean +/- SEM. (H) Donor LT-HSC engraftment in LT-HSC. T-test. Mean +/- 95% CI.

division (Figure 16B). In fact, CUX1 knockdown decreased cell cycle participation in vitro, confirming these transcriptional findings (Figure 16D). On the other hand, the $CUX1^{Dim}$ transcriptome, marked by innate immunity-associated genes, was reinstated upon CUX1 knockdown in $CUX1^{Bright}$ HSCs (Figure 16A, bottom right, 3C, 15D). All in all, the restoration of the HSC self-renewal program involves shedding the $CUX1^{Bright}$ transcriptional identity and redonning the $CUX1^{Dim}$ identity.

To explore the bounds of CUX1-regulated cellular reprogramming, we transplanted MPPs and ST-HSCs to determine if these more differentiated progenitors could also revert to HCS. As opposed to LT-HSCs, knockdown of CUX1 in either population did not increase the self-renewal capacity or promote HSC activity, highlighting the distinct plasticity of the LT-HSC compartment (Figure 16E-F, 15E-F). Further exploring this plasticity, we assessed the sufficiency of low levels of CUX1 to promote HSC activity and whether CUX1 knockdown induced a permanent effect in HSCs. $CUX1^{Bright}$ HSCs were transplanted, and CUX1 levels were depressed for six weeks post-transplant, at which point CUX1 was restored in half the mice, leading to the demise of donor HSCs in those recipients (Figure 16G). These findings indicate the observed HSC reprogramming upon CUX1 knockdown depends on continuous low levels of CUX1.

<u>CUX1 knockdown alters differentiation trajectories</u>

Based on our observation that the reversion of a $CUX1^{Bright}$ to a $CUX1^{Dim}$ HSC is incomplete at the cellular level, we sought to analyze the relationship between cells after CUX1 knockdown at the single-cell level. We performed scRNA-seq on segmented HSCs after five

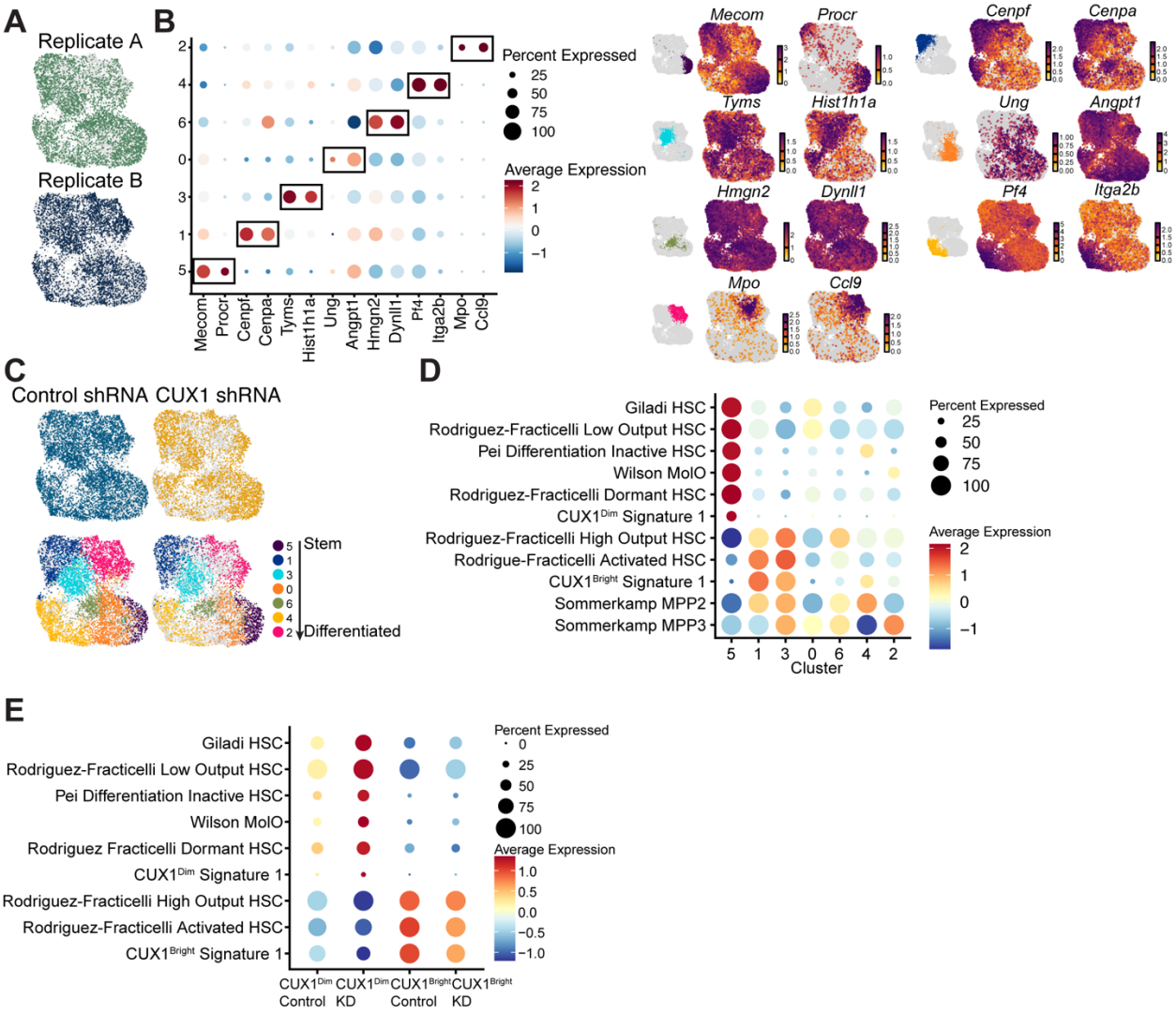

Figure 17. CUX1 knockdown does not induce an alternative cell state

(A) UMAP corresponding to Figure 18A showing two independent replicates of the experiment.
(B) Dot plot and UMAP of marker genes of two marker genes for each cluster. Marker genes were determined by pseudobulk RNA-seq analysis and the seurat feature FindMarkers.
(C) UMAP of cells pooled from the control shRNA and CUX1-targeting shRNA conditions, demonstrating that these two populations are intermingled.
(D) Dot plot of HSC and MPP gene signatures for each cluster.
(E) Dot plot of HSC gene signatures by condition.

days of shRNA induction *in vitro* (Figure 17A-B, 18A). In agreement with our previous assertion that CUX1 knockdown reverses the normal $CUX1^{Dim}$ to $CUX1^{Bright}$ transition, knockdown and control cells comingled in the same clusters, confirming that CUX1 knockdown did not induce an alternative cell state not transited during normal hematopoiesis (Figure 17C). We used published HSC gene signatures to explore the sampled cell states and rank clusters according to their approximation of the HSC transcriptome (Figure 18B, 17D). Knockdown of CUX1 in $CUX1^{Bright}$ and $CUX1^{Dim}$ shifted cells to less differentiated clusters (Figure 18C-D). Concordantly, using Monocle3 pseudotime trajectory inference, CUX1 knockdown repositioned cells to earlier points along the pseudotime axis, confirming that CUX1 deficiency directs HSCs to a more dormant stem cell identity (Figure 18E, 17E). This shift remained evident if we only considered the cells most enriched for the HSC gene signatures and if we removed the confounding proliferating cells (Figure 19A-F).

To infer the dynamic relationships between cells from this single timepoint experiment, we used TopicVelo, an approach that combines probabilistic topic modeling and stochastic transcriptional dynamics to estimate RNA velocity[322]. TopicVelo identified different trajectories for HSCs upon CUX1 knockdown, with streamlines pointed towards more stem states relative to more differentiation-prone control HSCs (Figure 20A-C). To validate these findings experimentally, we analyzed cultured HSCs daily by flow cytometry. In fact, CUX1 knockdown slowed the rate at which cells differentiate, as evidenced by the slower disappearance of $CD150^{+}$ cells in culture (Figure 18F, 20D). Overall, these data indicate that knockdown of CUX1 alters differentiation trajectories at the earliest stages of HSC differentiation.

Given the role of CUX1 in regulating chromatin accessibility[323], we performed ATAC-seq after five days of *in vitro* CUX1 knockdown. We identified 20,805 differentially accessible

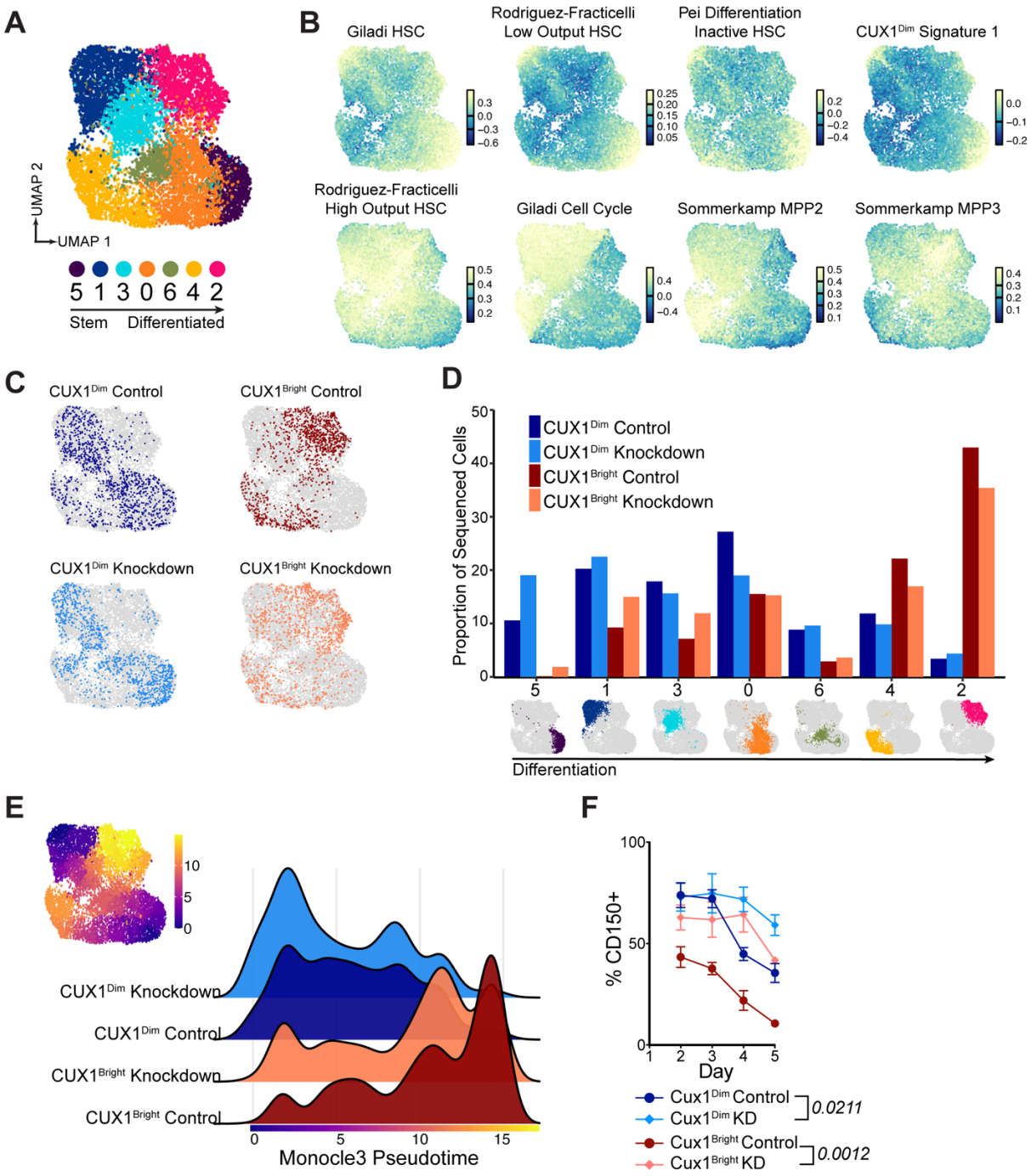

Figure 18. CUX1 knockdown promotes retention in less differentiated cell states

(A-E) scRNA-Seq (10X Chromium) of sorted HSCs cultured for five days in Loeffler et al. media. (A) UMAP of all cells. (B) UMAP of stem cell gene signatures. (C) UMAP randomly down-sampled cells from each condition. (D) Proportion of sequenced cells from each condition in each cluster. (E) UMAP and ridge plot of Monocle3 pseudotime.
(F) Flow cytometry-based analysis of sorted segmented HSCs cultured in Loeffler et al. media for 5 days. The percentage of $CD150^{+}$ cells is shown (n=4-5). Two-way ANOVA with Tukey's test. Error bars +/- SEM.

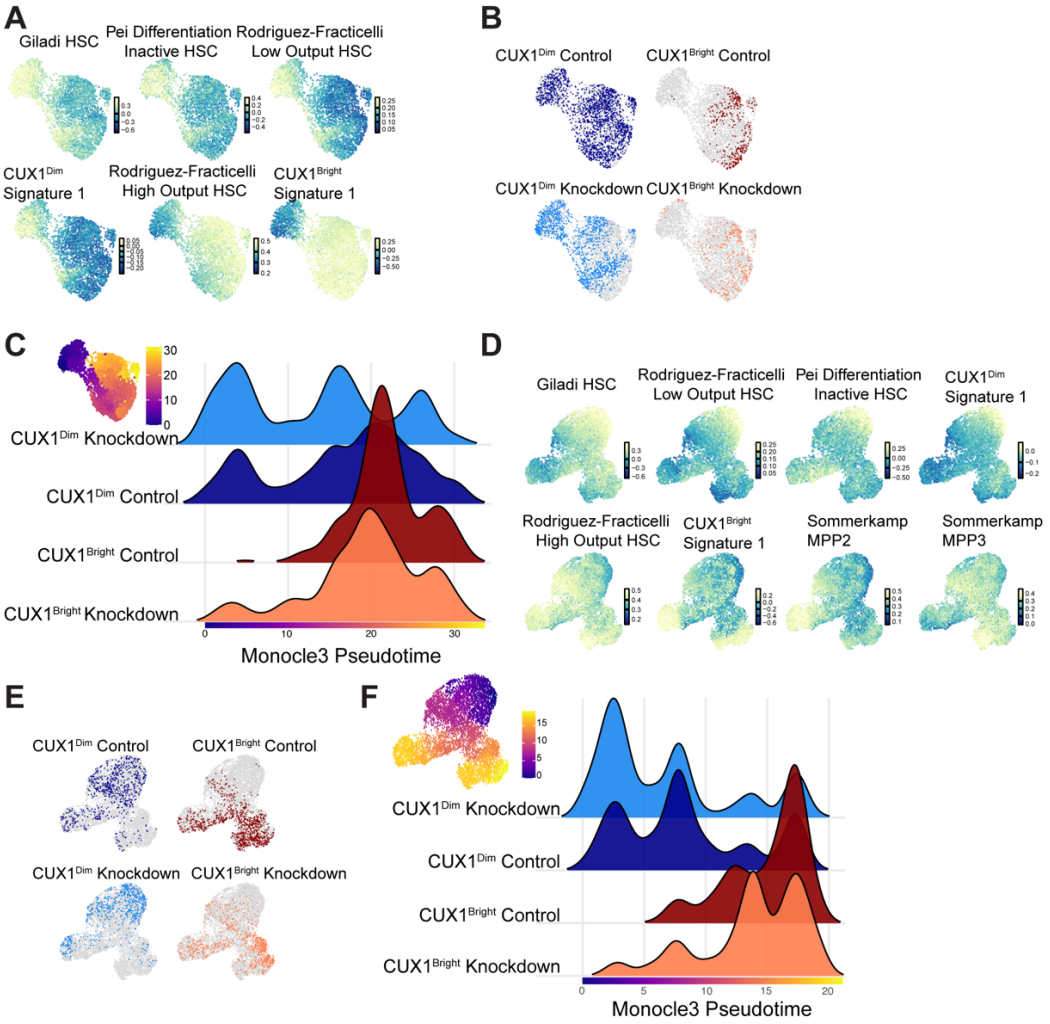

Figure 19. CUX1 knockdown alters differentiation trajectories

(A-C) To ensure the effects of CUX1 knockdown are not due to the greater propensity of CUX1^{Bright} HSCs to give rise to mature cells, cells from the clusters most enriched for HSC gene signatures were re-clustered (original clusters 5, 1, and 3) by re-running RunPCA, FindNeighbors, FindClusters, and RunUMAP. (A) UMAP of HSC gene signatures. (B) UMAP of cells from each condition. (C) UMAP and ridge plot of Monocle3 pseudotime for each condition.
(D-F) Cycling cells present a confounding variable in this analysis. We have previously demonstrated that the CUX1Bright identity involves a loss of dormancy and a commitment to cell cycle participation. However, it is nearly impossible to distinguish between the cell cycle commitment of CUX1Bright HSCs and normal cell cycle participation in this ex vivo assay. To remove this potentially confounding variable, HSC clusters marked by high levels of cycling (original clusters 0 and 3) were removed, and the remaining cells were re-clustered by re-running RunPCA, FindNeighbors, FindClusters, and RunUMAP. (D) UMAP of HSC gene signatures. (E) UMAP of cells from each condition. (F) UMAP and ridge plot of Monocle3 pseudotime for each condition.

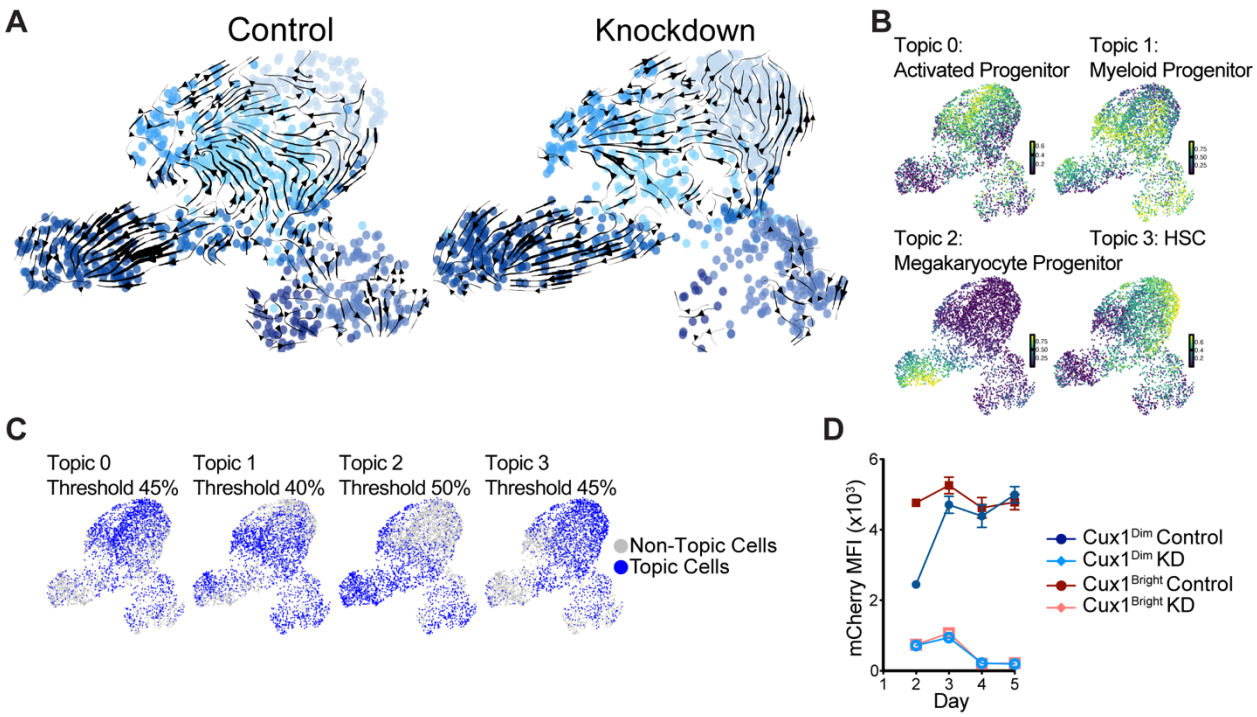

Figure 20. RNA velocity analysis reveals knockdown of CUX1 reverses earliest HSC differentiation trajectory

(A) RNA velocity streamlines from TopicVelo analysis. Control and knockdown cells were pooled prior to computing velocity streamlines. (B-C) UMAP of topic gene expression and topic thresholds for TopicVelo analysis corresponding to Figure 20A.
(D) mCherry MFI corresponding to Figure 18F.

sites between CUX1 knockdown and control and the initial conditions of CUX1^{Dim} and CUX1^{Bright} (Figure 21A-B). These differentially accessible sites are mainly found in introns and intergenic regions, suggesting they could be cis-regulatory elements (Figure 21C). As seen in bulk RNA-seq (Figure 16A), we observed a strong correlation between chromatin accessibility at initial levels of CUX1 and chromatin accessibility after CUX1 knockdown (Figure 22A). 32.5% differentially accessible sites were open in CUX1^{Bright} cells relative to CUX1^{Dim} cells and decreased accessibility upon CUX1 knockdown (Figure 22A, upper left, 22B-C). This is consistent with the pioneering activity of CUX1 mediated through interaction with the ATP-dependent chromatin remodeling BAF complex[323]. Assigning the nearest gene to these differentially accessible sites (±500kb from transcription-start-site) revealed that these genes are associated with proliferation and differentiation, consistent with the phenotype of CUX1^{Bright} HSCs (Figure 22E). Along these lines, CUX1 knockdown in CUX1^{Bright} and CUX1^{Dim} cells decreased accessibility at the BENC Myc enhancer[324] locus, leading to reduced expression of the Myc in CUX1^{Dim} HSCs (Figure 22G-H). In addition, these sites were enriched for motifs of the CEPB family of transcription factors potentially providing etiology to the defects in terminal myeloid differentiation we previously described after CUX1 knockdown leading to a fatal myelodysplastic syndrome[78] (Figure 22F).

14.1% of differentially accessible sites were open in CUX1^{Dim} HSCs (closed in CUX1^{Bright}) and increased accessibility upon CUX1 knockdown in a CUX1^{Bright} HSC (Figure 22A, bottom right, 22B, 22D). Nearest gene analysis for these peaks revealed that these genes are enriched for transcriptional regulation (Figure 22E). These peaks were enriched for motifs of transcription factors that regulate HSC identity, such as GFI1[298], HOXA10[325], HOXB5[326], and MECOM[297]. IRF1[327] and STAT1[328] motifs were also enriched and have recently been shown to

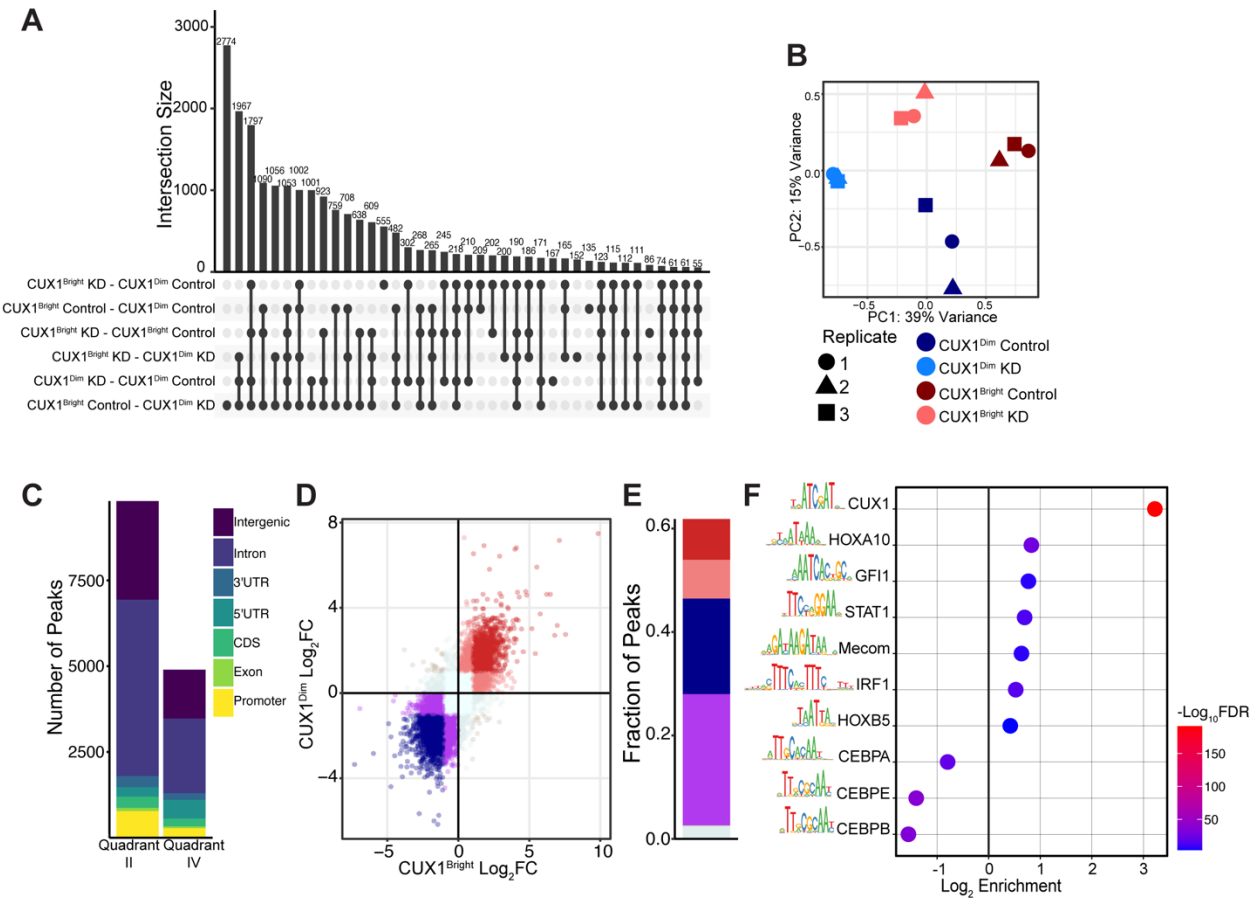

Figure 21. Knockdown of CUX1 induces similar chromatin accessibility landscapes in CUX1^{Dim} and CUX1^{Bright} HSCs

(A) Overlap of differentially accessible peaks from Figure 4G between each of the four conditions.
(B) PCA of the top 30,000 differentially accessible regions (n=3).
(C) Genomic region of each of the differentially accessible peaks in Quadrant II and Quadrant IV corresponding to Figure 22A
(D-F) Fold change by fold change plot showing differentially accessible peaks comparing knockdown in CUX1^{Bright} and CUX1^{Dim} (D) and quantification of peaks (E) with enriched motifs comparing Quadrant I vs Quadrant III (F).

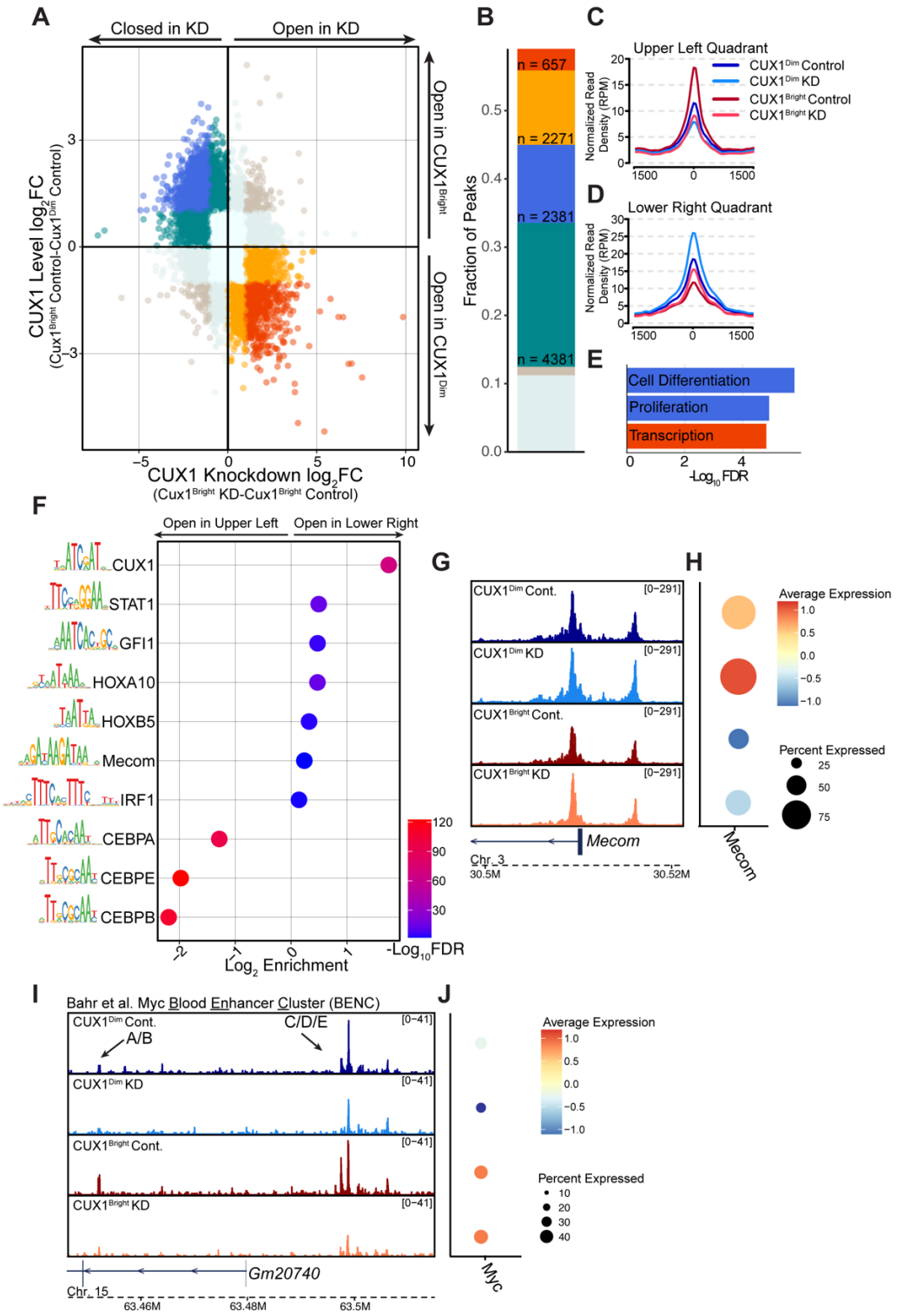

Figure 22. Knockdown of CUX1 in CUX1^{Bright} HSCs reinstates the chromatin accessibility landscape of CUX1^{Dim} HSCs

(A-J) ATAC-Seq of sorted HSCs cultured for five days in Loeffler et al. Media. (A-B) Fold change by fold change plot of differentially accessible peaks (n=20,805) and quantification. The

Figure 22, continued
y-axis represents the log2 fold change comparing CUX1 levels ($CUX1^{Bright}$ Control – $CUX1^{Dim}$ Control), and the x-axis represents the log2 fold change comparing CUX1 knockdown in a $CUX1^{Bright}$ cell ($CUX1^{Bright}$ KD -$CUX1^{Bright}$ Control). (C) Accessibility of profile of Quadrant II peaks (blue and green points from E) for each condition. (D) Accessibility profile of Quadrant IV peaks (red and orange points from E) for each condition. (E) GO terms for assigning Quadrant II peaks to nearest gene (blue bars) and GO term for assigning Quadrant IV peaks to nearest gene (red bar). (F) Differentially enriched transcription factor motifs between Quadrant II and IV. (G) Example of increased accessibility after CUX1 knockdown depicting the Mecom promoter element leading to increased *Mecom* expression after CUX1 knockdown (H). (I) Example of decreased accessibility after CUX1 knockdown depicting the Myc BENC enhancer and *Myc* expression (J).

regulate inflammatory gene expression and self-renewal in HSCs (Figure 22F). Interestingly, the CUX1 motif was also open in CUX1^{Dim} cells and closes upon CUX1 knockdown, suggesting an additional repressive role for CUX1[329] (Figure 22F). This leads us to a model wherein CUX1 sunsets the HSC program by promoting decreased accessibility at key loci for HSC identity. One such locus is the *Mecom* promoter, where we observed dose-dependent accessibility that increases upon CUX1 knockdown, leading to increased expression of Mecom and other HSC identity genes (Figure 22I-J, 17E). All in all, these findings demonstrate that CUX1 loss improves self-renewal in CUX1^{Bright} HSCs by invoking the chromatin landscape, as well as the transcriptome of a CUX1^{Dim} HSC, essentially reprogramming CUX1^{Bright} HSCs to an earlier cell state.

Searching for an explanation for the incomplete reprogramming of a CUX1^{Bright} HSC to a CUX1^{Dim} HSC, we compared the chromatin accessibility upon CUX1 knockdown in CUX1^{Dim} and CUX1^{Bright} HSCs. Here, too, we observed strong concordance in these chromatin landscapes and enrichment of transcription factor motifs in peaks (Figure 21D-F). This indicated that non-epigenetic factors could impose a limit on the extent of the reprogramming.

Loss of CUX1 results in Endogenous Retroelement de-repression

The re-emergence of the inflammatory gene expression program upon CUX1 knockdown (Figure 3I) prompted us to hypothesize it might be essential for HSC self-renewal and consider possible cell-intrinsic explanations. CUX1 interacts with the histone H3K9 methyltransferase, EHT2 (G9a), and loss of the repressive chromatin modification H3K9me2 activates the interferon response[321,330,331]. One function of EHMT2 is to repress endogenous retroelements (EREs) through the deposition of repressive H3K9me2 and the maintenance of DNA

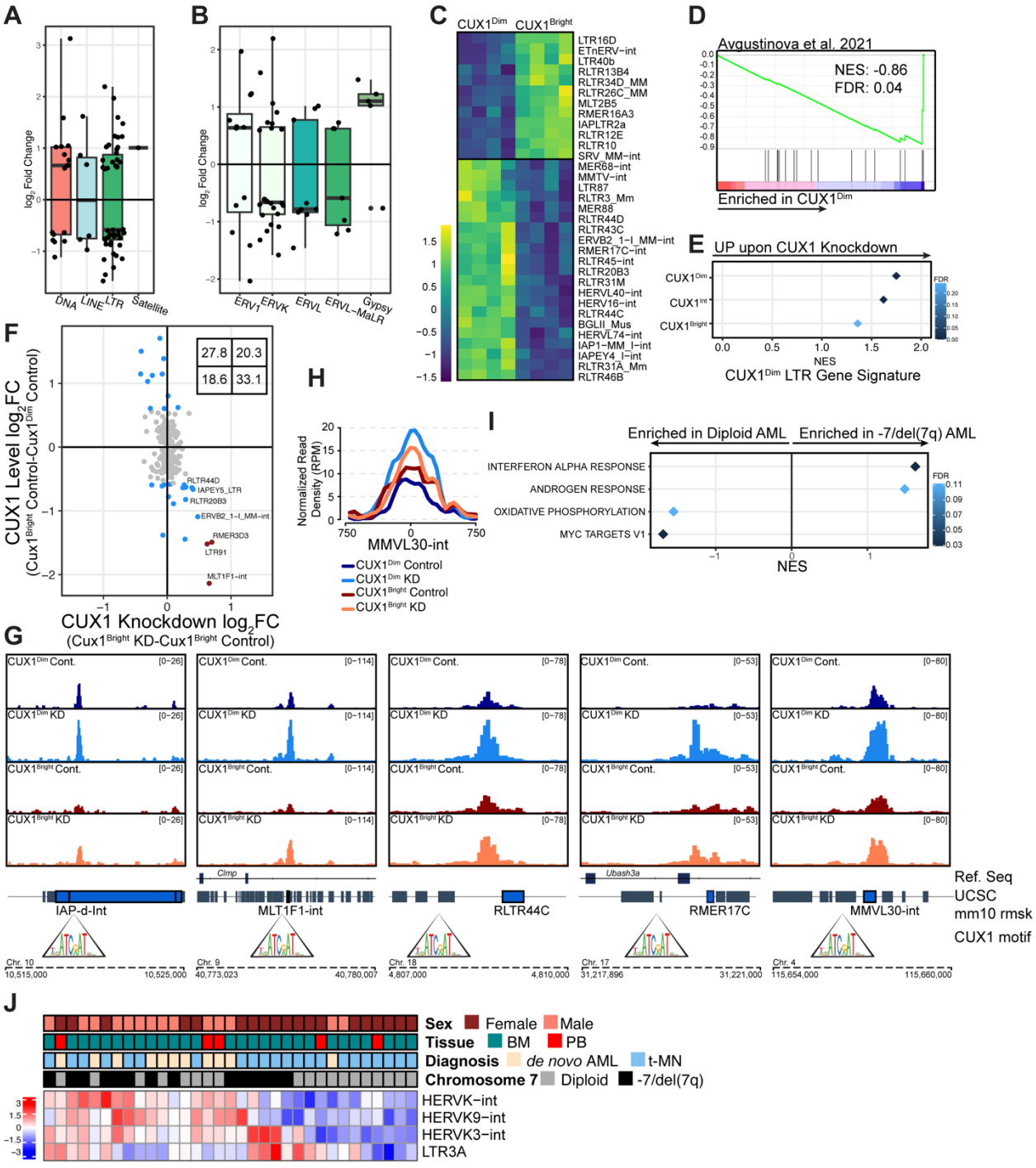

Figure 23. Low levels of CUX1 permit endogenous retroelement expression

(A-B) Differentially expressed repetitive elements (FDR < 0.01) grouped by class (A) and family of LTR (B) comparing CUX1^{Dim} and CUX1^{Bright} from Figure 12.
(C) Heatmap of differentially expressed ERVK and ERVL families.
(D) GSEA of Avgustinova et al.[332] differentially expressed repetitive elements.
(E) GSEA of CUX1^{Dim} LTR-containing repetitive element signature for each level of CUX1 expression.

Figure 23, continued
(F) Fold changes of differentially expressed repetitive elements (n=241) (FDR < 0.01). The y-axis represents the fold change comparing CUX1 levels ($CUX1^{Bright}$ Control – $CUX1^{Dim}$ Control), and the x-axis represents the fold change comparing CUX1 knockdown in a $CUX1^{Bright}$ cell ($CUX1^{Bright}$ KD -$CUX1^{Bright}$ Control). The proportion of genes in each quadrant is shown.
(G-H) Examples of differentially accessible EREs (G). Accessibility of all MMVL30-int elements in differentially accessible peaks (H).
(I-J) Analysis of McNerney et al. 2013[76] primary patient AML RNA-Seq. (I) Hallmark GSEA comparing -7/del(7q) to diploid chromosome 7. (J) Differentially expressed genes ($p < 0.05$) from ERVK class co-bound by IRF1 and STAT1 from Chuong et al. 2016.[333]

methylation[332,334-337]. EREs can ignite the antiviral response through multiple mechanisms, including triggering the dsRNA sensor MDA5 (RIG-I family receptor)[338] or the cytoplasmic dsDNA sensing cGAS-STING pathway[339,340]. Notably, similarly to the loss of CUX1, inhibition of EHMT2 preserves HSC identity in culture[341]. Therefore, we hypothesized that CUX1 represses EREs by recruitment of EHMT2 to shut down the antiviral response and push HSCs toward differentiation.

To assess if EREs could be the initiating factor of the interferon response, we re-analyzed the CUX1-regulated transcriptome (Figure 16C) to quantify the repetitive, multi-aligning RNAs emblematic of EREs. Most differentially expressed repetitive RNAs between $CUX1^{Dim}$ and $CUX1^{Bright}$ HSCs were LTR-containing (long terminal repeat) elements (Figure 23A). Most of these LTR-containing elements were from the ERVL and ERVK families, which are highly expressed in $CUX1^{Dim}$ HSCs and regulated by EHMT2[337,342] (Figure 23B-C). More specifically, the LTR-containing elements that increase expression upon EHMT2 deletion in murine mammary stem cells were enriched in $CUX1^{Dim}$ HSCs, implying that CUX1 recruits EHMT2 to silence these LTR-containing EREs[332] (Figure 23D). Moreover, upon loss of CUX1 at all three initial levels, expression of the $CUX1^{Dim}$ LTR signature increased (Figure 23E). Once again, we observed the same correlation between differentially expressed LTRs between CUX1 levels and the differentially expressed LTRs after CUX1 knockdown (Figure 23F). Returning to the CUX1-regulated chromatin landscape, some of the most differentially expressed EREs, including MLT1F1-int, RLTR44C, and RMER17C, exhibited increased chromatin accessibility after CUX1 knockdown in $CUX1^{Dim}$ and $CUX1^{Bright}$ HSCs (Figure 23G). Notably, IAP-d-int and MMVL30 also increased accessibility after CUX1 knockdown and de-repression of these elements, leads to the accumulation of cDNA in the cytoplasm[343] (Figure 23H). Moreover, the

CUX1 motif was enriched at de-repressed endogenous retroelements (Figure 24B-E). These data suggest that CUX1 is involved in silencing a subset of EREs, and the absence of CUX1 allows their continued accessibility and expression, providing etiology for the interferon response observed upon CUX1 knockdown.

To determine if EREs are also upregulated in human primary cells that have lost CUX1, we re-analyzed primary patient AMLs from McNerney et al. 2013[76]. The interferon alpha response is enriched in AMLs that have lost all or part of chromosome 7, including CUX1, relative to AMLs that are diploid for chromosome 7, suggesting a similar transcriptional phenotype for loss of CUX1 in this setting (Figure 23I). In line with the elevated interferon response in the absence of CUX1, HERV-K family EREs are also upregulated in these patients, and HERVK3-int is notably upregulated in -7/del(7q) patients (Figure 23J, 24A). HERVK elements encode a full provirus, including *Gag*, *Pro*, *Pol*, and *Env* genes, and induce a STING-dependent interferon response, suggesting the role of CUX1 in regulating EREs is conserved from mouse to human[344,345].

Innate immune activation is required for CUX1-mediated self-renewal

To determine if the re-expression of EREs leads contributes to the interferon response upon CUX1 knockdown, we cultured HSCs in the presence of nucleoside reverse transcriptase inhibitors (NRTIs) to inhibit the production of cDNA. Concordant with the induction of the antiviral response upon CUX1 knockdown, pIRF3 and IRF1 increased upon CUX1 knockdown (Figure 16C), which was inhibited by the NRTIs emtricitabine and tenofovir (EMT + TFV) (Figure 25A-B). To assess whether the re-expression of EREs contributes to the increase in self-renewal upon CUX1 loss, we measured the colony-forming potential of HSCs in the presence of

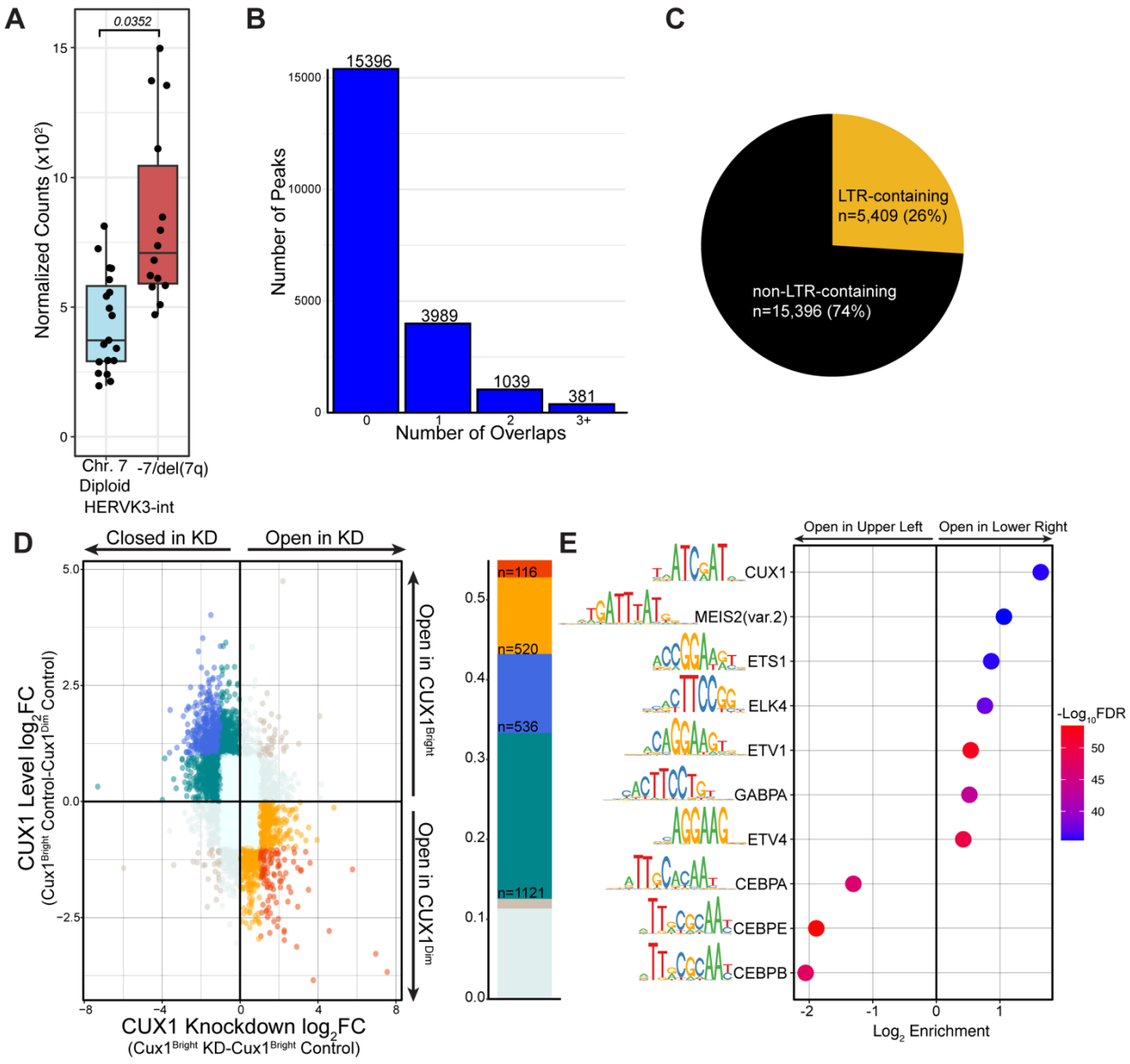

Figure 24. CUX1 motif enriched in open LTR-containing endogenous retroelements

(A) HERVK3-int expression relative to chromosome 7 status is primary patient AML from McNerney et al. 2013

(B-E) Analysis of peaks containing an LTR. (B) Histogram of the number of LTR-containing endogenous retroelements contained within each peak. (C) Pie chart of B. Peaks containing at least one LTR-containing ERE are considered for the remaining analysis. (D) Fold change by fold change plot of differentially accessible peaks containing at least one LTR-containing ERE (n=5,409) and quantification excluding peaks where neither fold change is greater than one. The y-axis represents the fold change comparing CUX1 levels ($CUX1^{Bright}$ Control – $CUX1^{Dim}$ Control), and the x-axis represents the fold change comparing CUX1 knockdown in a $CUX1^{Bright}$ cell ($CUX1^{Bright}$ KD -$CUX1^{Bright}$ Control). (F) Differentially enriched transcription factor motifs between Quadrant II and IV of Figure 24D.

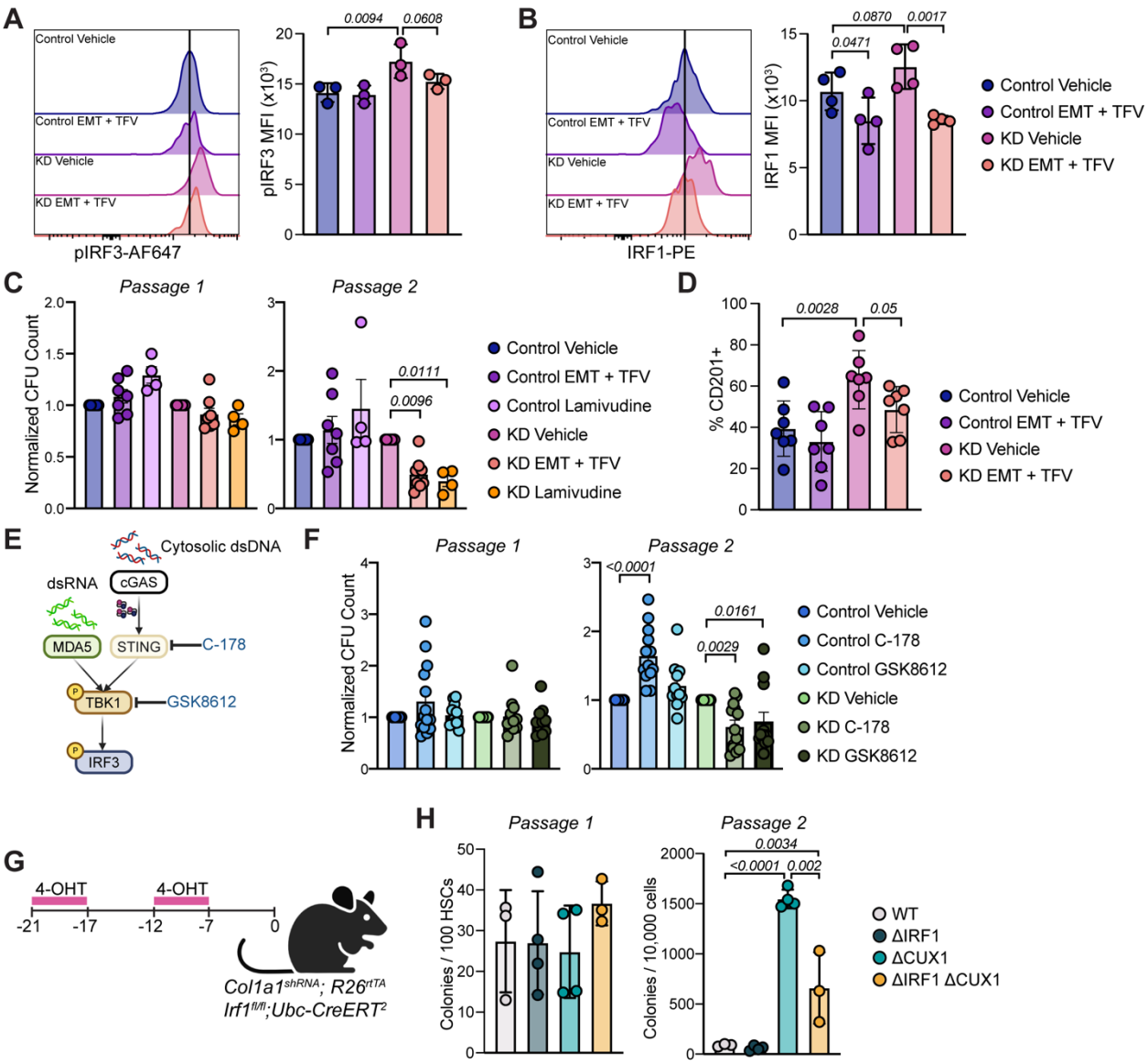

Figure 25. Innate immune response resulting from endogenous retroelement reverse transcription is required for CUX1-mediated HSC self-renewal

(A-B) Intracellular flow cytometry analysis of pIRF3 (A) and IRF1 (B) of sorted CUX1^{Bright} HSCs cultured in Wilkinson et al. media for one week with vehicle (DMSO), 5uM emtricitabine + tenofovir (n=3-4 per condition). One-way ANOVA with Fischer's LSD test. Mean +/- SD.
(C) CFU assay with sorted CUX1^{Dim} LT-HSCs cultured with Vehicle (DMSO), 5uM emtricitabine + tenofovir, or 100uM lamivudine (n=4-8 per condition). Colony number is normalized to the vehicle control for each genotype. One-way ANOVA with Fischer's LSD test. Mean +/- SEM.
(D) Flow cytometry analysis for CD201+ cells of sorted LT-HSCs cultured in Loeffler et al. media with vehicle or 5uM emtricitabine + tenofovir (n=7). One-way ANOVA and Fischer's LSD test. Mean +/- SD.
(E) Depiction of cellular dsDNA and dsRNA sensing pathways with respective inhibitors

Figure 25, continued
(F) CFU assay with sorted CUX1^{Dim} LT-HSCs cultured with Vehicle, 1.5uM C-178, or 1uM GSK8612 (n=11-13 per condition). Colony number is normalized to the vehicle control. One-way ANOVA with Fischer's LSD test. Mean +/- SEM.
(G) Depiction tamoxifen administration schedule of IRF1 and CUX1 double "knockout" experiment.
(H) CFU assay of sorted CD201^{High} LT-HSCs from each genotype indicated (n=3-4). One-way ANOVA and Fischer's LSD test. Mean +/- SD.

NRTIs. HSC re-plating capacity was diminished after CUX1 knockdown in the presence of EMT+TFV or lamivudine (Figure 25C). CD201+ HSCs were retained better in liquid culture upon CUX1 knockdown, and this is also abrogated in the presence of NRTIs, indicating that reverse transcription of EREs is necessary for the ability to retain stem cell identity (Figure 25D). Finally, HSC replating is similarly diminished after the loss of CUX1 in the presence of inhibitors of STING (C-178) or the downstream kinase TBK1 (GSK8612), demonstrating that it is not only reverse transcription essential for HSC self-renewal but also the resulting antiviral response is required to maintain stem cell activity (Figure 25E-F).

Given the upregulation of IRF1 in $CUX1^{Dim}$ HSCs and upon CUX1 knockdown and the role of IRF1 in regulating HSC self-renewal (Figure 26A-F), we interrogated whether IRF1 is required for HSC reprogramming upon loss of CUX1. We crossed the CUX1 shRNA knockdown mouse model to a conditional knockout of *Irf1* with tamoxifen-inducible Cre recombinase (Figure 25G). Given the high correlation between CUX1 protein level and CD201 expression (Figure 8B, 11I), we used $CD201^{High}$ cells as a proxy for $CUX1^{Dim}$ HSCs and $CD201^{Low}$ HSCs as a proxy for $CUX1^{Bright}$ HSCs. Loss of IRF1 in the absence of CUX1 attenuated the replating capacity of $CUX1^{Dim}$ HSCs after CUX1 knockdown (Figure 25H). These results indicate that IRF1 is required for self-renewal in the absence of CUX1.

Discussion

We sought to determine the function of the transcription factor CUX1 in hematopoiesis and identified its dose-dependent role in regulating innate immune activation and, ultimately, the fate of HSCs. CUX1 exhibits dynamic expression in hematopoietic progenitors. Explicitly, there is a gradient of CUX1 expression within the immunophenotypically defined LT-HSC compartment. This gradient constitutes a true continuum with HSCs at distinct points along the

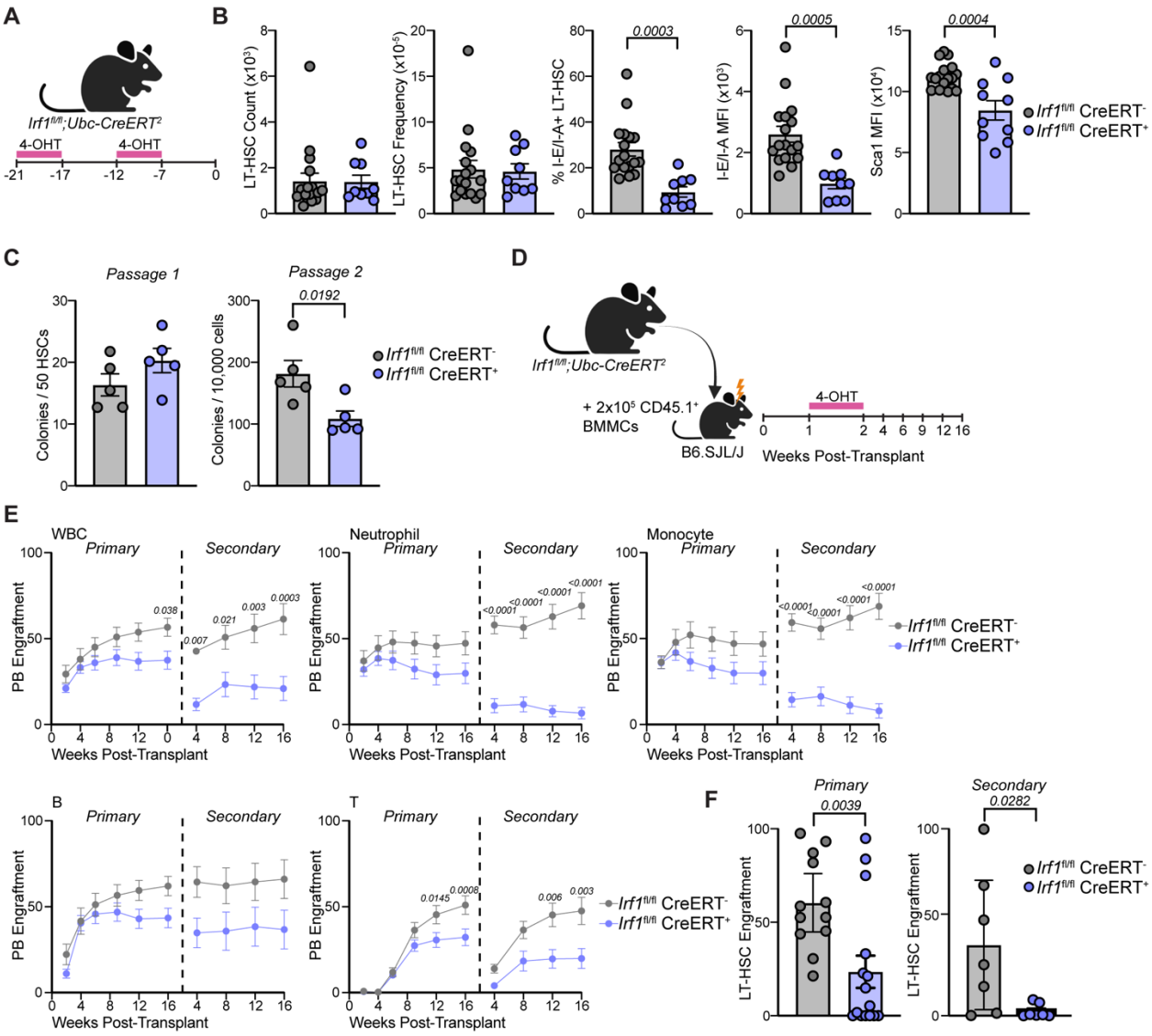

Figure 26. IRF1 is required for HSC self-renewal

(A-B) Irf1$^{fl/fl}$ Ubc-CreERT^{2} mice were maintained on an alternating diet of 500 mg/kg tamoxifen chow for 5 days, followed by 7 days of regular chow. One week following the second period of tamoxifen chow, bone marrow mononuclear cells were collected and subjected to analysis by flow cytometry. (B) Flow cytometry analysis of LT-HSC absolute count, frequency, percentage of I-E/I-A+ LT-HSCs, I-E/I-A MFI, and Sca1 MFI. Unpaired, two-tailed t-test. Mean +/- 95% CI. (C) CFU assay of LT-HSCs. Tamoxifen was administered according to the schedule in Figure 26A. Unpaired, two-tailed t-test. Mean +/- 95% CI.
(D-E) Transplant of 100 LT-HSCs with 200,000 competitor BMMCs into lethally irradiated recipients. Recipient mice were administered 500 mg/kg tamoxifen chow for one week beginning one week after transplant. (E) Analysis of peripheral blood (PB) engraftment of donor cells (CD45.2^{+}) in the competitive primary and secondary transplant of LT-HSCs from Irf1$^{fl/fl}$ Ubc-CreERT^{-} (1°n=15 2°n=9), Irf1$^{fl/fl}$ Ubc-CreERT^{+} (1°n=23, 2°n=9) mice. Two-way ANOVA with Sidak's test. Mean +/- SEM. (F) Flow cytometry analysis of donor LT-HSC engraftment

Figure 26, continued
after primary (left, n=12-15) and secondary (right, n=7) transplant from (J). Frequency of donor LT-HSC shown. Unpaired, two-tailed t-test. Mean +/- 95% CI.

continuum having distinct transcriptomes, chromatin accessibility landscapes, proliferation rates, lineage biases, and repopulating capacities (Figure 8). CUX1^{Dim} HSCs reside at the apex of hematopoiesis and reflect a highly pure population of HSCs. In the journey down to CUX1^{Bright} HSCs, genes associated with the antiviral response and HSC identity decrease expression with a concomitant gain in genes related to cell cycle and biosynthesis (Figure 12). CUX1^{Bright} HSCs fail to meet the functional definition of HSCs, but the reversal of CUX1 levels in this population extended the HSC potential from limited repopulation of the primary recipient to limited repopulation of the secondary recipient (Figure 13). Acquiring HSC activity involves approximating the transcriptome and chromatin landscape of the CUX1^{Dim} HSCs (Figure 13). This indicates low levels of CUX1 are required for HSC activity (Figure 13). HSCs with low levels of CUX1, either naturally or induced, are characterized by innate immune system activation. The antiviral response is a reaction to the re-expression of the EREs, and it is required for the increased self-renewal of HSCs after CUX1 knockdown (Figure 25). Taken together, these results provide a molecular mechanism for the increased self-renewal observed after loss of CUX1 and suggest how continuous activation of the viral inflammatory response may be exploited by myeloid neoplasms with -7/del(7q) (Figure 27).

Engraftment and quiescence

Quiescent cells have been demonstrated to engraft better in irradiated recipients[346-348]. One possible explanation for increased HSC potential after CUX1 knockdown is that the cells are more engraftable due to decreased cell cycle participation accompanying CUX1 knockdown. It is important to consider that CUX1^{Bright} HSCs have no defect in early engraftment relative to CUX1^{Dim} HSCs. In addition, CUX1 knockdown is not instantaneous; two days are required to

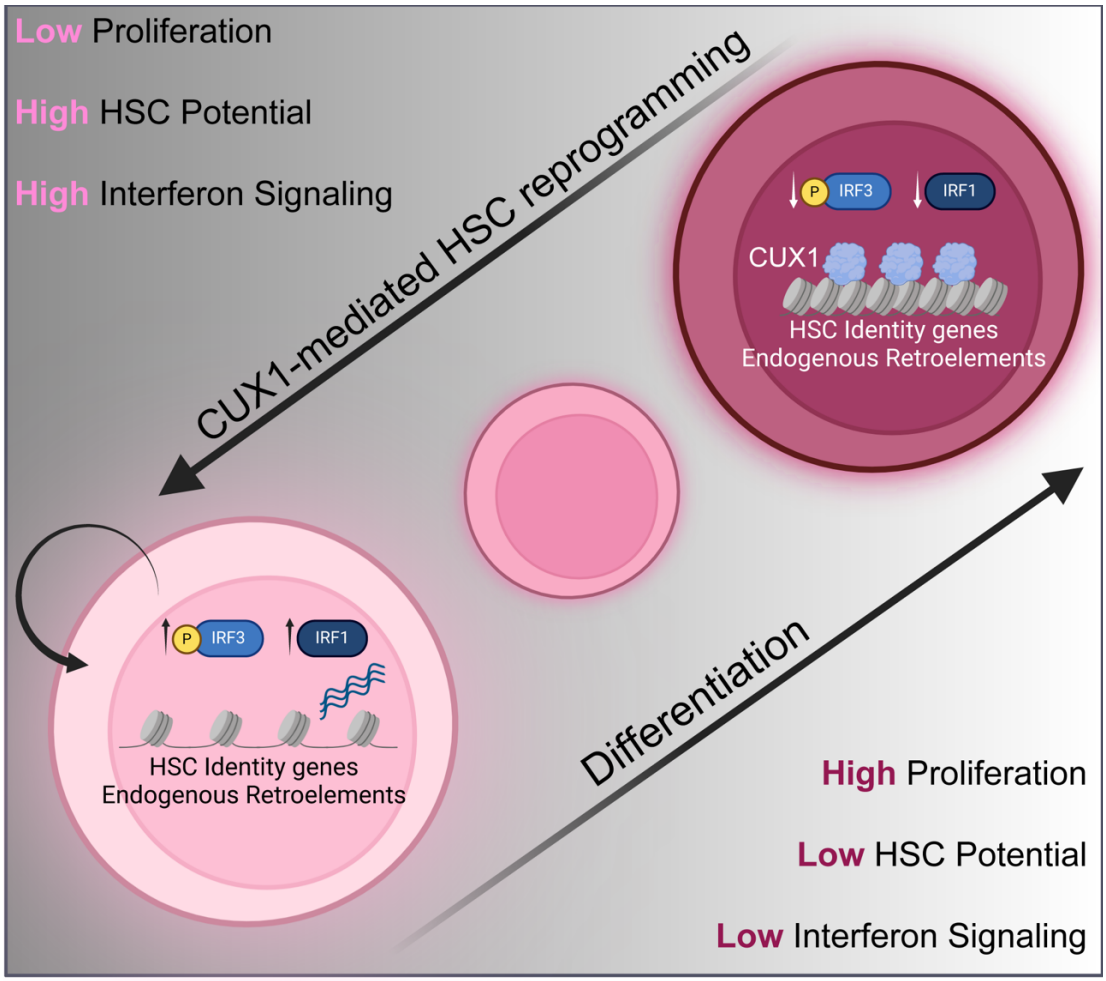

Figure 27. CUX1 controls HSC self-renewal through the regulation of innate immune activation

CUX1 levels segregate the LT-HSC compartment by cell cycle participation and HSC potential. As CUX1 levels increase, HSCs proliferate more and lose repopulating capacity. Low levels of CUX1 permit the expression of immunogenic endogenous retroelements, resulting in a cell-intrinsic interferon response. Knockdown of CUX1 reverses the progressive decline of HSCs and extends HSC potential, reinstating the transcriptome and chromatin accessibility landscape of HSCs with low levels of CUX1.

eliminate the protein, and three days are required to observe initial downstream gene expression changes. Therefore, it is unlikely that loss of CUX1 affects early engraftment. Finally, the increase in colony-forming potential of CUX1^{Bright} HSCs after knockdown *ex vivo* suggests that the *in vivo* self-renewal is not due to better engraftment.

HSCs count divisions

Our data strongly suggests that HSCs retain a memory of their division history, which is linked to CUX1 levels. HSCs can potentially divide without detectable loss of repopulating capacity, but stem cell activity in the daughter cells begins to decline with subsequent divisions[349-351]. Asymmetric inheritance of mitochondria and lysosomes throughout these cell divisions leads to the metabolic activation of subsequent daughter HSCs[287,292]. Considering our finding of dose-dependent cell cycle participation and corresponding loss of HSC activity linked to the transition from CUX1^{Dim} to CUX1^{Bright} HSCs, rising CUX1 levels could be a marker of HSC divisional history. Interestingly, CUX1 knockdown in CUX1^{Bright} HSCs could not reestablish the complete HSC activity of CUX1^{Dim} HSCs. Only knockdown of CUX1 in a CUX1^{Int} HSC achieved the HSC potential of CUX1^{Dim} HSCs. This finding was not explained by inefficient knockdown in cells that express more CUX1 (Figure S3F, S4D) or a difference in the gene expression and chromatin landscapes evoked by CUX1 knockdown relative to the starting dose. In general, we do not see as large of an effect size of CUX1 knockdown in a Bright HSC relative to a Dim HSC, so one explanation for the inefficiency of reprogramming could be that while the same genes are targeted, something inhibits recapitulating the magnitude of the CUX1^{Dim} state. However, the rapid downfall of formerly Bright HSCs upon restoration of CUX1 suggests that formerly Bright HSCs might remember that they were once Bright. This memory

might not be encoded in the epigenome but in the organelles they inherited through successive rounds of HSC division. Alternatively, the memory could be encoded in the shortened telomeres from previous cell divisions[352]. Forced low levels of CUX1 are permissive for extra self-renewal divisions. Still, the quality of inherited mitochondria and lysosomes or shortened telomeres could ultimately cause the exhaustion of formerly Bright HSCs.

Intrinsic and extrinsic viral response

Loss of CUX1 invokes the antiviral response, improving HSC reconstitution (Figure 3); however, chronic viral infection, exposure to viral ligands, or interferons induces HSC proliferation and a loss of HSC activity[351,353-359]. There are some important caveats to consider when comparing these results directly. First, the antiviral response resulting from the loss of CUX1 is cell-intrinsic, as opposed to viral infection or pIpC injection into an animal, which includes the direct effects of viral stimulation on HSCs and the inflammatory milieu created by other cells experiencing innate immune activation. Second, the duration, dynamics, and intensity of the viral stimulus can dictate cellular outcomes[134,163]. Exposure to IFNy or virus can have variable and dose-dependent cellular effects, mediated by specific genes exhibiting greater sensitivity to stimulation than others[355,358,360,361]. Furthermore, tonic interferon signaling is present normally[314,362,363] and required for tissue protection[314,364-369], cellular maturation[370,371], and maintenance of homeostasis[327,328,372,373]. Third, the antiviral response permitted by low levels of CUX1 is broadly suppressive of cell cycle participation, including genes such as *Ifitm1*[374], *Slfn2*[375], *Samhd1*[376], and *Ifi44*[377]. This program could give HSCs an additional mechanism to inhibit cell division and maintain HSC potential. In addition, the ISGs induced by the loss of CUX1 include negative regulators of innate signaling like *Parp14*[378] and *Oasl2*[379], suggesting

there might exist a level of signaling that optimally invigorates HSC potential and further emphasizes the role of dose in HSC fate. Indicative of an optimal level of innate immune signaling, HSCs carrying the DMT3A^{R882H} mutation exhibit hypomethylation at ERE loci and the accompanying innate immune activation, resulting in increased HSC potential. Treatment with the hypomethylating agent, azacytidine, results in further hypomethylation of EREs and a more pronounced interferon response, pushing mutant HSCs over the tipping point toward apoptosis[380]. Given the ubiquity of ISG expression in somatic stem cells, it is unsurprising that constitutive expression of this pathway has been manipulated to serve multiple functions: increased resistance to viral infection and maintenance of the stem cell state[314,381,382].

Dose-dependent control of HSCs

We demonstrated that CUX1 levels profoundly impact HSC fate decisions, highlighting the importance of transcription factor dose. Hematopoietic progenitors can normally adopt low levels of CUX1, but the inability to raise CUX1 levels forces progenitors to aberrantly self-renew and sets up the transition to malignancy. As an early event, loss of CUX1 expands the self-renewing pool of progenitors that can acquire additional mutations, such as a Ras family mutation leading to AML[25,383]. Even without additional mutations, our study demonstrates that the advantage to HSCs conferred by the loss of CUX1 is that they can undergo more self-renewal divisions than wild-type counterparts. Over time, this clone can expand so it is detectable in clonal hematopoiesis[384,385]. Or, as a late event in myeloid neoplasia, the loss of CUX1 can endow hematopoietic progenitors carrying mutations in *TET2*, *ASXL1*, or *BCOR* with additional capacity to self-renew[75]. This highlights that the dynamic regulation of CUX1 throughout hematopoiesis is essential for the continuous movement of hematopoietic progenitors toward

terminal differentiation. The inability to achieve the full dynamic range of CUX1 can be a step toward malignancy.

Somatic stem cells

Given the extra-hematopoietic defects in tissue homeostasis after CUX1 loss, CUX1 could play similar roles in somatic stem cells in other tissues[300-302]. Although the mechanism could be similar, the outcome is likely different. In mammary stem cells, the ensuing inflammation resulting from loss of EHMT2 and re-expression of EREs results in impaired stem cell behavior[332]. Similarly, loss of SETDB1 in muscle stem cells leads to c-GAS-STING mediated inflammation from ERE re-expression and impaired muscular regeneration[386]. This is more consistent with the defects observed in other tissues upon CUX1 loss.

Limitations of the study

In this study, we demonstrate the dose-dependent role of CUX1 in controlling HSC fate. While we have characterized some of the heterogeneity in the HSC, lineage tracing would have allowed us to characterize better the output of HSCs relative to the CUX1 dose for individual HSCs and resolve more heterogeneity in the HSC compartment. To interrogate the interaction between the starting CUX1 level and the effects of knockdown, we utilized *in vitro* models because, after knockdown, the original dose of CUX1 cannot be determined. This limited our abilities to discern the effects of CUX1 knockdown in the native setting without the stress of cell isolation and HSC transplant. In addition, there is a tradeoff when comparing the initial levels of CUX1 and the effects of knockdown. The control condition allows cells to naturally vary the levels of CUX1 after they have been sorted. As time increases, the deviation from the starting

dose of CUX1 increases, but the effects of CUX1 knockdown are more pronounced. Finally, we ascribe the inflammatory response upon CUX1 knockdown to the re-expression of EREs. Still, other inflammatory stimuli could contribute to the phenotype, such as DNA damage and mitochondrial DNA. Moreover, we ascribe the inflammatory response of EREs to be initiated by dsDNA recognition, but we have not ruled out a role for dsRNA recognition.

Methods

Experimental Models and Subject Details

Generation of CUX1 Reporter Mice

Cux1^{mCherry} and *Cux1*^{ATPN} reporter mice were generated using CRISPR-Cas9. Single cell-fertilized ova from B6.SJL/J mice were injected with a repair construct (12.5ng/μL) consisting of a 2.8kb 5' homology arm, the reporter sequence, and a 1 kb 3' homology arm along with a sgRNA targeting the stop codon of exon 24 of *Cux1* (5'-CCAUCGAAUGGGAGUUC-3') (25ng/μL) (IDT), Cas9 (100 ng/μL) (IDT). Repair constructs were purified from *E. coli* cultured for 8 hours in LB (Invitrogen) dissolved in cell culture-grade water (37°C, 250 rpm) using the EndoFree Plasmid Maxi Kit (Qiagen). Pups were screened for homologous repair of the locus by PCR and Sanger sequencing to confirm the entire sequence had been inserted into the genome without mutation. Transgenic mice were backcrossed with B6.SJL/J for one generation to generate F1 founders.

Forward	5'-GGGAGGAGCCCATCGAAT-3'
Reverse	5'-ATTGGGCTTAATGCTCCTTTGT-3'

Table 2. Genotyping Primers for *Cux1* Exon 24

Mouse models

All animal studies were approved by the University of Chicago Institutional Animal Care and Use Committee. The *Cux1*^{mCherry} reporter mice were crossed to the UBC-GFP mouse strain[312] (JAX00435) to track erythrocytes and platelets upon transplant into irradiated recipients. This model was crossed with previously published Col1a1^{Ren}; Rosa26^{rtTA}, Col1a1^{Mid}; Rosa26^{rtTA}, or Col1a1^{Low}; Rosa26^{rtTA} models[78,321]. C57BL6.SJL/J (CD45$^{.1/.1}$) mice were used as recipients in transplant studies. *Irf1*^{flox} mice were provided to us by the lab of Dr. Bana Jabri and crossed to the UBC-Cre-ERT2 mouse strain (JAX007001) as well as the CUX1 shRNA knockdown model. Mice of both sexes were used in all experiments. When necessary, mice of a single sex were

pooled to generate sufficient cell numbers. Mice used in these studies were 6-12 weeks of age at the start. CD45.1 mice were at least eight weeks old at the time of lethal irradiation. All mice were housed in specific-pathogen-free animal care facilities at the University of Chicago.

Experimental Diets

Control or *Cux1*-targeting shRNAs were induced by maintaining mice continuously on doxycycline-containing chow (TD.12006, 1 mg/kg, Envigo). Transplant recipient mice were started on doxycycline-containing chow two days prior to transplant. Cre-ERT2 was induced by placing mice on tamoxifen-containing chow (TD.130858, 500 mg/kg, Envigo) for five days at a time. Mice were allowed at least one week to recover prior to analysis.

Cell isolation

Mice were euthanized by carbon dioxide asphyxiation, and secondary euthanasia was performed by cervical dislocation. Hindlimb long bones, pelvic bones, and vertebrae were gently crushed by mortar and pestle in PBS (pH 7.2, Gibco) supplemented with 2 mM EDTA (Fisher) and 0.5% autoclaved BSA (Fisher) (FACS Buffer). Cells were then filtered through a 70uM filter (Fisher). Spleen and thymus were gently dissociated on a 40μM filter in FACS Buffer. Erythrocytes were lysed by incubation with Ammonium Chloride (Stem Cell Tech) for 5 minutes at room temperature. Cell number and viability were assessed by Cellometer (Nexcelom) and AOPI (Revity) staining.

Mice were bled at the indicated time points. At least 70μL of blood was collected by facial vein bleeding using a 5mm or 5.5mm lancet (Goldenrod) in tubes containing EDTA to prevent clotting (Sarstedt).

Primary cell culture

Primary murine HSCs were isolated by cell sorting and cultured in media from either Wilkinson et al.[387] [Ham's F-12 Nutrient Mix (Gibco), 1x Penicillin/Streptomycin/Glutamine (Gibco), 10mM HEPES (Gibco), 1mg/mL Polyvinyl alcohol (), 1x Insulin-Transferrin-Selenium-Ethanolamine (Gibco), 100ng/mL TPO (Peprotech), 10ng/mL SCF (Peprotech), 1mg/mL Gentamicin (Gibco), 1µg/mL doxycycline] or Loeffler et al.[292] [Phenol-free IMDM (Gibco), 20% BSA-Insulin-Transferrin 9500 (Stem Cell Tech), 50µM β-Mecaptoethanol (Sigma), 100ng/mL TPO (Peprotech), 100ng/mL SCF (Peprotech), 1x Penicillin/Streptomycin/Glutamine (2mM glutamine) (Gibco), 1mg/mL Gentamicin (Gibco), 1µg/mL doxycycline (Fisher)]. Cultures were maintained at 37°C and 5% CO_2.

Method Details

Flow Cytometry and Cell Sorting

Bone Marrow: We used the Direct Lineage Cell Depletion Kit (Miltenyi, 130-110-470) to enrich for hematopoietic stem and progenitor cells. Approximately 80 million BMMCs were incubated with 25µL of beads for 10 minutes on ice, and then the fractions were separated on an LS column (Miltenyi, 130-042-401). The lineage-negative fraction was incubated with True Stain FcX Plus (Biolegend), followed by incubation with anti-lineage antibodies: anti-CD11b—Biotin (M1/70), anti-CD19—Biotin (1D3), anti-CD3e—Biotin (145-211C), anti-CD45R—Biotin (RA3-6B2), anti-CD5—Biotin (53-7.3), anti-Ly6C/Ly6G—Biotin (RB6-8C5), anti-Ter119—Biotin (TER119). For cell sorting, cells were stained with the following antibodies: Streptavidin-BUV395, anti-CD150—BV421 (TC15-12F.2), anti-CD117—BV650 (ACK2), anti-CD135—PerCP-eFlour 710 (A2F10), anti-CD48—PE-Cy7 (HM48-1), and anti-Sca1—APC-Cy7 (D7). Monoclonal conjugate antibody staining was performed for 30 minutes on ice. Sorting was performed on a BD FACS Aria Fusion (BD) running BD FACSDiva (BD). Cells were sorted

into sterile FACS Buffer for transplants and CFU assays, media for in vitro differentiation assays, and lysis buffer for RNA isolation. Representative purity checks were acquired for each experiment to ensure proper separation of populations. For flow cytometry analysis, cells were stained with the following antibodies: anti-CD16/32—BUV395 (2.4G2), Streptavidin—BUV737, anti-CD150—BV421 (TC15-12F.2), anti-CD117—BV650 (ACK2), anti-CD127—BV785 (A7R34), anti-CD135—PerCP-eFlour 710 (A2F10), anti-CD48—PE-Cy7 (HM48-1), anti-CD34—eFluor 660 (RAM34), and anti-Sca1—APC-Cy7 (D7). In certain panels, the following antibodies were also used: anti-CD201—PE (RCR-16), anti-CD49b—FITC (HMa2), anti-CD41—FITC (MWReg30), anti-CD45.2—FITC (104), anti-CD45.1—BUV563 (A20), and anti-I-E/I-A—AlexaFluor 700 (M5/114.15.2). Samples were acquired on a five-laser Aurora (Cytek) running SpectroFlo (Cytek).

Spleen and peripheral blood: Cells were incubated with True Stain FcX Plus (Biolegend). For flow cytometry analysis, cells were stained with the following antibodies: anti-Ly6G—BUV395 (1A8), anti-CD3ε—BUV737 or anti-CD3ε—APC (145-2C11), anti-TCRβ—BV421 (H57-597), anti-FcεRIα—PacBlue (Mar-1), anti-CD11c—BV570 (N418), anti-CD49b—BV605 (HMa2), anti-Ly6C—BV650 (HK1.4), anti-NK1.1—BV711 (PK136), anti-SiglecF—BV786 (E50-2440), anti-CD45R—PerCp-Cy5.5 (RA3-6B2), anti-CD200R3—PE (Ba13), anti-CD8α—PE-Cy7 or anti-CD8α—PerCP (53-6.7), anti-CD4—APC or anti-CD4—BUV563 (GK1.5), anti-I-E/I-A—AlexaFluor 700 (M5/114.15.2), anti-CD11b—APC-Cy7 (M1/70), anti-TCRγδ—PE-Cy7 (GL3), anti-CD45.1—PE-Cy7 (A20), anti-CD41—AlexaFluor 700 (MWReg30) and anti-CD45.2—BUV737 (104). Prior to erythrocyte lysis for peripheral blood samples in FACS lysis buffer (BD), 0.5µL of whole blood was stained was stained with anti-Ter119—PE-Cy5 (TER119).

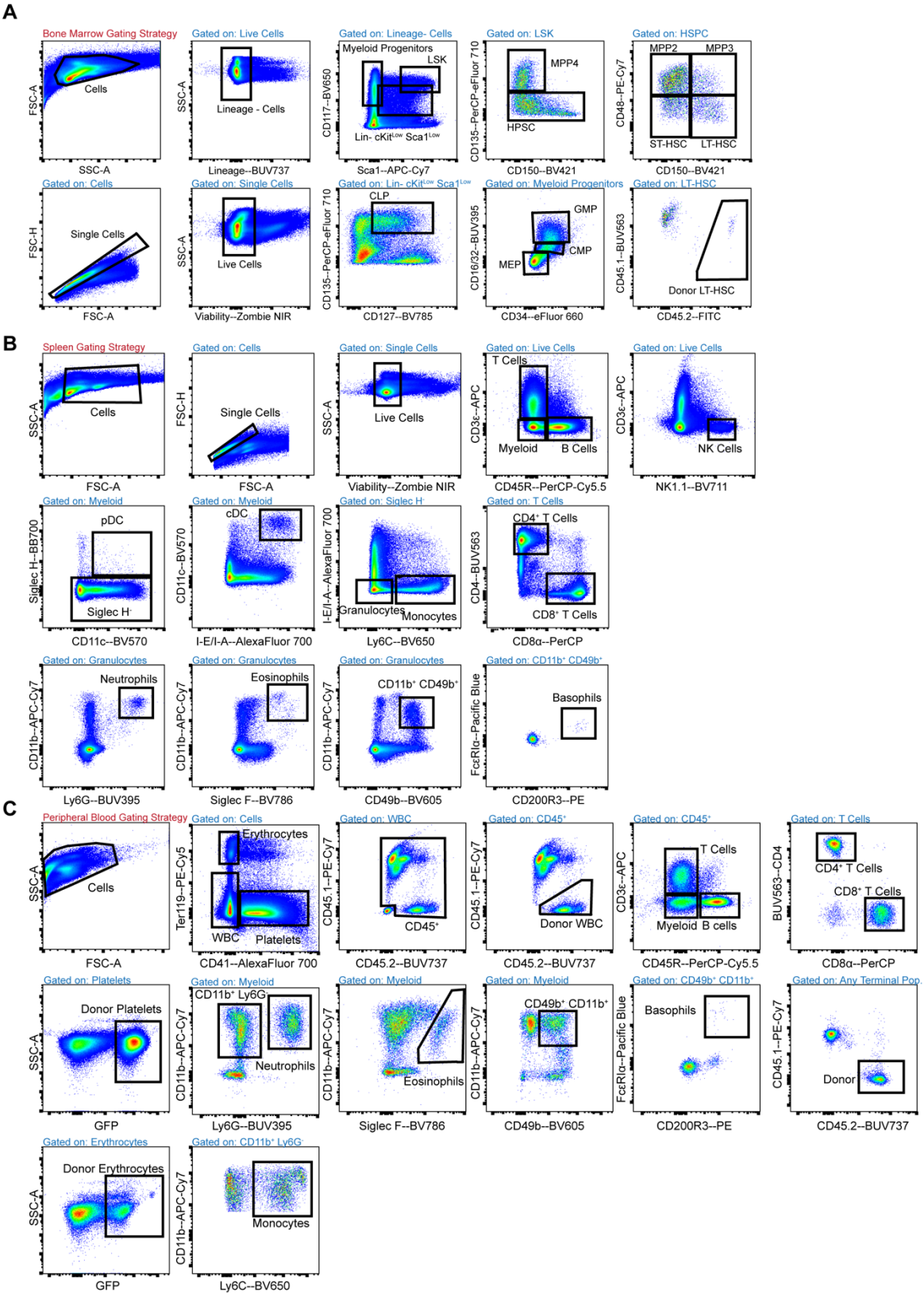

Figure 28. Bone marrow, spleen, and peripheral blood gating schemes

Figure 28, continued
(A) Bone marrow hematopoietic stem and progenitor gating scheme. All populations are gated on single cells > live cells > lineage negative cells. Populations are defined as the following:
Common Lymphoid Progenitor (CLP): $CD117^{Low}$ $Sca1^{Low}$ $CD127^{+}$ $CD135^{+}$
Common Myeloid Progenitor (CMP): $CD117^{High}$ $Sca1^{-}$ $CD34^{+}$ $CD16/32^{Low}$
Megakaryocyte Erythrocyte Progenitor (MEP): $CD117^{High}$ $Sca1^{-}$ $CD34^{-}$ $CD16/32^{-}$
Granulocyte Monocyte Progenitor (GMP): $CD117^{High}$ $Sca1^{-}$ $CD34^{+}$ $CD16/32^{High}$
MPP4: $CD117^{High}$ $Sca1^{High}$ $CD135^{+}$
MPP3: $CD117^{High}$ $Sca1^{High}$ $CD135^{-}$ $CD48^{+}$ $CD150^{-}$
MPP2: $CD117^{High}$ $Sca1^{High}$ $CD135^{-}$ $CD48^{+}$ $CD150^{+}$
ST-HSC: $CD117^{High}$ $Sca1^{High}$ $CD135^{-}$ $CD48^{-}$ $CD150^{-}$
LT-HSC: $CD117^{High}$ $Sca1^{High}$ $CD135^{-}$ $CD48^{-}$ $CD150^{+}$
(B) Spleen gating strategy. All populations are gated on single cells > live cells. Populations are defined as the following:
CD4 T cells: $CD3^{+}$ $CD4^{+}$
CD8 T cells: $CD3^{+}$ $CD8^{+}$
B cells: $CD45R^{+}$
NK cells: $NK1.1^{+}$
pDC: CD3- $CD45R^{-}$ $SiglecH^{+}$ $CD11c^{+}$
cDC: $CD3^{-}$ $CD45R^{-}$ $CD11c^{+}$ $I\text{-}E/I\text{-}A^{+}$
Monocyte: $CD3^{-}$ $CD45R^{-}$ $SiglecH^{-}$ $Ly6C^{+}$
Neutrophil: $CD3^{-}$ $CD45R^{-}$ $SiglecH^{-}$ $Ly6C^{-}$ $I\text{-}E/I\text{-}A^{-}$ $CD11b^{+}$ $Ly6G^{+}$
Eosinophil: $CD3^{-}$ $CD45R^{-}$ $SiglecH^{-}$ $Ly6C^{-}$ $I\text{-}E/I\text{-}A^{-}$ $CD11b^{+}$ $SiglecF^{+}$
Basophil: $CD3^{-}$ $CD45R^{-}$ $SiglecH^{-}$ $Ly6C^{-}$ $I\text{-}E/I\text{-}A^{-}$ $CD11b^{+}$ $CD49b^{+}$ $CD200R3^{+}$ $FceRIa^{+}$
(C) Peripheral Blood Gating Strategy: Populations are defined as follows:
Erythrocyte: $Ter119^{+}$
Platelet: $CD41^{+}$
CD4 T cell: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD3^{+}$ $CD4^{+}$
CD8 T cell: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD3^{+}$ $CD8^{+}$
B cell: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD45R^{+}$
Monocyte: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD45R^{-}$ $CD3^{-}$ $CD11b^{+}$ $Ly6G^{-}$ $Ly6C^{+}$
Neutrophil: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD45R^{-}$ $CD3^{-}$ $CD11b^{+}$ $Ly6G^{+}$
Eosinophil: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD45R^{-}$ $CD3^{-}$ $CD11b^{+}$ $SiglecF^{+}$
Basophil: $Ter119^{-}$ $CD41^{-}$ $CD45^{+}$ $CD45R^{-}$ $CD3^{-}$ $CD11b^{+}$ $CD49b^{+}$ $CD200R3^{+}$ $FceRIa^{+}$

Complete Blood Count

Complete blood counts were acquired from approximately 20µL of blood on a Hemavet 950.

Cell cycle analysis

For *in vivo* BrdU labeling, mice were injected 3.33mg BrdU (Biolegend) on day zero and then continuously maintained on water containing 0.8 mg/mL BrdU (Sigma) with 1% surcorse (Sigma). For *ex vivo* BrdU labeling, cells were labeled with 1mM BrdU in Wilkinson et al.[387] media for five hours. Lin- cells were stained as described above. Fixation, permeabilization, and intracellular staining were performed according to the BD Pharmingen Apoptosis, DNA Damage, and Proliferation Kit (BD). Briefly, cells were fixed in CytoFix/CytoPerm buffer at room temperature, followed by permeabilization in CytoPerm+ Buffer on ice and re-fixation in CytoFix/CytoPerm Buffer. Then, cells were treated with 300µg/mL DNase for one hour at 37°C. Finally, cells were stained with anti-BrdU—AlexaFlour700 or anti-BrdU—APC (3D4) and anti-Ki67—FITC or anti-Ki67—PE.

Serial colony-forming-unit assays (CFU)

150 LT-HSCs were sorted into FACS Buffer and cultured in Methocult M3434 (Stem Cell Tech) with 1µg/mL doxycycline and indicated inhibitor concentrations. Colonies were counted and scored ten days after primary plating. To passage, 10,000 cells (control shRNA) or 2,000 cells (CUX1-targetting shRNA) were re-plated in M3434. Colonies were quantified seven days after secondary plating.

In vitro differentiation assay

Approximately 100 LT-HSCs were sorted into Loeffler et al.[292] media with 1µg/mL doxycycline and indicated inhibitor concentrations and cultured for five days. Cells were stained along with approximately 500,000 thymocytes. For flow cytometry analysis, cells were stained

with the following antibodies: anti-CD11b—BUV395 (M1/70), anti-CD45.2—BUV737 (104), anti-CD150—BV421 (TC15-12F.2), anti-CD74—BV480 (In-1), anti-CD274—BV785 (10F.9G2), anti-Sca1—PE-Cy7 (D7), anti-CD201—APC(RCR-16), anti-CD41—AlexaFluor 700 (MWReg30), and anti-CD71—APC-Fire750 (RI7217).

Intracellular flow cytometry

For ATP5A staining, cells were fixed and permeabilized according to the intracellular staining protocol for BrdU and Ki67. Fixed cells were incubated with anti-ATP5A—FITC (15H4C4) for 30 minutes at room temperature.

For IRF1 and pIRF3 staining, sorted HSCs were expanded in Wilkinson et al. media. For staining, cultured HSCs were mixed with CD45.1^{+} thymocytes as carrier cells. Cells were fixed with 2% paraformaldehyde at room temperature for 15 minutes and permeabilized with 95% ice-cold methanol on ice for 30 minutes. Prior to antibody staining, cells were re-hydrated in FACS buffer for one hour at 4°C. Primary antibody incubation was performed for one hour at room temperature for pIRF3 (S396) (CST, D6O1M). After two stringent washes to remove excess antibody, cells were stained for one hour at room temperature with a combination of the following antibodies: donkey-anti-rabbit—AlexaFluor 647 (Millipore, SAB4600177), anti-CD45.2—BUV737 (104), anti-CD8α—PerCP (53-6.7), and anti-IRF1-PE (CST, D5E4).

Ex vivo OPP-incorporation assay

Protein Synthesis rates were quantified using the Click-iT Plus OPP Protein Synthesis Assay Kit (Life Technologies) with the protocol for hematopoietic stem and progenitor cells[388]. Breifly, Lin- cells were cultured in IMDM for 30 minutes at 37°C with 20μM OP-Puro. Cells were then fixed at room temperature for 15 minutes. Then, the Click reaction was performed

according to the kit parameters. Finally, cells were stained with appropriate surface markers for LT-HSCs as above.

HSC competitive transplantation

100 LT-HSCs were sorted into sterile FACS Buffer and combined with 200,000 CD45.1$^+$ BMMCs and transplanted into lethally irradiated (8.5Gy) CD45.1$^+$ recipients by retroorbital injection at least 4 hours and up to 24 hours post-irradiation. In primary transplants, peripheral blood engraftment was measured by flow cytometry at 2-, 4-, 6-, 9-, 12-, and 16 weeks post-transplant. At 16 weeks post-transplant, bone marrow from the two mice closest to the median level of engraftment per condition was harvested and combined 1:1. 1x10^6 BMMCs were then transplanted into lethally irradiated secondary recipients. This was repeated for tertiary transplants. In secondary and tertiary transplants, peripheral blood engraftment was measured by flow cytometry at 4-, 8-, 12-, and 16 weeks post-transplant.

Homing Assay

10,000 LSKs segmented on CUX1 level were transplanted into lethally irradiated recipients (8.5Gy) 24 hours after irradiation. 16 hours after transplant, mice were sacrificed and the bone marrow and spleen were harvested to analyze donor cell engraftment by flow cytometry.

Western Blot

An equivalent number of thymocytes from *Cux1*$^{+/+}$, *Cux1*$^{mCherry/+}$, and *Cux1*$^{mCherry/mCherry}$ mice were lysed in RIPA Buffer and prepared for immunoblotting as previously reported[78]. For protein detection, anti-CUX1 (ABE-217, 1:2000), and anti-mCherry (Abcam, 1:2000) were followed by anti-rabbit-HRP (Abcam, 1:4000) and anti-β-actin-HRP (Santa Cruz, 1:2500) was used to normalize protein levels.

Cycloheximide treatment

Thymocytes from *Cux1*$^{+/+}$ and *Cux1*$^{mCherry/mCherry}$ mice were cultured in RPMI with 50mg/mL cycloheximide (Sigma) for up to 12 hours. At the indicated intervals, cells were collected and lysed in RIPA buffer. All lysates were used for western blot probing for CUX1 and *β*-actin. Quantification was performed in Image Lab (BioRad). Briefly, the intensity of each CUX1 and *β*-actin band was quantified and normalized to the background. Relative CUX1 protein levels were determined by normalizing CUX1 bands to the corresponding *β*-actin band.

Immunofluorescence

To observe intracellular CUX1, thymocytes were isolated from WT *Cux1*^{mCherry} mice. Cells were allowed to settle on poly-D-lysine coated coverslips for 30 min at room temperature, then fixed in 4% paraformaldehyde for 20 min, quenched with 50mM Ammonium Chloride for 10 min and permeabilized with 0.2% Triton X-100 for 4 min. After washing three times with PBS, cells were blocked with PBS plus 0.1% bovine serum albumin and 1% donkey serum (blocking buffer) for 30 min. CUX1 was stained with a rabbit polyclonal anti-mCherry antibody for 60 min followed by a CF™647-conjugated anti-rabbit secondary antibody for 30 min. Cells were washed three time with PBS containing 10% blocking buffer after primary and secondary antibody stains. Coverslips were mounted with Prolong Diamond Antifade mountant with DAPI (Invitrogen). Images were acquired with a Zeiss AxioObserver 7 inverted widefield microscope using a 63X objective under oil immersion.

qPCR

The ratio of nuclear DNA to mitochondrial DNA was quantified using qPCR[313]. Briefly, 1000 segmented LT-HSCs were sorted into lysis buffer[389] and incubated at 56°C for two hours followed by 95°C for 15 minutes. The qPCR reaction was performed using SYBR Green qPCR

Master Mix (Applied Biosystems). The ΔΔCt method was used to calculate the ratio of nuclear DNA to mitochondrial DNA.

mCytB Forward	5'-CTTCATGTCGGACGAGGCTTA-3'
mCytB Reverse	5'-TGTGGCTATGACTGCGAACA-3'
mActB Forward	5'-CGGCTTGCGGGTGTTAAAAG-3'
mActB Reverse	5'-CGTGATCGTAGCGTCTGGTT-3'

Table 3. Primers from de Almeida et al. 2017[313]

Chemicals and Reagents

Tenofovir (Selleck) and Emtricitabine (Selleck) were reconstituted in sterile water at 5mM and 50mM stocks, respectively. Lamivudine (TCI America) was reconstituted in DMSO at 100mM stock. C-178 (Sigma) and GSK8612 (Selleck) were reconstituted in DMSO at 5mM and 10mM, respectively.

Bulk RNA-Seq

Generation

Cux1^{mCherry}: Segmented LT-HSCs were sorted directly into lysis buffer supplemented with *β*-ME from the Absolutely RNA Nanoprep Kit (Agilent) and combined with 80% sulfolane (Sigma). RNA was isolated according to the manufacturer's instructions and eluted in 14μL of elution buffer. RNA concentration and integrity were assessed using the Bioanalyzer RNAPico assay (Agilent). Libraries were prepared using the SMART-Seq v4 Ultra Low Input RNA Kit for sequencing (Takara) and sequenced to at least 60 million read depth 100bp paired-end on the Illumina NovaSeqX platform.

Cux1^{mCherry} + Knockdown: Segmented LT-HSCs were sorted into Wilkinson et al.[387] media with 1μg/mL doxycycline and cultured for 3 days in at least three biological replicates. RNA libraries were prepared and sequenced as above.

Low-level processing

Reads were aligned to the mm10 or hg19 genome using STAR[390] standard parameters for single aligning reads. Transcripts were counted using featureCounts[391]. For repetitive element analysis, we implemented the best practices for retroelement analysis[392] with previously described modifications[393]. For alignment, the following parameters were implemented in STAR to account for multi-mapping: outFilterMultimapNmax 1000, outSAMmultNmax 1, outMultimapperOrder Random, and winAnchorMultimapNmax 1000. For TE quantification, multimapping was allowed in featureCounts. Annotated GTF files of repeats from UCSC or Hammell Lab were used to aid in repetitive element quantification in featureCounts and assignment of repetitive elements to classes and families.

Downstream analysis

Differential expression testing was performed using DESeq2[394] using the LRT test for mouse datasets and the Wald Test for human AML samples from McNerney et al.[76], which were corrected for sex, diagnosis, and tissue. Genes were considered differentially expressed if FDR < 0.01 and log2FC > 0.58 for mouse datasets and p-value < 0.05 log2FC > 0.58 for human repetitive elements. Gene set enrichment analysis was performed using fgsea[395]. Pathways were considered significantly enriched at FDR < 0.1. K means clustering was performed in R using kmeans with default parameters. Heatmaps were generated using ComplexHeatmap[396] and pheatmap. DAVID[397] was used for Gene Ontology analysis.

ATAC-Seq

Generation

Segmented LT-HSCs were sorted in Loeffler et al.[292] media and cultured for five days in three biological replicates. Libraries were prepared according to the OMNI-ATAC-Seq protocol[398,399]. Briefly, nuclei were isolated from roughly 50,000 cells in cell lysis buffer (10mM Tris-HCl,

10mM NaCl, 3mM $MgCl_2$, 0.1% NP-40, 0.1% Tween-20, 0.01% Digitonin). After washing, the transposition reaction was carried out using the Nextera DNA Library Prep Kit (Nextera) for 30 minutes at 37°C in a 50µL reaction (25µL 2X TD buffer, 16.5µL PBS, 0.1% Tween-20, 0.01% Digitonin, 2.5µL Tn5 Transposase, 5µL water). DNA was isolated using the Qiagen MinElute Reaction Cleanup Kit (Qiagen) and amplified using NEBNext High-Fidelity PCR Master Mix (NEB) for five cycles of PCR and adjusted using SYBR Green qPCR Master Mix (Applied Biosystems). Left-sided size selection was performed on libraries using AMPure XP beads (Beckman Coulter). Libraries were sequenced using Illumina Sequencing (NovaSeqX) to at least 50 million read depth per sample, 100bp paired-end.

Low-level processing

FASTQ reads were aligned to the mm10 genome using bwa-mem[400]. Reads were filtered to exclude low quality (q<30), mitochondrial, blacklist (mm10 v2)[401], and PCR duplicate reads using samtools[402].

Downstream analysis

Differential accessibility was performed using the edgeR[403] implementation in csaw[404]. Briefly, read depth was calculated in non-overlapping, sliding 150bp windows across the genome. Low abundance windows (< 3-fold increase over global background) and those containing assembly gaps were removed from analysis. Libraries were normalized using normOffsets to account for foreground and background windows. After differential accessibility testing, significant windows were merged if they were within 300bp. Multiple testing correction was performed using combineTests. Peaks were considered differentially accessible if FDR <0.01 and log2FC > 1 for any of the six possible comparisons from the four conditions. Differential motifs were calculated using monaLisa[405] with binary motif analysis. Peaks were

assigned to the nearest genes using rGREAT[406] in oneClosest mode. DAVID[397] was used to generate significantly enriched pathways for assigned nearest genes.

scRNA-Seq

Generation

Cux1^{ATPN}: Prior to the published experiment, all antibody-DNA tags (ADTs) were titrated across two log-fold concentrations by scRNA-Seq (10X Genomics) to determine the concentration that achieved the best separation of positive and negative populations. For hNGFR, HSCs were combined from *Cux1*$^{+/+}$, *Cux1*$^{ATPN/+}$, and *Cux1*$^{ATPN/ATPN}$ mice and sequenced to ensure that scRNA-Seq was sensitive enough to separate negative cells from those expressing low levels of CUX1 (hNGFR) from high expressing cells.

In the presented experiment, Lin- cells were isolated from four separate female *Cux1*$^{ATPN/ATPN}$ mice. Lineage-negative cells from each mouse were brought to a concentration of $1x10^6$ cells/100µL and were stained with a unique cell multiplexing oligo (CMO) (Biolegend) as well as fluorescent antibodies for sorting and the following ADTs: anti-NGFR—TotalSeqB0816 (ME20.4) [2ng*µL^{-1}], anti-CD201—TotalSeqB0439 (RCR-16) [2ng*µL^{-1}], anti-CD74—TotalSeqB1043 (In-1) [2.5ng*µL^{-1}], anti-CD274—TotalSeqB0190 (MIH6) [2ng*µL^{-1}], anti-CD86—TotalSeqB0200 (GL-1) [1ng*µL^{-1}], anti-I-E/I-A—TotalSeqB0117 (M5/114.15.2) [2.5ng*µL^{-1}], and Mouse IgG1κ Isotype Control—TotalSeqB0090 (MOPC-21) [2ng*µL^{-1}]. Three stringent washes were performed to remove any additional unbound antibodies before combining all four samples for sorting. Two samples were sorted for library preparation. One sample contained only sorted immunophenotypic LT-HSCs. The other sample contained a pooled population of sorted ST-HSCs (LSK, CD135$^-$, CD48$^-$, CD150$^-$) MPPs (LSK, CD48$^+$), and MPs (Lin$^-$, cKit$^+$) to enrich for more rare cells. Samples were confirmed to have >94% viability by trypan blue exclusion before libraries were prepared according to the manufacturer's protocol

using the Chromium Single Cell 3' Reagent Kits (v3). Quality assessment of libraries was performed by Agilent Bioanalyzer (Agilent) and sequenced on the Illumina NovaSeqX platform using the 29-10-10-89 run type.

Cux1^{mCherry} + Knockdown: Segmented LT-HSCs were sorted into Loeffler et al.[292] media and cultured for five days in two biological replicates (one male and one female). Each condition—Cux1^{Dim} Control, CUX1^{Dim} Knockdown, CUX1^{Bright} Control, and CUX1^{Bright} Knockdown—was labeled with a unique CMO (Biolegend). Three stringent washes were performed to remove additional unbound antibodies before combining all four conditions from each biological replicate for library preparation. Each sample was confirmed to have >90% viability by trypan blue exclusion before libraries were prepared, as described above.

Low-level processing

CMO and ADT libraries were generated together; however, for disambiguation, libraries were duplicated, and the first four sequencing reads were removed as recommended by 10X Genomics. The cellranger 7.1.0 multi pipeline (10X Genomics) was used to align FASTQ reads and generate the Gene Expression, ADT, and CMO matrices. Velocyto was used to generate matrices of spliced and unspliced counts[407].

Downstream analysis

Seurat (v5.0.2)[408] was used for filtering, scaling, normalization, dimensionality reduction, and visualization. Cells were removed from downstream analysis if mitochondrial genes exceeded 8% of all mRNAs and if more than 8000 or fewer than 200 genes were detected for a given cell. Genes were removed if they were not detected in at least ten cells. SCTransform[409] was used to normalize and scale the gene expression matrix, regressing the fraction of mitochondrial gene expression. HTODmux[410] with the CLR normalization method was used to assign cells to CMOs. Cells assigned two more than CMO were removed from further analysis as

doublets. ADT data was normalized using the Relative Counts method and scaled regressing the fraction of mitochondrial gene expression. FindVariableFeatures was used to determine the 5000 most variable genes and linear dimensionality reduction was performed using RunPCA. Cell clusters were visualized using RunUMAP using the first 50 principal components. Two small clusters of mature B cells and dendritic cells were removed from the progenitor data set because they lacked continuous progenitors (Figure 1). Gene expression signatures were visualized using AddModuleScore and UCell[411] to confirm results were not due to the relative nature of AddModuleScore. Pearson correlations of gene expression signatures and surface CUX1 counts with cor.test in R with Benjamini-Hochberg correction for multiple hypothesis testing.

Monocle3[412] was used for pseudotime trajectory inference. The root was defined at the node where Giladi et al.[52] HSC signature was the highest. To run Palantir[307] pseudotime analysis, the count matrix and dimensionality reduction from Seurat were imported into scanpy[413]. The gene expression matrix was imputed with Magic[414]. The starting cell was set to the cell with the highest Giladi et al. HSC score. Four terminal cells were defined for each of the following lineages: monocyte, neutrophil, megakaryocyte, and erythrocyte. Terminal cells were selected as the cells of each lineage with the highest score of relevant gene signatures from Giladi et al[52].

Prior to implementing Hotspot[317] for module detection, ribosomal genes were removed from the matrix. Hotspot was initialized with the filtered counts matrix and dimensionality reduction matrix from Seurat. For module detection, the 1000 most correlated genes were used to assign modules with a minimum membership of 50 genes and all genes forced into a module (core_only=False). Module names were derived from gene ontology analysis from DAVID[397].

Prior to RNA velocity analysis with TopicVelo[322], tomotopy was used to calculate coherence and determine the number of topics in the data. FastTopics[415] was used to determine

which genes co-occur in topics. Smoothening of the spliced_unspliced count matrix and determination of neighbors was performed with scVelo[416]. Topics were included if they represented biologically meaningful representations of the data. Thresholds were determined by KL divergence. Transition matrices were calculated on the combined data set, as well as independently on control and knockdown cells.

Quantification and Statistical Analysis

The figure legends accompanying each figure indicate the sample size, statistical test, and definition of error bars. Sample size was determined empirically. Flow cytometry data were analyzed using FlowJo 10 (BD). In transplantation experiments, recipient animals were excluded on the basis of low engraftment at the first time point (all lineages < 2% engraftment). All statistical tests were performed using Prism 10 (GraphPad) or R (4.2.3). Each experiment combines at least two biological replicates unless otherwise noted.

Discussion

Endogenous Retroelements and Chromatin

Expression

First described by Barbara McClintock, transposable elements (TEs) are segments encoded in the genome that can insert themselves into another location in the genome[417]. Endogenous retroelements (EREs) are the remnants of exogenous retroviruses that infected vertebrate germ cells millions of years ago[418]. There are two mechanisms by which EREs can relocate. DNA transposons relocate through a cut-and-paste mechanism, in which the element excises and inserts itself in a new location. These elements do not duplicate themselves, and mammalian genomes have no active DNA transposons[419]. Non-LTR transposable elements, like LINE1 and Alu elements, use the copy-and-paste mechanism that progresses through an RNA intermediate to duplicate themselves in the genome. Similarly, LTR-containing transposable elements (endogenous retroelements) act through a similar copy-and-paste mechanism with an obligate dsDNA intermediate for genome duplication[420]. Millions of years of duplication have resulted in 40-50% of the human genome constituting insertions of transposable elements[421]. Over time and without positive selection, retroelement duplications acquire mutations and lose the ability to replicate[420]. So now, most retroelements are no longer replicating, but some are still actively duplicating, more so in mice than humans[422]. Once thought of as junk DNA, it is becoming increasingly clear that endogenous retroelements have a role in normal physiology and can be exploited in disease states[423,424].

Class I endogenous retroelements are TEs that use the copy-and-paste mechanism and, therefore, progress through an RNA intermediate and DNA intermediate[340]. LTR-containing EREs contain the same gene structure as a typical retroviral genome, comprising three open

reading frames (ORF). The *pro-pol* ORF is required for duplication because it encodes the polyprotein that is cleaved into reverse transcriptase and integrase[340]. The two remaining ORFs, *gag, and env,* are needed to form infectious viral particles, which is generally a rare and less successful duplication mechanism[425]. All stages of the viral lifecycle can trigger the innate immune system[340]. LTRs can promote bidirectional transcription, resulting in the annealing of repetitive RNAs, forming dsRNA, and activating the RIG-I receptor MDA5[338,426]. Reverse transcription of viral RNAs results in the formation of dsDNA in the cytoplasm, triggering the cGAS-STING pathway[427]. And, theoretically, incoming retrovirus could trigger TLR7/TLR8[428]. The induction of the active antiviral response by EREs has been termed viral mimicry and has been associated with many different cellular phenotypes[429,430]. The current paradigm on EREs states that silencing these genomic elements is pertinent for two aims: preventing their duplication and avoiding the antiviral response that their expression elicits. The following discussion on EREs will suggest that this paradigm misses important homeostatic exaptations of EREs.

Before considering the repression mechanisms of EREs, it is important to broach the idea that while repression is essential in many contexts, we live in homeostasis with EREs. EREs are ubiquitously expressed across human tissues[423]. Expression from up to 10,000 ERE loci can be detected across normal tissues, and ERE transcripts constitute up to 2% of all cellular transcripts[431,432]. The HERV-K proviruses with intact *gag*, *pro*, and *pol* open reading frames can be detected in all 54 assayed tissues of one study[433]. However, ERE expression is largely tissue-specific. Different EREs or proviruses deme are detected in various tissues, likely due to the different transcription factor binding sites in the 5'LTR and the differences in the expression of surrounding genes. Specific tissues may have different sensitivities to the expression level of

EREs, but they are ubiquitously expressed, nonetheless. For example, HML-2 provirus expression is highest in the testis and cerebellum[433]. These studies highlight the extent of ERE expression in healthy tissues and suggest that a view of EREs where they are entirely silenced likely lacks explanatory power.

As an example of the homeostatic nature of EREs, deficiency of the 3'-5' DNA exonuclease, *TREX1*, can result in systemic lupus erythematosus (SLE)[434] or Aicardi-Goutières Syndrome (AGS)[435]. One role of TREX1 is the degradation of endogenous retroelements, preventing innate immune system activation[436]. In the absence of this enzyme, persistent sterile inflammation leads to the development of autoimmunity[437,438]. Serendipitously, a combination of nucleoside reverse transcriptase inhibitors (NRTIs) can reduce the interferon signature in patients with a deficiency in *TREX1*[439,440]. The need for proper expression of TREX1 highlights the paradigm that avoiding the adverse effects of retroelement expression is ongoing and not entirely reliant on repression at the DNA level.

Another example of the ERE function in homeostasis highlights the exaptation of EREs. Exaptation describes when a trait acquires a new function outside of the original function. In the developing human embryo at the eight-cell stage, the pluripotency transcription factor OCT4 drives the expression of one of the most recently acquired EREs, HERVK (HML2). RNA transcripts of HERVK (HML2) are translated, and viral particles can be detected at the blastocyst stage. Expression of this ERE induces an antiviral response in the blastocyst, which is sufficient to prevent infection. The functions of this ERE extend further than the consequences of expression. Viral proteins modulate the translation of early expressed genes[441]. This example highlights the fact that EREs are expressed normally, even in the developing human embryo, and EREs have gained new purposes in development and gene regulation.

Further considering the functions of EREs, it would be enlightening to think about the contexts in which EREs are expressed. The primary mechanism of ERE expression is likely leaky expression. Given that the lifecycle of an ERE depends on having a good promoter to be expressed, 5' LTRs can be particularly strong promoters. In fact, LTRs have evolved to be binding sites for transcription factors to increase expression[442]. However, ERE expression can also be induced. Epstein-Bar Viral infection can induce the expression of EREs[443]. Along these lines, TLR ligands like LPS (TLR4) and flagellin (TLR5) can also induce ERE expression[372,444,445]. HIV infection can also induce the expression of HML-2 (HERKV) provirus, and proviral genes can interfere with HIV infection *in vitro*[446,447]. This highlights a prominent hypothesis in the benefits of EREs in the genome: endogenous retroviruses serve as restriction factors for exogenous retroviruses. It is even hypothesized that endogenous retroviruses were exapted to serve as restriction factors for exogenous retroviruses that are no longer extant[418].

The microbiota is another important source of stimulation for ERE expression[372,448]. In the context of keratinocytes, the skin microbiota induces ERE expression and a STING-dependent antiviral state. Inhibition of reverse transcriptase promotes inappropriate immune responses to commensal microbes and inhibits tissue repair. Interestingly, a high-fat diet further induces ERE expression and microbiota-induced inflammation[372]. This example highlights three critical points about ERE expression. First, EREs have been co-opted by the immune system for proper regulation of immune functions. EREs are also induced by the interferon response, and deletion of specific EREs inhibits essential immune functions such as induction of the AIM2 inflammasome[333]. In this setting, EREs serve as enhancers for immune genes. In fact, LTRs of EREs are enriched for transcription factor binding sites[442,449]. On average, 20% of transcription factor binding sites are associated with EREs[450]. This benefits the retroelement because it makes

it more likely to get expressed and benefits the cell by providing an additional layer of control over the expression of essential immune functions. Second, concurrent ERE expression with microbial ligands broadens, strengthens, and potentiates immune responses. Endogenous pathogen-associated molecular patterns (PAMPs) amplify the response to exogenous PAMPs. Another example of this relates to influenza infection, where cellular stress leads to the loss of the inhibitory function of the ERE repressor TRIM28, leading to the expression of EREs and an enhanced antiviral state[451]. Further evidence for the role of endogenous PAMPs in comes from another RNA virus family, paramyxoviridae, where loss of the cGAS-STING pathway results in increased lethality following infection[452]. Numerous RNA viruses, such as influenza[453], Dengue (DENV)[454], and Zika (ZIKV)[455], also inhibit the cGAS-STING pathway. Given that these are all RNA viruses, it is surprising that the cGAS-STING pathway is essential in controlling infection because exogenous viral products are not expected to trigger this pathway. However, the recurrent importance of this pathway for RNA viruses suggests that the utilization of ERE-derived endogenous PAMPs to augment antiviral response might be a conserved strategy for the immune system across many viral infections. From these examples, one can conceptualize EREs as natural adjuvants, but too much adjuvant in the wrong context or setting is liable to result in autoimmunity[456]. Finally, the level of ERE expression is important. Low levels of ERE expression, like those presented in this study, are beneficial, but excessive ERE expression and immune activation can be deleterious.

These beneficial roles of EREs discussed above can also be hijacked in disease. EREs have been conceptualized as the epitome of the double-edged sword, perfectly exemplified in the present study[456]. HSCs normally use EREs to drive a beneficial interferon response to maintain HSC potential. Still, the loss of CUX1 co-opts this system to drive aberrant self-renewal, which

can lead to myeloid malignancy. There is no selective pressure to stop HSCs from using EREs in this manner because the selective events are post-reproductive. This is true for a number of other examples, briefly mentioned below. First, the ability of the LTRs of EREs to serve as enhancers can also be used to drive oncogenic gene expression programs[457-459]. Second, unrestricted expression of EREs can contribute to aging and senescence[345]. In senescent cells, heterochromatin begins to lose focused boundaries, allowing EREs to be re-expressed[460]. Above-normal ERE expression might also contribute to the onset of motor neuron disease as the env gene is upregulated in a subset of patients with sporadic amyotrophic lateral sclerosis, and upregulation of env is sufficient to induce neuronal damage[461]. ERE expression has been linked to the pathogenesis of many diseases, emphasizing the need to study the repression of EREs.

Repression

Having considered the contexts in which EREs can be expressed and the importance of maintaining tissue-appropriate levels of ERE expression, the next section will highlight the important pathways for ERE repression. The repression of EREs relies jointly on DNA methylation and repressive histone methylation[424]. Some retroelements are more sensitive to the loss of DNA methylation, and some retroelements are more sensitive to the loss of repressive histone methylation[85,462]. Interestingly, there might be a divergence between the roles of these repressive epigenetic modifications. Young EREs tend to be more CpG-rich and rely more on DNA methylation for their repression. Opposingly, the silencing of older EREs depends more on repressive H3K9 methylation[463]. This could explain the synergistic effect of inhibiting both DNA methyltransferases in the induction of an antiviral state[335]. Nevertheless, three major, non-redundant H3K9 methyltransferases in the mammalian genome are tasked with ERE repression: SETDB1, SUVAR39H1/2, and EHMT2[464]. SETDB1 is recruited by TRIM28 (KAP1) to deposit

H3K9me3 at LINE elements and LTR-containing retroelements, thus suppressing the interferon response[393,465-468]. SUVAR39H1 has the unique task of re-establishing H3K9me after cell division. Loss of FBXO44, which recruits SUVAR39H1 to replication forks, or SUVAR39H1, results in depression of EREs and the interferon response[469].

The final H3K9 methyltransferase, EHMT2, is more mysterious than the other enzymes. First, EHMT2 only deposits H3K9me2 because it is sterically incapable of adding a third methyl group[470]. Some loci are by H3K9me2 and H3K9me3, suggesting that both EHMT2 and SETDB1 regulate these loci, and these two enzymes exist in equilibrium. Second, EHMT2 potentially has more cell-type-specific effects. The upregulated retroelements after EHMT2 loss change depending on the cell type. In murine mammary stem cells, ERVK elements are depressed upon EHMT2 loss[332], but it's ERVL elements in embryonic stem cells[337], ERVL and ERVL-MaLR elements in epiblasts[471], and LINE1 elements in spermatagonia[472]. This could be due to differences in EHMT2 targets. Alternatively, it could be that de-repression is not sufficient for expression and that the same elements are de-repressed in every cell upon loss of EHMT2, and different cell-type specific transcription factors drive ERE expression in the absence of repressive H3K9me2[336]. Plausibly, EHMT2 interacts with several cell-type-specific transcription factors that influence its activity[464]. Finally, EHMT2 has a non-enzymatic role in maintaining DNA methylation and genome organization surrounding EREs[334]. Loss of EHMT2 results in dysregulated CTCF and cohesin binding, leading to loss of three-dimensional genome architecture. This is consistent with other findings highlighting the ability of retroelements to hijack super-enhancers and disrupt the transcriptional condensates, altering cellular fates[457]. The disruption of CTCF binding is not trivial. The three-dimensional genomic architecture is dynamically regulated during hematopoiesis. The transition from LT-HSCs to ST-HSCs is

marked by the gain of CTCF binding sites and the enclosing of stemness genes in loops to shut down their transcription. Loss of CTCF prevents LT-HSC activation and promotes increased self-renewal[473]. Similarly, the inhibition of EHMT2 preserves HSC identity in culture[341]. Ultimately, EHMT2, in contrast to other H3K9 methyltransferases, serves multiple roles in the repression of EREs.

DNA methylation (DNAme) is also an important factor in ERE silencing. The evidence for this comes from two separate lines of investigation. First, DNA hypomethylating agents, like azacytidine and decitabine, allow for the de-repression and re-expression of EREs[429,430]. There are two main DNA methyltransferases in mice and humans[474]. The primary role of DNMT1 is the maintenance of existing DNAme. In comparison, the primary role of DNTM3A is the deposition of *de novo* DNAme[474]. Azacytidine complexes with the DNA methyltransferase DNMT1, marking it for early degradation, ultimately leading to hypomethylated regions of DNA[475]. It is worth mentioning that this is certainly only part of the mechanism of hypomethylating agents. DNMT1-DNA adducts disrupt mitosis and inhibit the proper segregation of chromosomes[476]. Nevertheless, genetic aberration of DNMT3A also leads to hypomethylated regions of DNA[477]. Accordingly, treatment with hypomethylating agents or loss of DNMT3A leads to an upregulation of the interferon response to accompany the de-repression of EREs, highlighting the role of DNAme in ERE repression[380,429,430,478,479].

The loss of DMT3A promotes HSC self-renewal, practically immortalizing them in the HSC state[293,477]. DNMT3A-deficient HSCs are maintained through 12 serial transplants without signs of exhaustion[293]. Similar to the loss of CUX1 in HSCs, the up-regulated genes after DNMT3A loss include important regulators of HSC identity[477]. There are other explanations for the immortalization of DNMT3A-deficient HSCs, including the role of DNMT3A in maintaining

telomeres. Loss of DNMT3A in murine embryonic stem cells results in longer telomeres using noncanonical telomere-lengthening strategies like sister chromatid telomere swapping[480]. It is also possible that telomere lengthening is independent of the methyltransferase activity of DNMT3A because the mammalian telomere repeat sequence does not contain CpG sites. This is all to say that there are potentially multiple mechanisms by which loss of DNMT3A leads to HSC immortalization. It will be necessary to tease apart each potential mechanism's contributions to understand HSC self-renewal better.

DNMT3A is not often deleted in myeloid malignancies. However, the dominant-negative DNMT3A^{R8882} mutation is widespread across the spectrum of myeloid disorders, from clonal hematopoiesis to acute myeloid leukemia[93,104]. Indicative of an optimal level of innate immune signaling, HSCs carrying the DMT3A^{R882H} mutation exhibit hypomethylation at ERE loci and the accompanying innate immune activation, resulting in increased HSC potential. Treatment with azacytidine results in further hypomethylation of EREs and a more pronounced interferon response, pushing mutant HSCs over the tipping point toward apoptosis[380]. This mechanism could potentially be exploited in myeloid malignancies that have lost CUX1 to treat the disease. It is worth investigating if CUX1 loss similarly sensitizes cells to hypomethylating agents.

DNMT3A mutant HSCs are constantly experiencing an internal inflammatory response[380]. Curiously, they are also resistant to the adverse effects of exogenous viral stimulation[355,481]. Mutant HSCs resist the potent effects of IFNy to start cycling and differentiate, allowing them to expand in the bone marrow. Mechanistically, they stabilize p53 and then up-regulate p21 to promote cell cycle arrest[481]. While the WT HSCs are lost to differentiation, DNMT3 mutant HSCs remain, allowing them to contribute more to hematopoiesis in the future. There is some benefit to having these mutant HSCs around. Mutant HSCs will remain after the

infection is cleared, fulfilling the fundamental goal of HSCs to produce blood for the entire lifespan of the organism. It remains unclear what about the pre-interferon challenge transcriptome pushes them to undergo cell cycle arrest rather than apoptosis. It could be that mutant HSCs can tolerate a higher tone of inflammatory signaling because they have been living with low levels of inflammation. This example indicates a distinct difference between an interferon challenge originating from within the cell and an interferon challenge originating from outside the cell.

DNA methylation is a dynamic process, and the TET family of enzymes, which catalyze the removal of DNAme, opposes the deposition of DNAme. Like DNMT3A, TET2 is frequently mutated across the spectrum of myeloid neoplasms. Loss of TET2 in HSCs similarly promotes increased HSC self-renewal and initiates the transformation to acute myeloid leukemia[295,482,483]. TET2 does not have a clear role in regulating EREs; however, loss of TET2 similarly results in an antiviral state[484,485]. In the setting of TET2 deficiency, unresolved R-loops, and DNA damage trigger the c-GAS-STING pathway, setting off the inflammatory cascade. Loss of STING restores the aberrant HSC self-renewal[485]. These results highlight the role of the antiviral state in HSC self-renewal, highlighting that self-renewal is not driven by EREs *per se* but by their activation of the innate immune system.

Revisiting the Quality Control Hypothesis

Given the importance of ERE repression for healthy tissue function, ERE re-expression has been proposed as a quality control indicator for unhealthy cells[424]. This hypothesis states that one of the exaptations of EREs is to serve as a checkpoint for cell health. By this reasoning, ERE expression indicates poor maintenance of heterochromatin and foreboding of ensuing genomic instability. The Quality Control Hypothesis fails to consider the normal expression of EREs and

levels of ERE repression. This study highlights that ERE expression is essential in CUX1^{Dim} HSCs to maintain HSC potential and repopulating capacity. An improved Quality Control Hypothesis would consider the magnitude of ERE expression and the sensitivity of the cell to ERE expression.

First, consider the magnitude of ERE expression. Loss of EHMT2 and SETDB1 in mammary and muscle stem cells, respectively, result in ERE de-repression, re-expression, and stem cell dysfunction[332,386]. Based on the findings of Avgustinova et al., the magnitude of ERE re-expression is likely much greater than the magnitude of ERE expression in HSCs with low levels of CUX1[332]. For example, loss of EHMT2 in mammary stem cells results in general genome opening. Furthermore, all the copies of RLTR10D2 that are found in differentially accessible regions following EHMT2 loss are open in EHMT2-deficient mammary stem cells. Similarly, in muscle stem cells, ERV1, ERVK, and ERVL elements broadly open upon loss of SETDB1[386]. This contrasts with CUX1^{Dim} HSCs, where only a small fraction of ERE loci are open. Intuitively, it is also easier to accept that the magnitude of ERE de-repression is much greater when losing a major enzyme of the ERE repression system than a transcription factor. The magnitude of the resulting interferon response is also likely proportional to the magnitude of ERE de-repression. The interferon response inhibits stem cell function in mammary and muscle stem cells. However, CUX1^{Dim} HSCs, which have the highest expression of EREs, exhibit the best stem cell function. As is the case, for DNMT3A mutant HSCs[380], it's all about levels.

Second, the sensitivity of cells to ERE expression should be considered. Here, it was demonstrated that only a small fraction of HSCs normally express high levels of EREs. Only CUX1^{Dim} HSCs express the immunogenic ERVK class. This gene program is shut down early in HSC differentiation. Therefore, it is not abnormal for CUX1^{Dim} HSCs to express EREs, but it is

abnormal if downstream progenitors express EREs, which could indicate heterochromatin dysfunction. This is to say that it is important to consider if the ERE expression of a given cell is outside of the physiologic set point.

In summary, ERE expression does not always indicate poor cell health and ought to be considered in the proper cellular context. There are numerous examples of positive exaptations of ERE, including the present study. In simplest terms, CUX1^{Dim} HSCs use low levels of ERE expression to serve as fodder for a small but beneficial interferon response. Further investigation will be needed to determine if this exaptation of EREs is relevant in other somatic stem cells.

Limitations

The main limitation of studying EREs is the lack of characterization of individual elements. Many studies, including the present study, use the relative number of upregulated or downregulated EREs to make conclusions about the burden of ERE expression[338,393]. This is even somewhat dubious given the limitations of quantifying ERE transcripts due to their repetitive and multi-mapping nature[432]. To better understand the consequences of ERE expression, there needs to be a better appreciation of the role of individual families. This is, of course, easier said than done, given that a given LTR-containing ERE can have up to 1000s of insertions in the genome; it would be naïve to assume that all insertions exhibit the same functional potential for transcription, reverse transcription, and reintegration. Individual retroelement insertions are most functionally capable after insertion into a new locus[420]. Over time, without positive selection to maintain functionality, deleterious mutations and recombination render new insertions inert. However, it is also the youngest retroelement families, which are thought to be most functional and difficult to study and align because they have the greatest sequence similarity. Along these lines, mice might not be the best model

system to study EREs as they relate to human disease due to the differences in the ERE repertoire that occurred post-endogenization[432]. Nevertheless, a greater functional understanding of individual retroelement families will help to reconcile the numerous seemingly contradictory roles of ERE expression.

To make the study of EREs more reproducible and reliable, tools akin to studying other genes would be beneficial. This includes antibodies raised against ERE proteins to move beyond the over-reliance on RNA expression and better establish correlations. Necessity and sufficiency studies of EREs remain challenging. Previous studies have relied on CRISPR activation and CRISPR inhibition systems to activate and silence specific retroelements. Most studies have not provided adequate characterization of their tools to make these methods more accessible to other scientists or results reproducible. For example, most studies have not characterized the off-target effects of these CRISPR systems, which seems especially important given the highly similar and repetitive nature of EREs. It remains near impossible to determine the locus from which a given ERE is expressed, given the RNA transcript, especially using short-read sequencing. Hence, an indication of the activation or inhibition of one insertion of an ERE is simplistic and correlative at best. It is clear that EREs play a complicated role in cellular physiology, and a complete understanding of the role of EREs in different contexts can only be unlocked by more reliable methods and tools.

Self-Nonself Discrimination

Self-nonself discrimination (SNS) is a central paradigm in immunology and underlies the basis of immunity to pathogens. SNS is so fundamental to immunity that even single-celled organisms require mechanisms for SNS[486]. The paradigm of SNS rests on the fundamental assertion that nonself has essential, evolutionarily conserved PAMPs that can be recognized by

germline-encoded pathogen recognition receptors (PRRs)[487]. Given the inefficiencies of central tolerance, PRRs are necessary to inform the adaptive immune system that a pathogen has been encountered. This changes the shape of the ensuing adaptive immune response upon peptide-MHC encounter to promote a targeted immune response. This study highlights another use of the SNS paradigm. The SNS system is used to recognize self as nonself to generate the same cellular response, not in the defense against pathogens but in the maintenance of homeostasis. In this study, HSCs allow EREs to generate PAMPs, trigger PRRs, and drive a low-grade, tonic interferon response. This response provides an additional mechanism to control two basic properties of HSCs: cell cycle participation and self-renewal. Indeed, it must be acknowledged that this response also renders cells more resistant to viral infection, but this is not the only purpose[314]. This section has explored multiple exaptations of EREs, including serving as enhancers for gene regulatory networks, propagating immune responses, and complementing immune regulation. This study highlights how HSCs have exapted a subset of the immunogenic EREs to regulate HSC self-renewal.

Interferon and homeostasis

Just as EREs are expressed in homeostasis, a tonic interferon response is important for normal cell physiology[488]. This contrasts with the traditional view of the interferon response, which is primarily seen as a reaction to acute viral infections. In such cases, infected cells secrete type I interferons to trigger the antiviral response and prime the adaptive immune system[488]. Challenging this paradigm, an interferon-stimulated response element (ISRE) reporter mouse revealed a detectable interferon response in the bone marrow of unchallenged SPF-maintained mice, indicating the ubiquitous nature of interferon signaling. This response is diminished but not eliminated in the absence of IFNAR1 and STAT2, suggesting that the primary source of

interferon signaling is cell extrinsic but leaves open the possibility of cell-intrinsic interferon signaling[362]. The low-level ISG expression is termed tonic interferon signaling. ISGs regulated by tonic interferon signaling are responsive to low doses of interferon and have slow expression kinetics[363]. STAT1 binds these genes at baseline and increases after activation, suggesting these genes are sensitive to small changes in the interferon tone of the cell[363]. In support of this hypothesis, even small levels of IFNβ can rescue ISG expression in interferon-deficient animals[364]. This aligns with other findings demonstrating that specific genes are more sensitive to interferon stimulation than others[361]. Differential sensitivity to interferon stimulation is likely mediated through heterogeneity in the response elements of the promoters; for example, some ISGs are more dependent on STAT1 than others[363]. All in all, different types of interferon responses portend different outcomes based on the duration and intensity of signaling.

There are three main themes of tonic interferon signaling: tissue protection, cellular maturation, and homeostatic maintenance. First, diverse tissues require tonic interferon signaling for pathogen evasion and immunity. Low levels of interferon signaling drive the expression of viral restriction factors and serve as the first line of defense against incoming pathogens[314,364-369]. This tends to be true in particularly vulnerable cells like cardiac myocytes, endothelial cells, cortical neurons, and somatic stem cells. Second, tonic interferon signaling is required in the final maturation steps of naïve T cells[371] and regulatory T cells[489], as well as cortical microglia[370]. Specifically, dsRNA sensing by microglia is required to license phagocytic functions to remove damaged neurons and allow for the final developmental steps in somatosensory function[370]. In contrast to T cell development, which requires extrinsic interferon signaling, microglia require cell-intrinsic interferon signaling.

Finally, the role of interferon signaling in homeostatic maintenance is the most under-explored and relevant to this study. Tonic interferon signaling is required for normal cellular physiology. The following examples from knockout mice will highlight this in the context of HSC self-renewal. First, loss of IFNAR, the interferon α/β receptor, results in decreased maintenance of HSC number under steady-state conditions, suggesting exogenous interferon is required for HSC maintenance[359]. Second, similarly, loss of the transcription factor IRF2 in murine HSCs results in a loss of HSC numbers due to increased HSC cycling and decreased HSC potential upon transplant[490]. Third, the same loss of HSC potential is observed in IRF1-deficient HSCs[327]. The loss of HSC potential with IRF1 knockout is coupled with a loss of ISG expression and decreased MHC Class II expression. And finally, STAT1-deficient HSCs have the exact same phenotype as IRF1-deficient HSCs[328]. These studies of knockout mice clearly demonstrate the role of tonic interferon signaling in hematopoietic stem cell maintenance. The finding that IRF1 deficiency and STAT1 deficiency phenocopy one another is intriguing. These two transcription factors jointly regulate ISGs and EREs[333]. It is, therefore, interesting to hypothesize that loss of either of these transcription factors results in a loss of ERE expression, leading to the observed decrease in cell-intrinsic interferon response and loss of HSC potential. More mechanistic work is needed to confirm this hypothesis in HSCs. In summary, the interferon response is not restricted to the settings of viral infection and plays an essential role in the homeostasis of multiple tissues, particularly in HSCs.

Trajectories of LT-HSC differentiation

The study demonstrated that within the LT-HSC compartment, there is a clear differentiation path controlled by CUX1 levels. HSCs at the apex of hematopoiesis have low levels of CUX1 protein. CUX1 protein levels increase as HSCs begin to differentiate into

downstream progenitors. This is accompanied by transcriptional changes that prepare the cell for greater cell cycle participation and biosynthesis. There is also a downregulation of genes important for HSC self-renewal, such as *Mecom* and *Hlf*. CUX1 controls both gene programs simultaneously. Additionally, as previously mentioned, the transition from $CUX1^{Dim}$ to $CUX1^{Bright}$ is accompanied by a decrease in ISG expression due to the repression of EREs. Our model of CUX1 protein levels and the transcriptional changes brought about by CUX1 represent the canonical model of differentiating hematopoietic progenitors. This model dictates that there is one path out of the stem cell compartment. HSCs start as $CUX1^{Dim}$ and progress to $CUX1^{Bright}$, then exit the HSC compartment. However, this model has several shortcomings. First, this model does not account for the alternative path of megakaryocyte differentiation whereby HSCs at the apex of the hematopoietic hierarchy differentiate directly into megakaryocyte progenitors without passing through downstream progenitor compartments like ST-HSCs, MPPs, and MEPs[58,66,68]. The alternative pathway of platelet differentiation becomes more active with age and contributes to age-related platelet hyperactivity[491]. It is unresolved if CUX1 levels have a role in this differentiation trajectory. New tools to understand the alternative pathway of platelet production will be important to disentangle the interactions between the canonical differentiation trajectory and the alternative pathway. There might also be more direct paths out of the HSC compartment that have been unappreciated until now, and it will remain to be seen if CUX1 levels impact these trajectories[60].

The second limitation of the canonical model, as presented in this study, is that it fails to account for the repopulating capacity of ST-HSCs. $CUX1^{Bright}$ HSCs have less repopulating capacity than the average ST-HSC. Even so, all the repopulating capacity of ST-HSCs is contained within the $CUX1^{Dim}$ segment of ST-HSCs (data not shown). This begs the question of

how CUX1^{Dim} ST-HSCs arise if they are not coming from Cux1^{Bright} LT-HSCs. The easiest way out of this problem is to admit that some CUX1^{Dim} LT-HSCs downregulate CD150 to become CD150- before upregulating CUX1. In support of this assertion, scRNA-seq reveals the transcriptional similarity of some LT-HSCs and some ST-HSCs with uniformly low levels of CUX1 (Figure 6). These cells all express high levels of HSC markers, suggesting these cells are very similar and could just differ in the expression of CD150. The differentiation steps in this part of hematopoiesis are murky, and more careful experimentation is required to assess the role CUX1 plays. In summary, the evidence presented in this study necessitates at least two paths out of the LT-HSC compartment into the ST-HSC compartment and beyond. In addition to the alternative megakaryocyte/platelet differentiation pathway, there must be at least two other ways that a CUX1^{Dim} HSC can differentiate and possibly more than that. Overall, these limitations will be challenging to overcome until we understand what controls CUX1 levels.

The limitations of the model as presented could be overcome with a greater understanding of how different levels of CUX1 are originally established and how CUX1 levels change with age. Understanding how different levels of CUX1 arise gets at one of the most fundamental questions in HSC biology: what is the first cause of HSC heterogeneity? This study clearly established that CUX1 levels control the progressive loss of HSC self-renewal and vast differences in cell cycle participation among HSCs (Figure 8). However, there are some HSC behaviors that CUX1 does not control, such as lineage bias. Even so, CUX1 levels in HSCs exist in an approximately normal distribution. It is not clear how or when this distribution is established and how it is maintained. HSC heterogeneity could be encoded before HSCs arrive in the bone marrow or established during the early bone marrow seeding[46]. It could be that CUX1 levels are controlled in a cell-extrinsic manner and that different bone marrow

microenvironments or different cellular interactions maintain CUX1 levels. It is worth noting that ex vivo culture conditions cannot maintain cells with low levels of CUX1. Therefore, two options are possible. Either the culture conditions are missing an essential factor to maintain low levels of CUX1, or the cytokine concentrations in the media are conducive to low levels of CUX1. Determining the factor or combination of factors that maintain low levels could unlock additional potential in maintaining HSCs *ex vivo*.

Understanding what controls CUX1 levels also presents an exciting opportunity to intervene on behalf of patients with myeloid neoplasms to treat their disease. Given that the phenotype of CUX1 loss is reversible upon allowing cells to express CUX1 again[78,191], this same strategy could be exploited to give CUX1 back in the setting of myeloid disease. Based on the experiments here, this treatment would turn all the CUX1-deficient blasts into $CUX1^{Bright}$ HSCs, forcing them to differentiate. However, this study also highlights a fundamental concern with this strategy. A treatment that raised CUX1 levels in every cell would also raise CUX1 levels in $CUX1^{Dim}$ cells, potentially sending them on a path toward premature terminal differentiation. Therefore, such a treatment strategy ought to seek to manipulate a unique pathway to transformed cells, which spares normal HSCs. It is also worth noting that almost all haploinsufficient genes are also triplosensitive[12], so such a treatment strategy ought to bring CUX1 levels back to normal without tipping to overexpression.

Given the potential problems with giving CUX1 back to cells that have lost it, it is worth considering situations where one might want to limit CUX1 protein levels. This study emphasizes a subtle point, illustrated by a quote from Douglas Adams, the author of "The Hitchhiker's Guide to the Galaxy": "It's not the fall that kills you; it's the sudden stop at the end." Cells can adopt different levels of CUX1 normally, even barely detectable levels. Loss of

one allele of *CUX1* is pathogenic because of the inability to maintain the dynamic regulation of CUX1. Potentially, inducing low levels of CUX1 temporarily in certain settings could be beneficial. Given that pathogenic HSC clones can expand the face of negative external stimuli, such as inflammation or chemotherapy, temporarily inducing low levels of CUX1 might stop this expansion. CUX1^{Dim} HSCs are less reactive to exogenous inflammation (data not shown) and have lower proliferation rates. The competitive advantage certain clones have under inflammatory settings could be diminished if the playing field was leveled by inducing all HSCs to be CUX1^{Dim} HSCs. One could imagine a setting where an individual with a mild infection takes some medication to lower CUX1 levels for the duration of the infection. The induction of low levels of CUX1 would be transient and potentially maintain clonal diversity in the bone marrow. This approach is not without concerns. Notably, low levels of CUX1 can also induce multilineage dysplasia and anemia. However, these effects are also temporary, as evidenced by prior studies[78,191].

Asymmetric division

The central finding of this study is that knockdown of restores and extends the HSC potential of immunophenotypic HSCs. On the molecular level, this reprogramming event is characterized by the restoration of the chromatin accessibility landscape and gene expression program of CUX1^{Dim} HSCs. It is unclear how this is brought about at the cellular level. There are three hypotheses as to how HSC reprogramming could be accomplished. First, the knockdown of CUX1 could induce an alternative cell state that is not transited by CUX1-sufficient HSCs during normal hematopoiesis. Single-cell RNA-seq ruled out this possibility and revealed that CUX1-mediated HSC reprogramming does not shuttle cells into an alternative state because all the cells clustered together regardless of genotype. The second hypothesis stated that CUX1 knockdown

could slow the normal maturation process. This hypothesis is supported by the slower disappearance of $CD150^{+}$ HSCs in culture after CUX1 knockdown. The main caveat of this experiment is that the culture conditions used in this experiment are not even conducive to the maintenance of the HSC state, so it would be unlikely to observe anything other than what was documented in this experiment. The final hypothesis is that CUX1 knockdown returns cells to an earlier cell state, and cells move backward through cellular differentiation. Using the available experiments, it is tough to parse the differences between hypothesis two and hypothesis three.

All these experiments utilize ex vivo culture conditions to infer what is happening in vivo. There are two major experimental limitations that make this question difficult to answer. First, one of the hallmarks of $CUX1^{Bright}$ HSCs is the commitment to cell cycle participation. This is hard to dissect in culture conditions that promote cell cycling. It becomes difficult to tell which cells are cycling because they have committed to cycling and which cells are cycling because environmental stimuli are forcing them to cycle. This is particularly relevant to the interpretations and caveats of scRNA-seq after CUX1 knockdown. Second, CUX1 knockdown is not instantaneous. When we assay the effects of CUX1 knockdown, the control HSCs have differentiated even if slightly. This makes it hard to assess the absolute cell state of cells after knockdown. For example, it is difficult to tell if CUX1 knockdown in a $CUX1^{Dim}$ HSC creates a state that is akin to $CUX1^{Dim}$ HSCs or a state that supersedes $CUX1^{Dim}$ akin to $CUX1^{Dim}$ on steroids. Even so, it is also impossible to know the full effects of CUX1 knockdown because cells differentiate before CUX1 levels are diminished, and the effects of the loss of CUX1 can be actualized. Along these lines, given the short window of time that CUX1 knockdown can potentiate HSC activity, it is likely that all of the experiments underestimate the effects of CUX1 knockdown. The cellular mechanism of CUX1-mediated reprogramming is almost an

unanswerable question because of the experimental limitations. Regardless, the best explanation of CUX1-mediated HSC reprogramming is explained by a combination of hypotheses two and three: CUX1 knockdown forces cells to slow down differentiation and remain at the same level of HSC potential by invoking the phenotype of $CUX1^{Dim}$ HSCs.

It is well established that $CUX1^{Bright}$ HSCs re-don the $CUX1^{Dim}$ identity, and $CUX1^{Dim}$ HSCs potentiate the $CUX1^{Dim}$ identity upon CUX1 knockdown. This satisfies the part of the model that describes the role of the $CUX1^{Dim}$ identity. However, as previously stated, there is something that $CUX1^{Bright}$ HSCs remember about the previous $CUX1^{Bright}$ identity. As soon as CUX1 levels are allowed to return to normal after CUX1 knockdown, they lose HSC potential. This indicates that CUX1-mediated HSC reprogramming is not permanent and relies on continuous low levels of CUX1. This is also evidence that CUX1-mediated reprogramming does turn $CUX1^{Bright}$ HSCs into true $CUX1^{Dim}$ HSCs. If this were the case, one would expect $CUX1^{Bright}$ HSCs after CUX1 knockdown followed by a relief of shRNA expression not to decay like $CUX1^{Bright}$ HSCs. This finding strongly, but not definitively, suggests that CUX1 knockdown allows HSCs to stay in place longer and likely does not fully recreate an earlier cell state. From here, the question becomes how HSCs remain in the same place. The most attractive hypothesis for this is that CUX1 knockdown alters the ratio of HSC division modes.

HSCs can divide in the three different ways. They can undergo symmetric division, where both daughter cells remain HSCs or become differentiating progenitors. Alternatively, HSCs can divide asymmetrically, resulting in one daughter cell remaining an HSC and the other becoming a differentiating progenitor.

Asymmetric cell division (ACD) is one mechanism to bring about different fates in daughter cells after a round of cell division. This mechanism is utilized in many different

systems, including during zygote development and CD8 T cell activation. HSCs also use ACD to maintain the HSC pool and ensure a steady flow of differentiating progenitors. HSCs divide lysosomes[292], mitochondria[287], and signaling proteins[492] asymmetrically. It could be that CUX1 is also subject to asymmetric division, and this is why there are always HSCs with low levels of CUX1 to maintain the HSC pool. Whether this is true or not, changing the rates of ACD remains an attractive hypothesis for the cellular mechanism of HSC re-programming because it allows for continued HSC retention without invoking a differentiation block. Increasing the rate of ACD relative to the rate of symmetric differentiating division upon CUX1 knockdown fits several previously mentioned observations that a mechanism must satisfy. ACD is difficult to measure, and there are only a few markers. It is also unclear if the existing markers are the only markers and whether the asymmetric division of these markers represents the only instances of ACD. Regardless, it is important to understand precisely how CUX1 knockdown allows HSCs to retain HSC potential to target this in disease settings, so the role of CUX1 and CUX1 levels in ACD ought to be investigated.

A paradigm for studying CUX1

A simple model of CUX1 function has been difficult to conceptualize for two reasons. First, CUX1 has strong cell-type-dependent effects. Second, CUX1 has at least three roles and likely performs all of them simultaneously. It is very likely that these two reasons are intertwined and that the pool of CUX1 molecules in a cell is utilized differently in a cell-type-dependent manner. What follows is an attempt to explicate the three roles of CUX1. The first role of CUX1, the Imgruet-Jotte role, does not involve transcriptional regulation, and CUX1 participates in the DNA damage response in a sequence-independent manner[191]. It has not been formally demonstrated that this role is sequence-independent. Still, it follows logically because DNA

damage is sequence-independent, even if certain regions of the genome are more likely to be damaged than others.

The second role of CUX1, the Liu-Kurkewich role, is the transcriptionally activating role of CUX1. In this role, CUX1 recruits the BAF complex to open specific loci to remodel the surrounding chromatin and activate gene expression[323]. This role is likely cell-type-dependent. CUX1 has roles in somatic stem cells of many different tissues, and while not formally demonstrated, it likely has pioneering activity in all of them. Given that CUX1 likely has pioneering activity in hair follicle stem cells and intestinal stem cells, as well as hematopoietic stem cells, CUX1 could interact with other transcription factors to promote chromatin accessibility at lineage-specific sites. In line with this hypothesis, the CUX1 motif was not enriched in the sites that close upon CUX1 knockdown, as demonstrated by ATAC-seq (Figure 22). Instead, the motif of the CEPB family of transcription factors was enriched at these sites. This potentially provides etiology to the myeloid dysplasia observed upon CUX1 knockdown. In the absence of CUX1, these sites remain closed, and CEPB family transcription factors cannot bind and drive proper terminal myeloid differentiation. An interesting implication of this hypothesis is that these sites are opened in HSCs well before they are needed downstream in differentiation. Potentially, CUX1 is absolutely required to open these sites, and in the absence of CUX1, they are never opened. This would suggest that the CUX1-related myeloid dysplasia is potentially a legacy effect. It is not necessarily the case that CUX1 is required in the cell type where these sites are utilized, but the fact that they are not open by the time they are used leads to improper differentiation and dysplasia.

The final role of CUX1 is most relevant to the current study. In this role, CUX1 recruits the histone methyltransferase EHMT2 to silence specific loci. Given all that has been discussed

about the importance of silence EREs, it follows that this role could be ubiquitous and cell-type independent. In support of this model, the CUX1 motif is highly enriched in the sites that open upon CUX1 knockdown (Figure 22). This role is required in the differentiation of potentially all somatic stem cells, and therefore, CUX1 might not need lineage-specific partners to perform this role.

Conceptualizing the function of CUX1 helps to arrive at a more unified view of how this transcription factor functions. This model also creates several testable hypotheses that can be used to generate a better model. It will be critical to separate these functions of CUX1 to study them in isolation. This will require understanding exactly which parts of the protein are important for each listed function and how a given molecule of CUX1 gets recruited to perform each function. Future models also ought to incorporate more transcription factor dynamics as the model presented here ignores biochemical properties of transcription factors like search time and on/off rates.

Conclusion

It is important to understand the fundamental properties of stem cells: multipotency and self-renewal. The exact molecular mechanisms supporting these homeostasis properties are exploited in disease states. This study demonstrates that CUX1 levels control HSC fate through the exaptation of EREs to drive a low-grade interferon response in addition to the regulation of canonical HSC regulators. This mechanism was only possible to uncover after first dissecting and trying to understand the HSC heterogeneity in the HSC compartment. Through this careful dissection, we challenged multiple accepted paradigms, including the plasticity of the hematopoietic system, the role of the interferon response in somatic stem cells, and the

implications of ERE expression. All of these contribute to a greater understanding of HSC self-renewal.

Remaining Questions/Future Directions

1. What factors control the levels of CUX1 *in vivo*? Are there conditions that better maintain low levels of CUX1 *ex vivo* than current culture conditions?

2. What is the impetus for CUX1 levels to rise *in vivo*?

3. How is the regulation of genes like *Mecom* and *Hlf* connected to the interferon response? Are these classically defined regulators regulated by IRF1 and the interferon response? How is the classical HSC self-renewal program related to interferon signaling?

4. What is the role of IRF1 in the regulation of HSC potential? Does IRF1, in combination with STAT1, drive the expression of EREs in the absence of CUX1? Or does IRF1 serve as the final effector of the interferon signaling pathway?

5. Is the expression of certain endogenous retroelements sufficient to promote HSC activity?

6. What is the interaction of exogenous inflammation with the high interferon tone of HSCs with low levels of CUX1? Does the initiating cytokine have an effect? Is there a difference between IL-1β and exogenous interferons? Does this create a therapeutic vulnerability?

7. What is the role of antigen presentation in HSC potential? Is MHC Class II expression induced by ERE expression? Do regulatory T cells serve a protective role in providing anti-inflammatory stimulus to HSCs with high interferon signaling?

References

1. Barski, G., and Cornefert, F. (1962). Characteristics of "Hybrid"-Type Clonal Cell Lines Obtained From Mixed Cultures In Vitro. JNCI: Journal of the National Cancer Institute *28*, 801-821. 10.1093/JNCI/28.4.801.
2. Harris, H., Miller, O.J., Klein, G., Worst, P., and Tachibana, T. (1969). Suppression of Malignancy by Cell Fusion. Nature *223*, 363-368. 10.1038/223363A0.
3. Knudson, A.G. (1971). Mutation and Cancer: Statistical Study of Retinoblastoma. Proceedings of the National Academy of Sciences *68*, 820-823. 10.1073/PNAS.68.4.820.
4. Fero, M.L., Randel, E., Gurley, K.E., Roberts, J.M., and Kemp, C.J. (1998). The murine gene p27Kip1 is haplo-insufficient for tumour suppression. Nature *396*, 177-180. 10.1038/24179.
5. Bartha, I., Di Iulio, J., Venter, J.C., and Telenti, A. (2017). Human gene essentiality. Nature Reviews Genetics *19*, 51-62. 10.1038/NRG.2017.75.
6. Bartha, I., Rausell, A., McLaren, P.J., Mohammadi, P., Tardaguila, M., Chaturvedi, N., Fellay, J., and Telenti, A. (2015). The Characteristics of Heterozygous Protein Truncating Variants in the Human Genome. PLOS Computational Biology *11*, e1004647-e1004647. 10.1371/JOURNAL.PCBI.1004647.
7. Delneri, D., Hoyle, D.C., Gkargkas, K., Cross, E.J.M., Rash, B., Zeef, L., Leong, H.S., Davey, H.M., Hayes, A., Kell, D.B., et al. (2007). Identification and characterization of high-flux-control genes of yeast through competition analyses in continuous cultures. Nature Genetics *40*, 113-117. 10.1038/NG.2007.49.
8. Ohnuki, S., and Ohya, Y. (2018). High-dimensional single-cell phenotyping reveals extensive haploinsufficiency. PLOS Biology *16*, e2005130-e2005130. 10.1371/JOURNAL.PBIO.2005130.
9. Deutschbauer, A.M., Jaramillo, D.F., Proctor, M., Kumm, J., Hillenmeyer, M.E., Davis, R.W., Nislow, C., and Giaever, G. (2005). Mechanisms of Haploinsufficiency Revealed by Genome-Wide Profiling in Yeast. Genetics *169*, 1915-1925. 10.1534/GENETICS.104.036871.
10. White, J.K., Gerdin, A.K., Karp, N.A., Ryder, E., Buljan, M., Bussell, J.N., Salisbury, J., Clare, S., Ingham, N.J., Podrini, C., et al. (2013). Genome-wide Generation and Systematic Phenotyping of Knockout Mice Reveals New Roles for Many Genes. Cell *154*, 452-464. 10.1016/J.CELL.2013.06.022.
11. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285-291. 10.1038/NATURE19057.
12. Collins, R.L., Glessner, J.T., Porcu, E., Lepamets, M., Brandon, R., Lauricella, C., Han, L., Morley, T., Niestroj, L.M., Ulirsch, J., et al. (2022). A cross-disorder dosage sensitivity map of the human genome. Cell *185*, 3041-3055.e3025. 10.1016/j.cell.2022.06.036.
13. Abel, H.J., Larson, D.E., Regier, A.A., Chiang, C., Das, I., Kanchi, K.L., Layer, R.M., Neale, B.M., Salerno, W.J., Reeves, C., et al. (2020). Mapping and characterization of structural variation in 17,795 human genomes. Nature *583*, 83-89. 10.1038/S41586-020-2371-0.
14. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational

constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434-443. 10.1038/s41586-020-2308-7.
15. Ni, Z., Zhou, X.Y., Aslam, S., and Niu, D.K. (2019). Characterization of Human Dosage-Sensitive Transcription Factor Genes. Frontiers in Genetics *10*, 1208-1208. 10.3389/FGENE.2019.01208/BIBTEX.
16. Davoli, T., Xu, A.W., Mengwasser, K.E., Sack, L.M., Yoon, J.C., Park, P.J., and Elledge, S.J. (2013). Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns and Shape the Cancer Genome. Cell *155*, 948-962. 10.1016/J.CELL.2013.10.011.
17. Solimini, N.L., Xu, Q., Mermel, C.H., Liang, A.C., Schlabach, M.R., Luo, J., Burrows, A.E., Anselmo, A.N., Bredemeyer, A.L., Li, M.Z., et al. (2012). Recurrent hemizygous deletions in cancers may optimize proliferative potential. Science *336*, 104-109. 10.1126/SCIENCE.1219580/SUPPL_FILE/SOLIMINI032612.XLS.
18. Chen, Z., Trotman, L.C., Shaffer, D., Lin, H.K., Dotan, Z.A., Niki, M., Koutcher, J.A., Scher, H.I., Ludwig, T., Gerald, W., et al. (2005). Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nature *436*, 725-730. 10.1038/NATURE03918.
19. Berger, A.H., Knudson, A.G., and Pandolfi, P.P. (2011). A continuum model for tumour suppression. Nature *476*, 163-169. 10.1038/NATURE10275.
20. Lee, J.Y., Nakada, D., Yilmaz, O.H., Tothova, Z., Joseph, N.M., Lim, M.S., Gilliland, D.G., and Morrison, S.J. (2010). MTOR activation induces tumor suppressors that inhibit leukemogenesis and deplete hematopoietic stem cells after pten deletion. Cell Stem Cell *7*, 593-605. 10.1016/J.STEM.2010.09.015/ATTACHMENT/8595DED1-7DCD-45A0-9322-CFDC6007596E/MMC1.PDF.
21. Alimonti, A., Carracedo, A., Clohessy, J.G., Trotman, L.C., Nardella, C., Egia, A., Salmena, L., Sampieri, K., Haveman, W.J., Brogi, E., et al. (2010). Subtle variations in Pten dose determine cancer susceptibility. Nature Genetics *42*, 454-458. 10.1038/NG.556.
22. Will, B., Vogler, T.O., Narayanagari, S., Bartholdy, B., Todorova, T.I., Da Silva Ferreira, M., Chen, J., Yu, Y., Mayer, J., Barreyro, L., et al. (2015). Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. Nature Medicine *21*, 1172-1181. 10.1038/nm.3936.
23. Rosenbauer, F., Wagner, K., Kutok, J.L., Iwasaki, H., Le Beau, M.M., Okuno, Y., Akashi, K., Fiering, S., and Tenen, D.G. (2004). Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. Nature Genetics *36*, 624-630. 10.1038/NG1361.
24. Poluri, R.T.K., and Audet-Walsh, É. (2018). Genomic deletion at 10q23 in prostate cancer: More than PTEN loss? Frontiers in Oncology *8*, 246-246. 10.3389/FONC.2018.00246/BIBTEX.
25. Mcnerney, M.E., Brown, C.D., Peterson, A.L., Banerjee, M., Larson, R.A., Anastasi, J., Le Beau, M.M., and White, K.P. (2014). The spectrum of somatic mutations in high-risk acute myeloid leukaemia with -7/del(7q). British Journal of Haematology *166*, 550-556. 10.1111/bjh.12964.
26. Neumann, E. Ueber die Bedeutung des Knochenmarkes für die Blutbildung, Vorläufige Mittheilung. Centralblatt für die medicinischen Wissenschaften. 1868; 6 (44). Neumann-

Redlin v Meding, E. Der Pathologe Ernst Neumann (1834–1918) und sein Beitrag zur Begründung.
27. Maximow, A. (1922). Untersuchungen über Blut und Bindegewebe. Archiv für Mikroskopische Anatomie *96*, 494-527. 10.1007/bf02977422.
28. Haeckel, E. (1868). Natürliche schöpfungsgeschichte (Deutsches Textarchiv (Kernkorpus)).
29. Ramalho-Santos, M., and Willenbring, H. (2007). On the Origin of the Term "Stem Cell". Cell Stem Cell *1*, 35-38. 10.1016/j.stem.2007.05.013.
30. Lazarus, D.A., and Ehrlich, P. (1900). Histology of the Blood, Normal and Pathological, by P. Ehrlich and A. Lazarus, Edited and Translated by W. Myers... with a Preface by G. Sims Woodhead (University Press).
31. Maehle, A.-H. (2011). Ambiguous cells: the emergence of the stem cell concept in the nineteenth and twentieth centuries. Notes and Records of the Royal Society *65*, 359-378. 10.1098/rsnr.2011.0023.
32. Brecher, G., Cronkite, E.P., and G. Brecher, E.P.C. (1951-06-01). Post-Radiation Parabiosis and Survival in Rats.∗. Proceedings of the Society for Experimental Biology and Medicine *77*. 10.3181/00379727-77-18754.
33. Ford, C.E., Hamerton, J.L., Barnes, D.W.H., and Loutit, J.F. (1956). Cytological Identification of Radiation-Chimæras. Nature *177*, 452-454. 10.1038/177452a0.
34. Becker, A.J., Mcculloch, E.A., and Till, J.E. (1963). Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells. Nature *197*, 452-454. 10.1038/197452a0.
35. Siminovitch, L., Mcculloch, E.A., and Till, J.E. (1963). The distribution of colony-forming cells among spleen colonies. Journal of Cellular and Comparative Physiology *62*, 327-336. 10.1002/jcp.1030620313.
36. Bonner, W.A., Hulett, H.R., Sweet, R.G., and Herzenberg, L.A. (1972). Fluorescence Activated Cell Sorting. Review of Scientific Instruments *43*. 10.1063/1.1685647.
37. Spangrude, G.J., Heimfeld, S., and Weissman, I.L. (1988). Purification and Characterization of Mouse Hematopoietic Stem Cells. Science. 10.1126/science.2898810.
38. Morrison, S.J., and Weissman, I.L. (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity *1*. 10.1016/1074-7613(94)90037-X.
39. Kiel, M.J., Yilmaz, Ö.H., Iwashita, T., Yilmaz, O.H., Terhorst, C., and Morrison, S.J. (2005). SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells. Cell *121*, 1109-1121. 10.1016/j.cell.2005.05.026.
40. Oguro, H., Ding, L., and Morrison, J., Sean (2013). SLAM Family Markers Resolve Functionally Distinct Subpopulations of Hematopoietic Stem Cells and Multipotent Progenitors. Cell Stem Cell *13*, 102-116. 10.1016/j.stem.2013.05.014.
41. Baum, C.M., Weissman, I.L., Tsukamoto, A.S., Buckle, A.M., and Peault, B. (1992). Isolation of a candidate human hematopoietic stem-cell population. Proceedings of the National Academy of Sciences *89*, 2804-2808. 10.1073/pnas.89.7.2804.
42. Challen, G.A., Boles, N.C., Chambers, S.M., and Goodell, M.A. (2010). Distinct Hematopoietic Stem Cell Subtypes Are Differentially Regulated by TGF-β1. Cell Stem Cell *6*, 265-278. 10.1016/j.stem.2010.02.002.

43. Goodell, M.A., Brose, K., Paradis, G., Conner, A.S., and Mulligan, R.C. (1996). Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. The Journal of experimental medicine *183*, 1797-1806. 10.1084/jem.183.4.1797.
44. Weksberg, D.C., Chambers, S.M., Boles, N.C., and Goodell, M.A. (2008). CD150− side population cells represent a functionally distinct population of long-term hematopoietic stem cells. Blood *111*, 2444-2451. 10.1182/blood-2007-09-115006.
45. Warr, M.R., Pietras, E.M., and Passegué, E. (2011). Mechanisms controlling hematopoietic stem cell functions during normal hematopoiesis and hematological malignancies. WIREs Systems Biology and Medicine *3*, 681-701. 10.1002/wsbm.145.
46. Crisan, M., and Dzierzak, E. (2016). The many faces of hematopoietic stem cell heterogeneity. Development *143*, 4571-4581. 10.1242/dev.114231.
47. Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nature Methods *6*, 377-382. 10.1038/nmeth.1315.
48. Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J., Teichmann, S.A., Marioni, J.C., and Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nature Biotechnology *33*, 155-160. 10.1038/nbt.3102.
49. Kolodziejczyk, A., Aleksandra, Kim, J.K., Svensson, V., Marioni, C., John, and Teichmann, A., Sarah (2015). The Technology and Biology of Single-Cell RNA Sequencing. Molecular Cell *58*, 610-620. 10.1016/j.molcel.2015.04.005.
50. Wilson, K., Nicola, Kent, G., David, Buettner, F., Shehata, M., Macaulay, C., Iain, Calero-Nieto, J., Fernando, Castillo, S., Manuel, Oedekoven, A., Caroline, 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. 10.1016/j.stem.2015.04.004.
51. Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nature Cell Biology *19*, 271-281. 10.1038/ncb3493.
52. Giladi, A., Paul, F., Herzog, Y., Lubling, Y., Weiner, A., Yofe, I., Jaitin, D., Cabezas-Wallscheid, N., Dress, R., Ginhoux, F., et al. (2018). Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. Nature Cell Biology *20*, 836-846. 10.1038/S41556-018-0121-4.
53. Pucella, J.N., Upadhaya, S., Reizis, B., Pucella, J.N., Upadhaya, S., and Reizis, B. (2020). The Source and Dynamics of Adult Hematopoiesis: Insights from Lineage Tracing. Annual Review of Cell and Developmental Biology *36*. 10.1146/annurev-cellbio-020520-114601.
54. Lemischka, I.R., Raulet, D.H., and Mulligan, R.C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. Cell *45*. 10.1016/0092-8674(86)90566-0.
55. Naik, S.H., Perié, L., Swart, E., Gerlach, C., Van Rooij, N., De Boer, R.J., and Schumacher, T.N. (2013). Diverse and heritable lineage imprinting of early haematopoietic progenitors. Nature *496*, 229-232. 10.1038/nature12013.
56. Pei, W., Feyerabend, T.B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., et al. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. Nature *548*. 10.1038/nature23653.

57. Li, L., Bowling, S., McGeary, S.E., Yu, Q., Lemke, B., Alcedo, K., Jia, Y., Liu, X., Ferreira, M., Klein, A.M., et al. (2023). A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. Cell *186*. 10.1016/j.cell.2023.09.019.
58. Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M., Sun, J., Calogero, R.A., Klein, A.M., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature *553*, 212-216. 10.1038/nature25168.
59. Pei, W., Shang, F., Wang, X., Fanti, A.-K., Greco, A., Busch, K., Klapproth, K., Zhang, Q., Quedenau, C., Sauer, S., et al. (2020). Resolving Fates and Single-Cell Transcriptomes of Hematopoietic Stem Cell Clones by PolyloxExpress Barcoding. Cell Stem Cell *27*, 383-395.e388. 10.1016/j.stem.2020.07.018.
60. Lin, D.S., Zhang, S., Schreuder, J., Tran, J., Sargeant, T., Metcalf, D., Ng, A.P., Weber, T.S., and Naik, S.H. (2024). Amulti-tracklandscape of haematopoiesis informed by cellular barcoding and agent-based modelling. Cold Spring Harbor Laboratory.
61. Rodriguez-Fraticelli, A.E., Weinreb, C., Wang, S.-W., Migueles, R.P., Jankovic, M., Usart, M., Klein, A.M., Lowell, S., and Camargo, F.D. (2020). Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. Nature *583*, 585-589. 10.1038/s41586-020-2503-6.
62. Bowling, S., Sritharan, D., Osorio, F.G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W.-C., Fujiwara, Y., Li, B.E., et al. (2020). An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. Cell *181*, 1410-1422.e1427. 10.1016/j.cell.2020.04.048.
63. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F.D., and Klein, A.M. (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science *367*, eaaw3381. 10.1126/science.aaw3381.
64. Liggett, L.A., and Sankaran, V.G. (2020). Unraveling Hematopoiesis through the Lens of Genomics. Cell *182*, 1384-1400. 10.1016/j.cell.2020.08.030.
65. Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature *553*, 418-426. 10.1038/nature25022.
66. Carrelha, J., Meng, Y., Kettyle, L.M., Luis, T.C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., et al. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. Nature *554*, 106-111. 10.1038/NATURE25455.
67. Morita, Y., Ema, H., and Nakauchi, H. (2010). Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. Journal of Experimental Medicine *207*, 1173-1182. 10.1084/jem.20091318.
68. Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Jones, T.B., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232-236. 10.1038/nature12495.
69. Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, L., Karl, Ema, H., and Nakauchi, H. (2013). Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. Cell *154*, 1112-1126. 10.1016/j.cell.2013.08.007.

70. Kucinski, I., Campos, J., Barile, M., Severi, F., Bohin, N., Moreira, P.N., Allen, L., Lawson, H., Haltalli, M.L.R., Kinston, S.J., et al. (2024). A time- and single-cell-resolved model of murine bone marrow hematopoiesis. Cell Stem Cell *31*. 10.1016/j.stem.2023.12.001.
71. Chen, X., Deng, H., Churchill, M.J., Luchsinger, L.L., Du, X., Chu, T.H., Friedman, R.A., Middelhoff, M., Ding, H., Tailor, Y.H., et al. (2017). Bone Marrow Myeloid Cells Regulate Myeloid-Biased Hematopoietic Stem Cells via a Histamine-Dependent Feedback Loop. Cell Stem Cell *21*, 747-760.e747. 10.1016/j.stem.2017.11.003.
72. Olson, O.C., Kang, Y.-A., and Passegué, E. (2020). Normal Hematopoiesis Is a Balancing Act of Self-Renewal and Regeneration. Cold Spring Harbor Perspectives in Medicine *10*. 10.1101/cshperspect.a035519.
73. Bowling, S., Sritharan, D., Osorio, F.G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W.C., Fujiwara, Y., Li, B.E., et al. (2020). An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. Cell *181*, 1410-1422.e1427. 10.1016/j.cell.2020.04.048.
74. Lu, F., and Lionnet, T. (2021). Transcription Factor Dynamics. Cold Spring Harbor Perspectives in Biology *13*, a040949-a040949. 10.1101/cshperspect.a040949.
75. Aly, M., Ramdzan, Z.M., Nagata, Y., Balasubramanian, S.K., Hosono, N., Makishima, H., Visconte, V., Kuzmanovic, T., Adema, V., Nazha, A., et al. (2019). Distinct clinical and biological implications of CUX1 in myeloid neoplasms. Blood Advances *3*, 2164-2178. 10.1182/bloodadvances.2018028423.
76. Mcnerney, M.E., Brown, C.D., Wang, X., Bartom, E.T., Karmakar, S., Bandlamudi, C., Yu, S., Ko, J., Sandall, B.P., Stricker, T., et al. (2013). CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia. Blood *121*, 975-983. 10.1182/blood-2012-04-426965.
77. Supper, E., Rudat, S., Iyer, V., Droop, A., Wong, K., Thomas, P., Sauvageau, G., Adams, D.J., Wong, C.C., and Spinella, J.-f. (2021). Cut-like homeobox1 (Cux1) tumor suppressor gene haploinsufficiency induces apoptosis evasion to sustain myeloid leukaemia. Nature Communications *2048*, 1-20. 10.1038/s41467-021-22750-8.
78. An, N., Khan, S., Imgruet, M.K., Gurbuxani, S.K., Konecki, S.N., Burgess, M.R., and Mcnerney, M.E. (2018). Gene dosage effect of CUX1 in a murine model disrupts HSC homeostasis and controls the severity and mortality of MDS. Blood *131*, 2682-2697. 10.1182/blood-2017-10-810028.
79. Gao, T., Ptashkin, R., Bolton, K.L., Sirenko, M., Fong, C., Spitzer, B., Menghrajani, K., Ossa, J.E.A., Zhou, Y., Bernard, E., et al. (2021). Interplay between chromosomal alterations and gene mutations shapes the evolutionary trajectory of clonal hematopoiesis. Nature Communications *12*, 1-11. 10.1038/s41467-020-20565-7.
80. Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. New England Journal of Medicine *371*, 2488-2498. 10.1056/NEJMOA1408617/SUPPL_FILE/NEJMOA1408617_DISCLOSURES.PDF.
81. Pich, O., Reyes-Salazar, I., Gonzalez-Perez, A., and Lopez-Bigas, N. (2022). Discovering the drivers of clonal hematopoiesis. Nature Communications *13*, 1-12. 10.1038/s41467-022-31878-0.

82. Bolton, K.L., Koh, Y., Foote, M.B., Im, H., Jee, J., Sun, C.H., Safonov, A., Ptashkin, R., Moon, J.H., Lee, J.Y., et al. (2021). Clonal hematopoiesis is associated with risk of severe Covid-19. Nature Communications *12*, 1-8. 10.1038/s41467-021-26138-6.
83. Gallì, A., Todisco, G., Catamo, E., Sala, C., Elena, C., Pozzi, S., Bono, E., Ferretti, V.V., Rizzo, E., Molteni, E., et al. (2021). Relationship between clone metrics and clinical outcome in clonal cytopenia. Blood *138*, 965-976. 10.1182/BLOOD.2021011323.
84. Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., Yoon, C.J., Ellis, P., Wedge, D.C., Pellagatti, A., et al. (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood *122*, 3616-3627. 10.1182/BLOOD-2013-08-518886.
85. Tang, W.C., Walfred, Dietmann, S., Irie, N., Leitch, G., Harry, Floros, I., Vasileios, Bradshaw, R., Charles, Hackett, A., Jamie, Chinnery, F., Patrick, and Surani, A., M. (2015). A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell *161*, 1453-1467. 10.1016/j.cell.2015.04.053.
86. Onida, F., Kantarjian, H.M., Smith, T.L., Ball, G., Keating, M.J., Estey, E.H., Glassman, A.B., Albitar, M., Kwari, M.I., and Beran, M. (2002). Prognostic factors and scoring systems in chronic myelomonocytic leukemia: a retrospective analysis of 213 patients. Blood *99*, 840-849. 10.1182/BLOOD.V99.3.840.
87. Such, E., Cervera, J., Costa, D., Solé, F., Vallespí, T., Luño, E., Collado, R., Calasanz, M.J., Hernández-Rivas, J.M., Cigudosa, J.C., et al. (2011). Cytogenetic risk stratification in chronic myelomonocytic leukemia. Haematologica *96*, 375-383. 10.3324/HAEMATOL.2010.030957.
88. Wong, T.N., Ramsingh, G., Young, A.L., Miller, C.A., Touma, W., Welch, J.S., Lamprecht, T.L., Shen, D., Hundal, J., Fulton, R.S., et al. (2014). Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. Nature *518*, 552-555. 10.1038/nature13968.
89. Marcellino, B.K., Hoffman, R., Tripodi, J., Lu, M., Kosiorek, H., Mascarenhas, J., Rampal, R.K., Dueck, A., and Najfeld, V. (2018). Advanced forms of MPNs are accompanied by chromosomal abnormalities that lead to dysregulation of TP53. Blood Advances *2*, 3581-3589. 10.1182/BLOODADVANCES.2018024018.
90. Grinfeld, J., Nangalia, J., Baxter, E.J., Wedge, D.C., Angelopoulos, N., Cantrill, R., Godfrey, A.L., Papaemmanuil, E., Gundem, G., MacLean, C., et al. (2018). Classification and Personalized Prognosis in Myeloproliferative Neoplasms. New England Journal of Medicine *379*, 1416-1430. 10.1056/NEJMOA1716614/SUPPL_FILE/NEJMOA1716614_DISCLOSURES.PDF.
91. Nangalia, J., Massie, C.E., Baxter, E.J., Nice, F.L., Gundem, G., Wedge, D.C., Avezov, E., Li, J., Kollmann, K., Kent, D.G., et al. (2013). SomaticCALRMutations in Myeloproliferative Neoplasms with NonmutatedJAK2. New England Journal of Medicine *369*, 2391-2405. 10.1056/nejmoa1312542.
92. Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., Kantarjian, H., Raza, A., Levine, R.L., Neuberg, D., and Ebert, B.L. (2011). Clinical Effect of Point Mutations in Myelodysplastic Syndromes. New England Journal of Medicine *364*, 2496-2506. 10.1056/NEJMOA1013343/SUPPL_FILE/NEJMOA1013343_DISCLOSURES.PDF.
93. Bernard, E., Tuechler, H., Greenberg, P.L., Hasserjian, R.P., Arango Ossa, J.E., Nannya, Y., Devlin, S.M., Creignou, M., Pinel, P., Monnier, L., et al. (2022). Molecular

International Prognostic Scoring System for Myelodysplastic Syndromes. NEJM Evidence *1*. 10.1056/evidoa2200008.
94. Yoshizato, T., Nannya, Y., Atsuta, Y., Shiozawa, Y., Iijima-Yamashita, Y., Yoshida, K., Shiraishi, Y., Suzuki, H., Nagata, Y., Sato, Y., et al. (2017). Genetic abnormalities in myelodysplasia and secondary acute myeloid leukemia: impact on outcome of stem cell transplantation. Blood *129*, 2347-2358. 10.1182/blood-2016-12-754796.
95. Rubin, C.M., Larson, R.A., Anastasi, J., Winter, J.N., Thangavelu, M., Vardiman, J.W., Rowley, J.D., and Le Beau, M.M. (1990). t(3;21)(q26;q22): A Recurring Chromosomal Abnormality in Therapy-Related Myelodysplastic Syndrome and Acute Myeloid Leukemia. Blood *76*, 2594-2598. 10.1182/BLOOD.V76.12.2594.2594.
96. Makishima, H., Yoshizato, T., Yoshida, K., Sekeres, M.A., Radivoyevitch, T., Suzuki, H., Przychodzen, B.J., Nagata, Y., Meggendorfer, M., Sanada, M., et al. (2016). Dynamics of clonal evolution in myelodysplastic syndromes. Nature Genetics *49*, 204-212. 10.1038/ng.3742.
97. Haferlach, T., Nagata, Y., Grossmann, V., Okuno, Y., Bacher, U., Nagae, G., Schnittger, S., Sanada, M., Kon, A., Alpermann, T., et al. (2013). Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia *28*, 241-247. 10.1038/leu.2013.336.
98. Wong, C.C., Martincorena, I., Rust, A.G., Rashid, M., Alifrangis, C., Alexandrov, L.B., Tiffen, J.C., Kober, C., Green, A.R., Massie, C.E., et al. (2014). Inactivating CUX1 mutations promote tumorigenesis. Nature Genetics *46*, 33-38. 10.1038/ng.2846.
99. Gaidzik, V.I., Bullinger, L., Schlenk, R.F., Zimmermann, A.S., Rock, J., Paschka, P., Corbacioglu, A., Krauter, J., Schlegelberger, B., Ganser, A., et al. (2011). RUNX1 mutations in acute myeloid leukemia: Results from a comprehensive genetic and clinical analysis from the AML study group. Journal of Clinical Oncology *29*, 1364-1372. 10.1200/JCO.2010.30.7926.
100. Grimwade, D., Walker, H., Harrison, G., Oliver, F., Chatters, S., Harrison, C.J., Wheatley, K., Burnett, A.K., and Goldstone, A.H. (2001). The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. Blood *98*, 1312-1320. 10.1182/BLOOD.V98.5.1312.
101. Tyner, J.W., Tognon, C.E., Bottomly, D., Wilmot, B., Kurtz, S.E., Savage, S.L., Long, N., Schultz, A.R., Traer, E., Abel, M., et al. (2018). Functional genomic landscape of acute myeloid leukaemia. Nature *562*, 526-531. 10.1038/s41586-018-0623-z.
102. Cancer Genome Atlas Research, N., Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., et al. (2013). Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. New England Journal of Medicine *368*, 2059-2074. 10.1056/NEJMOA1301689/SUPPL_FILE/NEJMOA1301689_DISCLOSURES.PDF.
103. Fröhling, S., Schlenk, R.F., Kayser, S., Morhardt, M., Benner, A., DöHner, K., and DöHner, H. (2006). Cytogenetics and age are major determinants of outcome in intensively treated acute myeloid leukemia patients older than 60 years: results from AMLSG trial AML HD98-B. Blood *108*, 3280-3288. 10.1182/blood-2006-04-014324.
104. Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V.I., Paschka, P., Roberts, N.D., Potter, N.E., Heuser, M., Thol, F., Bolli, N., et al. (2016). Genomic Classification and

Prognosis in Acute Myeloid Leukemia. New England Journal of Medicine *374*, 2209-2221. 10.1056/nejmoa1516192.
105. Andersen, M.T., Andersen, M.K., Christiansen, D.H., and Pedersen-Bjergaard, J. (2008). NPM1 mutations in therapy-related acute myeloid leukemia with uncharacteristic features. Leukemia *22*, 951-955. 10.1038/leu.2008.17.
106. Sperling, A.S., Guerra, V.A., Kennedy, J.A., Yan, Y., Hsu, J.I., Wang, F., Nguyen, A.T., Miller, P.G., McConkey, M.E., Quevedo Barrios, V.A., et al. (2022). Lenalidomide promotes the development of TP53-mutated therapy-related myeloid neoplasms. Blood *140*, 1753-1763. 10.1182/BLOOD.2021014956.
107. Christiansen, D.H., Andersen, M.K., and Pedersen-Bjergaard, J. (2004). Mutations ofAML1are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. Blood *104*, 1474-1481. 10.1182/BLOOD-2004-02-0754.
108. Lindsley, R.C., Mar, B.G., Mazzola, E., Grauman, P.V., Shareef, S., Allen, S.L., Pigneux, A., Wetzler, M., Stuart, R.K., Erba, H.P., et al. (2015). Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. Blood *125*, 1367-1376. 10.1182/BLOOD-2014-11-610543.
109. Smith, S.M., Le Beau, M.M., Huo, D., Karrison, T., Sobecks, R.M., Anastasi, J., Vardiman, J.W., Rowley, J.D., and Larson, R.A. (2003). Clinical-cytogenetic associations in 306 patients with therapy-related myelodysplasia and myeloid leukemia: the University of Chicago series. Blood *102*, 43-52. 10.1182/blood-2002-11-3343.
110. Larsson, A.J.M., Johnsson, P., Hagemann-Jensen, M., Hartmanis, L., Faridani, O.R., Reinius, B., Segerstolpe, Å., Rivera, C.M., Ren, B., and Sandberg, R. (2019). Genomic encoding of transcriptional burst kinetics. Nature *565*, 251-254. 10.1038/s41586-018-0836-1.
111. Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. Cell *166*, 358-368. 10.1016/J.CELL.2016.05.025.
112. Wheat, J.C., Sella, Y., Willcockson, M., Skoultchi, A.I., Bergman, A., Singer, R.H., and Steidl, U. (2020). Single-molecule imaging of transcription dynamics in somatic stem cells. Nature *583*, 431-436. 10.1038/s41586-020-2432-4.
113. Wheat, J.C., and Steidl, U. (2021). Gene expression at a single-molecule level: implications for myelodysplastic syndromes and acute myeloid leukemia. Blood *138*, 625-636. 10.1182/BLOOD.2019004261.
114. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. Science *347*. 10.1126/SCIENCE.1260419/SUPPL_FILE/1260419_UHLEN.SM.PDF.
115. Cook, D.L., Gerber, A.N., and Tapscott, S.J. (1998). Modeling stochastic gene expression: Implications for haploinsufficiency. Proceedings of the National Academy of Sciences of the United States of America *95*, 15641-15646. 10.1073/PNAS.95.26.15641/ASSET/C97DAA68-20DB-482E-872D-3B37FC8BB7BD/ASSETS/GRAPHIC/PQ2684138003.JPEG.
116. Kemkemer, R., Schrank, S., Vogel, W., Gruler, H., and Kaufmann, D. (2002). Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 in vitro. Proceedings of the National Academy of Sciences of the United States of

America *99*, 13783-13788. 10.1073/PNAS.212386999/ASSET/FF90FE4B-2309-4632-A72C-D5F64E23FA45/ASSETS/GRAPHIC/PQ2123869007.JPEG.
117. Shaffer, S.M., Dunagin, M.C., Torborg, S.R., Torre, E.A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P.A., Xiao, M., et al. (2017). Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. Nature *546*, 431-435. 10.1038/nature22794.
118. Shaffer, S.M., Emert, B.L., Reyes Hueros, R.A., Cote, C., Harmange, G., Schaff, D.L., Sizemore, A.E., Gupte, R., Torre, E., Singh, A., et al. (2020). Memory Sequencing Reveals Heritable Single-Cell Gene Expression Programs Associated with Distinct Cellular Behaviors. Cell *182*, 947-959.e917. 10.1016/J.CELL.2020.07.003.
119. Ryl, T., Kuchen, E.E., Bell, E., Shao, C., Flórez, A.F., Mönke, G., Gogolin, S., Friedrich, M., Lamprecht, F., Westermann, F., and Höfer, T. (2017). Cell-Cycle Position of Single MYC-Driven Cancer Cells Dictates Their Susceptibility to a Chemotherapeutic Drug. Cell Systems *5*, 237-250.e238. 10.1016/j.cels.2017.07.005.
120. Hansen, A.S., Amitai, A., Cattoglio, C., Tjian, R., and Darzacq, X. (2019). Guided nuclear exploration increases CTCF target search efficiency. Nature Chemical Biology *16*, 257-266. 10.1038/s41589-019-0422-3.
121. Izeddin, I., Récamier, V., Bosanac, L., Cissé, I.I., Boudarene, L., Dugast-Darzacq, C., Proux, F., Bénichou, O., Voituriez, R., Bensaude, O., et al. (2014). Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. eLife *2014*. 10.7554/ELIFE.02230.
122. Marklund, E., van Oosten, B., Mao, G., Amselem, E., Kipper, K., Sabantsev, A., Emmerich, A., Globisch, D., Zheng, X., Lehmann, L.C., et al. (2020). DNA surface exploration and operator bypassing during target search. Nature *583*, 858-861. 10.1038/s41586-020-2413-7.
123. Li, L., Liu, H., Dong, P., Li, D., Legant, W.R., Grimm, J.B., Lavis, L.D., Betzig, E., Tjian, R., and Liu, Z. (2016). Real-time imaging of Huntingtin aggregates diverting target search and gene transcription. eLife *5*. 10.7554/eLife.17056.001.
124. Castellanos, M., Mothi, N., and Muñoz, V. (2020). Eukaryotic transcription factors can track and control their target genes using DNA antennas. Nature Communications *11*, 1-13. 10.1038/s41467-019-14217-8.
125. Chen, J., Zhang, Z., Li, L., Chen, B.C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., et al. (2014). Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. Cell *156*, 1274-1285. 10.1016/J.CELL.2014.01.062.
126. Yu, M., Mazor, T., Huang, H., Huang, H.-T., Kathrein, K.L., Woo, A.J., Chouinard, C.R., Labadorf, A., Akie, T.E., Moran, T.B., et al. (2012). Article Direct Recruitment of Polycomb Repressive Complex 1 to Chromatin by Core Binding Transcription Factors. Molecular Cell *45*, 330-343. 10.1016/j.molcel.2011.11.032.
127. Minderjahn, J., Schmidt, A., Fuchs, A., Schill, R., Raithel, J., Babina, M., Schmidl, C., Gebhard, C., Schmidhofer, S., Mendes, K., et al. (2020). Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. Nature Communications *11*, 1-16. 10.1038/S41467-019-13960-2.
128. Levine, J.H., Lin, Y., and Elowitz, M.B. (2013). Functional roles of pulsing in genetic circuits. Science *342*, 1193-1200. 10.1126/SCIENCE.1239999/ASSET/37F7A53A-92C4-4FFA-A640-8342411C9B2F/ASSETS/GRAPHIC/342_1193_F4.JPEG.

129. Stavreva, D.A., Garcia, D.A., Fettweis, G., Gudla, P.R., Zaki, G.F., Soni, V., McGowan, A., Williams, G., Huynh, A., Palangat, M., et al. (2019). Transcriptional Bursting and Co-bursting Regulation by Steroid Hormone Release Pattern and Transcription Factor Mobility. Molecular Cell *75*, 1161-1177.e1111. 10.1016/J.MOLCEL.2019.06.042.
130. Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A.J., Elowitz, M.B., and Alon, U. (2004). Dynamics of the p53-Mdm2 feedback loop in individual cells. Nature Genetics *36*, 147-150. 10.1038/NG1293.
131. Purvis, J.E., Karhohs, K.W., Mock, C., Batchelor, E., Loewer, A., and Lahav, G. (2012). p53 dynamics control cell fate. Science *336*, 1440-1444. 10.1126/SCIENCE.1218351/SUPPL_FILE/PURVIS-SOM.PDF.
132. Shimojo, H., Ohtsuka, T., and Kageyama, R. (2008). Oscillations in Notch Signaling Regulate Maintenance of Neural Progenitors. Neuron *58*, 52-64. 10.1016/J.NEURON.2008.02.014.
133. Imayoshi, I., Isomura, A., Harima, Y., Kawaguchi, K., Kori, H., Miyachi, H., Fujiwara, T., Ishidate, F., and Kageyama, R. (2013). Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. Science *342*, 1203-1208. 10.1126/SCIENCE.1242366/SUPPL_FILE/IMAYOSHI.SM.PDF.
134. Kull, T., Wehling, A., Etzrodt, M., Auler, M., Dettinger, P., Aceto, N., and Schroeder, T. (2022). NfκB signaling dynamics and their target genes differ between mouse blood cell types and induce distinct cell behavior. Blood *140*, 99-111. 10.1182/blood.2021012918.
135. Natarajan, A., Yardimci, G.G., Sheffield, N.C., Crawford, G.E., and Ohler, U. (2012). Predicting cell-type–specific gene expression from regions of open chromatin. Genome Research *22*, 1711-1722. 10.1101/GR.135129.111.
136. Crump, N.T., Ballabio, E., Godfrey, L., Thorne, R., Repapi, E., Kerry, J., Tapia, M., Hua, P., Lagerholm, C., Filippakopoulos, P., et al. (2021). BET inhibition disrupts transcription but retains enhancer-promoter contact. Nature Communications *12*, 1-15. 10.1038/s41467-020-20400-z.
137. Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A., and Cavalli, G. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572.e524. 10.1016/J.CELL.2017.09.043.
138. Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P., and Bickmore, W.A. (2019). Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. Molecular Cell *76*, 473-484.e477. 10.1016/j.molcel.2019.07.038.
139. Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. Cell *149*, 1233-1244. 10.1016/J.CELL.2012.03.051.
140. Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A., and Blobel, G.A. (2014). Reactivation of Developmentally Silenced Globin Genes by Forced Chromatin Looping. Cell *158*, 849-860. 10.1016/J.CELL.2014.05.050.
141. Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., and Weiner, O.D. (2019). Live-cell imaging reveals enhancer-dependent sox2 transcription in the absence of enhancer proximity. eLife *8*. 10.7554/ELIFE.41769.

142. Bartman, C.R., Hsu, S.C., Hsiung, C.C.S., Raj, A., and Blobel, G.A. (2016). Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. Molecular Cell *62*, 237-247. 10.1016/J.MOLCEL.2016.03.007.
143. Allahyar, A., Vermeulen, C., Bouwman, B.A.M., Krijger, P.H.L., Verstegen, M.J.A.M., Geeven, G., van Kranenburg, M., Pieterse, M., Straver, R., Haarhuis, J.H.I., et al. (2018). Enhancer hubs and loop collisions identified from single-allele topologies. Nature Genetics *50*, 1151-1160. 10.1038/s41588-018-0161-5.
144. Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., et al. (2009). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nature Genetics *42*, 53-61. 10.1038/ng.496.
145. Li, J., Dong, A., Saydaminova, K., Chang, H., Wang, G., Ochiai, H., Yamamoto, T., and Pertsinidis, A. (2019). Single-Molecule Nanoscopy Elucidates RNA Polymerase II Transcription at Single Genes in Live Cells. Cell *178*, 491-506.e428. 10.1016/J.CELL.2019.05.029.
146. Mir, M., Stadler, M.R., Ortiz, S.A., Hannon, C.E., Harrison, M.M., Darzacq, X., and Eisen, M.B. (2018). Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos. eLife *7*. 10.7554/ELIFE.40497.
147. Gu, B., Swigut, T., Spencley, A., Bauer, M.R., Chung, M., Meyer, T., and Wysocka, J. (2018). Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. Science *359*, 1050-1055. 10.1126/SCIENCE.AAO3136/SUPPL_FILE/AAO3136S5.MP4.
148. Petrovic, J., Zhou, Y., Fasolino, M., Goldman, N., Schwartz, G.W., Mumbach, M.R., Nguyen, S.C., Rome, K.S., Sela, Y., Zapataro, Z., et al. (2019). Oncogenic Notch Promotes Long-Range Regulatory Interactions within Hyperconnected 3D Cliques. Molecular Cell *73*, 1174-1190.e1112. 10.1016/J.MOLCEL.2019.01.006.
149. Therizols, P., Illingworth, R.S., Courilleau, C., Boyle, S., Wood, A.J., and Bickmore, W.A. (2014). Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science *346*, 1238-1242. 10.1126/SCIENCE.1259587/SUPPL_FILE/THERIZOLS.SM.PDF.
150. Lundgren, M., Chow, C.M., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. (2000). Transcription Factor Dosage Affects Changes in Higher Order Chromatin Structure Associated with Activation of a Heterochromatic Gene. Cell *103*, 733-743. 10.1016/S0092-8674(00)00177-X.
151. Liu, J., Perumal, N.B., Oldfield, C.J., Su, E.W., Uversky, V.N., and Dunker, A.K. (2006). Intrinsic disorder in transcription factors. Biochemistry *45*, 6873-6888. 10.1021/BI0602718/ASSET/IMAGES/MEDIUM/BI0602718N00001.GIF.
152. Zou, H., Pan, T., Gao, Y., Chen, R., Li, S., Guo, J., Tian, Z., Xu, G., Xu, J., Ma, Y., and Li, Y. (2022). Pan-cancer assessment of mutational landscape in intrinsically disordered hotspots reveals potential driver genes. Nucleic Acids Research *50*, e49-e49. 10.1093/NAR/GKAC028.
153. Sagie Brodsky, A., Jana, T., Mittelman, K., Chapal, M., Krishna Kumar, D., Carmi, M., Brodsky, S., and Barkai, N. (2020). Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. Molecular Cell *79*, 459-471.e454. 10.1016/j.molcel.2020.05.032.

154. Ravarani, C.N.J., Erkina, T.Y., Baets, G.D., Dudman, D.C., Erkine, A.M., and Babu, M.M. (2018). High-throughput discovery of functional disordered regions: investigation of transactivation domains. Molecular Systems Biology *14*, e8190-e8190. 10.15252/MSB.20188190.
155. Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell *175*, 1842-1855.e1816. 10.1016/J.CELL.2018.10.042.
156. Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., and Tjian, R. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*. 10.1126/SCIENCE.AAR2555/SUPPL_FILE/AAR2555_CHONG_SM.PDF.
157. Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*. 10.1126/SCIENCE.AAR3958/SUPPL_FILE/AAR3958_SABARI_SM_TABLE_S3.XLSX.
158. Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science *361*, 1341-1345. 10.1126/SCIENCE.AAU0320/ASSET/412087A1-DE48-433F-BC76-9CA9B70B9A1B/ASSETS/GRAPHIC/361_1341_F4.JPEG.
159. Cho, W.K., Spille, J.H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412-415. 10.1126/SCIENCE.AAR4199/SUPPL_FILE/AAR4199_S3.MOV.
160. Deviri, D., and Safran, S.A. (2021). Physical theory of biological noise buffering by multicomponent phase separation. Proceedings of the National Academy of Sciences of the United States of America *118*, e2100099118-e2100099118. 10.1073/PNAS.2100099118/SUPPL_FILE/PNAS.2100099118.SAPP.PDF.
161. Ng, K.P., Hu, Z., Ebrahem, Q., Negrotto, S., Lausen, J., and Saunthararajah, Y. (2013). Runx1 deficiency permits granulocyte lineage commitment but impairs subsequent maturation. Oncogenesis *2*, e78-e78. 10.1038/oncsis.2013.41.
162. Sinha, T., Lammerts van Bueren, K., Dickel, D.E., Zlatanova, I., Thomas, R., Lizama, C.O., Xu, S.M., Zovein, A.C., Ikegami, K., Moskowitz, I.P., et al. (2022). Differential Etv2 threshold requirement for endothelial and erythropoietic development. Cell Reports *39*, 110881-110881. 10.1016/J.CELREP.2022.110881.
163. Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., Pasparakis, M., Milani, P., Bulyk, M.L., and Natoli, G. (2010). Noncooperative Interactions between Transcription Factors and Clustered DNA Binding Sites Enable Graded Transcriptional Responses to Environmental Inputs. Molecular Cell *37*, 418-428. 10.1016/j.molcel.2010.01.016.
164. Staber, P.B., Zhang, P., Ye, M., Welner, R.S., Nombela-Arrieta, C., Bach, C., Kerenyi, M., Bartholdy, B.A., Zhang, H., Alberich-Jordà, M., et al. (2013). Sustained PU.1 Levels Balance Cell-Cycle Regulators to Prevent Exhaustion of Adult Hematopoietic Stem Cells. Molecular Cell *49*, 934-946. 10.1016/j.molcel.2013.01.007.
165. Naqvi, S., Kim, S., Hoskens, H., Matthews, H.S., Spritz, R.A., Klein, O.D., Hallgrímsson, B., Swigut, T., Claes, P., Pritchard, J.K., and Wysocka, J. (2022). Precise

modulation of transcription factor levels reveals drivers of dosage sensitivity. bioRxiv, 2022.2006.2013.495964-492022.495906.495913.495964. 10.1101/2022.06.13.495964.
166. Goyama, S., Schibler, J., Cunningham, L., Zhang, Y., Rao, Y., Nishimoto, N., Nakagawa, M., Olsson, A., Wunderlich, M., Link, K.A., et al. (2013). Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. The Journal of Clinical Investigation *123*, 3876-3888. 10.1172/JCI68557.
167. Staber, P.B., Zhang, P., Ye, M., Welner, R.S., Levantini, E., Di Ruscio, A., Ebralidze, A.K., Bach, C., Zhang, H., Zhang, J., et al. (2014). The Runx-PU.1 pathway preserves normal and AML/ETO9a leukemic stem cells. Blood *124*, 2391-2399. 10.1182/BLOOD-2014-01-550855.
168. Jones, L.C., Lin, M.L., Chen, S.S., Krug, U., Hofmann, W.K., Lee, S., Lee, Y.H., and Phillip Koeffler, H. (2002). Expression of C/EBPβ from the C/ebpα gene locus is sufficient for normal hematopoiesis in vivo. Blood *99*, 2032-2036. 10.1182/BLOOD.V99.6.2032.
169. Hirai, H., Zhang, P., Dayaram, T., Hetherington, C.J., Mizuno, S.I., Imanishi, J., Akashi, K., and Tenen, D.G. (2006). C/EBPβ is required for 'emergency' granulopoiesis. Nature Immunology *7*, 732-739. 10.1038/ni1354.
170. Ye, M., Zhang, H., Yang, H., Koche, R., Staber, P.B., Cusan, M., Levantini, E., Welner, R.S., Bach, C.S., Zhang, J., et al. (2015). Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia. Cell Stem Cell *17*, 611-623. 10.1016/j.stem.2015.08.011.
171. Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and Predicting Haploinsufficiency in the Human Genome. PLOS Genetics *6*, e1001154-e1001154. 10.1371/JOURNAL.PGEN.1001154.
172. Assi, S.A., Imperato, M.R., Coleman, D.J.L., Pickin, A., Potluri, S., Ptasinska, A., Chin, P.S., Blair, H., Cauchy, P., James, S.R., et al. (2018). Subtype-specific regulatory network rewiring in acute myeloid leukemia. Nature Genetics *51*, 151-162. 10.1038/s41588-018-0270-1.
173. Jotte, M.R.M., and McNerney, M.E. (2022). The significance of CUX1 and chromosome 7 in myeloid malignancies. Current opinion in hematology *29*, 92-102. 10.1097/MOH.0000000000000699.
174. Ben-David, U., and Amon, A. (2019). Context is everything: aneuploidy in cancer. Nature Reviews Genetics *21*, 44-62. 10.1038/S41576-019-0171-X.
175. Saiki, R., Momozawa, Y., Nannya, Y., Nakagawa, M.M., Ochi, Y., Yoshizato, T., Terao, C., Kuroda, Y., Shiraishi, Y., Chiba, K., et al. (2021). Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis. Nature Medicine *27*, 1239-1249. 10.1038/s41591-021-01411-9.
176. Chunduri, N.K., Menges, P., Zhang, X., Wieland, A., Gotsmann, V.L., Mardin, B.R., Buccitelli, C., Korbel, J.O., Willmund, F., Kschischo, M., et al. (2021). Systems approaches identify the consequences of monosomy in somatic human cells. Nature Communications *12*, 1-17. 10.1038/s41467-021-25288-x.
177. Rücker, F.G., Schlenk, R.F., Bullinger, L., Kayser, S., Teleanu, V., Kett, H., Habdank, M., Kugler, C.M., Holzmann, K., Gaidzik, V.I., et al. (2012). TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. Blood *119*, 2114-2121. 10.1182/BLOOD-2011-08-375758.

178. Haase, D., Stevenson, K.E., Neuberg, D., Maciejewski, J.P., Nazha, A., Sekeres, M.A., Ebert, B.L., Garcia-Manero, G., Haferlach, C., Haferlach, T., et al. (2019). TP53 mutation status divides myelodysplastic syndromes with complex karyotypes into distinct prognostic subgroups. Leukemia *33*, 1747-1758. 10.1038/s41375-018-0351-2.
179. Beach, R.R., Ricci-Tam, C., Brennan, C.M., Silberman, R.E., Springer, M., and Amon, A. (2017). Aneuploidy Causes Non-genetic Individuality. Cell *169*, 229-242. 10.1016/j.cell.2017.03.021.
180. Farquhar, K.S., Charlebois, D.A., Szenk, M., Cohen, J., Nevozhay, D., and Balázsi, G. (2019). Role of network-mediated stochasticity in mammalian drug resistance. Nature Communications *10*, 1-14. 10.1038/s41467-019-10330-w.
181. Mileyko, Y., Joh, R.I., and Weitz, J.S. (2008). Small-scale copy number variation and large-scale changes in gene expression. Proceedings of the National Academy of Sciences of the United States of America *105*, 16659-16664. 10.1073/PNAS.0806239105/SUPPL_FILE/APPENDIX_PDF.PDF.
182. Schukken, K.M., and Sheltzer, J.M. (2022). Extensive protein dosage compensation in aneuploid human cancers. Genome Research *32*, 1254-1270. 10.1101/GR.276378.121.
183. Liu, Y., Chen, C., Xu, Z., Scuoppo, C., Rillahan, C.D., Gao, J., Spitzer, B., Bosbach, B., Kastenhuber, E.R., Baslan, T., et al. (2016). Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. Nature *531*, 471-475. 10.1038/nature17157.
184. Stoddart, A., Nakitandwe, J., Chen, S.C., Downing, J.R., and Le Beau, M.M. (2015). Haploinsufficient loss of multiple 5q genes may fine-tune Wnt signaling in del(5q) therapy-related myeloid neoplasms. Blood *126*, 2899-2901. 10.1182/BLOOD-2015-10-673228.
185. Joslin, J.M., Fernald, A.A., Tennant, T.R., Davis, E.M., Kogan, S.C., Anastasi, J., Crispino, J.D., and Le Beau, M.M. (2007). Haploinsufficiency of EGR1, a candidate gene in the del(5q), leads to the development of myeloid disorders. Blood *110*, 719-726. 10.1182/BLOOD-2007-01-068809.
186. Stoddart, A., Fernald, A.A., Davis, E.M., McNerney, M.E., and Le Beau, M.M. (2022). EGR1 Haploinsufficiency Confers a Fitness Advantage to Hematopoietic Stem Cells Following Chemotherapy. Experimental Hematology *115*, 54-67. 10.1016/J.EXPHEM.2022.08.003.
187. Pitsouli, C., and Perrimon, N. (2013). The Homeobox Transcription Factor Cut Coordinates Patterning and Growth During
Drosophila
Airway Remodeling. Science Signaling *6*, ra12-ra12. 10.1126/scisignal.2003424.
188. Grueber, W.B., Jan, L.Y., and Jan, Y.N. (2003). Different Levels of the Homeodomain Protein Cut Regulate Distinct Dendrite Branching Patterns of Drosophila Multidendritic Neurons. Cell *112*, 805-818. 10.1016/s0092-8674(03)00160-0.
189. Arthur, R.K., An, N., Khan, S., and Mcnerney, M.E. (2017). The haploinsufficient tumor suppressor, CUX1, acts as an analog transcriptional regulator that controls target genes through distal enhancers that loop to target promoters. Nucleic Acids Research *45*, 6350-6361. 10.1093/nar/gkx218.
190. Schwartz, J.R., Ma, J., Lamprecht, T., Walsh, M., Wang, S., Bryant, V., Song, G., Wu, G., Easton, J., Kesserwan, C., et al. (2017). The genomic landscape of pediatric myelodysplastic syndromes. Nature Communications *8*, 1-10. 10.1038/s41467-017-01590-5.

191. Imgruet, M.K., Lutze, J., An, N.N.A., Hu, B., Khan, S., Kurkewich, J., Martinez, T.C., Wolfgeher, D.J., Gurbuxani, S., Kron, S.J., and McNerney, M.E. (2021). Loss of a 7q gene, CUX1 , disrupts epigenetic-driven DNA repair and drives therapy-related myeloid neoplasms. Blood. 10.1182/blood.2020009195.
192. Baker, S.J., Ma'ayan, A., Lieu, Y.K., John, P., Reddy, M.V.R., Chen, E.Y., Duan, Q., Snoeck, H.W., and Reddy, E.P. (2014). B-myb is an essential regulator of hematopoietic stem cell and myeloid progenitor cell development. Proceedings of the National Academy of Sciences of the United States of America *111*, 3122-3127. 10.1073/PNAS.1315464111/SUPPL_FILE/SD01.XLS.
193. Clarke, M., Dumon, S., Ward, C., Jäger, R., Freeman, S., Dawood, B., Sheriff, L., Lorvellec, M., Kralovics, R., Frampton, J., and García, P. (2012). MYBL2 haploinsufficiency increases susceptibility to age-related haematopoietic neoplasia. Leukemia *27*, 661-670. 10.1038/leu.2012.241.
194. Heinrichs, S., Conover, L.F., Bueso-Ramos, C.E., Kilpivaara, O., Stevenson, K., Neuberg, D., Loh, M.L., Wu, W.S., Rodig, S.J., Garcia-Manero, G., et al. (2013). MYBL2 is a sub-haploinsufficient tumor suppressor gene in myeloid malignancy. eLife *2013*. 10.7554/ELIFE.00825.
195. Braulke, F., Müller-Thomas, C., Götze, K., Platzbecker, U., Germing, U., Hofmann, W.K., Giagounidis, A.A.N., Lübbert, M., Greenberg, P.L., Bennett, J.M., et al. (2015). Frequency of del(12p) is commonly underestimated in myelodysplastic syndromes: Results from a German diagnostic study in comparison with an international control group. Genes, Chromosomes and Cancer *54*, 809-817. 10.1002/GCC.22292.
196. Feurstein, S., Rücker, F.G., Bullinger, L., Hofmann, W., Manukjan, G., Göhring, G., Lehmann, U., Heuser, M., Ganser, A., Döhner, K., et al. (2014). Haploinsufficiency of ETV6 and CDKN1B in patients with acute myeloid leukemia and complex karyotype. BMC Genomics *15*, 1-11. 10.1186/1471-2164-15-784/FIGURES/5.
197. Green, S.M., Coyne, H.J., McIntosh, L.P., and Graves, B.J. (2010). DNA Binding by the ETS Protein TEL (ETV6) Is Regulated by Autoinhibition and Self-association. Journal of Biological Chemistry *285*, 18496-18504. 10.1074/JBC.M109.096958.
198. Churpek, J.E., and Bresnick, E.H. (2019). Transcription factor mutations as a cause of familial myeloid neoplasms. Journal of Clinical Investigation *129*, 476-488. 10.1172/JCI120854.
199. Fisher, M.H., Kirkpatrick, G.D., Stevens, B., Jones, C., Callaghan, M., Rajpurkar, M., Fulbright, J., Cooper, M.A., Rowley, J., Porter, C.C., et al. (2020). ETV6 germline mutations cause HDAC3/NCOR2 mislocalization and upregulation of interferon response genes. JCI Insight *5*. 10.1172/JCI.INSIGHT.140332.
200. Bernard, E., Nannya, Y., Hasserjian, R.P., Devlin, S.M., Tuechler, H., Medina-Martinez, J.S., Yoshizato, T., Shiozawa, Y., Saiki, R., Malcovati, L., et al. (2020). Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. Nature Medicine *26*, 1549-1556. 10.1038/s41591-020-1008-z.
201. Boettcher, S., Miller, P.G., Sharma, R., McConkey, M., Leventhal, M., Krivtsov, A.V., Giacomelli, A.O., Wong, W., Kim, J., Chao, S., et al. (2019). A dominant-negative effect drives selection of TP53 missense mutations in myeloid malignancies. Science *365*, 599-604. 10.1126/SCIENCE.AAX3649/SUPPL_FILE/AAX3649_BOETTCHER_SM.PDF.

202. Venkatachalam, S., Shi, Y.P., Jones, S.N., Vogel, H., Bradley, A., Pinkel, D., and Donehower, L.A. (1998). Retention of wild-type p53 in tumors from p53 heterozygous mice: Reduction of p53 dosage can promote cancer formation. EMBO Journal *17*, 4657-4667. 10.1093/EMBOJ/17.16.4657.
203. Davoli, T., Uno, H., Wooten, E.C., and Elledge, S.J. (2017). Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. Science *355*. 10.1126/SCIENCE.AAF8399/SUPPL_FILE/AAF8399-DAVOLI-SM.PDF.
204. Fischer, K., Fröhling, S., Scherer, S.W., Brown, J.M.A., Scholl, C., Stilgenbauer, S., Tsui, L.C., Lichter, P., and Döhner, H. (1997). Molecular Cytogenetic Delineation of Deletions and Translocations Involving Chromosome Band 7q22 in Myeloid Leukemias. Blood *89*, 2036-2041. 10.1182/BLOOD.V89.6.2036.
205. Wong, J.C., Weinfurtner, K.M., Alzamora, M.d.p., Kogan, S.C., Burgess, M.R., Zhang, Y., Nakitandwe, J., Ma, J., Cheng, J., Chen, S.-C., et al. (2015). Functional evidence implicating chromosome 7q22 haploinsufficiency in myelodysplastic syndrome pathogenesis. eLife *4*. 10.7554/ELIFE.07839.
206. Wong, J.C.Y., Zhang, Y., Lieuw, K.H., Tran, M.T., Forgo, E., Weinfurtner, K., Alzamora, P., Kogan, S.C., Akagi, K., Wolff, L., et al. (2010). Use of chromosome engineering to model a segmental deletion of chromosome band 7q22 found in myeloid malignancies. Blood *115*, 4524-4532. 10.1182/BLOOD-2009-07-232504.
207. Ouillette, P., Collins, R., Shakhan, S., Li, J., Li, C., Shedden, K., and Malek, S.N. (2011). The prognostic significance of various 13q14 deletions in chronic lymphocytic leukemia. Clinical Cancer Research *17*, 6778-6790. 10.1158/1078-0432.CCR-11-0785/84430/AM/THE-PROGNOSTIC-SIGNIFICANCE-OF-VARIOUS-13Q14.
208. Rego, E.M., Wang, Z.G., Peruzzi, D., He, L.Z., Cordon-Cardo, C., and Pandolfi, P.P. (2001). Role of Promyelocytic Leukemia (Pml) Protein in Tumor Suppression. Journal of Experimental Medicine *193*, 521-530. 10.1084/JEM.193.4.521.
209. Welch, J.S., Klco, J.M., Varghese, N., Nagarajan, R., and Ley, T.J. (2011). Rara haploinsufficiency modestly influences the phenotype of acute promyelocytic leukemia in mice. Blood *117*, 2460-2468. 10.1182/BLOOD-2010-08-300087.
210. Largaespada, D.A. (2001). Haploinsufficiency for Tumor SuppressionThe Hazards of Being Single and Living a Long Time. Journal of Experimental Medicine *193*, F15-F18. 10.1084/JEM.193.4.F15.
211. Gröschel, S., Sanders, M.A., Hoogenboezem, R., De Wit, E., Bouwman, B.A.M., Erpelinck, C., Van Der Velden, V.H.J., Havermans, M., Avellino, R., Van Lom, K., et al. (2014). A Single Oncogenic Enhancer Rearrangement Causes Concomitant EVI1 and GATA2 Deregulation in Leukemia. Cell *157*, 369-381. 10.1016/J.CELL.2014.02.019.
212. Katayama, S., Suzuki, M., Yamaoka, A., Keleku-Lukwete, N., Katsuoka, F., Otsuki, A., Kure, S., Engel, J.D., and Yamamoto, M. (2017). GATA2 haploinsufficiency accelerates EVI1-driven leukemogenesis. Blood *130*, 908-919. 10.1182/BLOOD-2016-12-756767.
213. Gröschel, S., Sanders, M.A., Hoogenboezem, R., Zeilemaker, A., Havermans, M., Erpelinck, C., Bindels, E.M.J., Beverloo, H.B., Döhner, H., Löwenberg, B., et al. (2015). Mutational spectrum of myeloid malignancies with inv(3)/t(3;3) reveals a predominant involvement of RAS/RTK signaling pathways. Blood *125*, 133-139. 10.1182/BLOOD-2014-07-591461.

214. Wang, Q., Stacy, T., Miller, J.D., Lewis, A.F., Gu, T.L., Huang, X., Bushweller, J.H., Bories, J.C., Alt, F.W., Ryan, G., et al. (1996). The CBFβ Subunit Is Essential for CBFα2 (AML1) Function In Vivo. Cell *87*, 697-708. 10.1016/S0092-8674(00)81389-6.
215. De Bruijn, M., and Dzierzak, E. (2017). Runx transcription factors in the development and function of the definitive hematopoietic system. Blood *129*, 2061-2069. 10.1182/BLOOD-2016-12-689109.
216. Yan, M., Burel, S.A., Peterson, L.F., Kanbe, E., Iwasaki, H., Boyapati, A., Hines, R., Akashi, K., and Zhang, D.E. (2004). Deletion of an AML1-ETO C-terminal NcoR/SMRT-interacting region strongly induces leukemia development. Proceedings of the National Academy of Sciences of the United States of America *101*, 17186-17191. 10.1073/PNAS.0406702101/SUPPL_FILE/06702FIG7.PDF.
217. Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S., and Liu, J.M. (1998). ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proceedings of the National Academy of Sciences of the United States of America *95*, 10860-10865. 10.1073/PNAS.95.18.10860/ASSET/C561F173-71ED-42F2-B115-33E459510D16/ASSETS/GRAPHIC/PQ1982691005.JPEG.
218. Al-Harbi, S., Aljurf, M., Mohty, M., Almohareb, F., and Ahmed, S.O.A. (2020). An update on the molecular pathogenesis and potential therapeutic targeting of AML with t(8;21)(q22;q22.1);RUNX1-RUNX1T1. Blood Advances *4*, 229-238. 10.1182/BLOODADVANCES.2019000168.
219. Pabst, T., Mueller, B.U., Harakawa, N., Schoch, C., Haferlach, T., Behre, G., Hiddemann, W., Zhang, D.E., and Tenen, D.G. (2001). AML1–ETO downregulates the granulocytic differentiation factor C/EBPα in t(8;21) myeloid leukemia. Nature Medicine *7*, 444-451. 10.1038/86515.
220. Choi, Y., Elagib, K.E., Delehanty, L.L., and Goldfarb, A.N. (2006). Erythroid Inhibition by the Leukemic Fusion AML1-ETO Is Associated with Impaired Acetylation of the Major Erythroid Transcription Factor GATA-1. Cancer Research *66*, 2990-2996. 10.1158/0008-5472.CAN-05-2944.
221. Vangala, R.K., Heiss-Neumann, M.S., Rangatia, J.S., Singh, S.M., Schoch, C., Tenen, D.G., Hiddemann, W., and Behre, G. (2003). The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. Blood *101*, 270-277. 10.1182/BLOOD-2002-04-1288.
222. Ptasinska, A., Assi, S.A., Martinez-Soria, N., Imperato, M.R., Piper, J., Cauchy, P., Pickin, A., James, S.R., Hoogenkamp, M., Williamson, D., et al. (2014). Identification of a Dynamic Core Transcriptional Network in t(8;21) AML that Regulates Differentiation Block and Self-Renewal. Cell Reports *8*, 1974-1988. 10.1016/J.CELREP.2014.08.024.
223. Loke, J., Assi, S.A., Imperato, M.R., Ptasinska, A., Cauchy, P., Grabovska, Y., Soria, N.M., Raghavan, M., Delwel, H.R., Cockerill, P.N., et al. (2017). RUNX1-ETO and RUNX1-EVI1 Differentially Reprogram the Chromatin Landscape in t(8;21) and t(3;21) AML. Cell Reports *19*, 1654-1668. 10.1016/J.CELREP.2017.05.005.
224. Ben-Ami, O., Friedman, D., Leshkowitz, D., Goldenberg, D., Orlovsky, K., Pencovich, N., Lotem, J., Tanay, A., and Groner, Y. (2013). Addiction of t(8;21) and inv(16) Acute Myeloid Leukemia to Native RUNX1. Cell Reports *4*, 1131-1143. 10.1016/J.CELREP.2013.08.020.

225. Lukasik, S.M., Zhang, L., Corpora, T., Tomanicek, S., Li, Y., Kundu, M., Hartman, K., Liu, P.P., Laue, T.M., Biltonen, R.L., et al. (2002). Altered affinity of CBFβ-SMMHC for Runx1 explains its role in leukemogenesis. Nature Structural Biology *9*, 674-679. 10.1038/nsb831.
226. Zhen, T., Cao, Y., Ren, G., Zhao, L., Hyde, R.K., Lopez, G., Feng, D., Alemu, L., Zhao, K., and Paul Liu, P. (2020). RUNX1 and CBFβ-SMMHC transactivate target genes together in abnormal myeloid progenitors for leukemia development. Blood *136*, 2373-2385. 10.1182/BLOOD.2020007747.
227. Hyde, R.K., Zhao, L., Alemu, L., and Liu, P.P. (2015). Runx1 is required for hematopoietic defects and leukemogenesis in Cbfb-MYH11 knock-in mice. Leukemia *29*, 1771-1778. 10.1038/leu.2015.58.
228. Qi, J., Singh, S., Hua, W.K., Cai, Q., Chao, S.W., Li, L., Liu, H., Ho, Y., McDonald, T., Lin, A., et al. (2015). HDAC8 Inhibition Specifically Targets Inv(16) Acute Myeloid Leukemic Stem Cells by Restoring p53 Acetylation. Cell Stem Cell *17*, 597-610. 10.1016/J.STEM.2015.08.004.
229. Estevez, B., Borst, S., Jarocha, D., Sudunagunta, V., Gonzalez, M., Garifallou, J., Hakonarson, H., Gao, P., Tan, K., Liu, P., et al. (2021). RUNX-1 haploinsufficiency causes a marked deficiency of megakaryocyte-biased hematopoietic progenitor cells. Blood *137*, 2662-2675. 10.1182/BLOOD.2020006389.
230. Ichikawa, M., Asai, T., Saito, T., Yamamoto, G., Seo, S., Yamazaki, I., Yamagata, T., Mitani, K., Chiba, S., Hirai, H., et al. (2004). AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. Nature Medicine *10*, 299-304. 10.1038/NM997.
231. Sun, W., and Downing, J.R. (2004). Haploinsufficiency of AML1 results in a decrease in the number of LTR-HSCs while simultaneously inducing an increase in more mature progenitors. Blood *104*, 3565-3572. 10.1182/BLOOD-2003-12-4349.
232. Bellissimo, D.C., and Speck, N.A. (2017). RUNX1 mutations in inherited and sporadic leukemia. Frontiers in Cell and Developmental Biology *5*, 111-111. 10.3389/FCELL.2017.00111/BIBTEX.
233. Kellaway, S.G., Keane, P., Edginton-White, B., Regha, K., Kennett, E., and Bonifer, C. (2021). Different mutant RUNX1 oncoproteins program alternate haematopoietic differentiation trajectories. Life Science Alliance *4*. 10.26508/LSA.202000864.
234. Matheny, C.J., Speck, M.E., Cushing, P.R., Zhou, Y., Corpora, T., Regan, M., Newman, M., Roudaia, L., Speck, C.L., Gu, T.L., et al. (2007). Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. The EMBO Journal *26*, 1163-1175. 10.1038/SJ.EMBOJ.7601568.
235. Harada, H., Harada, Y., Niimi, H., Kyo, T., Kimura, A., and Inaba, T. (2004). High incidence of somatic mutations in the AML1/RUNX1 gene in myelodysplastic syndrome and low blast percentage myeloid leukemia with myelodysplasia. Blood *103*, 2316-2324. 10.1182/BLOOD-2003-09-3074.
236. Antony-DebréDebr, I., Manchev, V.T., Balayn, N., Bluteau, D., Tomowiak, C., Legrand, C., Langlois, T., Bawa, O., Tosca, L., Tachdjian, G., et al. (2015). Level of RUNX1 activity is critical for leukemic predisposition but not for thrombocytopenia. 10.1182/blood-2014-06.
237. Tsai, S.C., Shih, L.Y., Liang, S.T., Huang, Y.J., Kuo, M.C., Huang, C.F., Shih, Y.S., Lin, T.H., Chiu, M.C., and Liang, D.C. (2015). Biological activities of RUNX1 mutants

predict secondary acute leukemia transformation from chronic myelomonocytic leukemia and myelodysplastic syndromes. Clinical Cancer Research *21*, 3541-3551. 10.1158/1078-0432.CCR-14-2203/126448/AM/BIOLOGICAL-ACTIVITIES-OF-RUNX1-MUTANTS-PREDICT.

238. McDevitt, M.A., Shivdasani, R.A., Fujiwara, Y., Yang, H., and Orkin, S.H. (1997). A "knockdown" mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. Proceedings of the National Academy of Sciences of the United States of America *94*, 6781-6785. 10.1073/PNAS.94.13.6781/ASSET/CB447AB1-0494-4686-8B1D-A558517A78E8/ASSETS/GRAPHIC/PQ1371330004.JPEG.
239. Gregory, T., Yu, C., Ma, A., Orkin, S.H., Blobel, G.A., and Weiss, M.J. (1999). GATA-1 and Erythropoietin Cooperate to Promote Erythroid Cell Survival by Regulating bcl-xL Expression. Blood *94*, 87-96. 10.1182/BLOOD.V94.1.87.413K41_87_96.
240. Yoshida, K., Toki, T., Okuno, Y., Kanezaki, R., Shiraishi, Y., Sato-Otsubo, A., Sanada, M., Park, M.J., Terui, K., Suzuki, H., et al. (2013). The landscape of somatic mutations in Down syndrome–related myeloid disorders. Nature Genetics *45*, 1293-1299. 10.1038/NG.2759.
241. Wechsler, J., Greene, M., McDevitt, M.A., Anastasi, J., Karp, J.E., Le Beau, M.M., and Crispino, J.D. (2002). Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. Nature Genetics *32*, 148-152. 10.1038/NG955.
242. Abdelfattah, A., Hughes-Davies, A., Clayfield, L., Menendez-Gonzalez, J.B., Almotiri, A., Alotaibi, B., Tonks, A., and Rodrigues, N.P. (2021). Gata2 haploinsufficiency promotes proliferation and functional decline of hematopoietic stem cells with myeloid bias during aging. Blood Advances *5*, 4285-4290. 10.1182/BLOODADVANCES.2021004726.
243. Rodrigues, N.P., Janzen, V., Forkert, R., Dombkowski, D.M., Boyd, A.S., Orkin, S.H., Enver, T., Vyas, P., and Scadden, D.T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. Blood *106*, 477-484. 10.1182/BLOOD-2004-08-2989.
244. Persons, D.A., Allay, J.A., Allay, E.R., Ashmun, R.A., Orlic, D., Jane, S.M., Cunningham, J.M., and Nienhuis, A.W. (1999). Enforced Expression of the GATA-2 Transcription Factor Blocks Normal Hematopoiesis. Blood *93*, 488-499. 10.1182/BLOOD.V93.2.488.
245. Collin, M., Dickinson, R., and Bigley, V. (2015). Haematopoietic and immune defects associated with GATA2 mutation. British Journal of Haematology *169*, 173-187. 10.1111/BJH.13317.
246. Ping, N., Sun, A., Song, Y., Wang, Q., Yin, J., Cheng, W., Xu, Y., Wen, L., Yao, H., Ma, L., et al. (2016). Exome sequencing identifies highly recurrent somatic GATA2 and CEBPA mutations in acute erythroid leukemia. Leukemia *31*, 195-202. 10.1038/leu.2016.162.
247. Di Genua, C., Valletta, S., Buono, M., Stoilova, B., Sweeney, C., Rodriguez-Meira, A., Grover, A., Drissen, R., Meng, Y., Beveridge, R., et al. (2020). C/EBPα and GATA-2 Mutations Induce Bilineage Acute Erythroid Leukemia through Transformation of a Neomorphic Neutrophil-Erythroid Progenitor. Cancer Cell *37*, 690-704.e698. 10.1016/j.ccell.2020.03.022.

248. Chen, S., Wang, Q., Yu, H., Capitano, M.L., Vemula, S., Nabinger, S.C., Gao, R., Yao, C., Kobayashi, M., Geng, Z., et al. (2019). Mutant p53 drives clonal hematopoiesis through modulating epigenetic pathway. Nature Communications *10*, 1-14. 10.1038/s41467-019-13542-2.
249. Loizou, E., Banito, A., Livshits, G., Ho, Y.J., Koche, R.P., Sánchez-Rivera, F.J., Mayle, A., Chen, C.C., Kinalis, S., Bagger, F.O., et al. (2019). A gain-of-function p53-mutant oncogene promotes cell fate plasticity and myeloid leukemia through the pluripotency factor foxh1. Cancer Discovery *9*, 962-979. 10.1158/2159-8290.CD-18-1391/42858/AM/A-GAIN-OF-FUNCTION-P53-MUTANT-ONCOGENE-PROMOTES.
250. Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E.A., Iwasaki-Arai, J., Mizuno, S.I., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., et al. (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood *106*, 1590-1600. 10.1182/BLOOD-2005-03-0860.
251. Kueh, H.Y., Champhekhar, A., Nutt, S.L., Elowitz, M.B., and Rothenberg, E.V. (2013). Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. Science *341*, 670-673. 10.1126/SCIENCE.1240831.
252. Chavez, J.S., Rabe, J.L., Loeffler, D., Higa, K.C., Hernandez, G., Mills, T.S., Ahmed, N., Gessner, R.L., Ke, Z., Idler, B.M., et al. (2021). PU.1 enforces quiescence and limits hematopoietic stem cell expansion during inflammatory stress. Journal of Experimental Medicine *218*. 10.1084/JEM.20201169/211996.
253. Scott, E.W., Simon, M.C., Anastasi, J., and Singh, H. (1994). Requirement of Transcription Factor PU.1 in the Development of Multiple Hematopoietic Lineages. Science *265*, 1573-1577. 10.1126/SCIENCE.8079170.
254. Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J.D., et al. (2007). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. Nature Genetics *40*, 51-60. 10.1038/NG.2007.7.
255. Trinh, B.Q., Ummarino, S., Zhang, Y., Ebralidze, A.K., Bassal, M.A., Nguyen, T.M., Heller, G., Coffey, R., Tenen, D.E., van der Kouwe, E., et al. (2021). Myeloid lncRNA LOUP mediates opposing regulatory effects of RUNX1 and RUNX1-ETO in t(8;21) AML. Blood *138*, 1331-1344. 10.1182/BLOOD.2020007920.
256. van der Kouwe, E., Heller, G., Czibere, A., Pulikkan, J.A., Agreiter, C., Castilla, L.H., Delwel, R., Di Ruscio, A., Ebralidze, A.K., Forte, M., et al. (2021). Core-binding factor leukemia hijacks the T-cell–prone PU.1 antisense promoter. Blood *138*, 1345-1358. 10.1182/BLOOD.2020008971.
257. Huang, G., Zhao, X., Wang, L., Elf, S., Xu, H., Zhao, X., Sashida, G., Zhang, Y., Liu, Y., Lee, J., et al. (2011). The ability of MLL to bind RUNX1 and methylate H3K4 at PU.1 regulatory regions is impaired by MDS/AML-associated RUNX1/AML1 mutations. Blood *118*, 6544-6552. 10.1182/BLOOD-2010-11-317909.
258. Mueller, B.U., Pabst, T., Fos, J., Petkovic, V., Fey, M.F., Asou, N., Buergi, U., and Tenen, D.G. (2006). ATRA resolves the differentiation block in t(15;17) acute myeloid leukemia by restoring PU.1 expression. Blood *107*, 3330-3338. 10.1182/BLOOD-2005-07-3068.
259. Mizuki, M., Schwäble, J., Steur, C., Choudhary, C., Agrawal, S., Sargin, B., Steffen, B., Matsumura, I., Kanakura, Y., Böhmer, F.D., et al. (2003). Suppression of myeloid transcription factors and induction of STAT response genes by AML-specific Flt3 mutations. Blood *101*, 3164-3173. 10.1182/BLOOD-2002-06-1677.

260. Yun, H., Narayan, N., Vohra, S., Giotopoulos, G., Mupo, A., Madrigal, P., Sasca, D., Lara-Astiaso, D., Horton, S.J., Agrawal-Singh, S., et al. (2021). Mutational synergy during leukemia induction remodels chromatin accessibility, histone modifications and three-dimensional DNA topology to alter gene expression. Nature Genetics *53*, 1443-1455. 10.1038/s41588-021-00925-9.
261. Yeamans, C., Wang, D., Paz-Priel, I., Torbett, B.E., Tenen, D.G., and Friedman, A.D. (2007). C/EBPα binds and activates the PU.1 distal enhancer to induce monocyte lineage commitment. Blood *110*, 3136-3142. 10.1182/BLOOD-2007-03-080291.
262. Gu, X., Ebrahem, Q., Mahfouz, R.Z., Hasipek, M., Enane, F., Radivoyevitch, T., Rapin, N., Przychodzen, B., Hu, Z., Balusu, R., et al. (2018). Leukemogenic nucleophosmin mutation disrupts the transcription factor hub that regulates granulomonocytic fates. The Journal of Clinical Investigation *128*, 4260-4279. 10.1172/JCI97117.
263. Pianigiani, G., Betti, C., Bigerna, B., Rossi, R., and Brunetti, L. (2020). PU.1 subcellular localization in acute myeloid leukaemia with mutated NPM1. British Journal of Haematology *188*, 184-187. 10.1111/BJH.16344.
264. Aivalioti, M.M., Bartholdy, B.A., Pradhan, K., Bhagat, T.D., Zintiridou, A., Jeong, J.J., Thiruthuvanathan, V.J., Pujato, M., Paranjpe, A., Zhang, C., et al. (2022). PU.1-dependent enhancer inhibition separates Tet2-deficient hematopoiesis from malignant transformation. Blood Cancer Discovery. 10.1158/2643-3230.BCD-21-0226.
265. Kaasinen, E., Kuismin, O., Rajamäki, K., Ristolainen, H., Aavikko, M., Kondelin, J., Saarinen, S., Berta, D.G., Katainen, R., Hirvonen, E.A.M., et al. (2019). Impact of constitutional TET2 haploinsufficiency on molecular and clinical phenotype in humans. Nature Communications *10*, 1-17. 10.1038/s41467-019-09198-7.
266. Tulstrup, M., Soerensen, M., Hansen, J.W., Gillberg, L., Needhamsen, M., Kaastrup, K., Helin, K., Christensen, K., Weischenfeldt, J., and Grønbæk, K. (2021). TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. Nature Communications *12*, 1-10. 10.1038/s41467-021-26093-2.
267. Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M.L., Dayaram, T., Owens, B.M., Shigematsu, H., Levantini, E., Huettner, C.S., Lekstrom-Himes, J.A., et al. (2004). Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBPα. Immunity *21*, 853-863. 10.1016/j.immuni.2004.11.006.
268. Zhang, D.E., Zhang, P., Wang, N.D., Hetherington, C.J., Darlington, G.J., and Tenen, D.G. (1997). Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein α-deficient mice. Proceedings of the National Academy of Sciences of the United States of America *94*, 569-574. 10.1073/PNAS.94.2.569/ASSET/731B9E31-1465-4690-BB8E-1CE4B22974C0/ASSETS/GRAPHIC/PQ0273423005.JPEG.
269. Ohlsson, E., Schuster, M.B., Hasemann, M., and Porse, B.T. (2015). The multifaceted functions of C/EBPα in normal and malignant haematopoiesis. Leukemia *30*, 767-775. 10.1038/leu.2015.324.
270. Calkhoven, C.F., Muller, C., and Leutz, A. (2000). Translational control of C/EBPα and C/EBPβ isoform expression. Genes & Development *14*, 1920-1932. 10.1101/GAD.14.15.1920.

271. Mannelli, F., Ponziani, V., Bencini, S., Bonetti, M.I., Benelli, M., Cutini, I., Gianfaldoni, G., Scappini, B., Pancani, F., Piccini, M., et al. (2017). CEBPA–double-mutated acute myeloid leukemia displays a unique phenotypic profile: a reliable screening method and insight into biological features. Haematologica *102*, 529-540. 10.3324/HAEMATOL.2016.151910.
272. Kirstetter, P., Schuster, M.B., Bereshchenko, O., Moore, S., Dvinge, H., Kurz, E., Theilgaard-Mönch, K., Månsson, R., Pedersen, T.Å., Pabst, T., et al. (2008). Modeling of C/EBPα Mutant Acute Myeloid Leukemia Reveals a Common Expression Signature of Committed Myeloid Leukemia-Initiating Cells. Cancer Cell *13*, 299-310. 10.1016/j.ccr.2008.02.008.
273. Bereshchenko, O., Mancini, E., Moore, S., Bilbao, D., Månsson, R., Luc, S., Grover, A., Jacobsen, S.E.W., Bryder, D., and Nerlov, C. (2009). Hematopoietic Stem Cell Expansion Precedes the Generation of Committed Myeloid Leukemia-Initiating Cells in C/EBPα Mutant AML. Cancer Cell *16*, 390-400. 10.1016/J.CCR.2009.09.036.
274. Khanna-Gupta, A., Abayasekara, N., Levine, M., Sun, H., Virgilio, M., Nia, N., Halene, S., Sportoletti, P., Jeong, J.Y., Pandolfi, P.P., and Berliner, N. (2012). Up-regulation of Translation Eukaryotic Initiation Factor 4E in Nucleophosmin 1 Haploinsufficient Cells Results in Changes in CCAAT Enhancer-binding Protein α Activity: IMPLICATIONS IN MYELODYSPLASTIC SYNDROME AND ACUTE MYELOID LEUKEMIA. Journal of Biological Chemistry *287*, 32728-32737. 10.1074/JBC.M112.373274.
275. Gialesaki, S., Bräuer-Hartmann, D., Bhayadia, R., Alejo-Valle, O., Issa, H., Regenyi, E., Beyer, C., Labuhn, M., Ihling, C., Sinz, A., et al. (2022). RUNX1 isoform disequilibrium in the development of trisomy 21 associated myeloid leukemia. bioRxiv, 2022.2003.2007.483334-482022.483303.483307.483334. 10.1101/2022.03.07.483334.
276. Sakurai, H., Harada, Y., Ogata, Y., Kagiyama, Y., Shingai, N., Doki, N., Ohashi, K., Kitamura, T., Komatsu, N., and Harada, H. (2017). Overexpression of RUNX1 short isoform has an important role in the development of myelodysplastic/myeloproliferative neoplasms. Blood Advances *1*, 1382-1386. 10.1182/BLOODADVANCES.2016002725.
277. Figueroa, M.E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P.J., Schifano, E., Booth, J., van Putten, W., Skrabanek, L., et al. (2010). DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia. Cancer Cell *17*, 13-27. 10.1016/j.ccr.2009.11.020.
278. Mulet-Lazaro, R., van Herk, S., Erpelinck, C., Bindels, E., Sanders, M.A., Vermeulen, C., Renkens, I., Valk, P., Melnick, A.M., de Ridder, J., et al. (2021). Allele-specific expression of GATA2 due to epigenetic dysregulation in CEBPA double-mutant AML. Blood *138*, 160-177. 10.1182/BLOOD.2020009244.
279. Lin, T.C., Hou, H.A., Chou, W.C., Ou, D.L., Yu, S.L., Tien, H.F., and Lin, L.I. (2010). CEBPA methylation as a prognostic biomarker in patients with de novo acute myeloid leukemia. Leukemia *25*, 32-40. 10.1038/leu.2010.222.
280. Mill, C.P., Fiskus, W., DiNardo, C.D., Qian, Y., Raina, K., Rajapakshe, K., Perera, D., Coarfa, C., Kadia, T.M., Khoury, J.D., et al. (2019). RUNX1-targeted therapy for AML expressing somatic or germline mutation in RUNX1. Blood *134*, 59-73. 10.1182/BLOOD.2018893982.
281. Krönke, J., Fink, E.C., Hollenbach, P.W., MacBeth, K.J., Hurst, S.N., Udeshi, N.D., Chamberlain, P.P., Mani, D.R., Man, H.W., Gandhi, A.K., et al. (2015). Lenalidomide

induces ubiquitination and degradation of CK1α in del(5q) MDS. Nature *523*, 183-188. 10.1038/nature14610.
282. Wilkinson, A.C., and Göttgens, B. (2013). Transcriptional Regulation of Haematopoietic Stem Cells. In Transcriptional and Translational Regulation of Stem Cells, (Springer Netherlands), pp. 187-212. 10.1007/978-94-007-6621-1_11.
283. Mantel, R., Charlie, O'Leary, A., Heather, Chitteti, R., Brahmananda, Huang, X., Cooper, S., Hangoc, G., Brustovetsky, N., Srour, F., Edward, Lee, R., Man, Messina-Graham, S., et al. (2015). Enhancing Hematopoietic Stem Cell Transplantation Efficacy by Mitigating Oxygen Shock. Cell *161*, 1553-1565. 10.1016/j.cell.2015.04.054.
284. Schönberger, K., Obier, N., Romero-Mulero, M.C., Cauchy, P., Mess, J., Pavlovich, P.V., Zhang, Y.W., Mitterer, M., Rettkowski, J., Lalioti, M.-E., et al. (2022). Multilayer omics analysis reveals a non-classical retinoic acid signaling axis that regulates hematopoietic stem cell identity. Cell Stem Cell *29*. 10.1016/j.stem.2021.10.002.
285. Cabezas-Wallscheid, N., Buettner, F., Sommerkamp, P., Klimmeck, D., Ladel, L., Thalheimer, F.B., Pastor-Flores, D., Roma, L.P., Renders, S., Zeisberger, P., et al. (2017). Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. Cell *169*, 807-823.e819. 10.1016/j.cell.2017.04.018.
286. Comazzetto, S., Cassidy, D.L., Devilbiss, A.W., Jeffery, E.C., Ottesen, B.R., Reyes, A.R., Muh, S., Mathews, T.P., Chen, B., Zhao, Z., and Morrison, S.J. (2024). Ascorbate depletion increases quiescence and self-renewal potential in hematopoietic stem cells and multipotent progenitors. Cold Spring Harbor Laboratory.
287. Hinge, A., He, J., Bartram, J., Javier, J., Xu, J., Fjellman, E., Sesaki, H., Li, T., Yu, J., Wunderlich, M., et al. (2020). Asymmetrically Segregated Mitochondria Provide Cellular Memory of Hematopoietic Stem Cell Replicative History and Drive HSC Attrition. Cell Stem Cell *26*, 420-430.e426. 10.1016/j.stem.2020.01.016.
288. Bonora, M., Morganti, C., Van Gastel, N., Ito, K., Calura, E., Zanolla, I., Ferroni, L., Zhang, Y., Jung, Y., Sales, G., et al. (2024). A mitochondrial NADPH-cholesterol axis regulates extracellular vesicle biogenesis to support hematopoietic stem cell fate. Cell Stem Cell *31*, 359-377.e310. 10.1016/j.stem.2024.02.004.
289. Kruta, M., Sunshine, M.J., Chua, B.A., Fu, Y., Chawla, A., Dillingham, C.H., Hidalgo San Jose, L., De Jong, B., Zhou, F.J., and Signer, R.A.J. (2021). Hsf1 promotes hematopoietic stem cell fitness and proteostasis in response to ex vivo culture stress and aging. Cell Stem Cell *28*, 1950-1965.e1956. 10.1016/j.stem.2021.07.009.
290. Signer, R.A.J., Magee, J.A., Salic, A., and Morrison, S.J. (2014). Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature *509*, 49-54. 10.1038/nature13035.
291. García-Prat, L., Kaufmann, K.B., Schneiter, F., Voisin, V., Murison, A., Chen, J., Chan-Seng-Yue, M., Gan, O.I., Mcleod, J.L., Smith, S.A., et al. (2021). TFEB-mediated endolysosomal activity controls human hematopoietic stem cell fate. Cell Stem Cell *28*, 1838-1850.e1810. 10.1016/j.stem.2021.07.003.
292. Loeffler, D., Wehling, A., Schneiter, F., Zhang, Y., Müller-Bötticher, N., Hoppe, P.S., Hilsenbeck, O., Kokkaliaris, K.D., Endele, M., and Schroeder, T. (2019). Asymmetric lysosome inheritance predicts activation of haematopoietic stem cells. Nature *573*, 426-429. 10.1038/s41586-019-1531-6.

293. Jeong, M., Park, H.J., Celik, H., Ostrander, E.L., Reyes, J.M., Guzman, A., Rodriguez, B., Lei, Y., Lee, Y., Ding, L., et al. (2018). Loss of Dnmt3a Immortalizes Hematopoietic Stem Cells In Vivo. Cell Reports *23*, 1-10. 10.1016/j.celrep.2018.03.025.
294. Calvanese, V., Nguyen, A.T., Bolan, T.J., Vavilina, A., Su, T., Lee, L.K., Wang, Y., Lay, F.D., Magnusson, M., Crooks, G.M., et al. (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. Nature *576*, 281-286. 10.1038/s41586-019-1790-2.
295. Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, E., Maria, 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. 10.1016/j.ccr.2011.06.001.
296. Kataoka, K., Sato, T., Yoshimi, A., Goyama, S., Tsuruta, T., Kobayashi, H., Shimabe, M., Arai, S., Nakagawa, M., Imai, Y., et al. (2011). Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity. Journal of Experimental Medicine *208*, 2403-2416. 10.1084/jem.20110447.
297. Voit, R.A., Tao, L., Yu, F., Cato, L.D., Cohen, B., Fleming, T.J., Antoszewski, M., Liao, X., Fiorini, C., Nandakumar, S.K., et al. (2023). A genetic disorder reveals a hematopoietic stem cell regulatory network co-opted in leukemia. Nature Immunology *24*, 69-83. 10.1038/s41590-022-01370-4.
298. Zeng, H., Yücel, R., Kosan, C., Klein-Hitpass, L., and Möröy, T. (2004). Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. The EMBO Journal *23*, 4116-4125. 10.1038/sj.emboj.7600419.
299. Komorowska, K., Doyle, A., Wahlestedt, M., Subramaniam, A., Debnath, S., Chen, J., Soneji, S., Van Handel, B., Mikkola, H.K.A., Miharada, K., et al. (2017). Hepatic Leukemia Factor Maintains Quiescence of Hematopoietic Stem Cells and Protects the Stem Cell Pool during Regeneration. Cell Reports *21*, 3514-3523. 10.1016/j.celrep.2017.11.084.
300. Ellis, T., Gambardella, L., Horcher, M., Tschanz, S., Capol, J., Bertram, P., Jochum, W., Barrandon, Y., and Busslinger, M. (2001). The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes & Development *15*, 2307-2319. 10.1101/gad.200101.
301. Luong, M.X., Van Der Meijden, C.M., Xing, D., Hesselton, R., Monuki, E.S., Jones, S.N., Lian, J.B., Stein, J.L., Stein, G.S., Neufeld, E.J., and Van Wijnen, A.J. (2002). Genetic Ablation of the CDP/Cux Protein C Terminus Results in Hair Cycle Defects and Reduced Male Fertility. Molecular and Cellular Biology *22*, 1424-1437. 10.1128/mcb.22.5.1424-1437.2002.
302. Tufarelli, C., Fujiwara, Y., Zapulla, D.C., and Neufeld, E.J. (1998). Hair Defects and Pup Loss in Mice with Targeted Deletion of the First Cut Repeat Domain of theCux/CDPHomeoprotein Gene. Developmental Biology *200*. 10.1006/dbio.1998.8950.
303. Zhao, B., Socha, J., Toth, A., Fernandes, S., Warheit-Niemi, H., Ruff, B., Khurana Hershey, G.K., Vandussen, K., Swarr, D., and Zacharias, W.J. (2024). The homeobox transcription factor Cux1 coordinates postnatal epithelial developmental timing but is dispensable for lung organogenesis and regeneration. Cold Spring Harbor Laboratory.
304. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. (2013). New England Journal of Medicine *368*, 2059-2074. 10.1056/nejmoa1301689.

305. Martinez, T.C., and McNerney, M.E. (2024). Haploinsufficient Transcription Factors in Myeloid Neoplasms. Annual Review of Pathology: Mechanisms of Disease *19*. 10.1146/annurev-pathmechdis-051222-013421.
306. Beckett, D., Kovaleva, E., and Schatz, P.J. (1999). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein Science *8*, 921-929. 10.1110/ps.8.4.921.
307. Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L., and Pe'Er, D. (2019). Characterization of cell fate probabilities in single-cell data with Palantir. Nature Biotechnology *37*, 451-460. 10.1038/s41587-019-0068-4.
308. Morcos, M.N.F., Schoedel, K.B., Hoppe, A., Behrendt, R., Basak, O., Clevers, H.C., Roers, A., and Gerbaulet, A. (2017). SCA-1 Expression Level Identifies Quiescent Hematopoietic Stem and Progenitor Cells. Stem Cell Reports *8*, 1472-1478. 10.1016/j.stemcr.2017.04.012.
309. Balazs, A.B., Fabian, A.J., Esmon, C.T., and Mulligan, R.C. (2006). Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. Blood *107*, 2317-2321. 10.1182/blood-2005-06-2249.
310. Shin, J.Y., Hu, W., Naramura, M., and Park, C.Y. (2014). High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. Journal of Experimental Medicine *211*. 10.1084/jem.20131128.
311. Benveniste, P., Frelin, C., Janmohamed, S., Barbara, M., Herrington, R., Hyam, D., and Iscove, N.N. (2010). Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential. Cell Stem Cell *6*, 48-58. 10.1016/j.stem.2009.11.014.
312. Schaefer, B.C., Schaefer, M.L., Kappler, J.W., Marrack, P., and Kedl, R.M. (2001). Observation of Antigen-Dependent CD8+ T-Cell/ Dendritic Cell Interactions in Vivo. Cellular Immunology *214*. 10.1006/cimm.2001.1895.
313. De Almeida, M.J., Luchsinger, L.L., Corrigan, D.J., Williams, L.J., and Snoeck, H.-W. (2017). Dye-Independent Methods Reveal Elevated Mitochondrial Mass in Hematopoietic Stem Cells. Cell Stem Cell *21*, 725-729.e724. 10.1016/j.stem.2017.11.002.
314. Wu, X., Dao Thi, V.L., Huang, Y., Billerbeck, E., Saha, D., Hoffmann, H.-H., Wang, Y., Silva, L.A.V., Sarbanes, S., Sun, T., et al. (2018). Intrinsic Immunity Shapes Viral Resistance of Stem Cells. Cell *172*, 423-438.e425. 10.1016/j.cell.2017.11.018.
315. Lengefeld, J., Cheng, C.-W., Maretich, P., Blair, M., Hagen, H., Mcreynolds, M.R., Sullivan, E., Majors, K., Roberts, C., Kang, J.H., et al. (2021). Cell size is a determinant of stem cell potential during aging. Science Advances *7*. 10.1126/sciadv.abk0271.
316. Sommerkamp, P., Romero-Mulero, M.C., Narr, A., Ladel, L., Hustin, L., Schönberger, K., Renders, S., Altamura, S., Zeisberger, P., Jäcklein, K., et al. (2021). Mouse multipotent progenitor 5 cells are located at the interphase between hematopoietic stem and progenitor cells. Blood *137*, 3218-3224. 10.1182/blood.2020007876.
317. Detomaso, D., and Yosef, N. (2021). Hotspot identifies informative gene modules across modalities of single-cell genomics. Cell Systems *12*, 446-456.e449. 10.1016/j.cels.2021.04.005.
318. Lehnertz, B., Chagraoui, J., Macrae, T., Tomellini, E., Corneau, S., Mayotte, N., Boivin, I., Durand, A., Gracias, D., and Sauvageau, G. (2021). HLFexpression defines the human hematopoietic stem cell state. Blood *138*, 2642-2654. 10.1182/blood.2021010745.

319. Luo, H., Cortés-López, M., Tam, C.L., Xiao, M., Wakiro, I., Chu, K.L., Pierson, A., Chan, M., Chang, K., Yang, X., et al. (2023). SON is an essential m6A target for hematopoietic stem cell fate. Cell Stem Cell *30*. 10.1016/j.stem.2023.11.006.
320. Lawrence, H.J., Christensen, J., Fong, S., Hu, Y.-L., Weissman, I., Sauvageau, G., Humphries, R.K., and Largman, C. (2005/12/01). Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. Blood *106*. 10.1182/blood-2005-05-2003.
321. Imgruet, M.K., Lutze, J., An, N., Hu, B., Khan, S., Kurkewich, J., Martinez, T.C., Wolfgeher, D., Gurbuxani, S.K., Kron, S.J., and Mcnerney, M.E. (2021). Loss of a 7q gene,CUX1, disrupts epigenetically driven DNA repair and drives therapy-related myeloid neoplasms. Blood *138*, 790-805. 10.1182/blood.2020009195.
322. Gao, C.F., Vaikuntanathan, S., and Riesenfeld, S.J. (2024). Dissection and integration of bursty transcriptional dynamics for complex systems. Proceedings of the National Academy of Sciences *121*. 10.1073/pnas.2306901121.
323. Liu, W., Kurkewich, J.L., Stoddart, A., Khan, S., Anandan, D., Gaubil, A.N., Wolfgeher, D.J., Jueng, L., Kron, S.J., and Mcnerney, M.E. (2024). CUX1 regulates human hematopoietic stem cell chromatin accessibility via the BAF complex. Cell Reports *43*, 114227. 10.1016/j.celrep.2024.114227.
324. Bahr, C., Von Paleske, L., Uslu, V.V., Remeseiro, S., Takayama, N., Ng, S.W., Murison, A., Langenfeld, K., Petretich, M., Scognamiglio, R., et al. (2018). A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. Nature *553*, 515-520. 10.1038/nature25193.
325. Magnusson, M., Brun, A.C.M., Miyake, N., Larsson, J., Ehinger, M., Bjornsson, J.M., Wutz, A., Sigvardsson, M., and Karlsson, S. (2007). HOXA10 is a critical regulator for hematopoietic stem cells and erythroid/megakaryocyte development. Blood *109*, 3687-3696. 10.1182/blood-2006-10-054676.
326. Chen, J.Y., Miyanishi, M., Wang, S.K., Yamazaki, S., Sinha, R., Kao, K.S., Seita, J., Sahoo, D., Nakauchi, H., and Weissman, I.L. (2016). Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. Nature *530*, 223-227. 10.1038/nature16943.
327. Rundberg Nilsson, S., Alexandra J., Xian, H., Shalapour, S., Cammenga, J., and Karin, M. (2023). IRF1 regulates self-renewal and stress responsiveness to support hematopoietic stem cell maintenance. Science Advances *9*. 10.1126/sciadv.adg5391.
328. Li, J., Williams, M.J., Park, H.J., Bastos, H.P., Wang, X., Prins, D., Wilson, N.K., Johnson, C., Sham, K., Wantoch, M., et al. (2022). STAT1 is essential for HSC function and maintains MHCIIhi stem cells that resist myeloablation and neoplastic expansion. Blood *140*, 1592-1606. 10.1182/blood.2021014009.
329. Corces, M.R., Granja, J.M., Shams, S., Louie, B.H., Seoane, J.A., Zhou, W., Silva, T.C., Groeneveld, C., Wong, C.K., Cho, S.W., et al. (2018). The chromatin accessibility landscape of primary human cancers. Science *362*, eaav1898. 10.1126/science.aav1898.
330. Nishio, H., and Walsh, M.J. (2004). CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription. Proceedings of the National Academy of Sciences *101*, 11257-11262. 10.1073/pnas.0401343101.
331. Fang, T.C., Schaefer, U., Mecklenbrauker, I., Stienen, A., Dewell, S., Chen, M.S., Rioja, I., Parravicini, V., Prinjha, R.K., Chandwani, R., et al. (2012). Histone H3 lysine 9 di-

methylation as an epigenetic signature of the interferon response. Journal of Experimental Medicine *209*, 661-669. 10.1084/jem.20112343.
332. Avgustinova, A., Laudanna, C., Pascual-García, M., Rovira, Q., Djurec, M., Castellanos, A., Urdiroz-Urricelqui, U., Marchese, D., Prats, N., Van Keymeulen, A., et al. (2021). Repression of endogenous retroviruses prevents antiviral immune response and is required for mammary gland development. Cell Stem Cell *28*, 1790-1804.e1798. 10.1016/j.stem.2021.04.030.
333. Chuong, E.B., Elde, N.C., and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science *351*, 1083-1087. 10.1126/science.aad5497.
334. Jiang, Q., Ang, J.Y.J., Lee, A.Y., Cao, Q., Li, K.Y., Yip, K.Y., and Leung, D.C.Y. (2020). G9a Plays Distinct Roles in Maintaining DNA Methylation, Retrotransposon Silencing, and Chromatin Looping. Cell Reports *33*, 108315. 10.1016/j.celrep.2020.108315.
335. Liu, M., Thomas, S.L., Dewitt, A.K., Zhou, W., Madaj, Z.B., Ohtani, H., Baylin, S.B., Liang, G., and Jones, P.A. (2018). Dual Inhibition of DNA and Histone Methyltransferases Increases Viral Mimicry in Ovarian Cancer Cells. Cancer Research *78*, 5754-5766. 10.1158/0008-5472.can-17-3953.
336. Leung, D.C., Dong, K.B., Maksakova, I.A., Goyal, P., Appanah, R., Lee, S., Tachibana, M., Shinkai, Y., Lehnertz, B., Mager, D.L., et al. (2011). Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing. Proceedings of the National Academy of Sciences *108*, 5718-5723. 10.1073/pnas.1014660108.
337. Maksakova, I.A., Thompson, P.J., Goyal, P., Jones, S.J., Singh, P.B., Karimi, M.M., and Lorincz, M.C. (2013). Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells. Epigenetics & Chromatin *6*, 15. 10.1186/1756-8935-6-15.
338. Clapes, T., Polyzou, A., Prater, P., Sagar, Morales-Hernández, A., Ferrarini, M.G., Kehrer, N., Lefkopoulos, S., Bergo, V., Hummel, B., et al. (2021). Chemotherapy-induced transposable elements activate MDA5 to enhance haematopoietic regeneration. Nature Cell Biology *23*, 704-717. 10.1038/s41556-021-00707-9.
339. Dhillon, P., Mulholland, K.A., Hu, H., Park, J., Sheng, X., Abedini, A., Liu, H., Vassalotti, A., Wu, J., and Susztak, K. (2023). Increased levels of endogenous retroviruses trigger fibroinflammation and play a role in kidney disease development. Nature Communications *14*. 10.1038/s41467-023-36212-w.
340. Kassiotis, G. (2023). The Immunological Conundrum of Endogenous Retroelements. Annual Review of Immunology *41*, 99-125. 10.1146/annurev-immunol-101721-033341.
341. Chen, X., Skutt-Kakaria, K., Davison, J., Ou, Y.-L., Choi, E., Malik, P., Loeb, K., Wood, B., Georges, G., Torok-Storb, B., and Paddison, P.J. (2012). G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. Genes & Development *26*, 2499-2511. 10.1101/gad.200329.112.
342. Zeng, T.-B., Pierce, N., Liao, J., Singh, P., Lau, K., Zhou, W., and Szabó, P.E. (2021). EHMT2 suppresses the variation of transcriptional switches in the mouse embryo. PLOS Genetics *17*, e1009908. 10.1371/journal.pgen.1009908.
343. Herquel, B., Ouararhni, K., Martianov, I., Le Gras, S., Ye, T., Keime, C., Lerouge, T., Jost, B., Cammas, F., Losson, R., and Davidson, I. (2013). Trim24-repressed VL30

retrotransposons regulate gene expression by producing noncoding RNA. Nature Structural & Molecular Biology *20*, 339-346. 10.1038/nsmb.2496.
344. Subramanian, R.P., Wildschutte, J.H., Russo, C., and Coffin, J.M. (2011). Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. Retrovirology *8*, 90. 10.1186/1742-4690-8-90.
345. Liu, X., Liu, Z., Wu, Z., Ren, J., Fan, Y., Sun, L., Cao, G., Niu, Y., Zhang, B., Ji, Q., et al. (2023). Resurrection of endogenous retroviruses during aging reinforces senescence. Cell *186*, 287-304.e226. 10.1016/j.cell.2022.12.017.
346. Pietras, E.M., Mirantes-Barbeito, C., Fong, S., Loeffler, D., Kovtonyuk, L.V., Zhang, S., Lakshminarasimhan, R., Chin, C.P., Techner, J.-M., Will, B., et al. (2016). Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. Nature Cell Biology *18*, 607-618. 10.1038/ncb3346.
347. Yang, L., Wang, L., Geiger, H., Cancelas, J.A., Mo, J., and Zheng, Y. (2007). Rho GTPase Cdc42 coordinates hematopoietic stem cell quiescence and niche interaction in the bone marrow. Proceedings of the National Academy of Sciences *104*, 5091-5096. 10.1073/pnas.0610819104.
348. Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, B., Daniel, Reyes, A., Wang, Q., Weichenhan, D., Lier, A., Paleske, V., Lisa, 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. 10.1016/j.stem.2014.07.005.
349. Bernitz, J.M., Kim, H.S., Macarthur, B., Sieburg, H., and Moore, K. (2016). Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. Cell *167*, 1296-1309.e1210. 10.1016/j.cell.2016.10.022.
350. Qiu, J., Papatsenko, D., Niu, X., Schaniel, C., and Moore, K. (2014). Divisional history and hematopoietic stem cell function during homeostasis - PubMed. Stem cell reports *2*. 10.1016/j.stemcr.2014.01.016.
351. Bogeska, R., Mikecin, A.-M., Kaschutnig, P., Fawaz, M., Büchler-Schäff, M., Le, D., Ganuza, M., Vollmer, A., Paffenholz, S.V., Asada, N., et al. (2022). Inflammatory exposure drives long-lived impairment of hematopoietic stem cell self-renewal activity and accelerated aging. Cell Stem Cell *29*, 1273-1284.e1278. 10.1016/j.stem.2022.06.012.
352. Allsopp, R.C., Morin, G.B., Depinho, R., Harley, C.B., and Weissman, I.L. (2003). Telomerase is required to slow telomere shortening and extend replicative lifespan of HSCs during serial transplantation. Blood *102*, 517-520. 10.1182/blood-2002-07-2334.
353. Baldridge, M.T., King, K.Y., Boles, N.C., Weksberg, D.C., and Goodell, M.A. (2010). Quiescent haematopoietic stem cells are activated by IFN-γ in response to chronic infection. Nature *465*, 793-797. 10.1038/nature09135.
354. Pietras, E.M., Lakshminarasimhan, R., Techner, J.-M., Fong, S., Flach, J., Binnewies, M., and Passegué, E. (2014). Re-entry into quiescence protects hematopoietic stem cells from the killing effect of chronic exposure to type I interferons. Journal of Experimental Medicine *211*, 245-262. 10.1084/jem.20131043.
355. Hormaechea-Agulla, D., Matatall, K.A., Le, D.T., Kain, B., Long, X., Kus, P., Jaksik, R., Challen, G.A., Kimmel, M., and King, K.Y. (2021). Chronic infection drives Dnmt3a-loss-of-function clonal hematopoiesis via IFNγ signaling. Cell Stem Cell *28*, 1428-1442.e1426. 10.1016/J.STEM.2021.03.002.

356. Chen, J., Feng, X., Desierto, M.J., Keyvanfar, K., and Young, N.S. (2015). IFN-γ-mediated hematopoietic cell destruction in murine models of immune-mediated bone marrow failure. Blood *126*, 2621-2631. 10.1182/blood-2015-06-652453.
357. Matatall, K.A., Jeong, M., Chen, S., Sun, D., Chen, F., Mo, Q., Kimmel, M., and King, K.Y. (2016). Chronic Infection Depletes Hematopoietic Stem Cells through Stress-Induced Terminal Differentiation. Cell Reports *17*. 10.1016/j.celrep.2016.11.031.
358. Hirche, C., Frenz, T., Haas, S.F., Döring, M., Borst, K., Tegtmeyer, P.-K., Brizic, I., Jordan, S., Keyser, K., Chhatbar, C., et al. (2017). Systemic Virus Infections Differentially Modulate Cell Cycle State and Functionality of Long-Term Hematopoietic Stem Cells In Vivo. Cell Reports *19*, 2345-2356. 10.1016/j.celrep.2017.05.063.
359. Essers, M.A.G., Offner, S., Blanco-Bose, W.E., Waibler, Z., Kalinke, U., Duchosal, M.A., and Trumpp, A. (2009). IFNα activates dormant haematopoietic stem cells in vivo. Nature *458*, 904-908. 10.1038/nature07815.
360. Song, M., Ping, Y., Zhang, K., Yang, L., Li, F., Zhang, C., Cheng, S., Yue, D., Maimela, N.R., Qu, J., et al. (2019). Low-Dose IFNγ Induces Tumor Cell Stemness in Tumor Microenvironment of Non–Small Cell Lung Cancer. Cancer Research *79*, 3737-3748. 10.1158/0008-5472.can-19-0596.
361. Naigles, B., Narla, A.V., Soroczynski, J., Tsimring, L.S., and Hao, N. (2023). Quantifying dynamic pro-inflammatory gene expression and heterogeneity in single macrophage cells. Journal of Biological Chemistry *299*, 105230. 10.1016/j.jbc.2023.105230.
362. Uccellini, M.B., and García-Sastre, A. (2018). ISRE-Reporter Mouse Reveals High Basal and Induced Type I IFN Responses in Inflammatory Monocytes. Cell Reports *25*, 2784-2796.e2783. 10.1016/j.celrep.2018.11.030.
363. Mostafavi, S., Yoshida, H., Moodley, D., Leboité, H., Rothamel, K., Raj, T., Ye, J., Chun, Chevrier, N., Zhang, S.-Y., Feng, T., et al. (2016). Parsing the Interferon Transcriptional Network and Its Disease Associations. Cell *164*, 564-578. 10.1016/j.cell.2015.12.032.
364. Liu, B.C., Sarhan, J., Panda, A., Muendlein, H.I., Ilyukha, V., Coers, J., Yamamoto, M., Isberg, R.R., and Poltorak, A. (2018). Constitutive Interferon Maintains GBP Expression Required for Release of Bacterial Components Upstream of Pyroptosis and Anti-DNA Responses. Cell Reports *24*, 155-168.e155. 10.1016/j.celrep.2018.06.012.
365. Bradley, K.C., Finsterbusch, K., Schnepf, D., Crotta, S., Llorian, M., Davidson, S., Fuchs, S.Y., Staeheli, P., and Wack, A. (2019). Microbiota-Driven Tonic Interferon Signals in Lung Stromal Cells Protect from Influenza Virus Infection. Cell Reports *28*, 245-256.e244. 10.1016/j.celrep.2019.05.105.
366. Dorrity, T.J., Shin, H., Wiegand, K.A., Aruda, J., Closser, M., Jung, E., Gertie, J.A., Leone, A., Polfer, R., Culbertson, B., et al. (2023). Long 3′UTRs predispose neurons to inflammation by promoting immunostimulatory double-stranded RNA formation. Science Immunology *8*. 10.1126/sciimmunol.adg2979.
367. Singhania, A., Graham, C.M., Gabryšová, L., Moreira-Teixeira, L., Stavropoulos, E., Pitt, J.M., Chakravarty, P., Warnatsch, A., Branchett, W.J., Conejero, L., et al. (2019). Transcriptional profiling unveils type I and II interferon networks in blood and tissues across diseases. Nature Communications *10*. 10.1038/s41467-019-10601-6.
368. Sun, X., Zeng, H., Kumar, A., Belser, J.A., Maines, T.R., and Tumpey, T.M. (2016). Constitutively Expressed IFITM3 Protein in Human Endothelial Cells Poses an Early

Infection Block to Human Influenza Viruses. Journal of Virology *90*, 11157-11167. 10.1128/jvi.01254-16.
369. Zurney, J., Howard, K.E., and Sherry, B. (2007). Basal Expression Levels of IFNAR and Jak-STAT Components Are Determinants of Cell-Type-Specific Differences in Cardiac Antiviral Responses. Journal of Virology *81*, 13668-13680. 10.1128/jvi.01172-07.
370. Escoubas, C.C., Dorman, L.C., Nguyen, P.T., Lagares-Linares, C., Nakajo, H., Anderson, S.R., Barron, J.J., Wade, S.D., Cuevas, B., Vainchtein, I.D., et al. (2024). Type-I-interferon-responsive microglia shape cortical development and behavior. Cell *187*, 1936-1954.e1924. 10.1016/j.cell.2024.02.020.
371. Xing, Y., Wang, X., Jameson, S.C., and Hogquist, K.A. (2016). Late stages of T cell maturation in the thymus involve NF-κB and tonic type I interferon signaling. Nature Immunology *17*, 565-573. 10.1038/ni.3419.
372. Lima-Junior, D.S., Krishnamurthy, S.R., Bouladoux, N., Collins, N., Han, S.-J., Chen, E.Y., Constantinides, M.G., Link, V.M., Lim, A.I., Enamorado, M., et al. (2021). Endogenous retroviruses promote homeostatic and inflammatory responses to the microbiota. Cell *184*, 3794-3811.e3719. 10.1016/j.cell.2021.05.020.
373. Maclauchlan, S., Kushwaha, P., Tai, A., Chen, S., Manning, C., Swarnkar, G., Abu-Amer, Y., Fitzgerald, K.A., Sharma, S., and Gravallese, E.M. (2023). STING-dependent interferon signatures restrict osteoclast differentiation and bone loss in mice. Proceedings of the National Academy of Sciences *120*. 10.1073/pnas.2210409120.
374. Yang, G., Xu, Y., Chen, X., and Hu, G. (2007). IFITM1 plays an essential role in the antiproliferative action of interferon-γ. Oncogene *26*, 594-603. 10.1038/sj.onc.1209807.
375. Katsoulidis, E., Carayol, N., Woodard, J., Konieczna, I., Majchrzak-Kita, B., Jordan, A., Sassano, A., Eklund, E.A., Fish, E.N., and Platanias, L.C. (2009). Role of Schlafen 2 (SLFN2) in the Generation of Interferon α-induced Growth Inhibitory Responses. Journal of Biological Chemistry *284*, 25051-25064. 10.1074/jbc.m109.030445.
376. Kodigepalli, K.M., Bonifati, S., Tirumuru, N., and Wu, L. (2018). SAMHD1 modulates in vitro proliferation of acute myeloid leukemia-derived THP-1 cells through the PI3K-Akt-p27 axis. Cell Cycle *17*, 1124-1137. 10.1080/15384101.2018.1480218.
377. Busse, D.C., Habgood-Coote, D., Clare, S., Brandt, C., Bassano, I., Kaforou, M., Herberg, J., Levin, M., Eléouët, J.-F., Kellam, P., and Tregoning, J.S. (2020). Interferon-Induced Protein 44 and Interferon-Induced Protein 44-Like Restrict Replication of Respiratory Syncytial Virus. Journal of Virology *94*. 10.1128/jvi.00297-20.
378. Iwata, H., Goettsch, C., Sharma, A., Ricchiuto, P., Goh, W.W.B., Halu, A., Yamada, I., Yoshida, H., Hara, T., Wei, M., et al. (2016). PARP9 and PARP14 cross-regulate macrophage activation via STAT1 ADP-ribosylation. Nature Communications *7*, 12849. 10.1038/ncomms12849.
379. Ghosh, A., Shao, L., Sampath, P., Zhao, B., Patel, N.V., Zhu, J., Behl, B., Parise, R.A., Beumer, J.H., O'Sullivan, R.J., et al. (2019). Oligoadenylate-Synthetase-Family Protein OASL Inhibits Activity of the DNA Sensor cGAS during DNA Virus Infection to Limit Interferon Production. Immunity *50*, 51-63.e55. 10.1016/j.immuni.2018.12.013.
380. Scheller, M., Ludwig, A.K., Göllner, S., Rohde, C., Krämer, S., Stäble, S., Janssen, M., Müller, J.-A., He, L., Bäumer, N., et al. (2021). Hotspot DNMT3A mutations in clonal hematopoiesis and acute myeloid leukemia sensitize cells to azacytidine via viral mimicry response. Nature Cancer *2*, 527-544. 10.1038/s43018-021-00213-9.

381. Altshuler, A., Amitai-Lange, A., Tarazi, N., Dey, S., Strinkovsky, L., Hadad-Porat, S., Bhattacharya, S., Nasser, W., Imeri, J., Ben-David, G., et al. (2021). Discrete limbal epithelial stem cell populations mediate corneal homeostasis and wound healing. Cell Stem Cell *28*, 1248-1261.e1248. 10.1016/j.stem.2021.04.003.
382. Kim, D., Chen, R., Sheu, M., Kim, N., Kim, S., Islam, N., Wier, E.M., Wang, G., Li, A., Park, A., et al. (2019). Noncoding dsRNA induces retinoic acid synthesis to stimulate hair follicle regeneration via TLR3. Nature Communications *10*. 10.1038/s41467-019-10811-y.
383. An, N., Khan, S., Imgruet, M.K., Jueng, L., Gurbuxani, S., and Mcnerney, M.E. (2023). Oncogenic RAS promotes leukemic transformation of CUX1-deficient cells. Oncogene *42*, 881-893. 10.1038/s41388-023-02612-x.
384. Gao, T., Ptashkin, R., Bolton, K.L., Sirenko, M., Fong, C., Spitzer, B., Menghrajani, K., Ossa, J.E.A., Zhou, Y., Bernard, E., et al. (2021). Interplay between chromosomal alterations and gene mutations shapes the evolutionary trajectory of clonal hematopoiesis. Nature Communications *12*. 10.1038/s41467-020-20565-7.
385. Pich, O., Reyes-Salazar, I., Gonzalez-Perez, A., and Lopez-Bigas, N. (2022). Discovering the drivers of clonal hematopoiesis. Nature Communications *13*. 10.1038/s41467-022-31878-0.
386. Garcia, P., Jarassier, W., Brun, C., Giordani, L., Agostini, F., Kung, W.H., Peccate, C., Ravent, J., Fall, S., Petit, V., et al. (2024). Setdb1 protects genome integrity in murine muscle stem cells to allow for regenerative myogenesis and inflammation. Developmental Cell. 10.1016/j.devcel.2024.05.012.
387. Wilkinson, A.C., Ishida, R., Nakauchi, H., and Yamazaki, S. (2020). Long-term ex vivo expansion of mouse hematopoietic stem cells. Nature Protocols *15*, 628-648. 10.1038/s41596-019-0263-2.
388. Hidalgo San Jose, L., and Signer, R.A.J. (2019). Cell-type-specific quantification of protein synthesis in vivo. Nature Protocols *14*, 441-460. 10.1038/s41596-018-0100-z.
389. Van Der Burg, M., Kreyenberg, H., Willasch, A., Barendregt, B.H., Preuner, S., Watzinger, F., Lion, T., Roosnek, E., Harvey, J., Alcoceba, M., et al. (2011). Standardization of DNA isolation from low cell numbers for chimerism analysis by PCR of short tandem repeats. Leukemia *25*, 1467-1470. 10.1038/leu.2011.118.
390. 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. 10.1093/bioinformatics/bts635.
391. 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. 10.1093/bioinformatics/btt656.
392. Teissandier, A., Servant, N., Barillot, E., and Bourc'His, D. (2019). Tools and best practices for retrotransposon analysis using high-throughput sequencing data. Mobile DNA *10*. 10.1186/s13100-019-0192-1.
393. Gu, Z., Liu, Y., Zhang, Y., Cao, H., Lyu, J., Wang, X., Wylie, A., Newkirk, S.J., Jones, A.E., Lee, M., et al. (2021). Silencing of LINE-1 retrotransposons is a selective dependency of myeloid leukemia. Nature Genetics *53*, 672-682. 10.1038/s41588-021-00829-8.

394. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*. 10.1186/s13059-014-0550-8.
395. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2016). Fast gene set enrichment analysis. Cold Spring Harbor Laboratory.
396. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847-2849. 10.1093/bioinformatics/btw313.
397. Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., and Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Research *50*, W216-W221. 10.1093/nar/gkac194.
398. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nature Methods *14*, 959-962. 10.1038/nmeth.4396.
399. Grandi, F.C., Modi, H., Kampman, L., and Corces, M.R. (2022). Chromatin accessibility profiling by ATAC-seq. Nature Protocols *17*, 1518-1552. 10.1038/s41596-022-00692-9.
400. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.
401. Amemiya, H.M., Kundaje, A., and Boyle, A.P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Scientific Reports *9*. 10.1038/s41598-019-45839-z.
402. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.
403. Robinson, M.D., Mccarthy, D.J., and Smyth, G.K. (2010). <tt>edgeR</tt>: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140. 10.1093/bioinformatics/btp616.
404. Lun, A.T.L., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Research *44*, e45-e45. 10.1093/nar/gkv1191.
405. Machlab, D., Burger, L., Soneson, C., Rijli, F.M., Schübeler, D., and Stadler, M.B. (2022). monaLisa: an R/Bioconductor package for identifying regulatory motifs. Bioinformatics *38*, 2624-2625. 10.1093/bioinformatics/btac102.
406. Gu, Z., and Hübschmann, D. (2023). rGREAT: an R/bioconductor package for functional enrichment on genomic regions. Bioinformatics *39*. 10.1093/bioinformatics/btac745.
407. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494-498. 10.1038/s41586-018-0414-6.
408. 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.e3529. 10.1016/j.cell.2021.04.048.
409. Choudhary, S., and Satija, R. (2022). Comparison and evaluation of statistical error models for scRNA-seq. Genome Biology *23*. 10.1186/s13059-021-02584-9.

410. Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biology *19*. 10.1186/s13059-018-1603-1.
411. Andreatta, M., and Carmona, S.J. (2021). UCell: Robust and scalable single-cell gene signature scoring. Computational and Structural Biotechnology Journal *19*. 10.1016/j.csbj.2021.06.043.
412. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496-502. 10.1038/s41586-019-0969-x.
413. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biology *19*. 10.1186/s13059-017-1382-0.
414. Van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A.J., Burdziak, C., Moon, K.R., Chaffer, C.L., Pattabiraman, D., et al. (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell *174*, 716-729.e727. 10.1016/j.cell.2018.05.061.
415. Carbonetto, P., Luo, K., Sarkar, A., Hung, A., Tayeb, K., Pott, S., and Stephens, M. (2023). GoM DE: interpreting structure in sequence count data with differential expression analysis allowing for grades of membership. Genome Biology *24*. 10.1186/s13059-023-03067-9.
416. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nature Biotechnology *38*, 1408-1414. 10.1038/s41587-020-0591-3.
417. McClintock, B. (1956). Controlling Elements and the Gene. Cold Spring Harbor Symposia on Quantitative Biology *21*. 10.1101/SQB.1956.021.01.017.
418. Johnson, W.E. (2019). Origins and evolutionary consequences of ancient endogenous retroviruses. Nature Reviews Microbiology *17*, 355-370. 10.1038/s41579-019-0189-2.
419. Goodier, J.L., and Kazazian, H.H. (2008). Retrotransposons Revisited: The Restraint and Rehabilitation of Parasites. Cell *135*, 23-35. 10.1016/j.cell.2008.09.022.
420. Lanciano, S., and Cristofari, G. (2020). Measuring and interpreting transposable element expression. Nature Reviews Genetics *21*, 721-736. 10.1038/s41576-020-0251-y.
421. Thompson, J., Peter, Macfarlan, S., Todd, and Lorincz, C., Matthew (2016). Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire. Molecular Cell *62*, 766-776. 10.1016/j.molcel.2016.03.029.
422. Gogvadze, E., and Buzdin, A. (2009). Retroelements and their impact on genome evolution and functioning. Cellular and Molecular Life Sciences *66*, 3727-3742. 10.1007/s00018-009-0107-2.
423. Larouche, J.-D., Trofimov, A., Hesnard, L., Ehx, G., Zhao, Q., Vincent, K., Durette, C., Gendron, P., Laverdure, J.-P., Bonneil, É., et al. (2020). Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues. Genome Medicine *12*. 10.1186/s13073-020-00740-7.
424. Chen, R., Ishak, C.A., and De Carvalho, D.D. (2021). Endogenous Retroelements and the Viral Mimicry Response in Cancer Therapy and Cellular Homeostasis. Cancer Discovery *11*, 2707-2725. 10.1158/2159-8290.cd-21-0506.

425. Contreras-Galindo, R., Kaplan, M.H., Dube, D., Gonzalez-Hernandez, M.J., Chan, S., Meng, F., Dai, M., Omenn, G.S., Gitlin, S.D., and Markovitz, D.M. (2015). Human Endogenous Retrovirus Type K (HERV-K) Particles Package and Transmit HERV-K–Related Sequences. Journal of Virology *89*, 7187-7201. 10.1128/jvi.00544-15.
426. Cañadas, I., Thummalapalli, R., Kim, J.W., Kitajima, S., Jenkins, R.W., Christensen, C.L., Campisi, M., Kuang, Y., Zhang, Y., Gjini, E., et al. (2018). Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. Nature Medicine *24*, 1143-1150. 10.1038/s41591-018-0116-5.
427. Gao, D., Wu, J., Wu, Y.-T., Du, F., Aroh, C., Yan, N., Sun, L., and Chen, Z.J. (2013). Cyclic GMP-AMP Synthase Is an Innate Immune Sensor of HIV and Other Retroviruses. Science *341*, 903-906. 10.1126/science.1240933.
428. Dembny, P., Newman, A.G., Singh, M., Hinz, M., Szczepek, M., Krüger, C., Adalbert, R., Dzaye, O., Trimbuch, T., Wallach, T., et al. (2020). Human endogenous retrovirus HERV-K(HML-2) RNA causes neurodegeneration through Toll-like receptors. JCI Insight *5*. 10.1172/jci.insight.131093.
429. Chiappinelli, B., Katherine, Strissel, L., Pamela, Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, S., Neal, Cope, M., Leslie, Snyder, A., et al. (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974-986. 10.1016/j.cell.2015.07.011.
430. Roulois, D., Yau, L., Helen, Singhania, R., Wang, Y., Danesh, A., Shen, Y., Shu, Han, H., Liang, G., Jones, A., Peter, Pugh, J., Trevor, et al. (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell *162*, 961-973. 10.1016/j.cell.2015.07.056.
431. She, J., Du, M., Xu, Z., Jin, Y., Li, Y., Zhang, D., Tao, C., Chen, J., Wang, J., and Yang, E. (2022). The landscape of hervRNAs transcribed from human endogenous retroviruses across human body sites. Genome Biology *23*. 10.1186/s13059-022-02804-w.
432. Dopkins, N., and Nixon, D.F. (2024). Activation of human endogenous retroviruses and its physiological consequences. Nature Reviews Molecular Cell Biology *25*, 212-222. 10.1038/s41580-023-00674-z.
433. Burn, A., Roy, F., Freeman, M., and Coffin, J.M. (2022). Widespread expression of the ancient HERV-K (HML-2) provirus group in normal human tissues. PLOS Biology *20*, e3001826. 10.1371/journal.pbio.3001826.
434. Lee-Kirsch, M.A., Gong, M., Chowdhury, D., Senenko, L., Engel, K., Lee, Y.-A., De Silva, U., Bailey, S.L., Witte, T., Vyse, T.J., et al. (2007). Mutations in the gene encoding the 3′-5′ DNA exonuclease TREX1 are associated with systemic lupus erythematosus. Nature Genetics *39*, 1065-1067. 10.1038/ng2091.
435. Rice, G., Newman, W.G., Dean, J., Patrick, T., Parmar, R., Flintoff, K., Robins, P., Harvey, S., Hollis, T., O'Hara, A., et al. (2007). Heterozygous Mutations in TREX1 Cause Familial Chilblain Lupus and Dominant Aicardi-Goutières Syndrome. The American Journal of Human Genetics *80*, 811-815. 10.1086/513443.
436. Yan, N., Regalado-Magdos, A.D., Stiggelbout, B., Lee-Kirsch, M.A., and Lieberman, J. (2010). The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. Nature Immunology *11*, 1005-1013. 10.1038/ni.1941.
437. Stetson, D.B., Ko, J.S., Heidmann, T., and Medzhitov, R. (2008). Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. Cell *134*, 587-598. 10.1016/j.cell.2008.06.032.

438. Doitsh, G., Cavrois, M., Lassen, K.G., Zepeda, O., Yang, Z., Santiago, M.L., Hebbeler, A.M., and Greene, W.C. (2010). Abortive HIV Infection Mediates CD4 T Cell Depletion and Inflammation in Human Lymphoid Tissue. Cell *143*, 789-801. 10.1016/j.cell.2010.11.001.
439. Rice, G.I., Meyzer, C., Bouazza, N., Hully, M., Boddaert, N., Semeraro, M., Zeef, L.A.H., Rozenberg, F., Bondet, V., Duffy, D., et al. (2018). Reverse-Transcriptase Inhibitors in the Aicardi–Goutières Syndrome. New England Journal of Medicine *379*, 2275-2277. 10.1056/nejmc1810983.
440. Beck-Engeser, G.B., Eilat, D., and Wabl, M. (2011). An autoimmune disease prevented by anti-retroviral drugs. Retrovirology *8*, 91. 10.1186/1742-4690-8-91.
441. Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature *522*, 221-225. 10.1038/nature14308.
442. Sundaram, V., Cheng, Y., Ma, Z., Li, D., Xing, X., Edge, P., Snyder, M.P., and Wang, T. (2014). Widespread contribution of transposable elements to the innovation of gene regulatory networks. Genome Research *24*, 1963-1976. 10.1101/gr.168872.113.
443. Sutkowski, N., Conrad, B., Thorley-Lawson, D.A., and Huber, B.T. (2001). Epstein-Barr Virus Transactivates the Human Endogenous Retrovirus HERV-K18 that Encodes a Superantigen. Immunity *15*, 579-589. 10.1016/s1074-7613(01)00210-2.
444. Dopkins, N., Singh, B., Michael, S., O'Mara, M.M., Marston, J.L., Fei, T., Bendall, M.L., and Nixon, D.F. (2023). Endogenous Reverse Transcriptase Inhibition Attenuates TLR5-Mediated Inflammation. mBio *14*. 10.1128/mbio.03280-22.
445. Rookhuizen, D.C., Bonte, P.-E., Ye, M., Hoyler, T., Gentili, M., Burgdorf, N., Durand, S., Aprahamian, F., Kroemer, G., Manel, N., et al. (2021). Induction of transposable element expression is central to innate sensing. Cold Spring Harbor Laboratory.
446. Young, G.R., Terry, S.N., Manganaro, L., Cuesta-Dominguez, A., Deikus, G., Bernal-Rubio, D., Campisi, L., Fernandez-Sesma, A., Sebra, R., Simon, V., and Mulder, L.C.F. (2018). HIV-1 Infection of Primary CD4
⁺
T Cells Regulates the Expression of Specific Human Endogenous Retrovirus HERV-K (HML-2) Elements. Journal of Virology *92*. 10.1128/jvi.01507-17.
447. Terry, S.N., Manganaro, L., Cuesta-Dominguez, A., Brinzevich, D., Simon, V., and Mulder, L.C.F. (2017). Expression of HERV-K108 envelope interferes with HIV-1 production. Virology *509*. 10.1016/j.virol.2017.06.004.
448. Young, G.R., Mavrommatis, B., and Kassiotis, G. (2014). Microarray analysis reveals global modulation of endogenous retroelement transcription by microbes. Retrovirology *11*, 59. 10.1186/1742-4690-11-59.
449. Manghera, M., and Douville, R.N. (2013). Endogenous retrovirus-K promoter: a landing strip for inflammatory transcription factors? Retrovirology *10*, 16. 10.1186/1742-4690-10-16.
450. Bourque, G., Leong, B., Vega, V.B., Chen, X., Lee, Y.L., Srinivasan, K.G., Chew, J.-L., Ruan, Y., Wei, C.-L., Ng, H.H., and Liu, E.T. (2008). Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Research *18*, 1752-1762. 10.1101/gr.080663.108.

451. Schmidt, N., Domingues, P., Golebiowski, F., Patzina, C., Tatham, M.H., Hay, R.T., and Hale, B.G. (2019). An influenza virus-triggered SUMO switch orchestrates co-opted endogenous retroviruses to stimulate host antiviral immunity. Proceedings of the National Academy of Sciences *116*, 17399-17408. 10.1073/pnas.1907031116.
452. Iampietro, M., Dumont, C., Mathieu, C., Spanier, J., Robert, J., Charpenay, A., Dupichaud, S., Dhondt, K.P., Aurine, N., Pelissier, R., et al. (2021). Activation of cGAS/STING pathway upon paramyxovirus infection. iScience *24*, 102519. 10.1016/j.isci.2021.102519.
453. Holm, C.K., Rahbek, S.H., Gad, H.H., Bak, R.O., Jakobsen, M.R., Jiang, Z., Hansen, A.L., Jensen, S.K., Sun, C., Thomsen, M.K., et al. (2016). Influenza A virus targets a cGAS-independent STING pathway that controls enveloped RNA viruses. Nature Communications *7*, 10680. 10.1038/ncomms10680.
454. Aguirre, S., Luthra, P., Sanchez-Aparicio, M.T., Maestre, A.M., Patel, J., Lamothe, F., Fredericks, A.C., Tripathi, S., Zhu, T., Pintado-Silva, J., et al. (2017). Dengue virus NS2B protein targets cGAS for degradation and prevents mitochondrial DNA sensing during infection. Nature Microbiology *2*, 17037. 10.1038/nmicrobiol.2017.37.
455. Zheng, Y., Liu, Q., Wu, Y., Ma, L., Zhang, Z., Liu, T., Jin, S., She, Y., Li, Y.P., and Cui, J. (2018). Zika virus elicits inflammation to evade antiviral response by cleaving <scp>cGAS</scp> via <scp>NS</scp> 1-caspase-1 axis. The EMBO Journal *37*, e99347. 10.15252/embj.201899347.
456. Chuong, E.B., Elde, N.C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. Nature Reviews Genetics *18*, 71-86. 10.1038/nrg.2016.139.
457. Asimi, V., Sampath Kumar, A., Niskanen, H., Riemenschneider, C., Hetzel, S., Naderi, J., Fasching, N., Popitsch, N., Du, M., Kretzmer, H., et al. (2022). Hijacking of transcriptional condensates by endogenous retroviruses. Nature Genetics *54*, 1238-1247. 10.1038/s41588-022-01132-w.
458. Karttunen, K., Patel, D., Xia, J., Fei, L., Palin, K., Aaltonen, L., and Sahu, B. (2023). Transposable elements as tissue-specific enhancers in cancers of endodermal lineage. Nature Communications *14*. 10.1038/s41467-023-41081-4.
459. Deniz, Ö., Ahmed, M., Todd, C.D., Rio-Machin, A., Dawson, M.A., and Branco, M.R. (2020). Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. Nature Communications *11*, 1-14. 10.1038/s41467-020-17206-4.
460. De Cecco, M., Criscione, S.W., Peckham, E.J., Hillenmeyer, S., Hamm, E.A., Manivannan, J., Peterson, A.L., Kreiling, J.A., Neretti, N., and Sedivy, J.M. (2013). Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. Aging Cell *12*, 247-256. 10.1111/acel.12047.
461. Li, W., Lee, M.-H., Henderson, L., Tyagi, R., Bachani, M., Steiner, J., Campanac, E., Hoffman, D.A., Von Geldern, G., Johnson, K., et al. (2015). Human endogenous retrovirus-K contributes to motor neuron disease. Science Translational Medicine *7*, 307ra153-307ra301. 10.1126/scitranslmed.aac8201.

462. Wang, Z., Fan, R., Russo, A., Cernilogar, F.M., Nuber, A., Schirge, S., Shcherbakova, I., Dzhilyanova, I., Ugur, E., Anton, T., et al. (2022). Dominant role of DNA methylation over H3K9me3 for IAP silencing in endoderm. Nature Communications *13*. 10.1038/s41467-022-32978-7.
463. Ohtani, H., Liu, M., Zhou, W., Liang, G., and Jones, P.A. (2018). Switching roles for DNA and histone methylation depend on evolutionary ages of human endogenous retroviruses. Genome Research *28*, 1147-1157. 10.1101/gr.234229.118.
464. Padeken, J., Methot, S.P., and Gasser, S.M. (2022). Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. Nature Reviews Molecular Cell Biology *23*, 623-640. 10.1038/s41580-022-00483-w.
465. Robbez-Masson, L., Tie, C.H.C., Conde, L., Tunbak, H., Husovsky, C., Tchasovnikarova, I.A., Timms, R.T., Herrero, J., Lehner, P.J., and Rowe, H.M. (2018). The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. Genome Research *28*, 836-845. 10.1101/gr.228171.117.
466. Tunbak, H., Enriquez-Gasca, R., Tie, C.H.C., Gould, P.A., Mlcochova, P., Gupta, R.K., Fernandes, L., Holt, J., Van Der Veen, A.G., Giampazolias, E., et al. (2020). The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. Nature Communications *11*. 10.1038/s41467-020-19170-5.
467. Griffin, G.K., Wu, J., Iracheta-Vellve, A., Patti, J.C., Hsu, J., Davis, T., Dele-Oni, D., Du, P.P., Halawi, A.G., Ishizuka, J.J., et al. (2021). Epigenetic silencing by SETDB1 suppresses tumour intrinsic immunogenicity. Nature *595*, 309-314. 10.1038/s41586-021-03520-4.
468. Cuellar, T.L., Herzner, A.-M., Zhang, X., Goyal, Y., Watanabe, C., Friedman, B.A., Janakiraman, V., Durinck, S., Stinson, J., Arnott, D., et al. (2017). ---Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia--. Journal of Cell Biology *216*, 3535-3549. 10.1083/jcb.201612160.
469. Shen, J.Z., Qiu, Z., Wu, Q., Finlay, D., Garcia, G., Sun, D., Rantala, J., Barshop, W., Hope, J.L., Gimple, R.C., et al. (2021). FBXO44 promotes DNA replication-coupled repetitive element silencing in cancer cells. Cell *184*, 352-369.e323. 10.1016/j.cell.2020.11.042.
470. Collins, R.E., Tachibana, M., Tamaru, H., Smith, K.M., Jia, D., Zhang, X., Selker, E.U., Shinkai, Y., and Cheng, X. (2005). In Vitro and in Vivo Analyses of a Phe/Tyr Switch Controlling Product Specificity of Histone Lysine Methyltransferases. Journal of Biological Chemistry *280*, 5563-5570. 10.1074/jbc.m410483200.
471. Zylicz, J.J., Dietmann, S., Günesdogan, U., Hackett, J.A., Cougot, D., Lee, C., and Surani, M.A. (2015). Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. 10.7554/eLife.09571.
472. Di Giacomo, M., Comazzetto, S., Sampath, S.C., Sampath, S.C., and O'Carroll, D. (2014). G9a co-suppresses LINE1 elements in spermatogonia. Epigenetics & Chromatin *7*, 24. 10.1186/1756-8935-7-24.
473. Takayama, N., Murison, A., Takayanagi, S.-I., Arlidge, C., Zhou, S., Garcia-Prat, L., Chan-Seng-Yue, M., Zandi, S., Gan, O.I., Boutzen, H., et al. (2021). The Transition from Quiescent to Activated States in Human Hematopoietic Stem Cells Is Governed by Dynamic 3D Genome Reorganization. Cell Stem Cell *28*, 488-501.e410. 10.1016/j.stem.2020.11.001.

474. Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nature Reviews Genetics *19*, 81-92. 10.1038/nrg.2017.80.
475. Christman, J.K. (2002). 5-Azacytidine and 5-aza-2′-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. Oncogene *21*, 5483-5495. 10.1038/sj.onc.1205699.
476. Yabushita, T., Chinen, T., Nishiyama, A., Asada, S., Shimura, R., Isobe, T., Yamamoto, K., Sato, N., Enomoto, Y., Tanaka, Y., et al. (2023). Mitotic perturbation is a key mechanism of action of decitabine in myeloid tumor treatment. Cell Reports *42*, 113098. 10.1016/j.celrep.2023.113098.
477. Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., et al. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. Nature Genetics *44*, 23-31. 10.1038/ng.1009.
478. Karpf, A.R., Peterson, P.W., Rawlins, J.T., Dalley, B.K., Yang, Q., Albertsen, H., and Jones, D.A. (1999). Inhibition of DNA methyltransferase stimulates the expression of signal transducer and activator of transcription 1, 2, and 3 genes in colon tumor cells. Proceedings of the National Academy of Sciences *96*, 14007-14012. 10.1073/pnas.96.24.14007.
479. Unnikrishnan, A., Papaemmanuil, E., Beck, D., Deshpande, N.P., Verma, A., Kumari, A., Woll, P.S., Richards, L.A., Knezevic, K., Chandrakanthan, V., et al. (2017). Integrative Genomics Identifies the Molecular Basis of Resistance to Azacitidine Therapy in Myelodysplastic Syndromes. Cell Reports *20*, 572-585. 10.1016/j.celrep.2017.06.067.
480. Gonzalo, S., Jaco, I., Fraga, M.F., Chen, T., Li, E., Esteller, M., and Blasco, M.A. (2006). DNA methyltransferases control telomere length and telomere recombination in mammalian cells. Nature Cell Biology *8*, 416-424. 10.1038/ncb1386.
481. Zhang, C.R., Ostrander, E.L., Kukhar, O., Mallaney, C., Sun, J., Haussler, E., Celik, H., Koh, W.K., King, K.Y., Gontarz, P., and Challen, G.A. (2022). Txnip Enhances Fitness of Dnmt3a-Mutant Hematopoietic Stem Cells via p21. Blood Cancer Discovery *3*, 220-239. 10.1158/2643-3230.bcd-21-0132.
482. 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. Proceedings of the National Academy of Sciences *108*, 14566-14571. 10.1073/pnas.1112317108.
483. Cimmino, L., Dolgalev, I., Wang, Y., Yoshimi, A., Martin, G.H., Wang, J., Ng, V., Xia, B., Witkowski, M.T., Mitchell-Flack, M., et al. (2017). Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. Cell *170*, 1079-1095.e1020. 10.1016/j.cell.2017.07.032.
484. Mclemore, A.F., Hou, H.-A., Meyer, B.S., Lam, N.B., Ward, G.A., Aldrich, A.L., Rodrigues, M.A., Vedder, A., Zhang, L., Padron, E., et al. (2022). Somatic gene mutations expose cytoplasmic DNA to co-opt the cGAS/STING/NLRP3 axis in myelodysplastic syndromes. JCI Insight *7*. 10.1172/jci.insight.159430.
485. Xie, J., Sheng, M., Rong, S., Zhou, D., Wang, C., Wu, W., Huang, J., Sun, Y., Wang, Y., Chen, P., et al. (2023). STING activation in TET2-mutated hematopoietic stem/progenitor cells contributes to the increased self-renewal and neoplastic transformation. Leukemia *37*, 2457-2467. 10.1038/s41375-023-02055-z.

486. Marraffini, L.A., and Sontheimer, E.J. (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature *463*, 568-571. 10.1038/nature08703.
487. Janeway, C.A. (1989). Approaching the Asymptote? Evolution and Revolution in Immunology. Cold Spring Harbor Symposia on Quantitative Biology *54*, 1-13. 10.1101/sqb.1989.054.01.003.
488. Gough, J., Daniel, Messina, L., Nicole, Clarke, J.P., Christopher, Johnstone, W., Ricky, and Levy, E., David (2012). Constitutive Type I Interferon Modulates Homeostatic Balance through Tonic Signaling. Immunity *36*, 166-174. 10.1016/j.immuni.2012.01.011.
489. Metidji, A., Rieder, S.A., Glass, D.D., Cremer, I., Punkosdy, G.A., and Shevach, E.M. (2015). IFN-α/β Receptor Signaling Promotes Regulatory T Cell Development and Function under Stress Conditions. The Journal of Immunology *194*, 4265-4276. 10.4049/jimmunol.1500036.
490. Sato, T., Onai, N., Yoshihara, H., Arai, F., Suda, T., and Ohteki, T. (2009). Interferon regulatory factor-2 protects quiescent hematopoietic stem cells from type I interferon–dependent exhaustion. Nature Medicine *15*, 696-700. 10.1038/nm.1973.
491. Poscablo, D.M., Worthington, A.K., Smith-Berdan, S., Rommel, M.G.E., Manso, B.A., Adili, R., Mok, L., Reggiardo, R.E., Cool, T., Mogharrab, R., et al. (2024). An age-progressive platelet differentiation path from hematopoietic stem cells causes exacerbated thrombosis. Cell *187*, 3090-3107.e3021. 10.1016/j.cell.2024.04.018.
492. Ugale, A., Shunmugam, D., Pimpale, L.G., Rebhan, E., and Baccarini, M. (2024). Signaling proteins in HSC fate determination are unequally segregated during asymmetric cell division. Journal of Cell Biology *223*. 10.1083/jcb.202310137.