Epigenetic mechanisms controlling human leukemia stem cells and therapy resistance

Sumiko Takao[1,2], Victor Morell[1,2], Fiona C Brown[1], Richard Koche[3] and Alex Kentsis[1,2,4]

[1]Molecular Pharmacology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[2]Tow Center for Developmental Oncology, Department of Pediatrics, Memorial Sloan Kettering Cancer Center, NY, USA.

[3]Center for Epigenetics Research, Sloan Kettering Institute, New York, NY, USA

[4]Departments of Pediatrics, Pharmacology, and Physiology & Biophysics, Weill Medical College of Cornell University, NY, USA.

Email: kentsisresearchgroup@gmail.com

Abstract

Many human cancers, including acute myeloid leukemia (AML), arise from mutations of stem and progenitor cells. Immunophenotypic profiling has shown that leukemia develops hierarchically, with mutations in leukemia stem cells associated with disease propagation and relapse[1,2]. Although leukemia initiating cells can be enriched using cell surface markers, their frequency tends to be variable and low, obscuring mechanisms and hindering effective therapies[3,4]. To define AML stem cells in human patients, we performed functional genomic profiling of diverse leukemias using label tracing techniques designed to preserve hematopoietic stem cell (HSC) function in vivo. We found that propagation of human AML is mediated by a rare but distinct quiescent label-retaining cell (LRC) population that evades detection by currently known immunophenotypic markers. We show that human AML LRC quiescence is reversible, sparing genetic clonal competition that maintains its epigenetic inheritance. LRC quiescence is defined by distinct promoter-centered chromatin and gene expression dynamics, and controlled by a novel AP-1/ETS transcription factor network, which is associated with disease persistence and chemotherapy resistance in diverse patients. These results enable prospective isolation and functional genetic manipulation of immunophenotypically-varied leukemia stem cells in human patient specimens, as well as establish the key functions of epigenetic plasticity in leukemia development and therapy resistance. We anticipate that these findings will lead to the elucidation of essential properties of leukemia stem cell quiescence and the design of therapeutic strategies for their clinical identification and control.

Main

Although the treatment of AML continues to improve, most patients develop disease that is refractory to intensive chemotherapy[5–7]. Leukemia development, evolution, and chemotherapy resistance have been attributed to leukemia stem cells, originally identified as rare cells that can propagate human leukemias upon transplantation in immunodeficient mice[1,2]. While leukemia initiating cells can be enriched using cell surface markers, their frequency tends to be variable and low (less than 1 in 1000 cells), obscuring mechanisms and hindering effective therapies[3,4]. Molecular mechanisms of leukemia stem cell development and therapy resistance have been investigated in genetically-engineered leukemias in mice[8–11], but there remains uncertainty about their relevance to human AML.

Prospective isolation of human patient AML stem cells

To define molecular mechanisms of human AML stem cells, we used orthotopic transplantation in immunodeficient mice in combination with chemical label tracing and functional genetics of primary human patient leukemia specimens. First, we assessed 48 human patient leukemias obtained at disease diagnosis or relapse by orthotopic transplantation in *NOD.Cg-Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice. Overall engraftment efficiency was 46 and 70% in primary and secondary transplants, respectively (Extended Data Table 1). Consistent with prior observations[12,13], serial engraftment and transplantation were higher for specimens obtained from patients with chemotherapy resistance, as compared to those who achieved remission upon standard-of-care chemotherapy treatment (80 versus 60%; Extended Data Table 1).

Based on prior studies using genetic and chemical label tracing approaches for the isolation of quiescent or dormant cells[14–20], we optimized carboxyfluorescein succinimidyl ester (CFSE) chemical label tracing technique by maximizing covalent cellular protein labeling while avoiding any measurable effects on the viability of hematopoietic cells (Extended Data Figure 1a-b). When applied to primary hematopoietic cells isolated from the bone marrow of healthy mice,

CFSE labeling preserved normal stem cell function, as evident from complete rescue of lethally irradiated wild-type mice upon transplantation of CFSE-labeled hematopoietic stem and progenitor cells (Extended Data Figure 1c). We confirmed that these labeling conditions also preserved primary human leukemia initiating cells in their ability to initiate disease upon orthotopic transplantation in NSG mice by CFSE-labeled bulk patient leukemia cells as compared to their unlabeled controls (Extended Data Figure 1d).

We hypothesized that CFSE label retention would specifically identify leukemia stem cells, given its ability to detect non-dividing cells with quiescent proteome turnover (Figure 1a). To test this idea, we selected genetically diverse patient AML specimens that exhibited short latency in serial orthotopic mouse transplants (Extended Data Table 2). First, we confirmed that CFSE labeling and orthotopic transplantation of human patient leukemia cells in NSG mice identified non-proliferating label retaining cells (LRCs), as measured by 5-ethynyl-2'-deoxyuridine (EdU) incorporation and fluorescence-activated cell sorting (FACS; Extended Data Figure 1e). Consistent with this, LRCs also exhibited little to no detectable apoptosis, as measured using cleaved caspase 3 intracellular FACS staining, in contrast to proliferating non-LRCs some of which undergo apoptosis (Extended Data Figure 1g).

Initiation of human patient leukemias is mediated by quiescent and chemoresistant LRCs

To determine whether patient LRCs have leukemia stem cell properties, we transplanted equal numbers of LRCs and non-LRCs into secondary recipient mice. All MSK011 LRC-transplanted mice developed leukemias (900 LRCs/mouse), while most non-LRC transplanted mice remained disease free (log-rank p = 0.021; Figure 1b). Whereas both MSK162 LRCs and non-LRCs caused disease initially, only LRCs caused leukemias in tertiary recipients and non-LRCs did not (log-rank p = 0.013; Figure 1c and Extended Data Figure 2c). This is consistent with varied rates of CFSE label attenuation and LRC frequency in different patient leukemias (18 and 0.3% at 14 days post-transplant with total 0.2 and 22% bone marrow human CD45+ cells,

Figure 1.

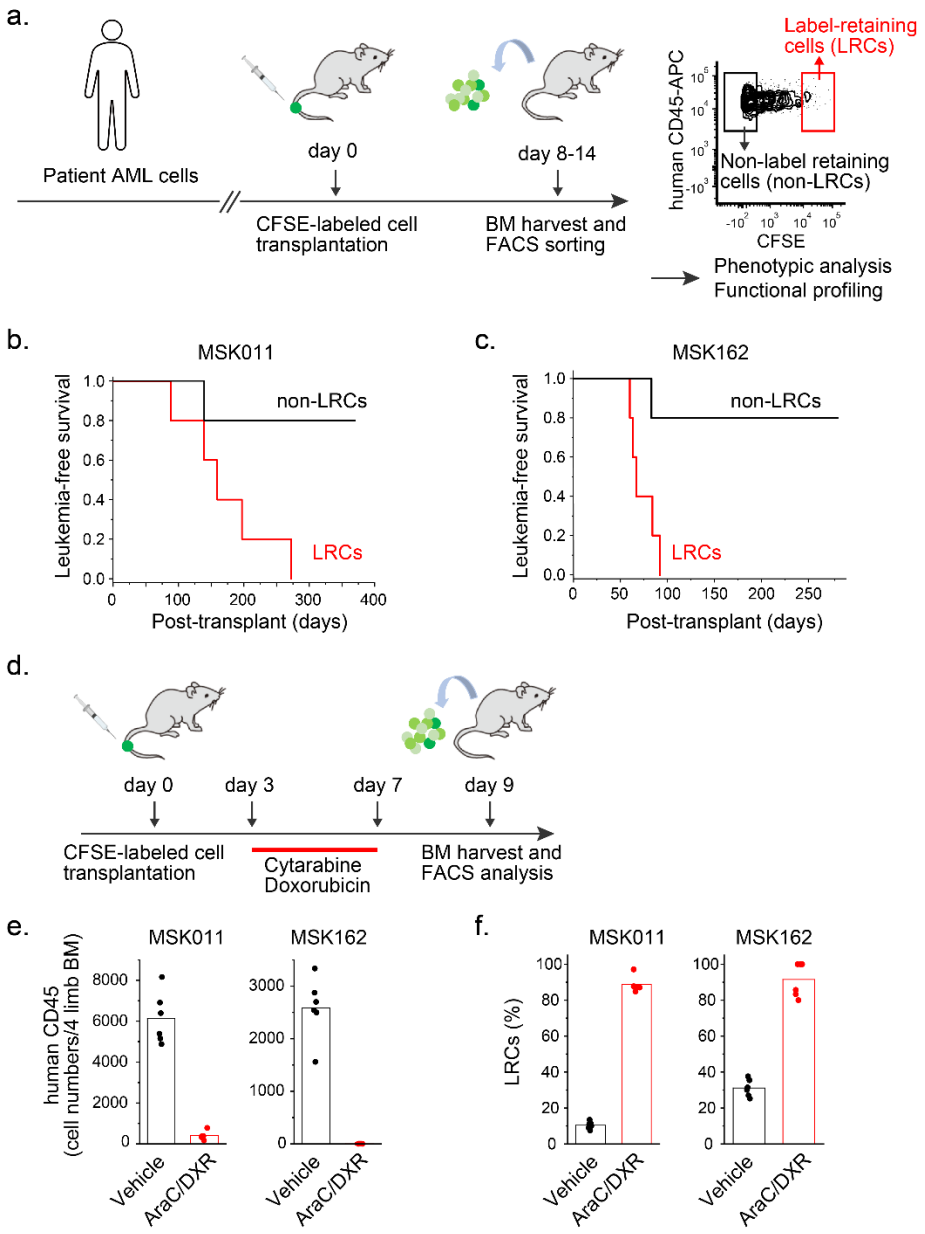

Figure 1. Quiescent human AML patient cells maintain leukemia initiation, propagation and chemotherapy resistance.
a. Experimental design to prospectively isolate quiescent AML cells from human patients using optimized CFSE labeling and orthotopic transplantation in immunodeficient mice (left panel). Representative FACS analysis of quiescent label-retaining human CD45-positive AML cells (LRCs, red box) with high CFSE fluorescence, as compared to their non-label retaining cells (non-LRC, black box) that have lost CFSE fluorescence through cell division and proteome turnover (right panel).
b. Leukemia-free survival of mice secondarily transplanted with equal numbers of human MSK011 patient AML LRCs (red) or non-LRCs (black), where LRCs initiate fully penetrant leukemia, and non-LRCs do not (log-rank $p = 0.021$).
c. Leukemia-free survival of mice tertiarily transplanted with equal numbers of human MSK162 patient LRCs (red) or non-LRCs (black), where LRCs propagate and initiate leukemia, whereas non-LRCs largely do not (log-rank $p = 0.013$).
d. Experimental design for the analysis of mice transplanted with CFSE-labeled human patient AMLs and treated with cytarabine (AraC) and doxorubicin (DXR).
e. Combined AraC and DXR chemotherapy treatment reduces bone marrow disease burden of human CD45-positive MSK011 (left panel) and MSK162 (right panel) AML cells in mice (t-test $p = 6.8 \times 10^{-5}$, and 1.2×10^{-4}, respectively). Bars represent mean values of 6 biological replicates.
f. Human patient MSK011 and MSK162 LRCs are resistant to AraC and DXR chemotherapy (red) as compared to vehicle-treated controls (black), in contrast to non-LRCs that are largely eradicated by chemotherapy treatment (t-test $p = 2.5 \times 10^{-7}$ and 1.4×10^{-6}, respectively). Bars represent mean values of measurement of 6 biological replicates.

respectively; Extended Data Figure 2a-b). Though MSK162 LRCs showed relatively high and low expression of CD34 and CD38, respectively, in agreement with prior observations[2], no established surface marker combinations, including lymphoid-primed multipotent progenitor (LMPP)-like (CD34+CD38-CD45RA+) and granulocyte-monocyte progenitor (GMP)-like (CD34+CD38+CD123+CD45RA+), could be used to reliably discriminate LRCs from non-LRCs (Expanded Data Figure 3a-c). We found similar results with MSK011 and MSK165 leukemias. Thus, initiation and maintenance of diverse human patient leukemias are mediated by a rare but distinct quiescent LRC population with varied cell surface markers.

Are LRCs chemotherapy resistant? To determine this, we used a combined cytarabine (AraC) and doxorubicin (DXR) treatment regimen of NSG mice developed to model induction chemotherapy used clinically for human patients[21]. Analysis of bone marrow of AraC- and DXR-treated mice transplanted with CFSE-labeled human patient AMLs showed that while chemotherapy significantly reduced total leukemia disease burden, more than 89 and 92% of chemotherapy-resistant cells were LRCs (t-test $p = 2.5 \times 10^{-7}$ and 1.4×10^{-6} for MSK011 and MSK162 patient specimens, respectively; Figure 1d-f and Extended Data Figure 2d). Thus, LRCs comprise a chemotherapy-resistant reservoir, consistent with their cellular quiescence.

Human patient AML cell quiescence is reversible

We next investigated whether patient LRC quiescence is associated with specific genetic clones. First, we identified disease-defining mutations for MSK011, MSK162, and MSK165 patient leukemias using high-coverage DNA sequencing of 585 genes recurrently mutated in hematologic malignancies (Extended Data Table 2)[22]. This analysis also identified 18, 21, and 52 single nucleotide and short insertion and deletion variants with varied allele frequencies (VAF), reflecting the specific clonal architectures of these leukemias (Extended Data Table 3; Figure 2a-c). These genetic variants included pathogenically cooperating mutations, such as *KRAS G12C* and

KMT2A-translocation in MSK011, *KAT6A*-translocation and *FLT3 D835Y* in MSK162, and *WT1 S381** and *FLT3-ITD E604_F605ins* in MSK165 leukemias. We reasoned that if LRC quiescence is caused by genetic clonal evolution, then specific subclonal mutations should segregate in LRCs versus non-LRCs, as measured by their relative VAFs (Extended Data Table 3; Figure 2a-c). No subclonal mutations exhibited significant association with patient LRCs. In contrast, several subclones were significantly enriched in non-LRCs, such as *FLT3 D835Y* in MSK162 and *FLT3-ITD E604_F605ins* in MSK165 patient leukemias (non-LRC 76 versus 0%, and 66 versus 31%, respectively; Extended Data Table 3; Figure 2a-c). This explains the tendency of *FLT3*-mutant subclones to be depleted by chemotherapy, and is consistent with recent measurements of clonal evolution using single-cell sequencing[23–25]. Thus, genetic clonal competition is not required for LRC quiescence.

To directly investigate the origin of human LRC quiescence, we conducted re-labeling experiments to measure LRC reversibility. Upon isolating non-LRCs from the primary recipient mice transplanted with CFSE-labeled patient AML cells, we re-labeled these cells and measured their ability to regenerate LRCs as compared to parental bulk patient AML cells in secondary recipient mice (Figure 2d). Approximately half of initially proliferating non-LRCs as compared to parental bulk leukemia cells were found to re-acquire LRC quiescence upon secondary transplantation (Figure 2e-f). Thus, human patient AML cell quiescence is a cell state which is reversibly accessible by proliferating cells.

Distinct promoter-centered chromatin and gene expression dynamics of human LRCs

We reasoned that the reversibility of LRC quiescence can be regulated epigenetically. To elucidate this, we used recently developed low-input assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) to define LRC-regulated chromatin regions. This identified 777, 3437, and 2036 differentially accessible regions in LRCs of MSK011, MSK162, and MSK165 patient leukemias, respectively (Extended Figure 4a-f), including commonly shared

Figure 2.

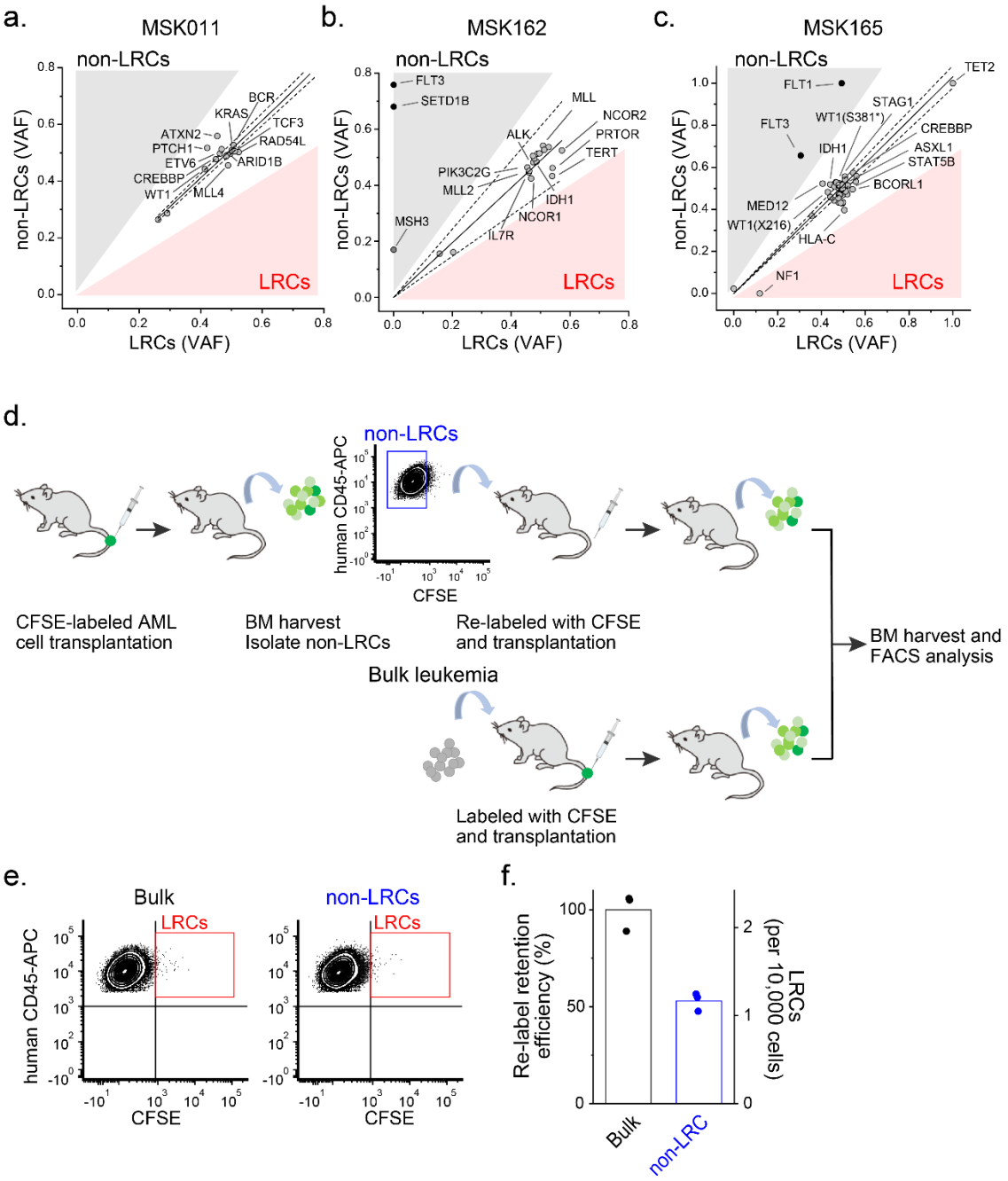

Figure 2. Human patient AML quiescence is reversible evading genetic clonal competition that maintains disease propagation in serial orthografts in vivo.

a-c. Comparison of variant allele frequencies (VAF) of genomic DNA sequencing of LRCs (red) versus non-LRCs (gray) in MSK011 (a), MSK162 (b) and MSK165 (c) human patient AMLs, demonstrating multiple genetic subclones that are enriched in non-LRCs as compared to LRCs. Dashed lines mark 95% confidence intervals.

d. Experimental design to examine quiescence reversibility of human AML patient cells. Bulk leukemia cells and purified non-LRCs (blue) are re-labeled and sequentially transplanted into recipient mice.

e. Representative FACS plots of human leukemia cells harvested from bone marrow of mice transplanted with bulk leukemia cells (left) or non-LRCs (right).

f. Initially proliferating, human patient non-LRCs exhibit significant ability to regenerate quiescent LRCs (blue) as compared to parental bulk leukemia cells (black; t-test p = 0.0016). Data show efficiencies of regenerating LRCs (right y-axis) relative to bulk cells (left y-axis). Bars represent mean values of measurement biological triplicates.

Figure 3.

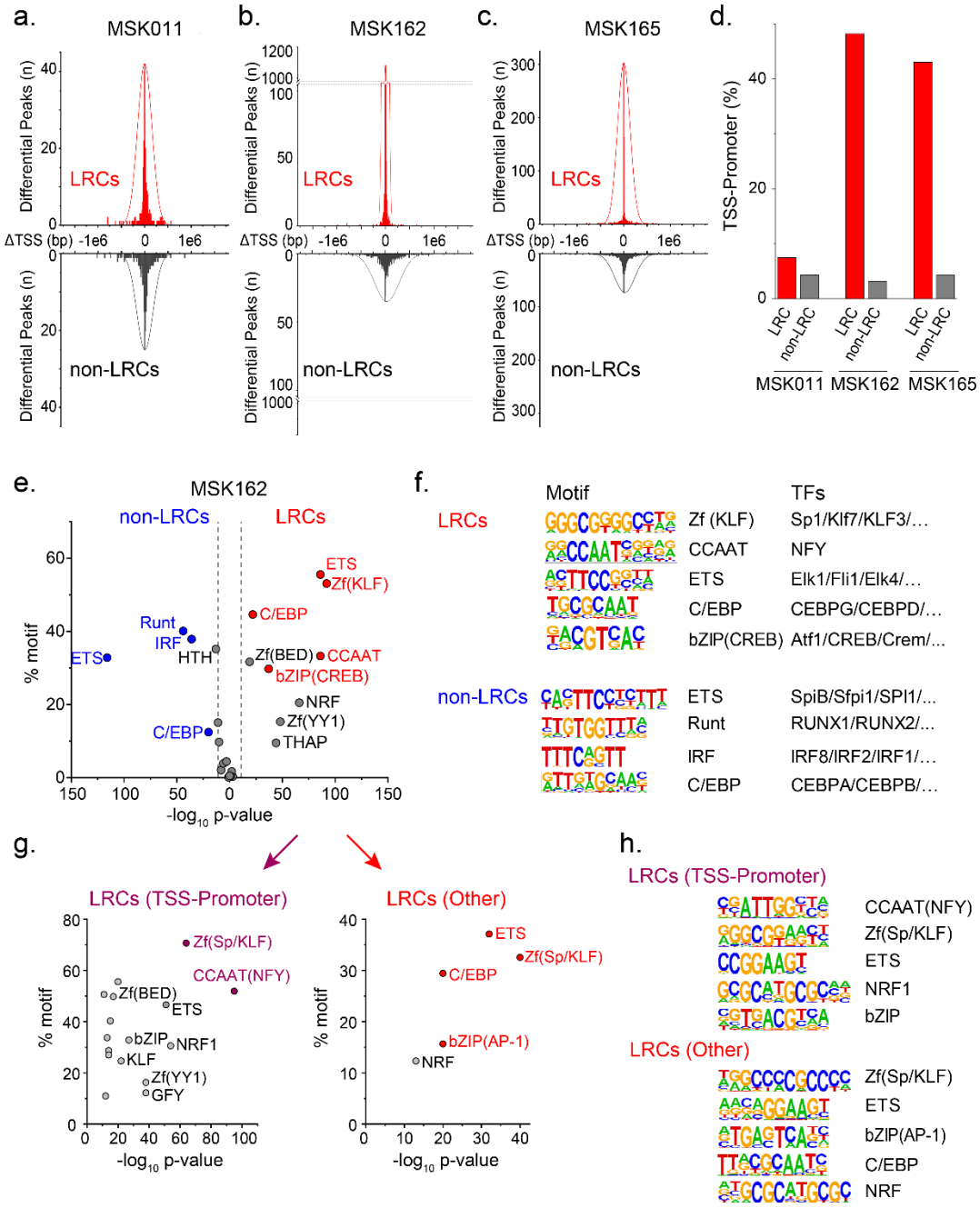

Figure 3. Human patient AML quiescence is associated with promoter-centered chromatin accessibility dynamics.
a-c. Histograms of differentially accessible chromatin regions in quiescent LRCs (red) as compared to non-LRCs (black) as a function of their distance from transcription start sites (ΔTSS) in MSK011 (a), MSK162 (b), and MSK165 (c) human patient leukemias.
d. Human quiescent LRCs (red) exhibit increased chromatin accessibility of transcriptional start promoter regions as compared to non-LRCs (gray) in MSK011, MSK162, and MSK165 leukemias (Fisher's exact test p = 6.9 x 10^{-2}, 2.2 x 10^{-16}, and 2.2 x 10^{-16}, respectively). Data represent measurement triplicates.
e-f. Transcription factor binding sequence motifs enriched in differentially accessible chromatin regions in MSK162 LRCs (red) as compared to non-LRCs (blue) as a function of their statistical significance of enrichment (e), with specific motif sequences shown (f). Dashed lines mark p-values of 1.0 x 10^{-11}.

loci (Extended Figure 4g). Whereas most non-LRC chromatin dynamics were distributed across the genome relatively uniformly, changes in LRC chromatin accessibility were concentrated near promoter regions (Figure 3a-d and Extended Data Figure 4h). We observed specific transcription factor DNA binding sequence motifs in LRC-accessible chromatin regions (Figure 3e-h, Extended Data Figure 5a-d and e-h), including those corresponding to ETS, KLF, CCAAT, and AP-1 (e.g. JUN) binding motifs[26–31].

Using RNA sequencing, we also identified 749, 567, and 489 significantly differentially LRC-expressed genes in MSK011, MSK162, and MSK165 patient leukemias, respectively (Figure 4a-b, Extended Data Figure 6a-f, Extended Data Table 4 and 5). This included *JUN*, *FLI1*, *ETS1* and *KLF2*, consistent with sequence motif analysis of differentially accessible LRC chromatin (Figure 4c). We also identified several other potential LRC regulators such as *ZFP36L1* for example (Figure 4b), whose paralog *ZFP36L2* was recently identified as a regulator of AML cell differentiation[32]. In all, these findings defined distinct promoter-centered chromatin and gene expression dynamics of human LRCs.

In addition, these data provided an opportunity to determine whether LRC quiescence contributes to therapy resistance and disease persistence in patients with AML. Gene set enrichment analysis (GSEA) of patient LRCs revealed 260 significantly upregulated and 94 significantly depleted gene sets in LRCs (Figure 4d and Extended Data Table 6). For example, we detected downregulation of MYC activity and ribosome biogenesis, similar to prior observations of dormant hematopoietic stem cells[33–36]. We combined GSEA with recent single-cell RNA sequencing (scRNA-seq) of bone marrow cells isolated from AML patients before and after chemotherapy treatment (Figure 5a-b, Extended Data Table 7), leveraging simultaneous detection of differentially expressed and mutant genes to specifically identify specific differential gene expression in leukemia cells[37]. Supervised comparison of genetically varied leukemia cells upon chemotherapy treatment of two different patients, AML329 and AML707B, showed multiple gene sets significantly upregulated and downregulated with those specifically expressed in human

Figure 4.

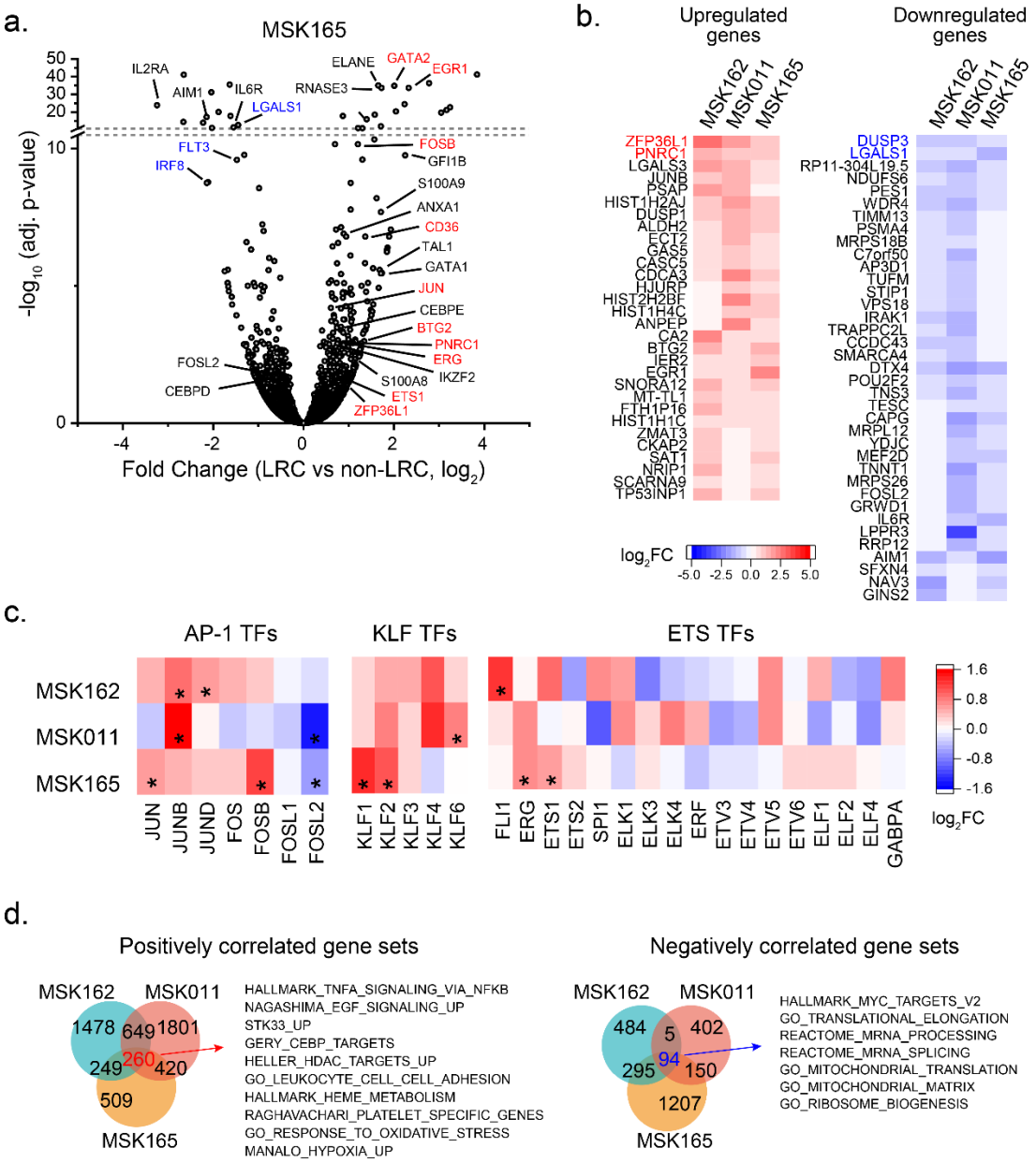

Figure 4. Coherent gene expression dysregulation in quiescent patient AML LRC cells.
a. Representative gene expression of MSK165 human patient AML cells with statistical significance of measurement triplicates as a function of differential gene expression of LRCs versus non-LRCs. Notable upregulated and downregulated genes are labeled in red and blue, respectively. Results of MSK011 and MSK162 LRC gene expression analysis are described in Extended Data Figure 6d-e.
b. Heatmaps showing commonly differentially regulated genes between LRCs and non-LRCs which are significantly upregulated (left panel) and downregulated (right panel) in at least 2 human patient leukemias (adjusted $p < 0.1$, fold change > 1.5), with blue to red color gradient representing relative decrease and increase of fold change of gene expression, respectively.
c. Heatmap showing differential gene expression of specific AP-1, KLF and ETS transcription factors (TFs) between LRCs and non-LRCs in MSK011, MSK162, and MSK165 human patient leukemias with blue to red color gradient representing relative decrease and increase of fold change of gene expression, respectively (* denotes adjusted $p < 0.1$, fold change > 1.5).
d. Gene set enrichment analysis for differentially expressed genes in MSK011, MSK162, and MSK165 LRCs versus non-LRCs, showing commonly dysregulated expression of genes regulating specific cellular signaling pathways. Venn diagrams show the numbers of significantly positively correlated (left) and negatively correlated (right) gene sets in LRCs ($p < 0.01$, false discovery rate < 0.25). Selected significantly enriched gene sets are listed.

Figure 5.

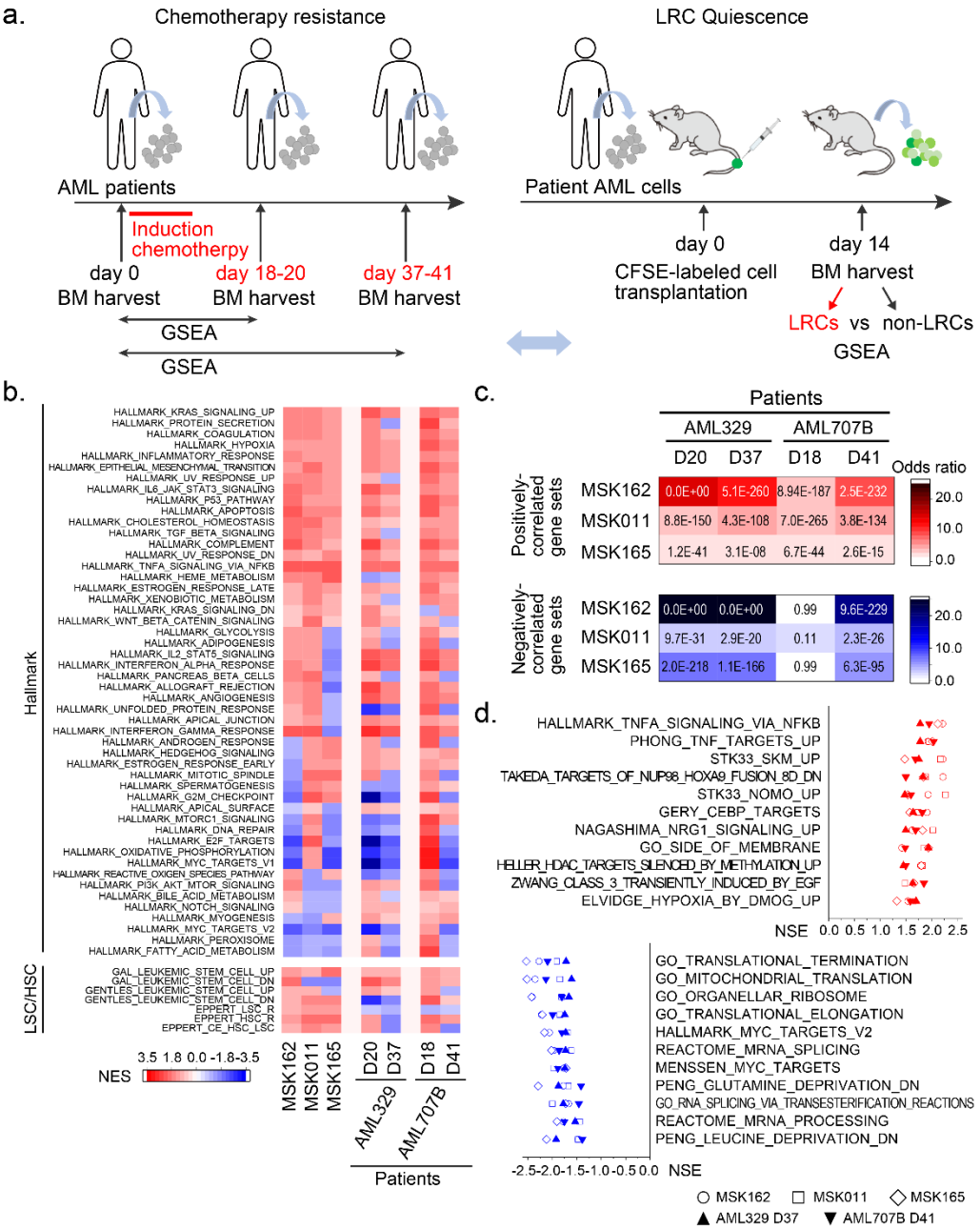

Figure 5. Shared gene expression dynamics associated with chemotherapy resistance and quiescence in diverse human AML patient specimens.

a. Schematic of comparative gene expression analyses of AML cells isolated from the bone marrow (BM) of patients before and after treatment with induction chemotherapy (left) on indicated days[37], and LRC quiescence of CFSE label retention in mouse orthografts (right).

b. Heatmaps showing specific enrichment of distinct gene sets between LRCs in MSK162, MSK011, and MSK165 specimens and chemotherapy-resistant cells in AML329 and AML707B specimens analyzed after induction chemotherapy treatment. Red to blue color gradient represents positive and negative normalized enrichment scores (NES), respectively.

c. Heatmaps of significance of similarity in gene expression between MSK162, MSK011, and MSK165 LRCs and chemotherapy-resistant cells in AML329 and AML707B patient specimens (numbers indicate hypergeometric test *p* values). White to red and white to blue color gradients represent positive and negative odds ratios, respectively.

d. Commonly dysregulated gene sets shared by both LRCs and residual leukemia cells after chemotherapy were listed. For example, TNF-alpha and STK33 signaling pathways were positively correlated, whereas MYC, RNA processing and mitochondrial biogenesis pathways were negatively correlated in both LRCs and residual leukemia cells after chemotherapy.

135 LRCs (Figure 5b-d). Among them, we identified TNF signaling and hypoxia response pathways for example, which have recently been nominated in AML stem cell pathogenesis[38–41], and multiple other pathways which provide important hypotheses for future studies (Figure 5b-d). LRCs and residual leukemia cells after chemotherapy also showed similarity to some, but not all, of the previously reported AML stem cell signatures identified by cell surface marker enrichment[42–
140 45], as also reported recently[38–40]. In all, these findings suggest that gene expression programs induced by LRC quiescence contribute to therapy resistance and leukemia persistence in patients.

AML LRC quiescence is controlled by a novel transcription factor network

145 Lastly, we sought to develop a functional genetic system to identify regulators of human patient LRC quiescence. To enable this, we designed a lentiviral vector for the production of ultrahigh virus titers necessary for the transduction of primary human leukemia cells, encoding mCherry fluorescent protein and doxycycline-inducible cDNAs marked by specific barcode sequences (Figure 6a, Extended Data Table 8). We selected 20 transcription factors with
150 differential expression in diverse LRCs with concordant markers of chromatin motif accessibility at LRC-regulated genes (Figures 3-4, Extended Data Table 9). We confirmed appropriate doxycycline-inducible cDNA expression by Western immunoblotting (Extended Data Figure 7a), prepared equimolar cDNA plasmid pools (Extended Data Figure 7b), transduced primary human MSK162 patient leukemia cells to achieve single-copy cDNA integration as assessed by mCherry
155 expression by FACS (<20% cells; Extended Data Figure 7c), and confirmed their stable representation of at least 5,000 cells/cDNA by quantitative DNA barcode sequencing after primary transplantation in NSG mice (10^6 cells/mouse) in the absence of doxycycline (Extended Data Figure 7d).

To identify regulators of human LRC quiescence, we CFSE-labeled MSK162 patient
160 leukemia cells genetically-modified with LRC regulator cDNA library, and transplanted them into

NSG mice treated with doxycycline to enforce specific cDNA expression in individual leukemia cells (Figure 6a and Extended Data Figure 7e). Analysis of abundance of cDNA barcodes in LRCs versus non-LRCs revealed positive and negative regulators of LRC quiescence, in contrast to control TagBFP which showed no significant effect (p = 0.54; Figure 6b and Extended Data Figure
165 7f). For example, enforced expression of *JUN* or *ZFP36L1* caused nearly 4-fold increase in LRCs (p = 1.4 x 10^{-3} and 2.9 x 10^{-3}, respectively), whereas enforced expression of *ETS1*, *FLI1* or *KLF2* suppressed LRC growth by more than 4-fold (p = 2.8 x 10^{-3}, 6.4 x 10^{-3}, and 2.1 x 10^{-3}, respectively; Figure 6b and Extended Data Figure 7f). Since many of the identified LRC regulators can interact physically, they may comprise an integrated LRC regulatory transcription factor network (Figure
170 6c).

Given evidence of transcriptional activity of *JUN* in LRCs and its induction of LRC quiescence suggest that *JUN* activity may be also associated with chemotherapy resistance and disease persistence in patients. To investigate this, we analyzed *JUN* expression in individual AML cells in diverse patients before and after chemotherapy treatment (Figure 6d, Extended Data
175 Table 7)[37]. We found *JUN* to be highly expressed in persistent AML cells upon induction chemotherapy, especially in leukemia cells with features of hematopoietic stem cells (Figure 6d). High *JUN* expression in HSC-like cells was also maintained in patients with monocytic AML (Figure 6d). These results argue that aberrant *JUN* activity may be required for chemotherapy resistance, LRC persistence, and ultimate AML relapse. In all, human AML quiescence and
180 therapy response can be dynamically regulated by a distinct transcription factor network.

Discussion

Cancer relapse occurs in the majority of patients after chemotherapy and frequently signifies incurable disease. In acute myeloid leukemia, a common blood cancer that affects both
185 children and adults, disease relapse is largely due to the persistence of leukemia initiating stem cells. Here, we implemented label tracing functional genomic techniques to human patient AML

Figure 6.

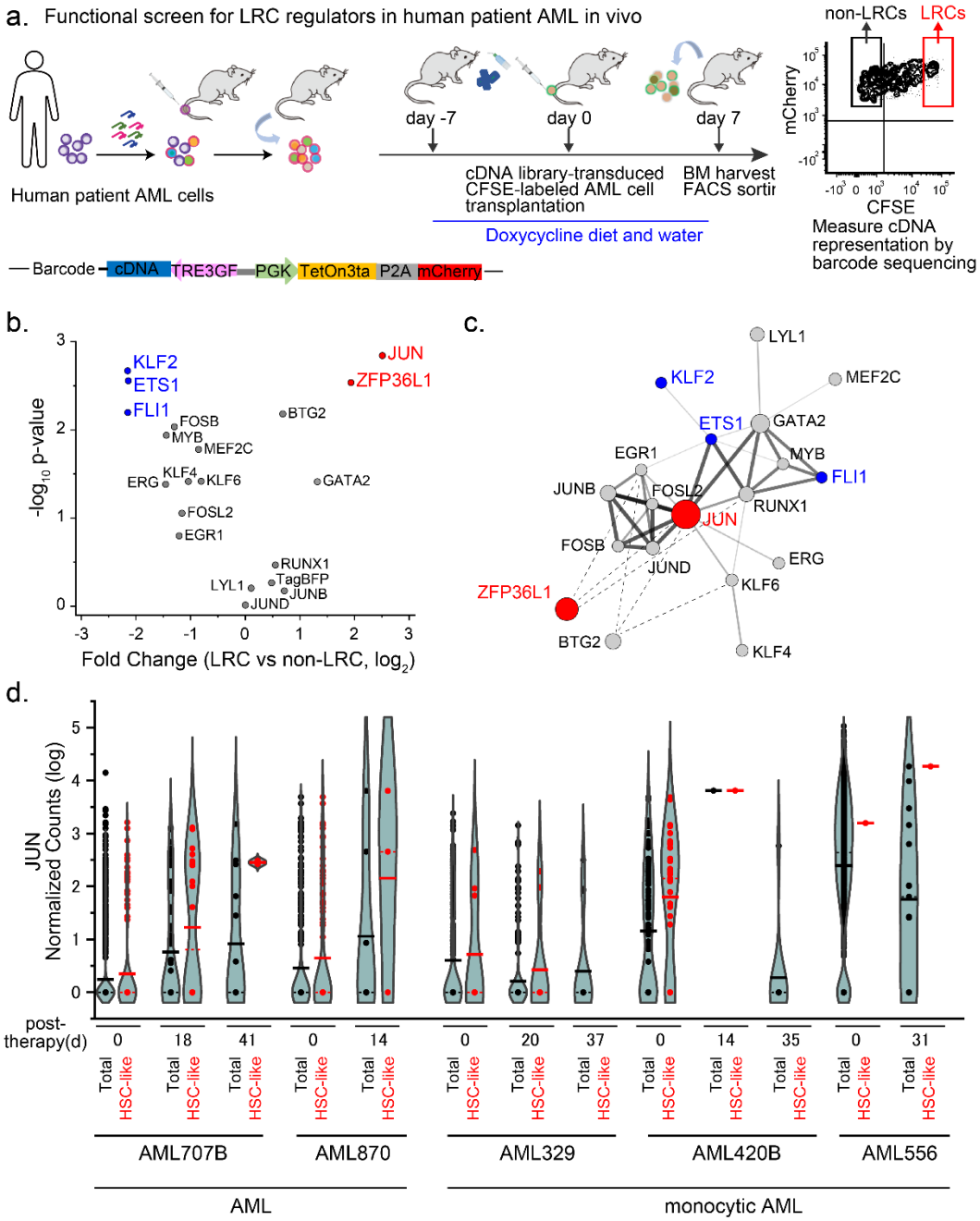

Figure 6. AML stem cell quiescence is controlled by a distinct transcription factor network, associated with chemotherapy resistance and disease persistence in patients.

A. Experimental design to identify regulatory factors controlling AML stem cell quiescence using lentiviral transduction of doxycycline-inducible cDNA library in primary human patient AML cells, followed by CFSE labeling to isolate quiescent LRCs (red) and expanding non-LRCs (black). Days of induction of cDNA expression using doxycycline treatment are indicated. Each lentiviral doxycycline-inducible cDNA vector includes a specific barcode sequence, enabling quantitative identification by DNA sequencing (bottom).

B. Volcano plot showing genes whose enforced expression induces (red) or depletes (blue) quiescent LRCs in MSK152 patient AML specimen. TagBFP serves as negative control.

C. Schematic of the LRC regulatory interaction network. Red and blue circles denote LRC activators repressors, respectively, with circle size proportional to effect size. Solid lines indicate physical interactions with line thickness corresponding to STRING confidence scores.

D. *JUN* mRNA expression level in total and HSC-like cells before and after chemotherapy treatment in diverse AML and monocytic AML patients, as measured using scRNA-seq[34]. Violin plots show normalized *JUN* read counts per cell including cells with zero reads. Bars and dashed lines represent mean and median values, respectively. Black and red denote total and HSC-like cells, respectively.

specimens. This unbiased approach enables molecular and functional analyses of diverse leukemia initiating and quiescent LRC stem cells that evade detection by currently known cell
190 surface markers, as also proposed by others[19,20,46].

Recent observations of altered protein homeostasis in AML and hematopoietic stem cells[47–49], as at least in part induced by dysregulation of mRNA translation[50], may also explain the specific ability of chemical protein label retention to identify functional leukemia initiating and disease persisting cells. Future studies will be important to define the molecular mechanisms of
195 LRC proteome quiescence.

Importantly, our findings explain the frequent and seemingly paradoxical observations how AML relapse in patients can occur both from rare leukemia stem cells as well as immunophenotypically committed or differentiated subclones[8,51]. Our findings also indicate that drug resistance is a consequence of dormancy[52]. Prior work by Bhatia and colleagues and Melnick
200 and colleagues have implicated chemotherapy, and cytarabine in particular, as an inducer of altered and senescent-like state responsible for AML relapse[38,39]. Our work indicates that many of the features of quiescence and dormancy observed in this state are already accessible to leukemia cells as LRCs. The similarity of gene expression profiles of LRCs, combined with the apparent kinetics of their growth upon transplantation in mice, suggest that this population exists
205 before onset of therapy and not as a consequence of treatment.

Do LRCs occupy different bone marrow niches from normal HSCs? Upregulation of cell adherent molecules or hypoxic response in LRCs suggest that LRCs may function in the context of a distinct bone marrow niche (Figure 4d), as previously observed for endosteal cells[53]. The interactions of LRCs with tissue niches might represent new therapeutic targets. For example, our
210 work helps to corroborate distinct metabolic requirements of leukemia stem cells[54], including the function of fatty acid oxidation in the promotion of LRC quiescence and therapy resistance, as evident from the high LRC expression of CD36 (Figure 4a, Extended Data Figure 6g and Extended Data Table 5). LRC-expressed molecules, such as CD36, which may be dispensable

for healthy hematopoietic stem and progenitor cells, represent compelling therapeutic targets to
215 selectively eradicate LRCs.

The organization of the LRC transcription factor network suggests that acquisition of LRC quiescence may be accomplished by the coordinated dysregulation of cell differentiation and dormancy. The apparently divergent functions of *JUN* and *ETS1* in controlling LRC quiescence, while both being required for normal hematopoietic stem and progenitor cell functions (Expanded
220 Data Figure 9), suggest that the LRC transcription factor network is aberrantly organized, as recently also observed for other aberrantly organized transcription factor complexes in genetically diverse human leukemias[55,56]. While our studies indicate that LRC quiescence is reversible, sparing genetic competition that maintains its epigenetic inheritance, clonal genetic evolution is an important part of leukemia development in patients, given its contributions to pre-leukemia
225 clonal hematopoiesis and leukemia therapy resistance[23].

Finally, our study is limited by transplantation in immunodeficient mice, with its distinct cytokine and cellular environments, which may incompletely model human physiology and thereby obscure additional mechanisms of human leukemia stem cells. We also cannot exclude the possibility of LRC genetic mutations that are not captured by target gene sequencing, even
230 though we targeted most known recurrently mutated genes in hematologic malignancies[22]. We anticipate that similar studies of pre-malignant and cancer cells will provide essential insights into the mechanisms of cancer initiation, evolution, and therapeutic targeting. This should lead to new therapeutic strategies aimed at restoring normal cell development and therapeutic control of cancer cell quiescence.

235

240 **References**

1. Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645–648 (1994).
2. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that
245 originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737 (1997).
3. Goardon, N. *et al.* Coexistence of LMPP-like and GMP-like Leukemia Stem Cells in Acute Myeloid Leukemia. *Cancer Cell* **19**, 138–152 (2011).
4. Sarry, J.-E. *et al.* Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2Rγc-deficient mice. *J. Clin. Invest.* **121**,
250 384–395 (2011).
5. Bewersdorf, J. P. & Abdel-Wahab, O. Translating recent advances in the pathogenesis of acute myeloid leukemia to the clinic. *Genes Dev.* **36**, 259–277 (2022).
6. Shallis, R. M., Wang, R., Davidoff, A., Ma, X. & Zeidan, A. M. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev.* **36**, 70–87 (2019).
255 7. Papaemmanuil, E. *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).
8. Fennell, K. A. *et al.* Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature* **601**, 125–131 (2022).
9. Iacobucci, I. *et al.* Modeling and targeting of erythroleukemia by hematopoietic genome
260 editing. *Blood* **137**, 1628–1640 (2021).
10. Tothova, Z. *et al.* Multiplex CRISPR/Cas9-Based Genome Editing in Human Hematopoietic Stem Cells Models Clonal Hematopoiesis and Myeloid Neoplasia. *Cell Stem Cell* **21**, 547-555.e8 (2017).
11. You, X. *et al.* Asxl1 loss cooperates with oncogenic Nras in mice to reprogram the immune
265 microenvironment and drive leukemic transformation. *Blood* **139**, 1066–1079 (2022).
12. Díaz de la Guardia, R. *et al.* Engraftment characterization of risk-stratified AML in NSGS mice. *Blood Adv.* **5**, 4842–4854 (2021).
13. Woiterski, J. *et al.* Engraftment of low numbers of pediatric acute lymphoid and myeloid leukemias into NOD/SCID/IL2Rcγnull mice reflects individual leukemogenecity and highly
270 correlates with clinical outcome. *Int. J. Cancer* **133**, 1547–1556 (2013).
14. Griessinger, E. *et al.* Frequency and Dynamics of Leukemia-Initiating Cells during Short-term Ex Vivo Culture Informs Outcomes in Acute Myeloid Leukemia Patients. *Cancer Res.* **76**, 2082–2086 (2016).
15. Foudi, A. *et al.* Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic
275 stem cells. *Nat. Biotechnol.* **27**, 84–90 (2009).
16. Qiu, J., Papatsenko, D., Niu, X., Schaniel, C. & Moore, K. Divisional history and hematopoietic stem cell function during homeostasis. *Stem Cell Rep.* **2**, 473–490 (2014).
17. Takizawa, H., Regoes, R. R., Boddupalli, C. S., Bonhoeffer, S. & Manz, M. G. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J. Exp.*
280 *Med.* **208**, 273–284 (2011).
18. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
19. Ebinger, S. *et al.* Characterization of Rare, Dormant, and Therapy-Resistant Cells in Acute Lymphoblastic Leukemia. *Cancer Cell* (2016) doi:10.1016/j.ccell.2016.11.002.
285 20. Ebinger, S. *et al.* Plasticity in growth behavior of patients' acute myeloid leukemia stem cells growing in mice. *Haematologica* **105**, 2855–2860 (2020).
21. Wunderlich, M. *et al.* AML cells are differentially sensitive to chemotherapy treatment in a human xenograft model. *Blood* **121**, e90–e97 (2013).
22. Grommes, C. *et al.* Ibrutinib Unmasks Critical Role of Bruton Tyrosine Kinase in Primary
290 CNS Lymphoma. *Cancer Discov.* **7**, 1018–1029 (2017).

23. Miles, L. A. *et al.* Single-cell mutation analysis of clonal evolution in myeloid malignancies. *Nature* **587**, 477–482 (2020).
24. Sun, Q.-Y. *et al.* Ordering of mutations in acute myeloid leukemia with partial tandem duplication of MLL (MLL-PTD). *Leukemia* **31**, 1–10 (2017).
25. Schmalbrock, L. K. *et al.* Clonal evolution of acute myeloid leukemia with FLT3-ITD mutation under treatment with midostaurin. *Blood* **137**, 3093–3104 (2021).
26. Wilson, N. K. *et al.* Combinatorial Transcriptional Control In Blood Stem/Progenitor Cells: Genome-wide Analysis of Ten Major Transcriptional Regulators. *Cell Stem Cell* **7**, 532–544 (2010).
27. Diffner, E. *et al.* Activity of a heptad of transcription factors is associated with stem cell programs and clinical outcome in acute myeloid leukemia. *Blood* **121**, 2289–2300 (2013).
28. Passegué, E., Wagner, E. F. & Weissman, I. L. JunB Deficiency Leads to a Myeloproliferative Disorder Arising from Hematopoietic Stem Cells. *Cell* **119**, 431–443 (2004).
29. Santaguida, M. *et al.* JunB Protects against Myeloid Malignancies by Limiting Hematopoietic Stem Cell Proliferation and Differentiation without Affecting Self-Renewal. *Cancer Cell* **15**, 341–352 (2009).
30. Park, C. S. *et al.* A KLF4-DYRK2–mediated pathway regulating self-renewal in CML stem cells. *Blood* **134**, 1960–1972 (2019).
31. Bungartz, G., Land, H., Scadden, D. T. & Emerson, S. G. NF-Y is necessary for hematopoietic stem cell proliferation and survival. *Blood* **119**, 1380–1389 (2012).
32. Wang, E. *et al.* Surface antigen-guided CRISPR screens identify regulators of myeloid leukemia differentiation. *Cell Stem Cell* **28**, 718-731.e6 (2021).
33. Cabezas-Wallscheid, N. *et al.* Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell* **169**, 807-823.e19 (2017).
34. Wilson, A. *et al.* c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev.* **18**, 2747–2763 (2004).
35. García-Prat, L. *et al.* TFEB-mediated endolysosomal activity controls human hematopoietic stem cell fate. *Cell Stem Cell* **28**, 1838-1850.e10 (2021).
36. Signer, R. A. J., Magee, J. A., Salic, A. & Morrison, S. J. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature* **509**, 49–54 (2014).
37. Galen, P. van *et al.* Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. *Cell* **176**, 1265-1281.e24 (2019).
38. Duy, C. *et al.* Chemotherapy Induces Senescence-Like Resilient Cells Capable of Initiating AML Recurrence. *Cancer Discov.* **11**, 1542–1561 (2021).
39. Boyd, A. L. *et al.* Identification of Chemotherapy-Induced Leukemic-Regenerating Cells Reveals a Transient Vulnerability of Human AML Recurrence. *Cancer Cell* **34**, 483-498.e5 (2018).
40. Farge, T. *et al.* Chemotherapy-Resistant Human Acute Myeloid Leukemia Cells Are Not Enriched for Leukemic Stem Cells but Require Oxidative Metabolism. *Cancer Discov.* **7**, 716–735 (2017).
41. Morris, V. *et al.* Hypoxic, glycolytic metabolism is a vulnerability of B-acute lymphoblastic leukemia-initiating cells. *Cell Rep.* **39**, (2022).
42. Eppert, K. *et al.* Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* **17**, 1086–1093 (2011).
43. Ng, S. W. K. *et al.* A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437 (2016).
44. Gentles, A. J., Plevritis, S. K., Majeti, R. & Alizadeh, A. A. Association of a Leukemic Stem Cell Gene Expression Signature With Clinical Outcomes in Acute Myeloid Leukemia. *JAMA* **304**, 2706–2715 (2010).

45. Gal, H. *et al.* Gene expression profiles of AML derived stem cells; similarity to hematopoietic stem cells. *Leukemia* **20**, 2147–2154 (2006).
46. Xie, X. P. *et al.* Quiescent human glioblastoma cancer stem cells drive tumor initiation, expansion, and recurrence following chemotherapy. *Dev. Cell* **57**, 32-46.e8 (2022).
47. Signer, R. A. J. *et al.* The rate of protein synthesis in hematopoietic stem cells is limited partly by 4E-BPs. *Genes Dev.* **30**, 1698–1703 (2016).
48. Goncalves, K. A. *et al.* Angiogenin Promotes Hematopoietic Regeneration by Dichotomously Regulating Quiescence of Stem and Progenitor Cells. *Cell* **166**, 894–906 (2016).
49. Jose, L. H. S. *et al.* Modest Declines in Proteome Quality Impair Hematopoietic Stem Cell Self-Renewal. *Cell Rep.* **30**, 69-80.e6 (2020).
50. Galen, P. van *et al.* Integrated Stress Response Activity Marks Stem Cells in Normal Hematopoiesis and Leukemia. *Cell Rep.* **25**, 1109-1117.e5 (2018).
51. Shlush, L. I. *et al.* Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature* **547**, 104–108 (2017).
52. Blatter, S. & Rottenberg, S. Minimal residual disease in cancer therapy – Small things make all the difference. *Drug Resist. Updat.* **21–22**, 1–10 (2015).
53. Ishikawa, F. *et al.* Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. *Nat. Biotechnol.* **25**, 1315–1321 (2007).
54. Jones, C. L., Inguva, A. & Jordan, C. T. Targeting Energy Metabolism in Cancer Stem Cells: Progress and Challenges in Leukemia and Solid Tumors. *Cell Stem Cell* **28**, 378–393 (2021).
55. Takao, S. *et al.* Convergent organization of aberrant MYB complex controls oncogenic gene expression in acute myeloid leukemia. *eLife* **10**, e65905 (2021).
56. Smeenk, L. *et al.* Selective Requirement of MYB for Oncogenic Hyperactivation of a Translocated Enhancer in Leukemia. *Cancer Discov.* **11**, 2868–2883 (2021).
57. Cheng, D. T. *et al.* Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J. Mol. Diagn.* **17**, 251–264 (2015).
58. Muench, D. E. *et al.* Mouse models of neutropenia reveal progenitor-stage-specific defects. *Nature* **582**, 109–114 (2020).
59. Shen L, Sinai ISoMaM (2022). GeneOverlap: Test and visualize gene overlaps. R package version 1.32.0, http://shenlab-sinai.github.io/shenlab-sinai/.

Methods

Human AML specimens

395 Primary AML specimens were obtained from the bone marrow of patients upon obtaining written informed consent and approval by the Institutional Review Board of Memorial Sloan Kettering Cancer Center, MD Anderson Cancer Center and the Children's Oncology Group in accordance with the Declaration of Helsinki.

CFSE-labeling

400 5(6)-Carboxyfluorescein diacetate N-hydroxysuccinimidyl ester (CFSE, Abcam) was dissolved in DMSO at 5 mM and stored at -20 °C. To prepare CFSE-labeled cells, cells were washed in PBS once and incubated with 1 μM CFSE in PBS supplemented with 1% FBS at 37 °C for 5 minutes, unless otherwise indicated. After the reaction was quenched by adding15 ml of Iscove's Modified
405 Dulbecco's Medium (IMDM) supplemented with 15% FBS, cells were washed twice with IMDM supplemented with 15% FBS using centrifugation.

Detection of fluorescently-labeled proteins

Cells were lysed in RIPA buffer (140 mM NaCl, 0.4% SDS, 0.1% sodium deoxycholate, 1% Triton
410 X-100 and 20 mM Tris-HCl) with sonication using the Covaris S220 instrument (Covaris). Lysates were cleared by centrifugation at 14,000 x g for 15 minutes at 4 °C, and the collected lysates were denatured in Laemmli sample buffer supplemented with 50 mM DTT at 95 °C for 5 minutes. Lysates were separated by sodium dodecyl sulfate polyacrylamide gels electrophoresis (SDS-PAGE, Novex) at 120 V for 12 minutes. Gels were fixed in 45% methanol and 10% acetic acid in
415 water and imaged using Typhoon laser-scanning fluorescence imager (Cytiva).

Patient-derived AML mouse xenografts

All mouse studies were conducted with approval from the Memorial Sloan Kettering Cancer Center Institutional Animal Care and Use Committee. NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG,
420 Jackson Laboratory) mice were used for transplantation of primary human patient AML cells. NSG mice were irradiated with 200 cGy and transplanted with 1,000,000 cells per mouse via tail vein injection. For analysis, bone marrow cells were isolated by dissecting and crushing femoral and humeral bones using mortar and pestle in PBS supplemented with 2.5% FBS, and isolated by filtration through 70-μm mesh, followed by lysis of red cells using the RBC Lysis Buffer
425 (BioLegend). In the case of subsequently performing fluorescence-activated cell sorting (FACS), bone marrow mononuclear cells were purified by density gradient centrifugation using Ficoll-Hypaque Plus, according to manufacturer's instructions (GE Healthcare). For serial transplantation, purified cells were incubated overnight in StemSpan (STEMCELL Technologies) supplemented with 100 ng/μl each of human SCF, FLT3 ligand, and TPO (PeproTech) at 37 °C
430 with 5% CO_2. Subsequently, cells were transplanted into 200 cGy-irradiated NSG mice with 1,000 cells per mouse via tail vein injection. Transplanted mice were fed with Sulfatrim-supplemented chow.

Mouse bone marrow transplantation

435 C57BL/6J mice were used as donors (The Jackson Laboratory). Upon dissection and crushing of femoral and humeral bones, bone marrow was filtered through 70-μm mesh, and red cells were lysed using the RBC Lysis Buffer (BioLegend). Hematopoietic stem and progenitor cells were isolated using magnetic purification with the mouse Lineage Cell Depletion Kit, according to the manufacturer's instructions (Miltenyi Biotec). Recipient C57BL/6J mice were irradiated with 900
440 cGy and transplanted with isolated bone marrow hematopoietic stem and progenitor cells with or without CFSE-labeling using 500,000 cells per mouse via tail vein injection. Transplanted mice were fed with Sulfatrim-supplemented chow.

Fluorescence-activated cell sorting

445 Isolated bone marrow cells were purified by density gradient centrifugation using Ficoll-Hypaque Plus, according to manufacturer's instructions (GE Healthcare). Purified mononuclear cells were suspended in PBS with supplemented with 10 μg/ml mouse gamma globulin (Jackson ImmunoResearch) and incubated on ice for 20 minutes. Subsequently, cells were resuspended in MACS buffer (PBS containing 2 mM ETDA and 2.5% FBS) supplemented with human FcR 450 Blocking Reagent (Miltenyi Biotec) at 1:5 dilution and anti-human CD45-APC (BioLegend 304012) at 1:10 dilution, and incubated on ice for 20 minutes. After washing 2 times in MACS buffer, cells were resuspended in MACS buffer containing 1 μM SYTOX Blue (Invitrogen) and processed by fluorescence-activated cell sorting (FACS) using LSRFortessa FACS analyzer or FACSAria II cell sorter (BD Biosciences). Unless otherwise indicated, cells were gated by selecting SYTOX Blue- 455 negative and human CD45-positive cells. For multi-color staining for cell surface markers, the following antibodies were used additionally: PerCP/Cy5.5 anti-human CD34 (BioLegend 343612), PE/Cy7 anti-human CD38 (BIoLegend 303516), APC/Cy7 human-CD45RA (BioLegend 304128), PE anti-human CD90 (BioLegend 328110), PE anti-human CD123 (BioLegend 306006), PE anti-human CD117 (BioLegend 313204), and APC/Cy7 anti-human CD244 (BioLegend 329518). Raw 460 data were processed and quantified using FCS Express 7 (De Novo).

Cell cycle profiling in vivo

5-ethynyl-2'-deoxyuridine (EdU) was obtained from baseclick (Munich, Germany), and dissolved in PBS at 5 mg/ml and stored at -20 °C. Twenty four hours before analysis, mice were treated 465 with 50 mg/kg EdU in PBS via intraperitoneal injection. Upon isolation of bone marrow mononuclear cells as described above, cells were fluorescently labeled using EdU 647 Kit, according to the manufacturer's instructions with the following modifications (baseclick). Cells were stained with Fixable Viability Dye eFluor 780 (eBiocience) in PBS on ice for 30 minutes. After washing cells with PBS supplemented with 1% BSA, cells were fixed using 4% 470 paraformaldehyde in PBS, and incubated at room temperature for 15 minutes. After washing with PBS supplemented with 1% BSA, cells were permeabilized using Saponin-based permeabilization buffer at room temperature for 20 minutes. Fluorophore labeling was performed using Eterneon-Red 645 azide and catalyst solution for 30 minutes at room temperature, according to manufacturer's instructions (baseclick). Labeled cells were subsequently stained 475 with anti-human CD45-PE (BioLegend 304008) or anti-human CD45-APC (BioLegend 304012) and anti-cleaved caspase-3-PE (BD) at 4 °C for 20 minutes.

Chemotherapy treatment

For mouse studies, cytosine β-D-arabinofuranoside hydrochloride (AraC, Sigma) and doxorubicin 480 hydrochloride (DXR, Sigma) were dissolved in PBS at 20 mg/ml and 0.6 mg/ml, respectively, and used the same day. Mice were treated with 100 mg/kg AraC in PBS daily for 5 days and 3 mg/kg DXR in PBS daily for 3 days via intraperitoneal injection. For cell culture experiments, AraC and DXR were dissolved at 10 mM in PBS and DMSO, respectively.

485 **Genomic DNA sequencing analysis**

DNA isolation, sequencing library preparation, and paired-end sequencing were performed using HiSeq and NovaSeq (Illumina) as described previously[57]. Hybridization capture and variant allele analysis were performed using the MSK-HemePACT panel of 585 genes recurrently mutated in hematological malignancies, as previously described[22].

490

Chromatin accessibility analysis

Purified cells (3,000 cells per sample) were lysed by incubation in 10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP-40 for 2 minutes at 4 °C, followed by sedimentation at 1500 g to

isolate nuclei. Tagmentation was performed using Nextera DNA sample prep kit (Illumina) at 37 °C for 30 minutes and subsequently stopped by addition of SDS to a final concentration of 0.2%. Tagmented DNA was purified using Agencourt AMPure XP beads (Beckman Coulter), and barcoded libraries were generated using the NEBNext Q5 Hot Start HiFi PCR Master Mix (New England Biolabs) and Nextera index primers (Illumina). Paired-end 50-bp sequencing (50 million reads per sample) was performed using HiSeq (Illumina). Sequencing reads were filtered for Q>15 and trimmed of adapter sequences using TrimGalore (v0.4.5), and aligned to hg19 using bowtie2 (v2.2.2). Peak calling was performed using MACS2 and filtered for blacklisted regions (http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/). Signal was sequencing depth normalized and motif signatures were identified using Homer ('findMotifsGenome.pl').

RNA sequencing

RNA from 5,000 cells per sample was extracted using the Quick-RNA MicroPrep kit, according to manufacturer's instructions (Zymo Research). Barcoded libraries were constructed using QuantSeq 3'mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen) with ERCC RNA spike-in Mix, according to the manufacturer's instructions (Thermo Fisher). Single-end 50-bp sequencing was performed using HiSeq (Illumina) with 40 million reads per sample. Sequencing reads were filtered for Q>15 and adapter trimmed using TrimGalore (v0.4.5) before aligning to human assembly hg19 with STAR v2.5 using the default parameters. The raw counts matrix was built using HTSeq v0.6.1 and normalization and differential gene expression were performed using DESeq2 with default parameters.

Single cell RNA-seq analysis

Raw read counts were downloaded from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/, accession GSE116256) together with metadata, cell classification and clustering scores, as published[37]. Total read counts per cell were normalized to 10,000 reads per cell. Normalized data were log-transformed using Seurat. Differential gene expression was performed using DESeq2 with the default parameters.

Gene set enrichment analysis

Gene set enrichment analysis was performed using GSEAPreranked v4.4.0 with default parameters using DESeq2 output and MSigDB gene set version 6.0 (https://www.gsea-msigdb.org/gsea/msigdb/).

Lentiviral cDNA library preparation

Doxycycline-inducible cDNA-expressing barcoded lentiviral plasmids (V191, Extended Data Table 8) were designed and synthesized by Transomic Technologies (Huntsville, Alabama). (One was synthesized by Custom DNA Constructs, Islandia, NY). Individual plasmids were sequence verified and quantified, and then pooled for lentivirus production. In total, 1.8 billion HEK293T cells grown in CellStack-5 (Corning) (30 million cells per CellStack-5) were transfected with pooled plasmids (600 μg per CellStack-5) and packaging plasmids, pMD2.G, and psPAX2 (300 μg and 300 μg per CellStack-5, respectively), using PEIpro according to the manufacturer's instructions (Polyplus). Collected virus supernatant (7.5 liters) was 0.45-μm mesh filtered, and concentrated using Lenti-X concentrators, according to manufacturer's instructions (Takara). Lentivirus titers were determined by infection of HEK293T cells as biological infectious units.

Lentiviral transduction of primary human leukemia cells

Cells were infected with lentivirus preparations at multiplicity of infection of less than 0.2 by spinoculation at 800 g for 90 min at room temperature in StemSpan media supplemented with 100 ng/μl each of human SCF, FLT3 ligand and TPO and 12.5 μl of LentiBOOST (SIRION Biotech) per well of 12-well plates, followed by incubation overnight at 37°C with 5% CO_2. Upon

545 replacement of media with fresh StemSpan supplemented with 100 ng/μl each of human SCF, FLT3 ligand and TPO, cells were cultured for 3 days. Transduced cells were isolated using FAC sorting by gating on SYTOX Blue-negative, human CD45- and mCherry-positive cells, and transplanted as described above.

550 **Inducible cDNA expression in human patient leukemia cells in vivo**

Recipient NSG mice were fed with 625 mg/kg doxycycline hyclate chow (Envigo) and water containing 2 mg/ml doxycycline hyclate (Sigma) for one week before transplantation, followed by the same treatment upon transplantation. Transduced cells (1 million cells per mouse) were transplanted into NSG mice irradiated with 200 cGy via tail vein injection. Each experiment used
555 45 NSG mice divided into three groups of 15 mice as biological triplicates. Bone marrow cells were isolated and pooled from 15 mice per group and purified using FAC sorting as described above. Genomic DNA was extracted from each sample using Quick-DNA Microprep Plus kit, according to manufacturer's instructions (Zymo Research). Plasmid cDNA barcode sequences were amplified using primers containing indexed Illumina sequencing adaptors (Extended Data
560 Table 10) using KOD Hot Start DNA Polymerase (Novagen). Amplicons were purified using Agencourt AMPure XP beads (Beckman Coulter). Single-end 50-bp sequencing was performed using NovaSeq (Illumina) with at least 10 million reads per sample. To quantify the read counts for each library sequence, a custom script was written based on the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Each samples fastq file was processed as follows: The
565 3' adapters were clipped and then reads were filtered to have a q value of 20 or more for all bases. The results sequences were then compared to the library sequence file to retain only sequences that match the library. Finally, the matching sequences were counted to give a count per library sequence. Total read counts were normalized to 10 million reads per sample for comparison.

570 **Western blot analysis**

OCI-AML3 cells transduced with cDNA library were lysed in Laemmli sample buffer (BioRad) supplemented with cOmplete protease inhibitors (Roche) and 100 mM DTT (BioRad) at a ratio of 100 µL sample buffer per 1 million cells. Cell suspensions were incubated at 95 °C for 7 min, and lysates were clarified by centrifugation for 5 min at 1500 rpm. Fourteen µL of clarified lysates were
575 resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 10% polyacrylamide Bis-Tris gels (Invitrogen) and transferred onto Immobilon FL PVDF membranes (Millipore, Billerica, MA, USA) at 30V for 90 min at 4 °C. Membranes were blocked using Odyssey Blocking buffer (Li-Cor, Lincoln, Nebraska, USA). Blots were incubated with primary antibodies for JUN (1:1000, CST60A8), ETS1 (1:1000, CST808A), or beta-actin (1:5000, CST3700). Blotted
580 membranes were visualized using secondary antibodies conjugated to IRDye 800CW (goat anti-mouse IgG, 1:15,000) or IRDye 680RD (goat anti-rabbit IgG, 1:15,000) and the Odyssey CLx fluorescence scanner (Li-Cor), according to manufacturer's instructions (Li-Cor, Lincoln, Nebraska, USA). After visualization, signal intensity of the bands of interest was quantified using the Image Studio Lite (Li-Cor).

585

Cell lines

Human AML cell lines, OCI-AML2 and OCI-AML3 were obtained from DSMZ and cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum 100 U/ml penicillin and 100 µg/ml streptomycin in a humidified atmosphere at 37 °C and 5% CO2. HEK293T cells were obtained
590 from ATCC and cultured in Dulbecco's Modified Eagle medium (DMEM) supplemented with 10% FBS and 100 U/ml penicillin and 100 µg/ml streptomycin. All cell lines were authenticated by STR genotyping (Integrated Genomics Operation, Center for Molecular Oncology, MSKCC). The absence of Mycoplasma species contamination was verified using MycoAlert Mycoplasma detection kit, according to manufacturer's instruction (Lonza).

595

Statistical methods

Statistical analyses were performed using OriginPro 2018 and 2022 (Microcal). Survival analyses were performed using Kaplan-Meier log-rank test. Statistical significance values were determined using two-tailed Welch's t-tests for continuous variables, and two-tailed Fisher exact tests for discrete variables. Confidence intervals were calculated using OriginPro 2018. Significance of gene set comparisons was assessed using hypergeometric tests as implemented in GeneOverlap v4.1[59]. Network diagrams were constructed using Cytoscape v3.9.1.

Data and material availability

All data are available openly via Zenodo, with raw RNA-seq and ATAC-seq data available via NCBI Gene Expression Omnibus. All plasmids are available via Addgene and Transomic.

Code Availability

All computational code is available via Zenodo.

Acknowledgements. We thank Michael Kharas, Alejandro Gutierrez, Peter van Galen, and Marc Mansour for critical discussions. We thank the MSKCC Integrated Genomics and Bioinformatics Cores, Flow Cytometry Core Facility, Antitumor Assessment Core Facility, Gene Editing and Screening Core, and Hematologic Oncology Tissue Bank, which are funded by MSKCC Support Grant NIH P30 CA008748. AK is a Scholar of the Leukemia & Lymphoma Society and acknowledges support from NIH R01 CA204396, R21 CA235285, Starr Cancer Consortium, St. Baldrick's Foundation, Burroughs Wellcome Fund, Pershing Square Sohn Cancer Research Alliance, Mathers Foundation, Damon Runyon-Richard Lumsden Foundation, Mr. William H. and Mrs. Alice Goodwin and the Commonwealth Foundation for Cancer Research and the Center for Experimental Therapeutics at MSKCC.

Author Contributions. AK and ST conceptualized the study; AK and ST designed and optimized experimental methods; ST, FB and VM performed experiments; ST and RK analyzed data; AK and ST wrote the original manuscript, which was edited by all authors.

Competing Interests. The authors have no conflicts of interest to declare. AK is a consultant to Rgenta, Novartis, and Blueprint Medicines.