

Supplementary Materials for

Targeting ZMIZ1 induces differentiation of acute myeloid leukemia via disruption of nuclear condensates and chromatin remodeling

Huimin Li^{1, *}, Yabin Liu^{1, *, #}, Jiachen Cui^{1, 2, *}, Ziyu Fu^{1, *}, Yiyang Xu^{1, *}, Sizhe Yang^{1, 3, *}, Kai Zhu⁴, Yuhao Xue¹, Yi Zhang¹, Yiheng Zhao³, Guiqiyang Xiang¹, Xiaoguang Xu¹, Fulun Jiao¹, Shuting Yu¹, Qiang Wang¹, Yuting Dai¹, Qianqian Zhang¹, Jianfeng Li¹, Ruihong Zhang¹, Fangyi Dong¹, Shanhe Yu¹, Shengyue Wang¹, Saijuan Chen¹, Bo Jiao^{1, #}, Shuo Chen^{1, 2, 5, #}, Yun Tan^{1, 2, #}

Correspondence to: Yun Tan (ty12260@rjh.com.cn), Shuo Chen (shuochen@sjtu.edu.cn), Bo Jiao (bjiao@sibs.ac.cn), and Yabin Liu (lyb13246@rjh.com.cn)

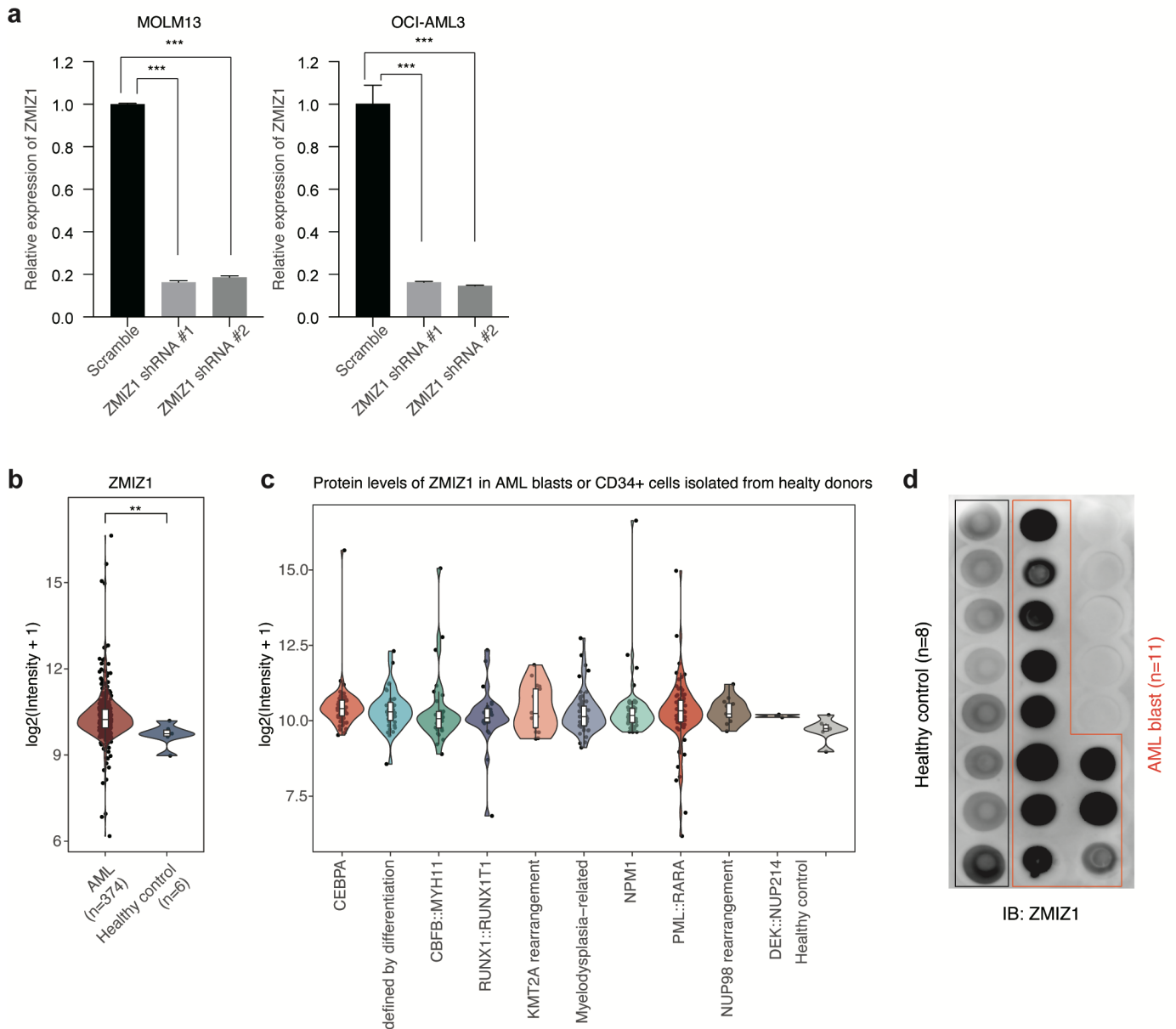
This PDF file includes:

Figures. S1 to S13
Original films of immunoblots

Other Supplementary Materials for this manuscript include the following:

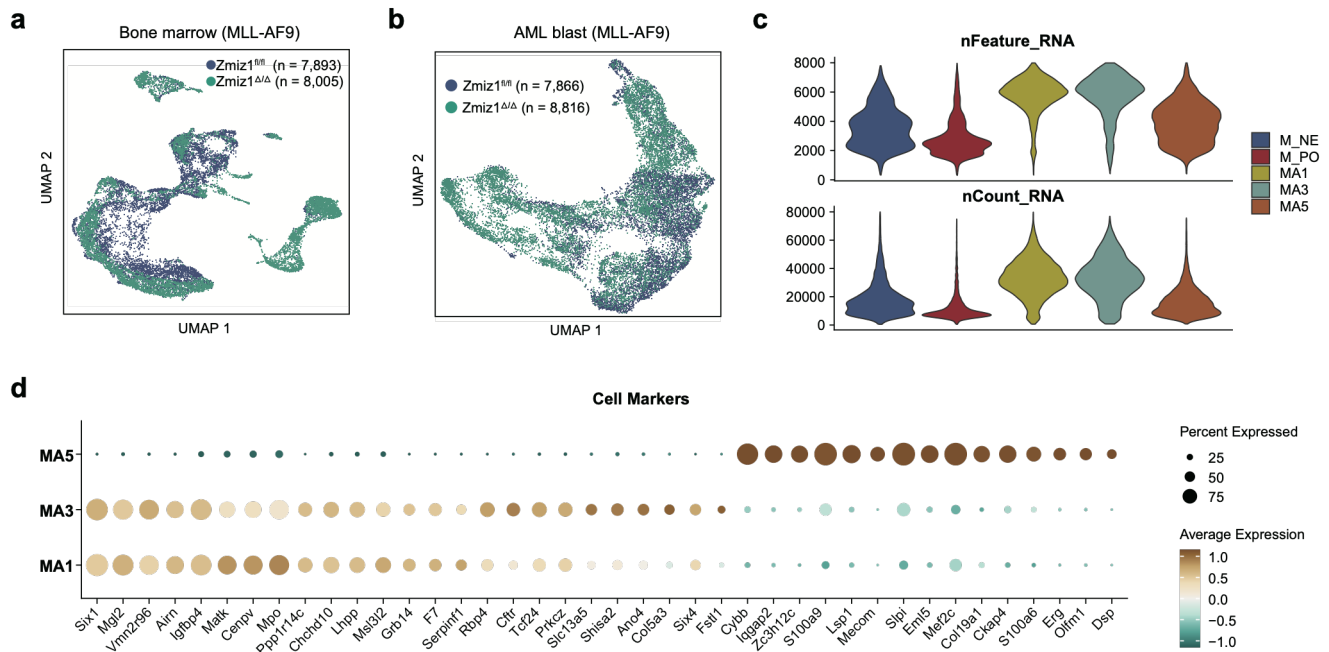
Supplementary Tables S1 to S19

Figure. S1.



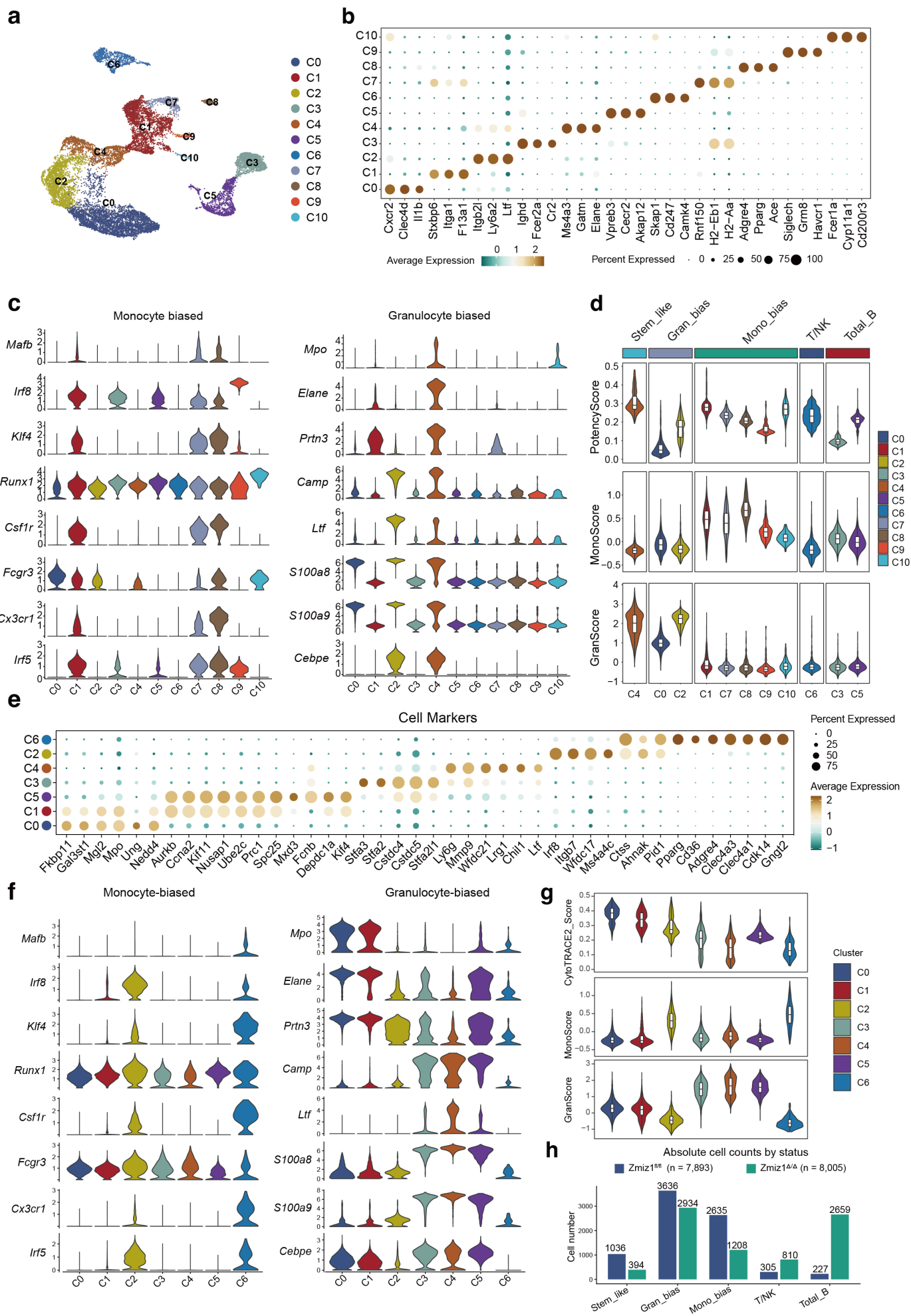
Supplementary Fig. 1. ZMIZ1 is highly expressed in AML. **a**, Efficient knockdown of ZMIZ1 by shRNA in leukemic cells. Quantitative analysis of ZMIZ1 expression following lentiviral transduction with two independent ZMIZ1-targeting short hairpin RNAs (shRNA #1 and shRNA #2), compared with a scrambled control. Relative ZMIZ1 expression levels were normalized to the scramble group. Data are presented as mean $\pm$ s.d. from biological replicates. Statistical significance was determined using a two-tailed Student's t-test; *** $P < 0.001$. **b**, Violin plot showing the ZMIZ1 protein levels in AML patients ($n = 374$) compared to healthy controls ($n = 6$). Data are represented as $\log_2(\text{intensity} + 1)$, with a significant difference indicated by ** ($p < 0.01$). **c**, Protein levels of ZMIZ1 in blast of AML patients with different genetic alterations or CD34+ cells isolated from healthy donors. The various groups include CEBPA, CBFB::MYH11, RUNX1::RUNX1T1, KMT2A rearrangement, Myelodysplasia-related, NPM1, PML::RARA, NUP98 rearrangement, and DEK::NUP214. Data are presented as $\log_2(\text{intensity}+1)$. **d**, Representative immunoblot (IB) showing ZMIZ1 protein expression in AML blasts ($n = 11$) compared to healthy controls ($n = 8$).

Figure. S2.



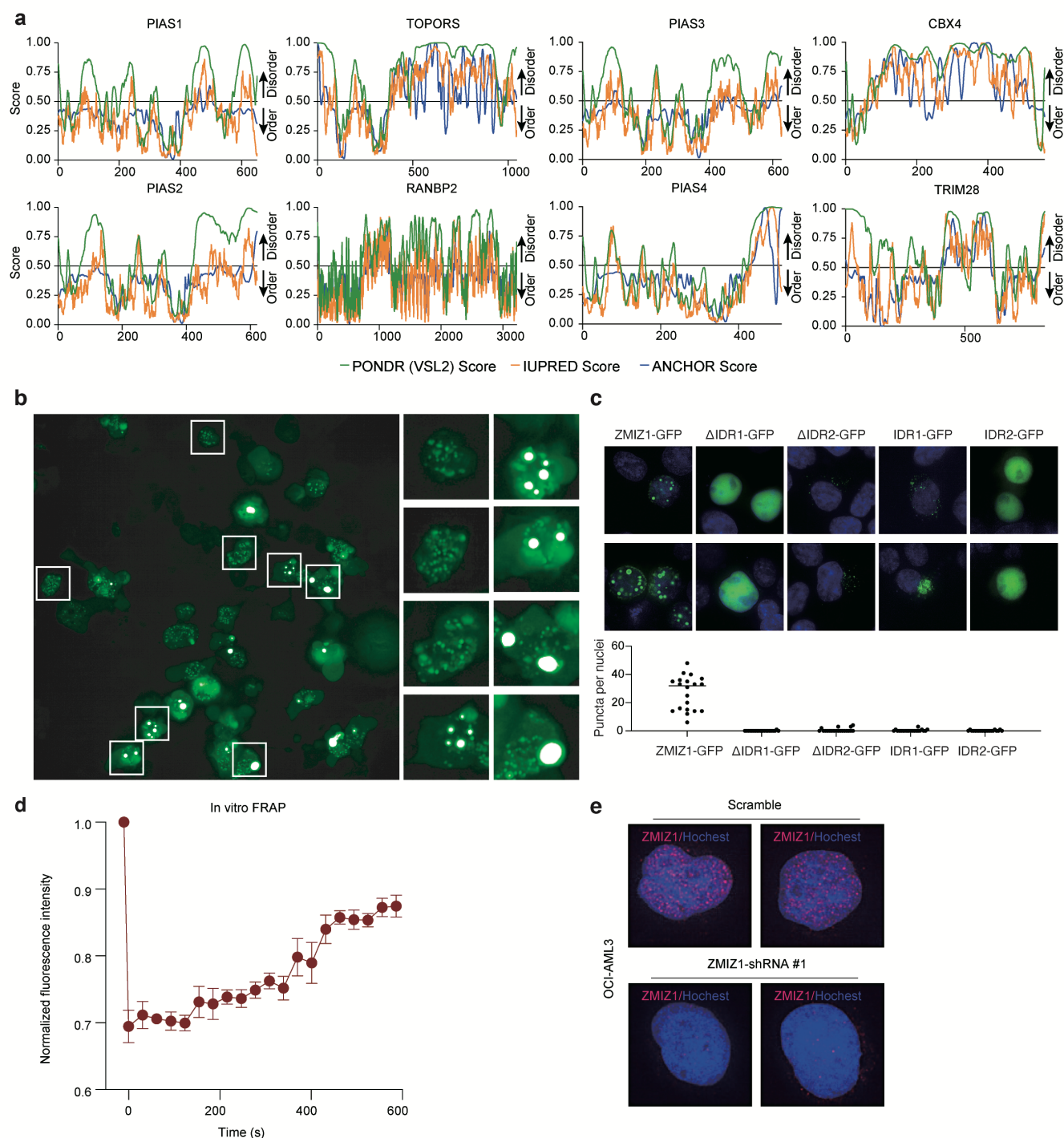
Supplementary Fig. 2. Quality control of scRNA-seq data. **a-b**, UMAP plots visualizing the integration and absolute cell numbers of scRNA-seq libraries from two independent batches. **a**, Whole bone marrow samples from Zmiz1^{fl/fl} (M_NE, n = 7,893) and Zmiz1^{Δ/Δ} (M_PO, n = 8,005) mice. **b**, Blast-enriched samples from Zmiz1^{fl/fl} (MA1 + MA3, n = 7,866) and Zmiz1^{Δ/Δ} (MA5, n = 8,816) mice. **c**, Violin plots showing distributions of sequencing/feature metrics across indicated samples or conditions, including nFeature_RNA (top) and nCount_RNA (bottom). **d**, Dot plot showing the expression levels of top marker genes across the blast-enriched biological replicates (MA1, MA3, MA5). Dot size represents the percentage of cells expressing the gene, and color intensity represents average expression. Note the highly consistent expression patterns between the two control replicates (MA1 and MA3), demonstrating robust biological reproducibility.

Figure. S3.



Supplementary Fig. 3. Single-cell transcriptomic analysis delineates monocyte- and granulocyte-biased populations in AML. a-d, scRNA-seq analysis of bone marrow cells in *Zmiz1*^{fl/fl} and *Zmiz1*^{Δ/Δ} mice. **a**, UMAP projection of bone marrow-derived single-cell transcriptomes, showing clustering of cells into 11 transcriptionally distinct populations (C0–C10). **b**, Dot plot showing expression of representative marker genes across clusters. Dot size indicates the proportion of cells expressing each gene, and color intensity reflects average expression level. **c**, Violin plots of selected monocyte-biased (e.g., *Mafb*, *Irf8*, *Klf4*, *Runx1*, *Csf1r*, *Fcgr3*, *Cx3cr1*, *Irf5*) and granulocyte-biased (e.g., *Mpo*, *Ela*, *Prtn3*, *Camp*, *Ltf*, *S100a8*, *S100a9*, *Cebpe*) genes across clusters, highlighting lineage-associated transcriptional programs. **d**, Violin plots of stemness (PotencyScore), monocyte bias (MonoScore), and granulocyte bias (GranScore) signatures across the identified clusters, showing differential lineage priming. **e-g**, scRNA-seq analysis of GFP sorted bone marrow cells (blast cells) in *Zmiz1*^{fl/fl} and *Zmiz1*^{Δ/Δ} mice. **e**, Dot plot illustrating additional lineage-defining cell markers in clusters C1, C2, C3, C4, C5, and C6. **f**, Violin plots of monocyte- and granulocyte-biased gene expression across these clusters, further supporting lineage assignment. **g**, Violin plots of CytoTRACE2.0 score, MonoScore, and GranScore across clusters, indicating differences in differentiation potential and lineage bias. **h**, Absolute cell counts of hematopoietic populations in *Zmiz1*^{fl/fl} and *Zmiz1*^{Δ/Δ} mice. Bar chart showing the absolute number of cells in distinct hematopoietic clusters: Stem-like, Granulocyte-biased (Gran_bias), Monocyte-biased (Mono_bias), T/NK cells, and Total B cells.

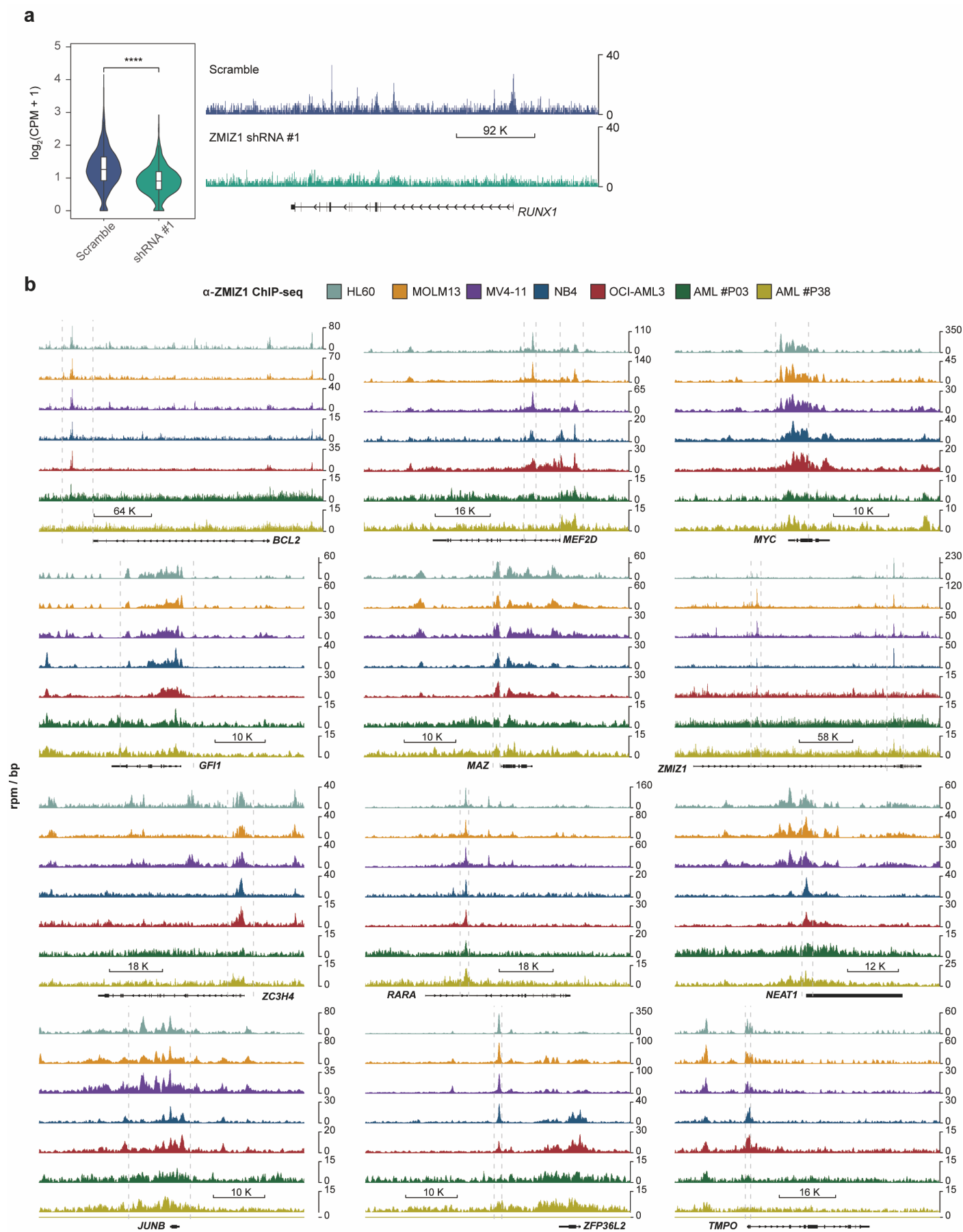
Figure. S4.



Supplementary Fig. 4. Intrinsically disordered regions of SUMO E3 ligase and condensate properties of ZMIZ1. **a**, Disorder tendency prediction of representative SUMO E3 ligases (PIAS1, PIAS2, PIAS3, PIAS4, TOPORS, RANBP2, CBX4, TRIM28). Scores were generated using three independent algorithms (PONDRL-VSL2, IUPRED, and ANCHOR). A score above 0.5 indicates intrinsically disordered regions. **b**, Representative fluorescence microscopy images showing nuclear puncta formation of GFP-tagged ZMIZ1 ligases in cells. White boxes highlight individual cells; magnified images of selected puncta are shown on the right. **c**, Distinct roles of ZMIZ1 IDR1 and IDR2 in puncta formation. Upper panel: Representative fluorescence microscopy images of cells expressing GFP-tagged full-length ZMIZ1 or the indicated truncation constructs (Δ IDR1-GFP, Δ IDR2-GFP, IDR1-GFP, IDR2-GFP). GFP signals are shown in green, and nuclei

are counterstained with Hoechst (blue). Lower panel: Quantification of puncta number per nucleus for cells expressing the indicated GFP-tagged constructs. Each dot represents an individual nucleus, and horizontal lines indicate the mean values. **d**, In vitro fluorescence recovery after photobleaching (FRAP) analysis of ZMIZ1 condensates. Normalized fluorescence intensity was measured over time, showing partial recovery and indicating dynamic molecular exchange within condensates. Data are presented as mean $\pm$ s.d. **e**, Representative immunofluorescence images of OCI-AML3 cells transduced with a scramble control or ZMIZ1-targeting shRNA (#1). ZMIZ1 is shown in magenta and nuclei are counterstained with Hoechst (blue).

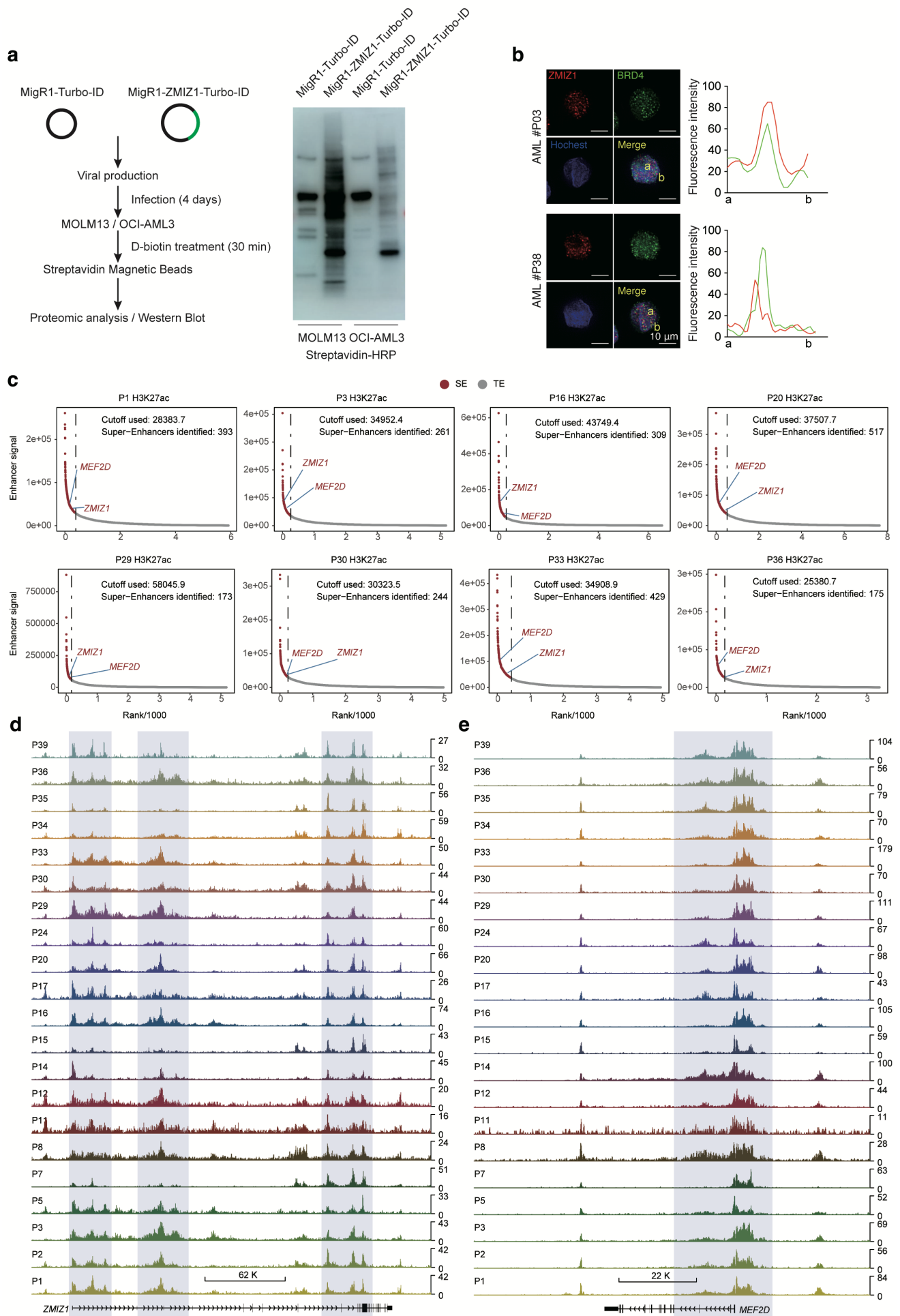
Figure. S5.



Supplementary Fig. 5. Genome browser tracks of ZMIZ1 occupancy at key leukemic target loci. a, ZMIZ1 knockdown reduces transcriptional output at the RUNX1 locus. Violin plots (left) show normalized ChIP-seq signals in ZMIZ1 binding sites in cells transduced with a scramble control or ZMIZ1 shRNA #1.

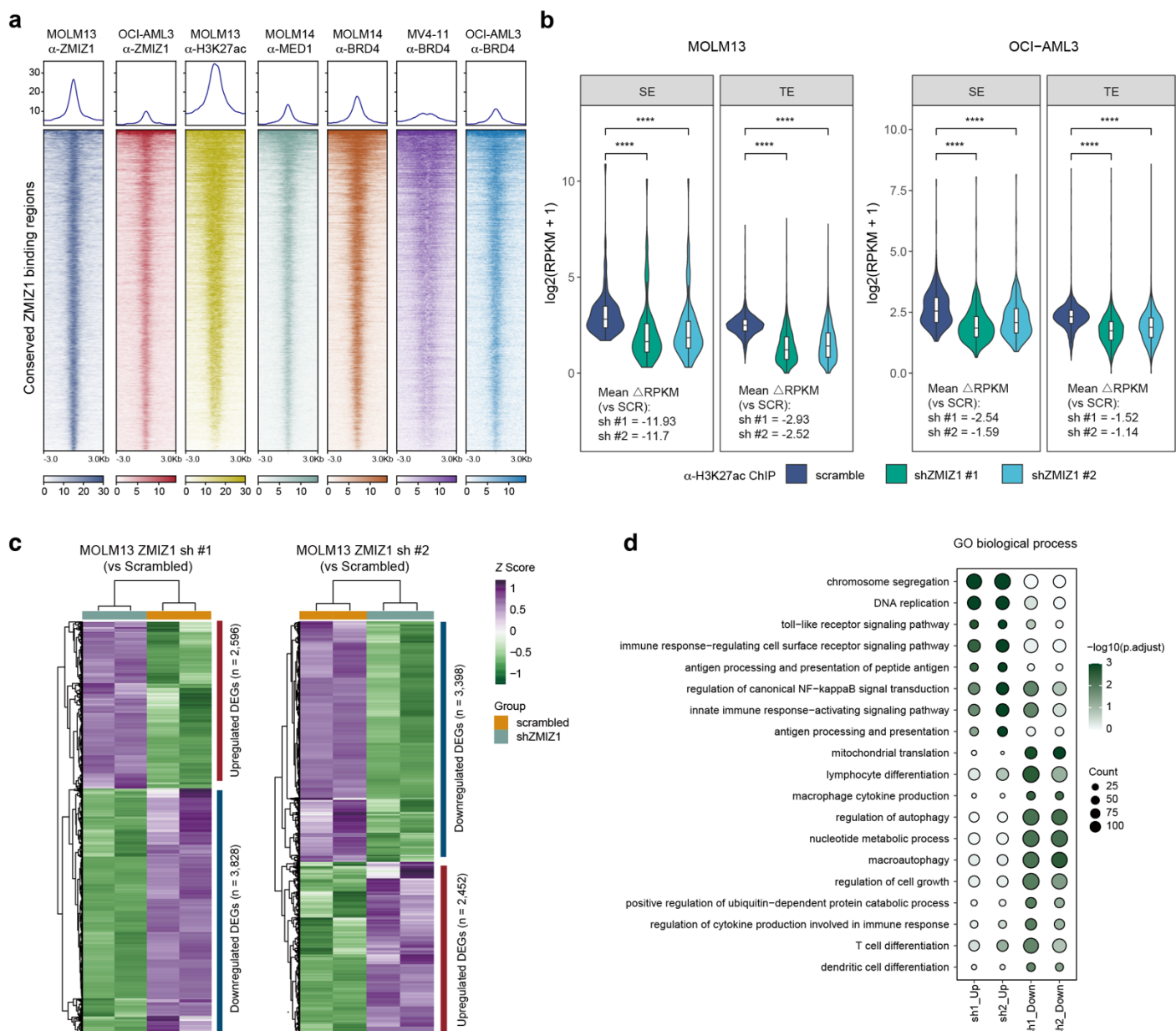
Boxes indicate the median and interquartile range; whiskers denote the distribution range. Statistical significance was assessed using a two-sided test; **** denotes $P < 0.0001$. Right panels display representative ChIP-seq signal tracks across the RUNX1 genomic locus in scramble control (top) and ZMIZ1 shRNA #1 (bottom) conditions. **b**, ChIP-seq tracks showing ZMIZ1 binding signals (rpm/bp) across a panel of AML cell lines (HL-60, MOLM13, MV4-11, NB4, OCI-AML3) and two primary AML patient samples (AML #P03, AML #P38). The tracks display consistent ZMIZ1 enrichment at regulatory regions of critical leukemia-associated genes, including *BCL2*, *MEF2D*, *MYC*, *GF11*, *MAZ*, *ZMIZ1* (auto-regulation), *ZC3H4*, *RARA*, *NEAT1*, *JUNB*, *ZFP36L2*, and *TMPO*. Gray dashed boxes highlight the specific peak regions conserved across multiple cell lines and patient samples, demonstrating the robustness of ZMIZ1 binding to these targets. Gene structures and orientations are indicated below the tracks.

Figure. S6.



Supplementary Fig. 6. ZMIZ1 is recurrently associated with broad H3K27ac domains and ranks as a super-enhancer in primary AML. **a**, TurboID-based proximity labeling identifies ZMIZ1-associated proteins in AML cells. Left: Schematic illustration of the TurboID proximity-labeling workflow. MOLM13 and OCI-AML3 cells were transduced with MigR1-TurboID control or MigR1-ZMIZ1-TurboID constructs, followed by D-biotin treatment for 30 min to induce proximity-dependent biotinylation. Biotinylated proteins were enriched using streptavidin magnetic beads and analyzed by proteomics and Western blotting. Right: Representative streptavidin-HRP Western blot showing enhanced biotinylation profiles in cells expressing ZMIZ1-TurboID compared with TurboID control in both MOLM13 and OCI-AML3 cells, confirming efficient ZMIZ1-centered proximity labeling. **b**, Representative immunofluorescence images of ZMIZ1 (red) and BRD4 (green) in primary acute myeloid leukemia (AML) samples (#P03 and #P38). Nuclei were counterstained with Hoechst (blue). Merged images show the combined fluorescence signals of ZMIZ1 and BRD4 within the nucleus. Line-scan fluorescence intensity profiles (right) were extracted along the indicated nuclear regions to display the spatial distribution of ZMIZ1 and BRD4 signals. Scale bars, 10 μ m. **c**, Rank-ordered H3K27ac enhancer landscapes from representative primary AML specimens. The vertical dashed line marks the inflection-point cutoff used to separate super-enhancers (SEs) from typical enhancers (TEs); the number above each plot indicates the total SEs called in that sample. ZMIZ1 and MEF2D are highlighted at their cohort-specific ranks; red labels denote loci classified as SEs, grey labels denote TEs. **d-e**, Genome browser views of H3K27ac ChIP-seq signal across multiple independent AML samples at the *ZMIZ1* (**d**) and *MEF2D* (**e**) loci. Each horizontal track is one patient sample (IDs at left; y-axis, normalized ChIP-seq signal). Shaded regions indicate the stitched enhancer clusters used for SE calling; gene models are shown below with scale bars (kb). Broad, high H3K27ac occupancy consistent with SE architecture is evident at *ZMIZ1* in many cases, comparable to the canonical SE at *MEF2D*.

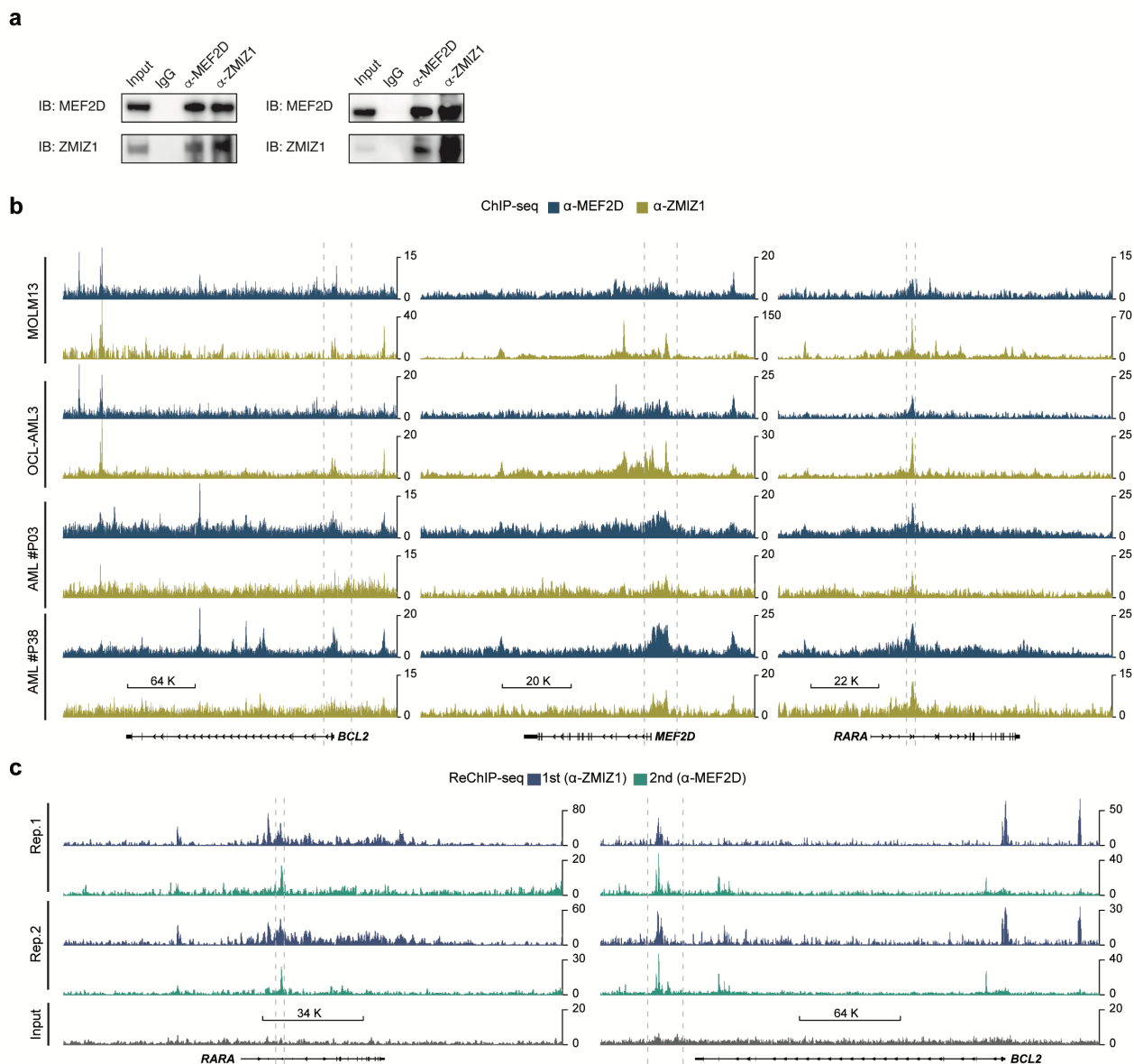
Figure. S7.



Supplementary Fig. 7. ZMIZ1 co-localizes with super-enhancer markers and regulates enhancer chromatin accessibility and leukemic transcriptional programs. a, ChIP-seq density heatmaps centered on conserved ZMIZ1 binding regions (± 3 kb). Tracks display ZMIZ1 occupancy alongside key enhancer and super-enhancer markers (H3K27ac, MED1, BRD4) across multiple AML cell lines (MOLM13, OCI-AML3, MOLM14, MV4-11), demonstrating strong co-localization. **b,** Violin plots quantifying H3K27ac ChIP-seq signal intensity ($\log_2 \text{RPKM} + 1$) at Super-Enhancers (SE) and Typical Enhancers (TE) in MOLM13 and OCI-AML3 cells. ZMIZ1 knockdown (sh#1 and sh#2) leads to a significant global reduction in H3K27ac levels compared to scrambled control, with a more pronounced effect at SEs. **** $P < 0.0001$. **c,** Heatmaps of differentially expressed genes (DEGs) after ZMIZ1 knockdown with two independent shRNAs (SH1 and SH2) in MOLM-13 cells versus scrambled control (SCR). Rows are genes (row-wise Z-score of normalized expression; purple, higher; green, lower); columns are samples hierarchically clustered. Side bars indicate the number of DEGs per comparison (left panel, SH1: upregulated $n = 2,596$; downregulated $n = 3,828$; right panel, SH2: upregulated $n = 2,452$; downregulated $n = 3,398$). Group annotation (top) denotes SCR (orange)

and shZMIZ1 (teal). **d**, Gene Ontology (GO, Biological Process) enrichment for up- and down-regulated gene sets from. Dot size reflects the number of genes contributing to each term; colour denotes $-\log_{10}(\text{adjusted } P)$. Immune and antigen-presentation pathways are preferentially induced on ZMIZ1 silencing, whereas DNA replication and chromosome-segregation programs are diminished.

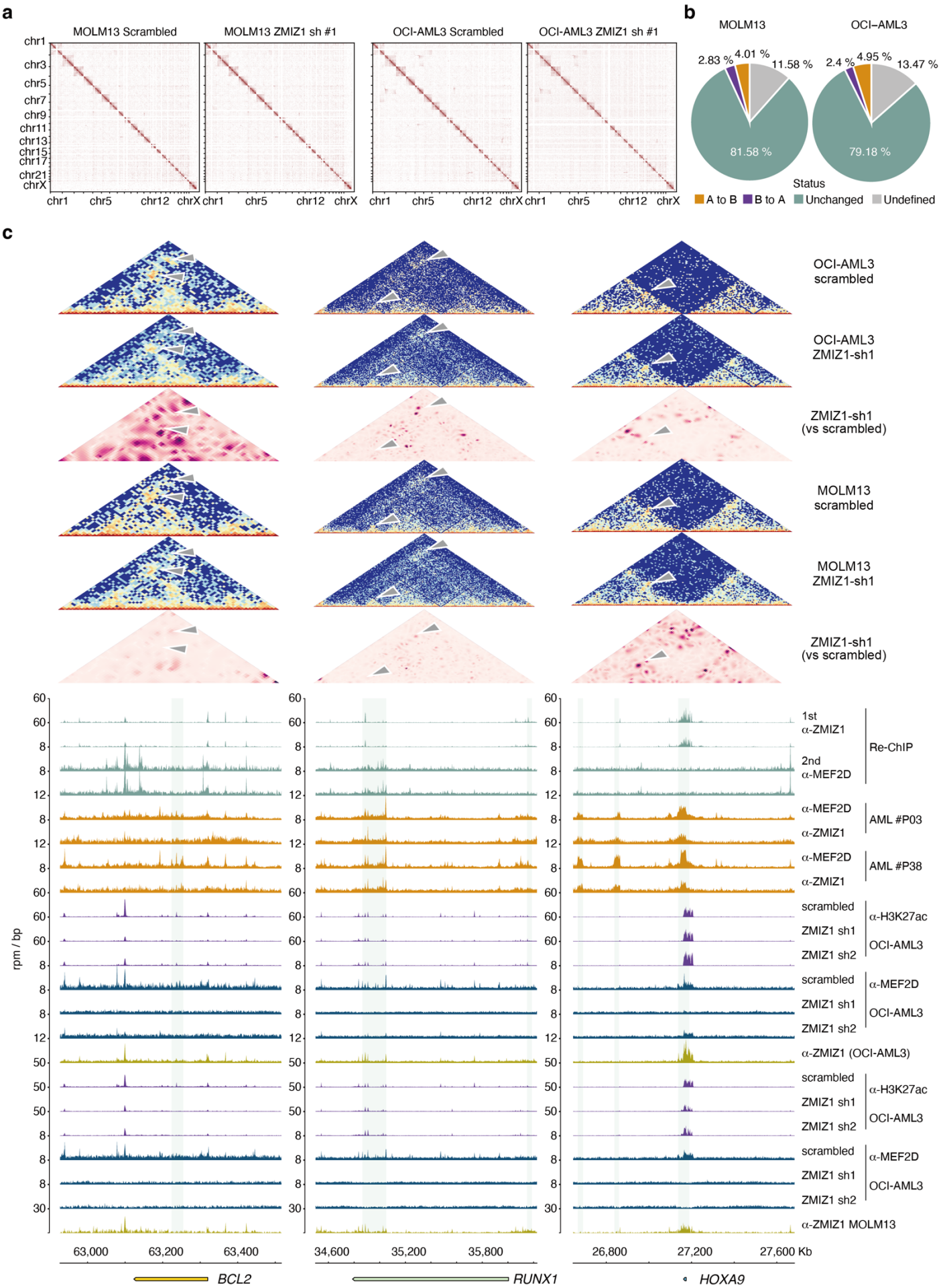
Figure. S8.



Supplementary Fig. 8. ZMIZ1 and MEF2D co-occupy chromatin at key leukemic driver genes.

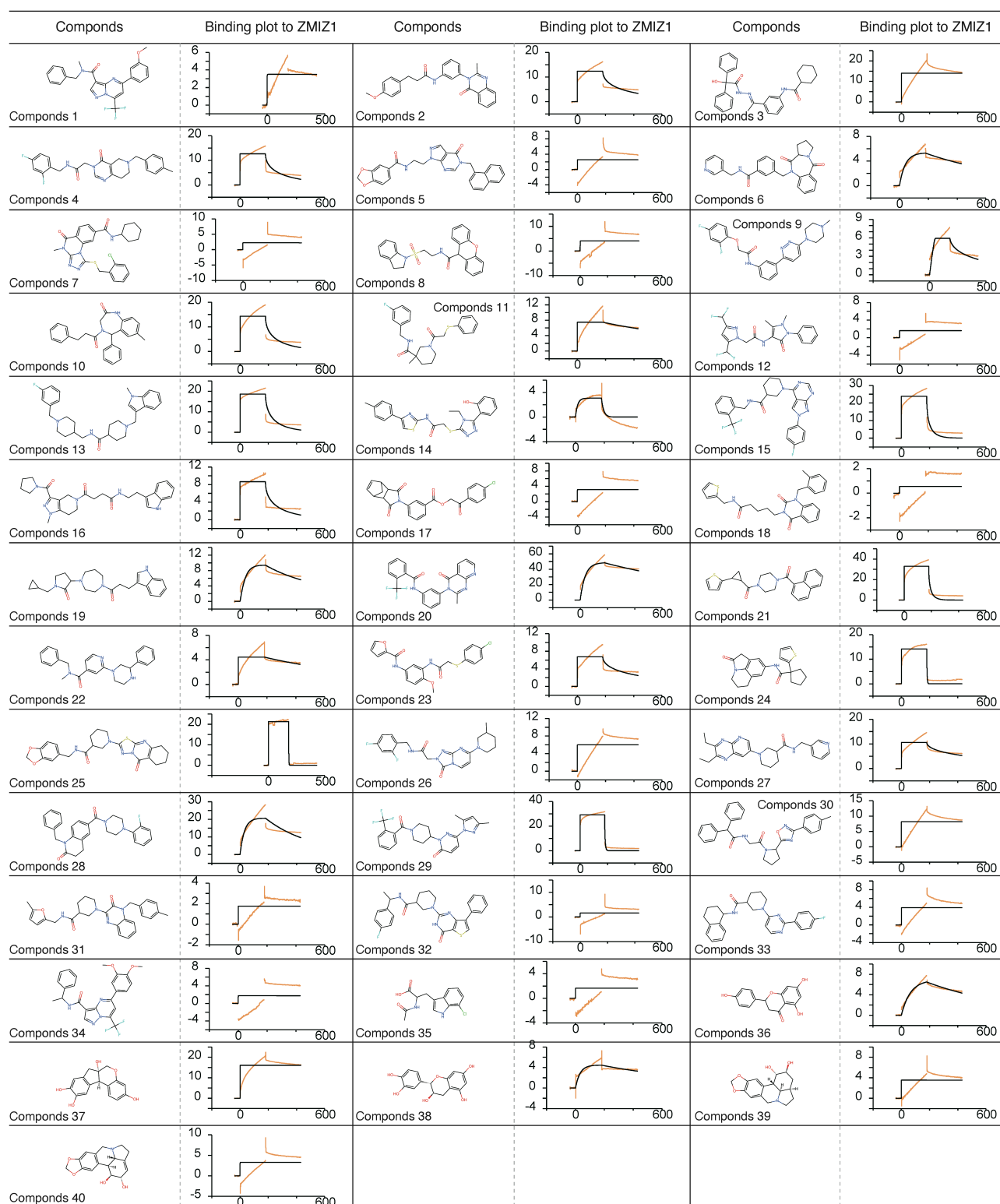
a, Interaction between ZMIZ1 and MEF2D detected by co-immunoprecipitation. MOLM13 (left) or OCI-AML3 (right) cell lysates were subjected to immunoprecipitation with antibodies against MEF2D (α -MEF2D), ZMIZ1 (α -ZMIZ1), or control IgG, followed by immunoblotting (IB) with the indicated antibodies. Input samples are shown as loading controls. **b**, Genome browser tracks displaying ChIP-seq occupancy for MEF2D (blue) and ZMIZ1 (yellow) at the *BCL2*, *MEF2D*, and *RARA* loci. Data is shown for AML cell lines (MOLM13, OCI-AML3) and primary AML patient samples (#P03, #P38). Gray dashed lines highlight the conserved co-localization of ZMIZ1 and MEF2D peaks across diverse AML contexts. **c**, Re-ChIP-seq validation of ZMIZ1/MEF2D co-occupancy at the *RARA* and *BCL2* loci. Tracks show the signal from the 1st IP (anti-ZMIZ1, blue) and the sequential 2nd IP (anti-MEF2D, green) for two biological replicates, alongside the matched Input control (grey). The robust enrichment of the 2nd IP signal specifically at the co-occupied regions (dashed boxes) confirms the physical interaction of the ZMIZ1-MEF2D complex on chromatin, while the Input track demonstrates low background noise.

Figure. S9.



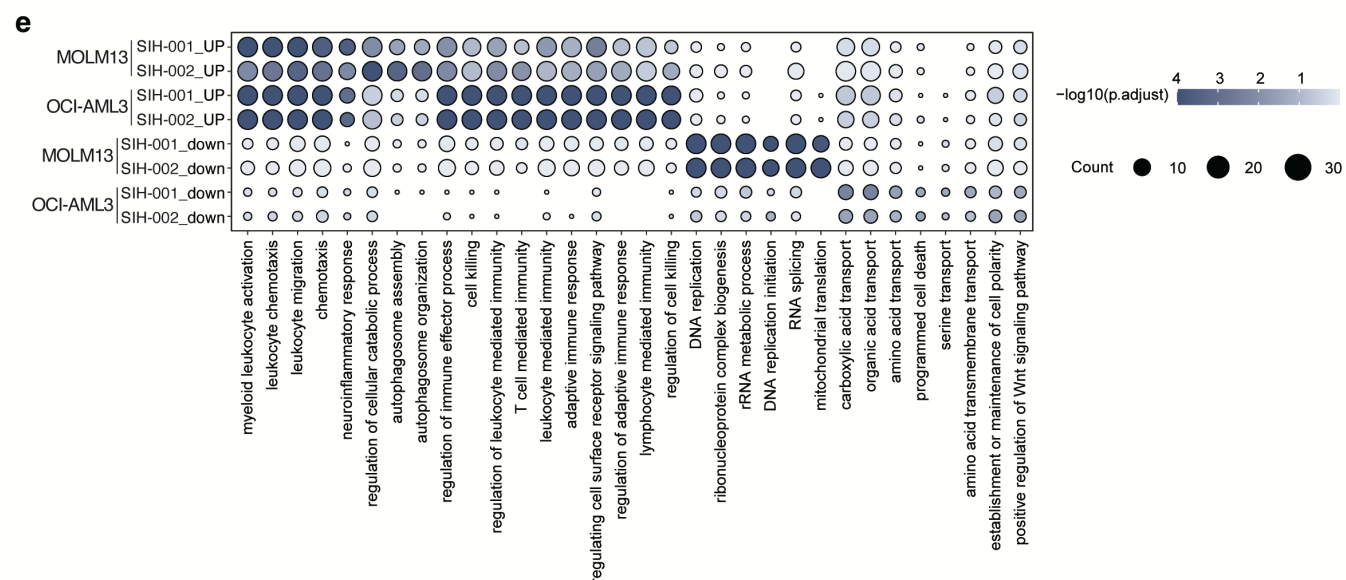
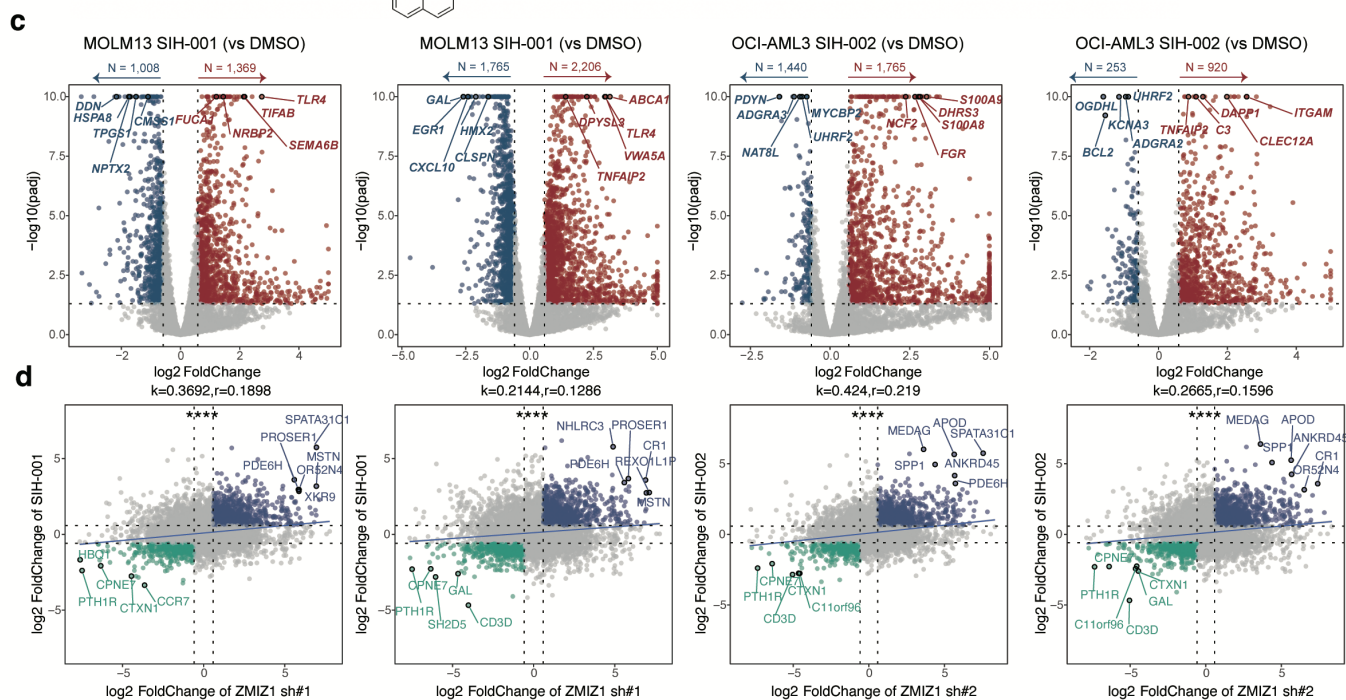
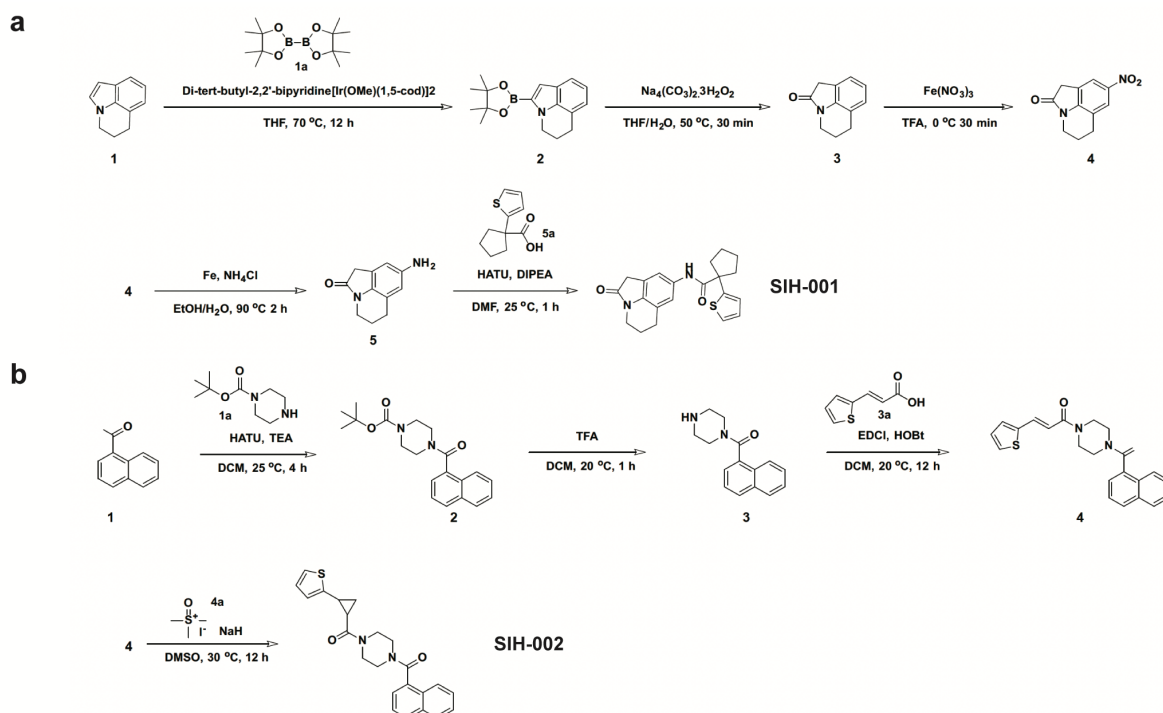
Supplementary Fig. 9. ZMIZ1 depletion preserves global chromosomal architecture. **a**, Genome-wide Hi-C contact maps for MOLM13 and OCI-AML3 cells transduced with scrambled control (scrambled) or an shRNA targeting ZMIZ1 (sh1). Axes denote chromosomes; colour scale indicates normalized contact frequency (observed/expected, log scale). Comparable plaid patterns across conditions indicate largely unchanged large-scale chromosome organization. **b**, Quantification of A/B-compartment status after ZMIZ1 knockdown, shown as the fraction of fixed-size genomic bins classified by eigenvector decomposition. ‘Unchanged’ denotes bins retaining their compartment; ‘Undefined’ indicates bins with insufficient signal for assignment. **c**, Hi-C interaction heatmaps centered on the *MEF2D*, *BCL2*, and *HOXA9* super-enhancer locus in MOLM13 and OCI-AML3 cells under control and ZMIZ1 knockdown conditions. Corresponding ChIP-seq tracks (below) for MEF2D, ZMIZ1, and H3K27ac were shown in the lower panel.

Figure. S10.



Supplementary Fig. 10. Primary biophysical screen identifies small-molecule binders of ZMIZ1. For each entry (Compounds 1–40), the chemical structure (left) and a representative binding trace to recombinant ZMIZ1 (right) are shown. Binding was assessed by label-free biosensing (biolayer interferometry/surface plasmon resonance) using association followed by dissociation over ~600 s; traces are reference-subtracted, and solid lines denote model fits. Responses span robust, kinetically resolvable interactions to negligible or non-specific signals, illustrating the diversity of chemotypes and binding behaviours observed in the primary screen.

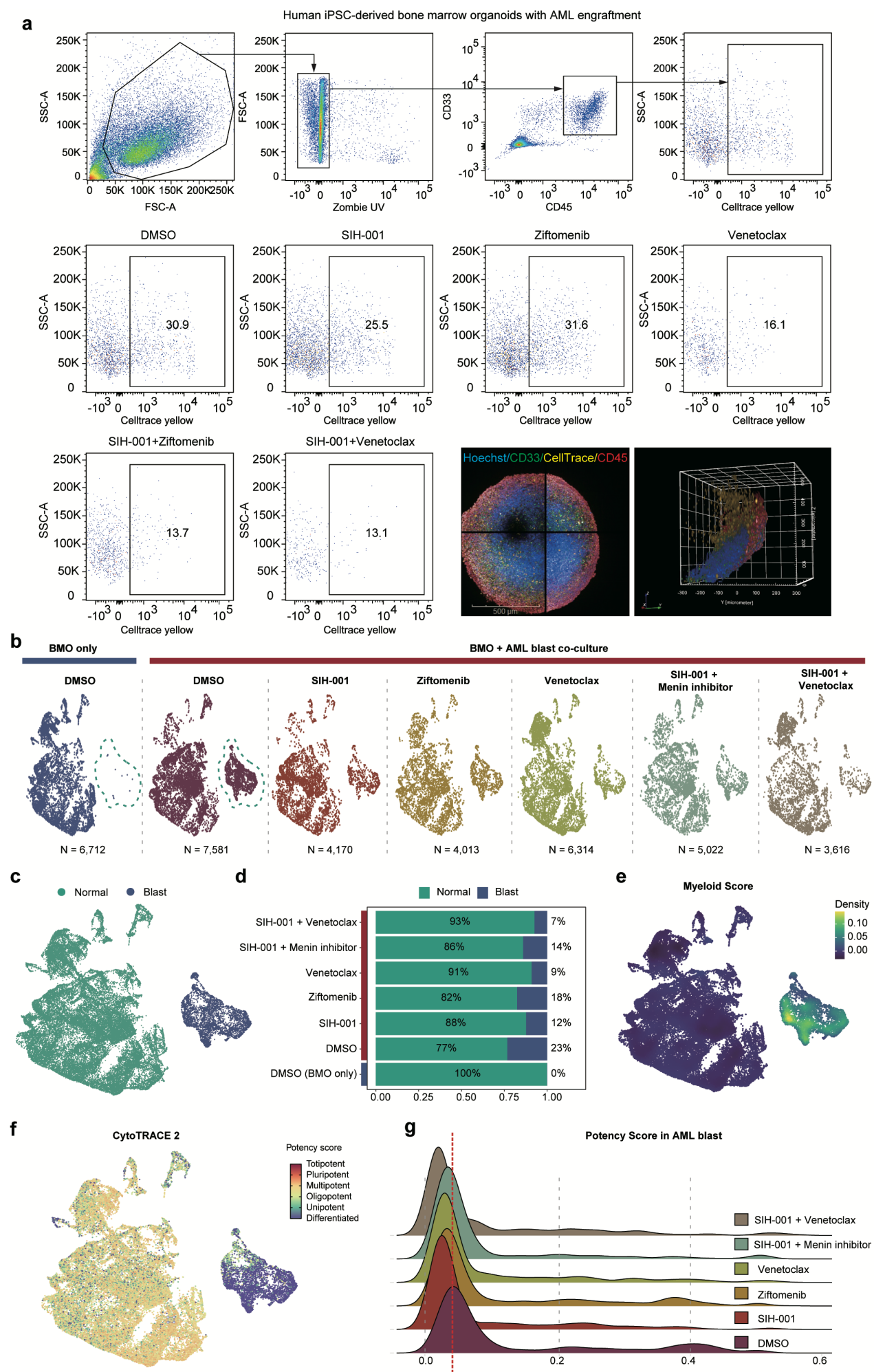
Figure. S11.



Supplementary Fig. 11. SIH-001 and SIH-002 induce transcriptional programs that converge with ZMIZ1 loss in AML cells.

a-b, Chemical synthesis of SIH-001 and SIH-002. Synthetic route for the preparation of SIH-001 (**a**). The synthesis starts from compound **1**, which undergoes Ir-catalyzed C–H borylation using di-tert-butyl-2,2'-bipyridine [Ir (OMe)(1,5-cod)] and reagent **1a** to afford intermediate **2**. Subsequent oxidative transformation with $\text{Na}_4(\text{CO}_3)_2 \cdot 3\text{H}_2\text{O}_2$ yields lactam **3**, followed by nitration with $\text{Fe}(\text{NO}_3)_3$ to give nitro compound **4**. Reduction of the nitro group using $\text{Fe}/\text{NH}_4\text{Cl}$ furnishes an aniline intermediate **5**, which is finally coupled with carboxylic acid **5a** via HATU/DIPEA-mediated amide formation to generate SIH-001. Synthetic route for SIH-002 (**b**). Starting from compound **1**, amide coupling with protected piperazine derivative **1a** using HATU/TEA affords intermediate **2**, which is deprotected under acidic conditions (TFA) to give amine **3**. Subsequent EDCI/HOBt-mediated coupling with carboxylic acid **3a** yields intermediate **4**, followed by sulfonium-ylide-mediated cyclopropanation (using **4a** and NaH) to afford the final compound SIH-002. **c**, Volcano plots showing differentially expressed genes (DEGs) after treatment with compound SIH-001 or SIH-002 versus DMSO in MOLM13 and OCI-AML3 cells (comparisons indicated above each panel). The x-axis shows $\log_2$ fold-change ($\log_2\text{FC}$); the y-axis shows $-\log_{10}(\text{adjusted } P)$. Vertical and horizontal dashed lines mark the significance thresholds ($|\log_2\text{FC}| \geq 1$ and $\text{FDR} \leq 0.05$). Numbers at the top denote counts of significantly down- (blue) and up-regulated (red) genes. Representative immune/inflammatory genes induced (for example *TLR4*, *TNFAIP2*, *CXCL10*) and other markers are annotated. **d**, Concordance between the transcriptional effects of SIH-001/SIH-002 and ZMIZ1 knockdown. Scatter plots relate $\log_2\text{FC}$ upon 4E (left) or 4H (right) treatment to $\log_2\text{FC}$ upon ZMIZ1 silencing with shRNAs (axes as labelled). Blue points indicate genes significantly up-regulated in both conditions; green points indicate genes significantly down-regulated in both (same thresholds as in **a**). Solid lines show linear fits; k and r report Kendall's τ and Pearson's correlation coefficients, respectively (****, $P < 0.0001$). Shared induction of immune-response genes and repression of select transcripts underscore that SIH-001/ SIH-002 phenocopy key features of ZMIZ1 loss. **e**, Gene Ontology (Biological Process) enrichment of differentially expressed genes (DEGs) after treatment with SIH-001 or SIH-002 in MOLM13 and OCI-AML3 cells (vs DMSO; thresholds in Methods). Rows denote condition and direction of change (UP/DOWN). Dot size indicates the fraction of DEGs annotated to the term (gene ratio); color encodes statistical significance (adjusted P value).

Figure. S12.



Supplementary Fig. 12. Single-cell RNA sequencing (scRNA-seq) analysis of AML cell response to drug treatments in bone marrow organoid (BMO) models.

a, Drug response profiling of AML cells in human iPSC-derived bone marrow organoids. Representative flow-cytometric gating strategy for the analysis of AML engraftment and treatment response in human iPSC-derived bone marrow organoids (upper panel). Cells were first gated by forward- and side-scatter (FSC-A/SSC-A) to exclude debris, followed by exclusion of dead cells using Zombie UV. Human AML cells were identified as CD45⁺CD33⁺ events and further assessed for CellTrace™ Yellow signal to quantify proliferative/retained AML populations within the organoids. Middle panels were SSC-A versus CellTrace™ Yellow plots showing AML cell abundance under the indicated treatment conditions. Representative whole-mount confocal image and three-dimensional reconstruction of an AML-engrafted bone marrow organoid stained with Hoechst (nuclei), CD33 (AML cells), CellTrace™ Yellow (cell tracking), and CD45 (human hematopoietic cells) (lower right panel). **b**, UMAP visualization of single-cell transcriptomes across different experimental conditions. Panels display cell distributions for BMO (normal background control), DMSO (vehicle control), and treatment groups including single agents SIH-001 (ZMIZ1 inhibitor), Ziftomenib, and Venetoclax, as well as combination therapies (SIH-001 + Ziftomenib and SIH-001 + Venetoclax). **c**, UMAP projection identifying cell subpopulations, distinguishing between leukemic Blast cells (blue) and Normal BMO cells (green). **d**, Bar charts showing the relative frequency (%) of Blast versus Normal cells across all experimental conditions, highlighting the reduction in leukemic burden. **e**, Density plot of Myeloid Lineage Scores comparing Blast and Normal populations. Scores were calculated using the Seurat AddModuleScore function based on a specific myeloid gene signature (*LYZ*, *S100A8*, *S100A9*, *LGALS3*, *CTSS*, *FCN1*, *FCER1G*, *LST1*, *TYMP*, *AIF1*, *MNDA*, *VCAN*, *MS4A7*, *CTSD*, *CST3*, *SPI1*). High myeloid scores in the Blast cluster validate the cell type annotation. **f**, Density plot of CytoTRACE 2 Potency scores, assessing the differentiation potential of cells (ranging from Differentiated to Totipotent). **g**, RidgePlot displaying the distribution of Potency Scores specifically within the AML Blast population across the indicated treatment conditions, assessing the impact of treatment on leukemic stemness/potential.

a

b

c

d

e

a, UMAP of all profiled bone-marrow cells showing unsupervised clusters C0–C7. **b**, Condition overlays on the same UMAP for DMSO, Venetoclax, and SIH-001. **c**, Dot plot showing marker gene expression across transcriptionally defined cell clusters (C0–C7). Annotation of each cluster were listed in the right panel. Dot size indicates the proportion of cells expressing the marker gene, and color intensity reflects average expression. **d**, UMAP projections and pie charts showing the distribution of major cell states across treatment groups (DMSO, Venetoclax, and SIH-001). Therapy-induced states and stem-like2 are enriched while the stem-like1 were reduced following SIH-001 treatment, highlighting its distinct transcriptional impact compared to Venetoclax and control. The stem-like1 were more malignant while the stem-like2 were relatively differentiated. **e**, UMAP visualization of representative marker genes (B2m, H2-D1, H2-Q7, Tap2, H2-K1, and Tap1) across the single-cell transcriptomes, with color gradients indicating expression density.

Original films of immunoblots

Fig. 2a

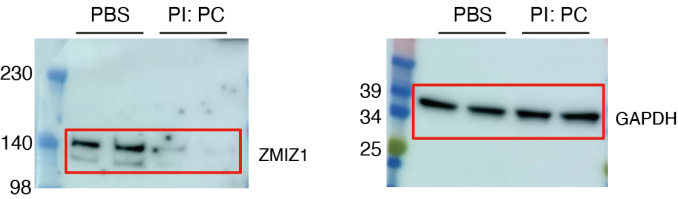


Fig. 5h

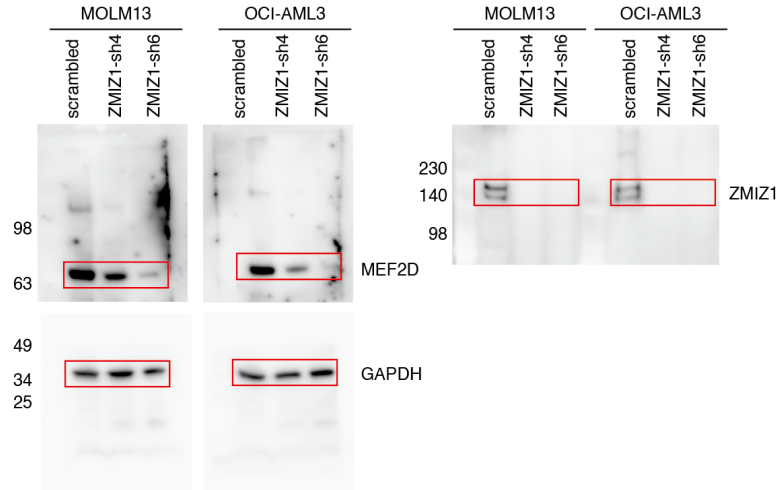


Fig. 6c

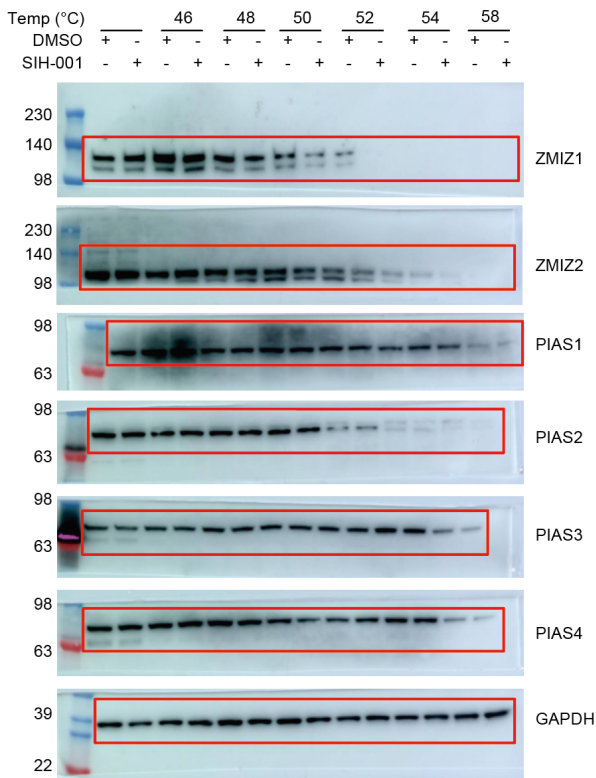


Fig. 6e

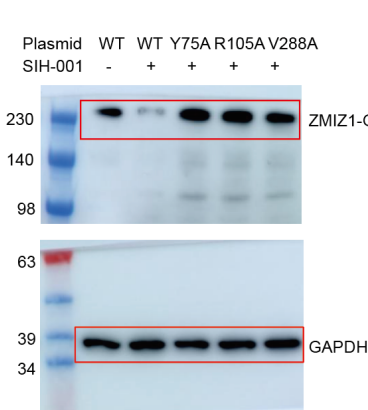
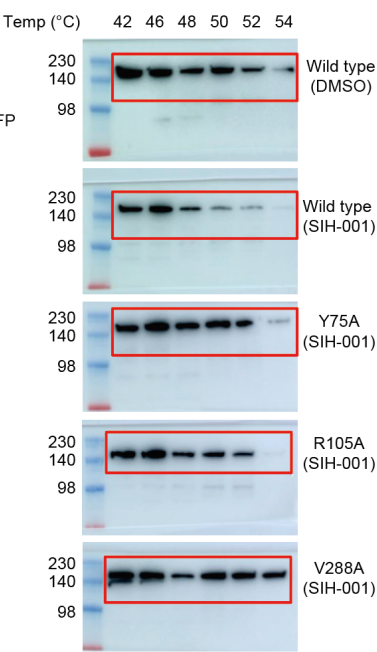
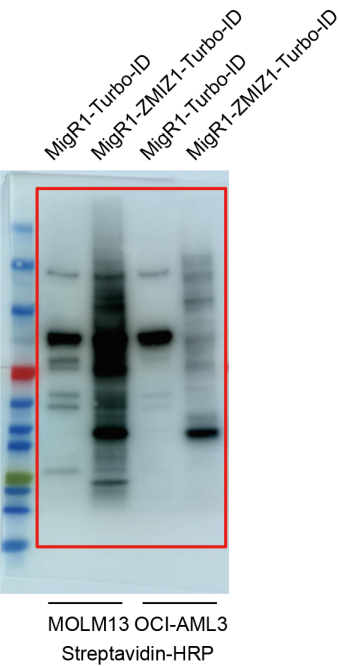


Fig. 6f



Supplementary Fig. 6a



Supplementary Fig. 8a

