



## Article

<https://doi.org/10.1038/s41588-024-01911-7>

# Pharmacological restriction of genomic binding sites redirects PU.1 pioneer transcription factor activity

Received: 25 July 2023

Accepted: 14 August 2024

Published online: 18 September 2024

Check for updates

Samuel J. Taylor<sup>1</sup>, Jacob Stauber<sup>1</sup>, Oliver Bohorquez<sup>1</sup>, Goichi Tatsumi<sup>1</sup>, Rajni Kumari<sup>1</sup>, Joyeeta Chakraborty<sup>1</sup>, Boris A. Bartholdy<sup>1</sup>, Emily Schwenger<sup>1</sup>, Sriram Sundaravel<sup>1</sup>, Abdelbasset A. Farahat<sup>2,3</sup>, Justin C. Wheat<sup>1</sup>, Mendel Goldfinger<sup>4,5,6</sup>, Amit Verma<sup>4,5,6,7,8,9</sup>, Arvind Kumar<sup>10</sup>, David W. Boykin<sup>10</sup>, Kristy R. Stengel<sup>1,5,6,8</sup>, Gregory M. K. Poon<sup>10,11</sup> & Ulrich Steidl<sup>1,4,5,6,8,9</sup>✉

Transcription factor (TF) DNA-binding dynamics govern cell fate and identity. However, our ability to pharmacologically control TF localization is limited. Here we leverage chemically driven binding site restriction leading to robust and DNA-sequence-specific redistribution of PU.1, a pioneer TF pertinent to many hematopoietic malignancies. Through an innovative technique, 'CLICK-on-CUT&Tag', we characterize the hierarchy of de novo PU.1 motifs, predicting occupancy in the PU.1 cistrome under binding site restriction. Temporal and single-molecule studies of binding site restriction uncover the pioneering dynamics of native PU.1 and identify the paradoxical activation of an alternate target gene set driven by PU.1 localization to second-tier binding sites. These transcriptional changes were corroborated by genetic blockade and site-specific reporter assays. Binding site restriction and subsequent PU.1 network rewiring causes primary human leukemia cells to differentiate. In summary, pharmacologically induced TF redistribution can be harnessed to govern TF localization, actuate alternate gene networks and direct cell fate.

Transcription factor (TF) networks regulate gene expression and are a key determinant of cell fate decisions and identity<sup>1</sup>. In cancer and other disorders, these networks are frequently corrupted and are often driven by genomic mutations that alter the function, expression or genomic localization of key TFs<sup>2</sup>. However, despite the importance of aberrant TF networks in disease development, our understanding of the intricacies of TF dynamics is limited and the repertoire of small molecules available to specifically target TF networks is inadequate<sup>3</sup>.

Given the importance of TF DNA binding site occupancy in dictating cell fate<sup>4–7</sup>, we decided to investigate a chemical approach to selectively block TF occupancy by using DNA-binding heterocyclic diamidines. These compounds are designed to non-covalently bind to adenine–thymine (AT)-rich minor grooves adjacent to ETS family core binding motifs and destabilize specific TF–DNA interactions, leading to TF displacement in a sequence-selective manner<sup>8,9</sup>. This pharmacological approach is unique in that it specifically inhibits the

<sup>1</sup>Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY, USA. <sup>2</sup>Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University, Mansoura, Egypt. <sup>3</sup>Master of Pharmaceutical Sciences Program, California Northstate University, Elk Grove, CA, USA. <sup>4</sup>Department of Oncology, Albert Einstein College of Medicine – Montefiore Medical Center, Bronx, NY, USA. <sup>5</sup>Blood Cancer Institute, Albert Einstein College of Medicine – Montefiore Medical Center, Bronx, NY, USA. <sup>6</sup>Montefiore Einstein Comprehensive Cancer Center, Albert Einstein College of Medicine – Montefiore Medical Center, Bronx, NY, USA. <sup>7</sup>Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, NY, USA. <sup>8</sup>Ruth L. and David S. Gottesman Institute for Stem Cell Research and Regenerative Medicine, Albert Einstein College of Medicine, Bronx, NY, USA. <sup>9</sup>Department of Medicine, Albert Einstein College of Medicine – Montefiore Medical Center, Bronx, NY, USA. <sup>10</sup>Department of Chemistry, Georgia State University, Atlanta, GA, USA. <sup>11</sup>Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, GA, USA. ✉e-mail: [ulrich.steidl@einsteinmed.edu](mailto:ulrich.steidl@einsteinmed.edu)

TF–DNA interaction, distinct from other approaches whereby the TF protein itself is inhibited or targeted for degradation<sup>10</sup>. We leverage select versions of these compounds to gain insight into TF dynamics and cell fate, with a particular focus on the hematopoietic TF PU.1 (also known as SPI1). This TF is an ideal candidate for investigation as it directs a well-characterized transcriptional program that governs cell fate<sup>4,11,12</sup>, has demonstrated proof-of-concept sensitivity to binding site inhibition, specifically compared to other ETS family TFs<sup>8,9,13,14</sup>, and is commonly dysregulated in leukemia and other diseases<sup>15–26</sup>.

We uncover that chemically driven binding site restriction drives a phenomenon that we refer to as ‘pharmacologically induced TF redistribution’. Binding site restriction caused a redistribution of PU.1 genomic binding and a seemingly paradoxical increase in PU.1-driven transcription at a subset of target genes. As a consequence, these chemical agents drive a robust myeloid differentiation program in both cell lines and primary samples from patients with acute myeloid leukemia (AML). We use CLICK chemistry-tagged versions of these PU.1-redistributing agents to catalog the sensitivity of the entire PU.1 cistrome to canonical binding site restriction and provide insight into the fast biology of endogenous PU.1, particularly into its pioneering ability and chromatin occupancy kinetics. In principle, the chemically induced TF redistribution we describe herein can be used not only to broaden our understanding of transcriptional processes but also as a mechanism to designate a new class of agents—‘TF redistributors’—which can be used in the fight against transcriptional aberrations in disease.

## Results

### Pharmacological binding site restriction repositions PU.1

Proof-of-concept studies into the inhibition of TF binding using synthetic heterocyclic diamidines in cell-free systems have established pharmacological displacement of PU.1 as a viable pharmacological strategy<sup>8,27</sup> but have provided little insight into the molecular and cellular consequences of such treatment. To characterize the consequences of PU.1 binding site restriction in a cellular context, we investigated the genomic localization of PU.1 following exposure to the tool compound DB2115 (Fig. 1a). A 5 μM dose of DB2115 was administered to PU.1-expressing MOLM13 cells to exert maximal functional effect (~90% growth inhibition) without altering the levels of PU.1 transcript or protein (Fig. 1b,c) before PU.1-directed CUT&Tag was performed.

Strikingly, robust redistribution of PU.1 binding rather than global PU.1 displacement was uncovered. CUT&Tag identified 34,904 total PU.1 binding sites, consistent with the number of PU.1 peaks found in prior localization studies<sup>28–30</sup>, with 19% of sites being lost and 15% being gained following DB2115 exposure ( $n = 3$ , Diffbind false discovery rate (FDR) < 0.1; Fig. 1d,g and Extended Data Fig. 1a,b,e). Log odds ratio scoring of a canonical PU.1 motif<sup>30</sup> and Homer analyses confirmed the enrichment of high-affinity PU.1 motifs (Class III ETS) within PU.1 CUT&Tag peaks, indicating a direct interaction between PU.1 and the DNA at gained, lost and unchanged sites (Fig. 1e,f and Supplementary Data 1). Interestingly, fewer lost PU.1 sites were found in promoter regions (4.4%) compared to gained sites (21.9%) (Fig. 1h).

**Fig. 1 | PU.1 binding site blockade causes a GC content-biased repositioning of genomic PU.1 binding.** **a**, Schematic of the experimental approach to preparing MOLM13 cells for CUT&Tag following a 12 h 5 μM DB2115 treatment. **b,c**, Representative western blot displaying protein (**b**) and RNA expression of PU.1 following 12 h of 5 μM DB2115 treatment (mean ± s.e.m.,  $n = 3$  experimental replicates) (**c**). **d**, Proportions of lost, gained and unchanged PU.1 peaks following DB2115 treatment. Differential binding was calculated by Diffbind with an FDR < 0.1,  $n = 3$ . **e**, Z-score heatmap of ETS motif enrichments. Known ETS motifs were identified from Homer analysis of lost, gained and unchanged PU.1 peaks. **f**, Log odds ratio score for the PU.1 consensus sequence (from a previous publication<sup>30</sup>) in lost, gained and unchanged PU.1 peaks. PWM, position weight matrix. **g**, Representative viewer tracks of genomic loci displaying lost (blue boxes), gained (red boxes) and unchanged (unmarked) PU.1 binding. **h**, Annotation in relation to

Given that DB2115 was designed to inhibit canonical PU.1 sites by blocking key AT-rich minor groove interactions surrounding the core GGAA<sup>8,9,13</sup>, we examined the guanine–cytosine (GC) content of the central 100 bp of lost, gained and unchanged PU.1 sites. As expected, the GC content was significantly less in lost PU.1 binding sites (Fig. 1i) compared to gained and unchanged sites. Indeed, the majority of PU.1 peak losses occurred in regions with <50% GC content (Fig. 1j), suggesting that DB2115 specifically displaced PU.1 from AT-rich binding sites.

To understand the AT bias surrounding PU.1 motifs, we aligned all peaks to a central motif (GAGGAAGT) and examined the GC content of the adjacent nucleotides (±25 bp). There was robust depletion of GC content surrounding the core GGAA ETS motif for the lost PU.1 sites compared to gained or unchanged sites (Fig. 1k). Further examination of the nucleotide composition identified that adenines were the predominant nucleotide enriched in the upstream and downstream regions of lost sites (Fig. 1l). This was reflected in a de novo motif generated from DB-sensitive (lost) PU.1 sites containing more than six adenines upstream of the core motif (Fig. 1m). Although heterocyclic diamidines do not have a preference for poly-A or poly-T minor grooves<sup>31</sup>, thymines were not equivalently enriched. This is partially explained by overall motif occurrence, with poly-A flanked ETS sites being more common than poly-T flanked sites (Extended Data Fig. 1f,g).

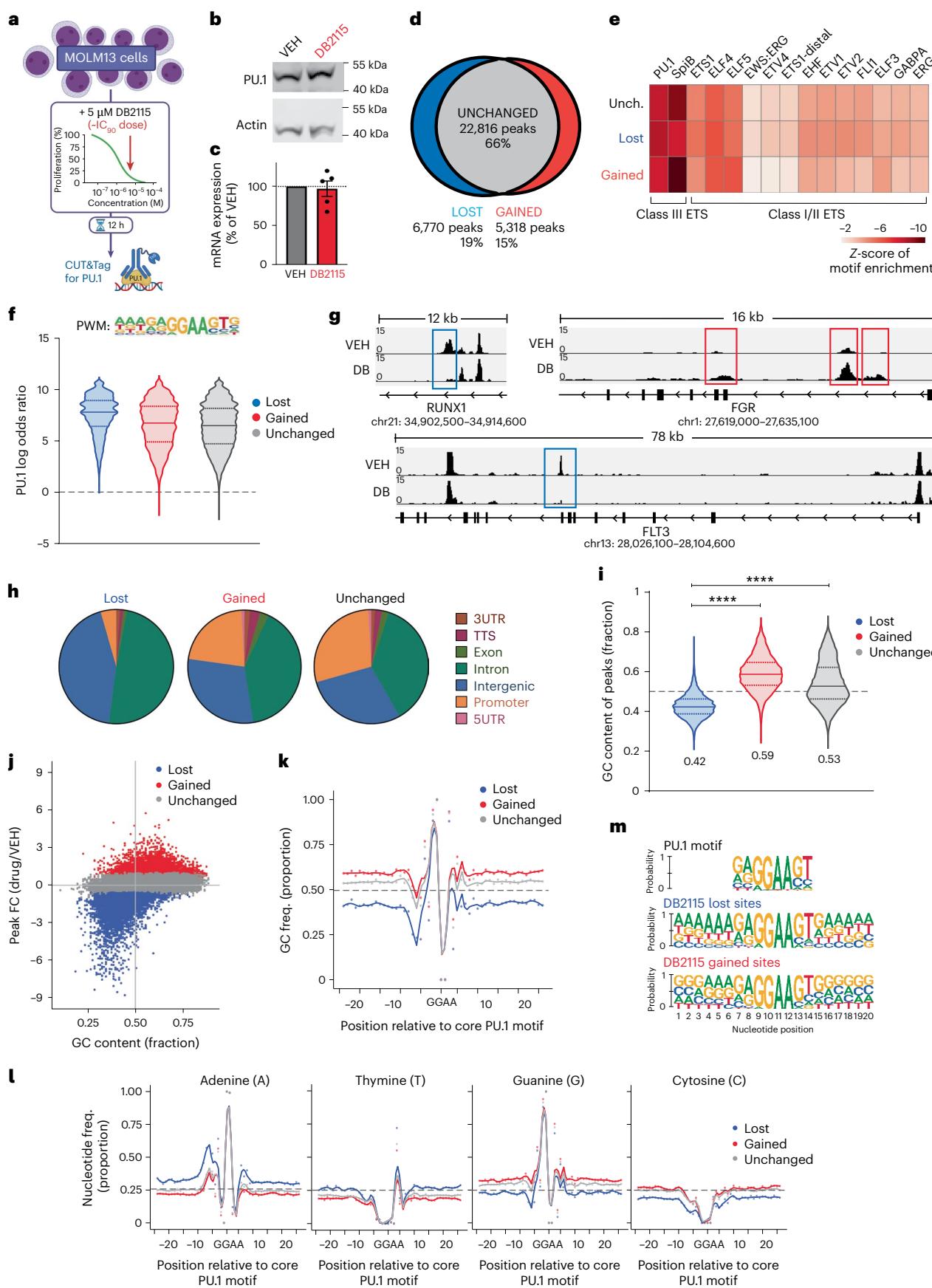
To confirm that pharmacological PU.1 redistribution was not an artefact of DB2115-mediated Trs5 transposase inhibition, we performed classical chromatin immunoprecipitation (ChIP). Despite some differences in peak number and distribution, the ChIP dataset corroborated the PU.1 redistributive phenomenon, characterized by profound losses and gains of PU.1 binding sites occurring with similar biases in GC content (Extended Data Fig. 2).

### PU.1 repositioning is a class-specific and selective phenomenon

Pharmacological redistribution of PU.1 occurred consistently across several PU.1-expressing cell lines, THP1, HL60 and MV411, with similar GC content biases (Fig. 2a and Extended Data Fig. 3a–c). Despite differences in baseline PU.1 localization, drug-induced PU.1 redistribution showed similarity across cell lines (Extended Data Fig. 3d,e). This was most profound amongst losses (71% of peaks lost in MOLM13 were commonly lost in ≥1 other cell line) compared to gains (32% of peaks gained in MOLM13 were commonly gained in ≥1 other cell line; Fig. 2b). Examination of these common sites identifies AT-richness and PU.1 motif score as contributing factors for consistent redistribution (Extended Data Fig. 2f–i). We generated a reference list of 347 high-confidence PU.1 binding sites that are consistently displaced with DB2115 (Supplementary Data 2). Overall, we posit that PU.1 loss is context-independent in nature, driven predominantly by sequence-specific drug affinity, whereas PU.1 relocalization is less predictable by sequence and more dependent on cellular context.

Encouragingly, primary AML or myelodysplastic syndrome (MDS)/myeloproliferative neoplasm cells exhibited robust TF redistribution following 12 h of DB2115 treatment in liquid culture (Fig. 2c,e,f).

distance to gene transcription start site of lost, gained and unchanged PU.1 peaks. UTR, untranslated region; TTS, transcription termination site. **i**, Proportion of GC content from the central 100 bp of lost, gained and unchanged PU.1 peaks. \*\*\* $P < 0.0001$ , two-sided Mann–Whitney test. **j**, Comparison of log<sub>2</sub> fold change (FC) in PU.1 peak score (DB2115-treated/vehicle) versus GC content of central 100 bp of all peaks (colored according to peak groups). **k**, GC content position frequency matrix of lost, gained and unchanged peaks that have been centered on a short consensus ETS motif (GAGGAAGT) and examined ±25 bp. **l**, Single-nucleotide position frequency matrices for adenine, thymine, guanine and cytosine, comparing lost, gained and unchanged PU.1-centered peaks. **m**, Motif generation from PU.1-centered, lost (DB-sensitive) and gained (DB-resistant) peaks displaying a robust and extended upstream enrichment of A nucleotides in lost peaks. Also shown is the PU.1 motif used for centering peaks. VEH, vehicle.



Similar biases in lost PU.1 peak GC content, including the enrichment for upstream A nucleotides, were evident (Fig. 2d and Extended Data Fig. 3j). Comparison of primary AML against the MOLM13 dataset revealed that there was more similarity between lost regions (33%) than gained regions (6%) (Extended Data Fig. 3k).

To evaluate the selectivity of the compounds for causing PU.1 redistribution, we performed CUT&Tag on other TFs (non-ETS family: RUNX1; other ETS family factors: ELF1 and GABPA). DB2115 exerted minimal (ELF1, GABPA) to mild (RUNX1) TF redistribution in MOLM13 cells (Fig. 2g and Extended Data Fig. 4a). However, a significant proportion of these minor TF redistribution events could be attributable to the redistribution of PU.1 itself, with 39–52% co-occurring with PU.1 (Fig. 2h). Furthermore, we examined drug-induced TF redistribution (RUNX1, GATA3, ELF1 and FLI1) in JURKAT cells (which express minimal to no PU.1). Minor redistribution was evident (Fig. 2i and Extended Data Fig. 4b) but this was ~fivefold less than the redistribution observed for PU.1 in MOLM13. Overall, these data indicate that PU.1 is uniquely sensitive to DB2115-mediated binding site inhibition compared to these other TFs.

To determine whether PU.1 redistribution was a drug class effect, we examined newly developed heterocyclic diamidine PU.1 binding site inhibitors DB2373 and DB2826 plus one previously described compound, DB2313 (ref. 9) (Extended Data Fig. 3l). Encouragingly, all compounds caused PU.1 redistribution, characterized by robust GC-content-specific gains and losses (Fig. 2j). Furthermore, to confirm that this is a class-specific effect, we examined PU.1 localization in MOLM13 cells after 12 h of exposure to 400 nM cytarabine or 200 nM daunorubicin and found that neither triggered PU.1 redistribution (Fig. 2k).

### CLICK-on-CUT&Tag identifies A-rich PU.1 site drug targeting

To test whether TF redistributors directly displace PU.1 and which genomic loci experience direct displacement, we generated an alkyne-linker-tagged version of DB2115 that is amenable to click chemistry (DB2750). This linker-tagged compound (Fig. 3a) was viable for copper-catalyzed cycloaddition with azide-AF488 and localized to the nucleus (Fig. 3b). Furthermore, the antiproliferative activity of the molecule was similar to DB2115 (inhibitory concentration 50 (IC<sub>50</sub>) DB2750, 270 ± 20 nM versus DB2115, 630 ± 8 nM; Fig. 3c) and we found that DB2750 caused GC-content-dependent PU.1 redistribution (Fig. 3d) with >70% equivalence in peak gains and losses compared to DB2115 treatment (Extended Data Fig. 5a–c).

To investigate the sequence specificity of TF redistributors, we developed an *in vitro* click chemistry approach to screen affinity in select synthesized PU.1 binding sites or the entire PU.1 cistrome (Fig. 3e). Initially, to test the viability of this approach, we mixed two 68-bp DNA fragments from either an AT-rich, DB2115-sensitive PU.1 binding site (SENP2 intergenic, TAAAAGGAGGAAGTG) or an AT-poor, DB2115-resistant PU.1 binding site (SPL1 upstream regulatory element, GCTGTTAGGAAGGG) and performed a pulldown with DB2750-coated magnetic beads (Fig. 3e and Extended Data Fig. 5d,e). There was a strong enrichment of AT-rich over AT-poor fragments, demonstrating

that the drug is indeed binding preferably to AT-rich DNA and that the CLICK-pulldown approach is viable (Fig. 3f).

To identify drug-binding sites amongst the entire PU.1 cistrome, we developed ‘CLICK-on-CUT&Tag’, in which the input DNA for the pull-down is the output from PU.1 CUT&Tag (Fig. 3e). CLICK-on-CUT&Tag identified drug binding at many PU.1 binding sites (16,447 out of 38,549 sites >log(0.5) enrichment; Fig. 3g,h). The strongest enrichment occurred at DB2115-lost PU.1 sites, denoting direct drug-mediated displacement of PU.1 (Fig. 3i). Furthermore, de novo motif enrichment of drug-bound sites identified an upstream A-rich ETS motif similar to the motif identified from DB2115-lost PU.1 sites (Figs. 3j and 1e). Conversely, motif enrichment of non-drug-bound sites generated a motif without upstream A-bias and predicted low-affinity PU.1 binding (Fig. 3j). There was an inverse relationship between CLICK-pulldown enrichment and both GC content and fold change of PU.1 peaks following DB2115 treatment (Fig. 3k,l). Notably, there was a pattern of lost PU.1 sites with low GC content corresponding with high CLICK enrichment compared to gained PU.1 sites (Fig. 3m). Overall, these findings support a model of direct compound-mediated displacement of PU.1 from AT-rich flanking sites.

### PU.1 repositioning restructures chromatin and transcriptome

To determine whether pharmacological TF repositioning could alter the accessible chromatin landscape, we performed assay for transposase-accessible chromatin with sequencing (ATAC-seq) following 12 h of DB2115 exposure (Fig. 4a). Profound changes to chromatin accessibility were evident; 33% of the total 149,462 accessible regions were closing and 20% were opening (FDR < 0.1; Fig. 4b). ETS motifs were strongly enriched in closing, opening and unchanged regions, suggesting that PU.1/ETS factors are key chromatin accessibility-controlling factors in these cells (Fig. 4c). Closing sites displayed a lower occurrence in promoter regions than unchanged and opening sites (Fig. 4e). Similarly, closed sites displayed a profound deficiency of GC content compared to gained or unchanged sites (Fig. 4d).

Next, we filtered PU.1 peaks on redistribution status and examined chromatin accessibility. The majority of PU.1-lost sites were found in closing areas of the chromatin (84%), whereas the majority of PU.1-gained sites were found in opening areas (78%; Fig. 4f,g and Extended Data Fig. 6a). Overall, these data support the notion that PU.1 exerts powerful pioneering control over chromatin accessibility<sup>29</sup>. Given that some sites with PU.1 loss remained unchanged in accessibility (Fig. 4f), we investigated whether other chromatin factors were cooperating with PU.1 to mediate chromatin accessibility changes. We examined enriched motifs in the four different classes of PU.1 changed sites (Fig. 4h). Interestingly, lost PU.1 and closing sites, as well as gained PU.1 and unchanged sites were enriched for A-rich motifs such as SOX5, IRF4 and ZNF384. An opposite effect was observed in the lost PU.1 and unchanged chromatin, and gained PU.1 and opening chromatin, with a strong enrichment of G-rich motifs such as ETS1, BORIS and SMAD3. These data suggest that surrounding co-factor binding may have a

**Fig. 2 | TF redistributors mediate selective, class-specific PU.1 repositioning in cell lines and primary AML samples.** **a**, Comparison of log<sub>2</sub>fold change of PU.1 peak score (12 h of 5 μM DB2115-treated/vehicle) versus GC content of the central 100 bp for all peaks for THP1, HL60 and MV411, showing the number and percentage of peaks redistributed in each cell line ( $n = 2$  for each cell line). **b**, Similarity between MOLM13 PU.1 peak changes with the other three cell lines. Peaks were classed as either common change to MOLM13 in one, two or three other cell lines or not commonly changed to MOLM13. Lost peaks were first filtered to exclude peaks not detected in other cell lines. **c**, Scatter plots comparing log<sub>2</sub>fold change of PU.1 peak score (DB2115-treated/vehicle) versus GC content of the central 100 bp for all peaks from two AML primary patient samples, including the number and percentage of peaks that are lost and gained in each sample. **d**, Motif generation from PU.1-centered, pooled primary AML sample redistributed peaks displaying a robust and extended upstream enrichment of A nucleotides in lost

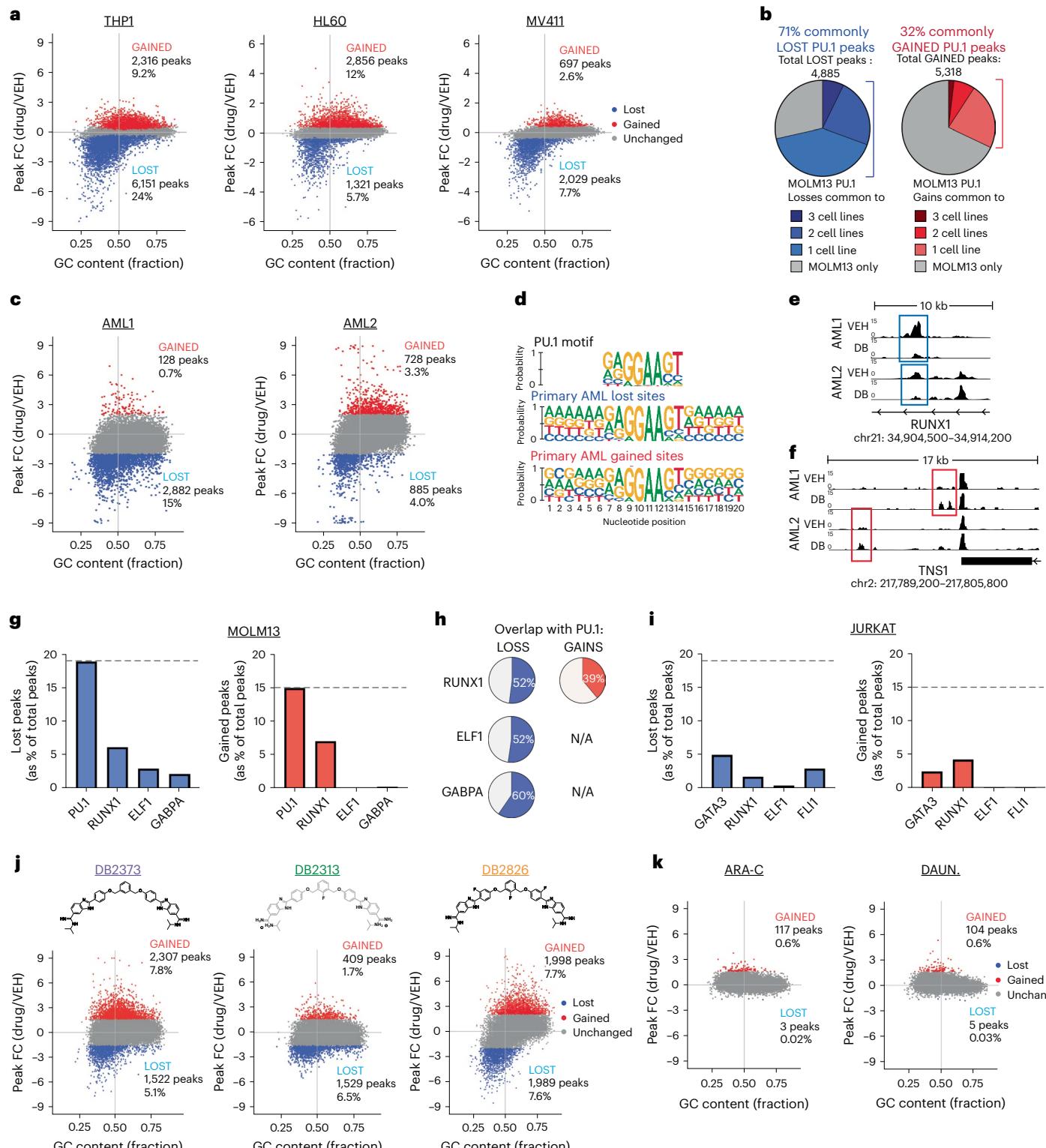
peaks. Also shown is the PU.1 motif used for centering peaks. **e,f**, Representative genomic loci displaying lost (e; blue boxes), gained (f; red boxes) and unchanged (unmarked) PU.1 binding in primary AML samples. **g**, Percentage of lost and gained CUT&Tag peaks of the TFs PU.1, RUNX1, ELF1 and GABPA in MOLM13 cells treated for 12 h with 5 μM DB2115. The dashed line represents the percentage of PU.1 lost or gained peaks. **h**, The percentage of lost or gained TF peaks that overlap with redistributed PU.1. N/A, not applicable. **i**, Percentage of lost and gained CUT&Tag peaks of the TFs GATA3, RUNX1, ELF1 and FLI1 in JURKAT cells treated for 12 h with 5 μM DB2115. The dashed line represents the percentage of PU.1 lost or gained peaks found in MOLM13 cells. **j,k**, Scatter plots comparing log<sub>2</sub>fold change of PU.1 peaks (drug-treated/vehicle) versus GC content of the central 100 bp following 12 h of 5 μM DB2373, DB2313 and DB2826 (j), or 400 nM cytarabine (ARA-C) and 200 nM daunorubicin (DAUN.; k). Shown as an inset in each graph are the numbers and percentage of lost or gained PU.1 peaks,  $n = 1$ .

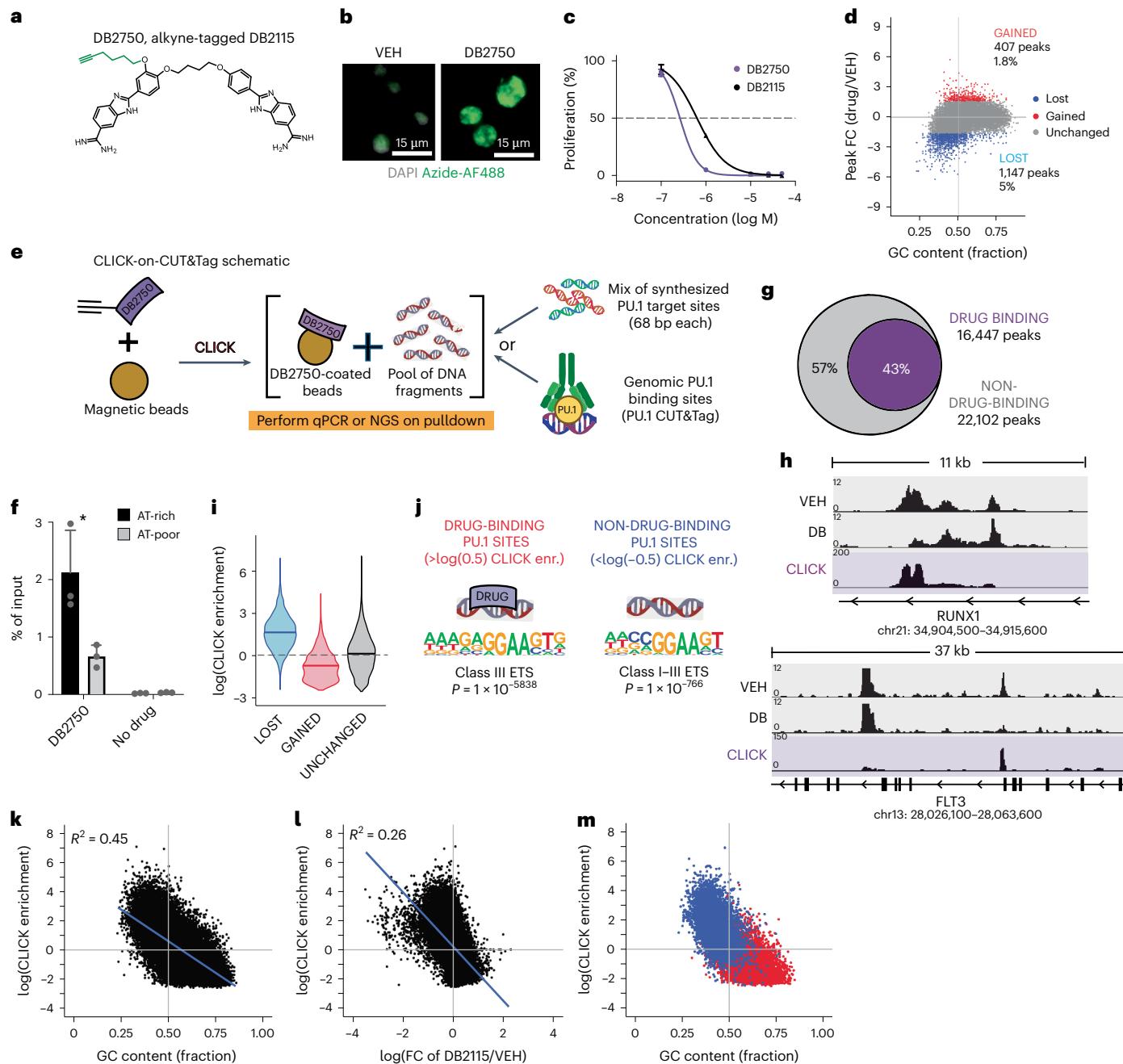
role in determining the chromatin accessibility changes of regions experiencing PU.1 redistribution.

To investigate the effects of PU.1 redistribution on the transcriptome, we performed RNA sequencing 20 h after exposure to DB2115 (Fig. 4a). Gene set enrichment analysis of differentially expressed genes identified the PU.1/ETS pathways as predominant upstream networks driven by DB2115 treatment (Fig. 4i,j). In addition, there was positive enrichment of differentiation signatures and negative enrichment of proliferative and stem cell signatures, which would be expected during myeloid differentiation driven by high levels of PU.1 (Fig. 4k).

Overlap of transcriptomic changes with PU.1 CUT&Tag was evaluated by examining gene expression near annotated CUT&Tag peaks (Extended Data Fig. 6b). PU.1-gained sites were enriched for up-regulation of gene expression, whereas PU.1-lost sites had an equivalent level of up-regulation and down-regulation (Fig. 4l). The low enrichment of down-regulated genes in PU.1-lost sites is probably a result of the time required to degrade pre-existing transcripts, with long half-life transcripts unlikely to show reductions at this time point.

To identify high-confidence redistributed PU.1 target genes, we filtered on PU.1-gained and PU.1-lost sites that occurred within close



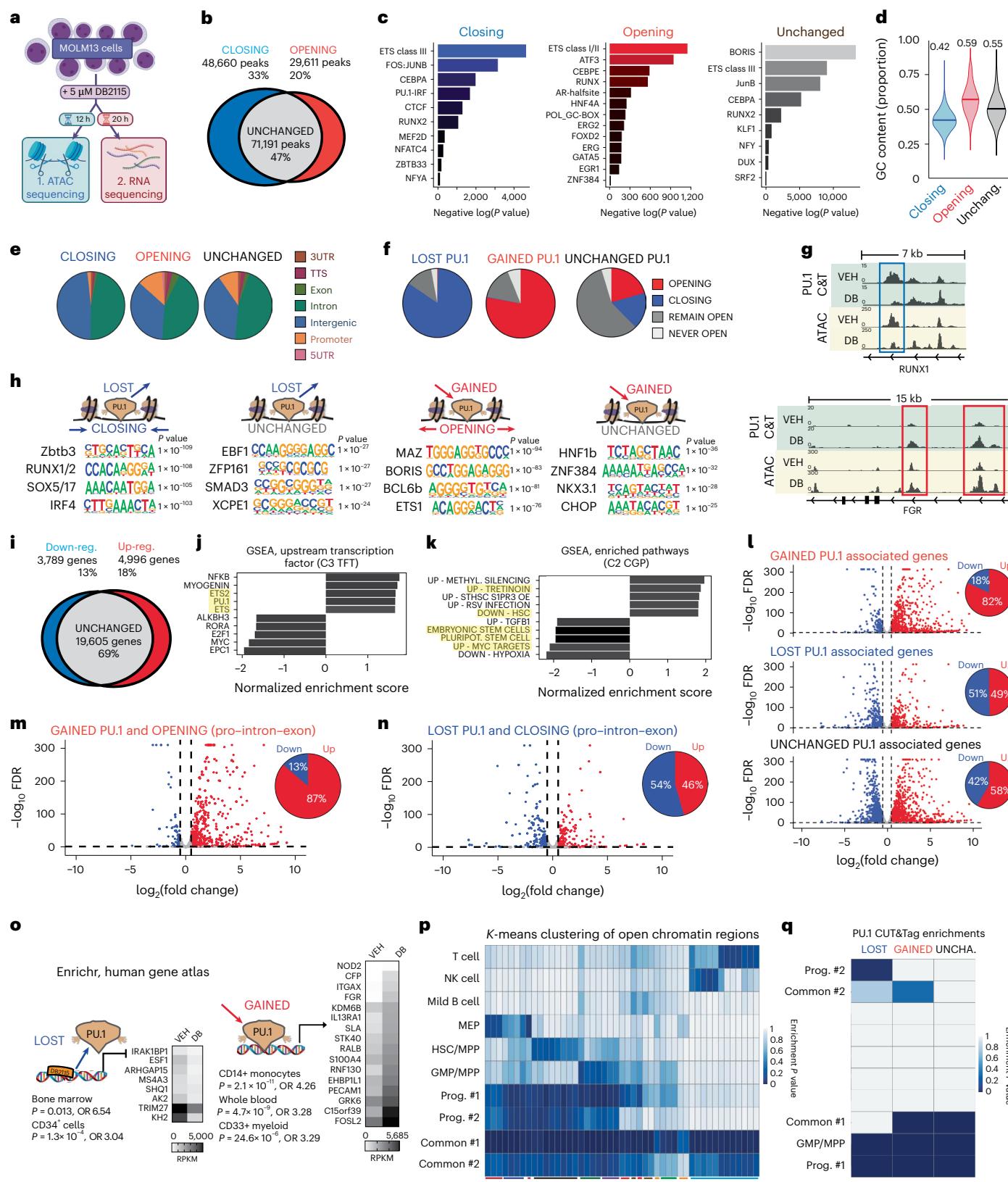


**Fig. 3 | CLICK-on-CUT&Tag identifies sensitive, A-rich PU.1 sites as direct targets of drug binding.** **a**, Chemical structure of DB2750, an alkyne-linker-tagged version of DB2115. **b**, Representative immunofluorescence image of MOLM13 cells treated with vehicle or 5  $\mu$ M DB2750, and CLICK-chemistry stained with azide-AF488 (green) and DAPI (gray). **c**, Dose-response curve of MOLM13 proliferation (cell titer blue assay) for DB2750 and DB2115;  $n = 3$  experimental replicates displaying mean  $\pm$  s.e.m. **d**, Scatter plots comparing  $\log_2$  fold change of PU.1 peaks (DB2115-treated/vehicle) with GC content of the central 100 bp following 12 h of DB2750. Shown as an inset in the graph are the number and percentage of lost or gained PU.1 peaks;  $n = 1$ . **e**, Schematic of CLICK-on-CUT&Tag and artificial fragment CLICK pulldown experimental procedures. NGS, next-generation sequencing. **f**, Enrichment of the two artificial DNA fragments (AT-rich and AT-poor; see Extended Data Fig. 2) following pulldown with DB2750-coated or non-drug-coated magnetic beads. Data are represented as a percentage of input cycle threshold score  $\pm$  s.e.m.;  $n = 3$  experimental replicates. \* $P < 0.05$  ( $P = 0.0312$ ) with two-sided Student's  $t$ -test. **g**, Proportion of DB2750 binding sites out of all PU.1 binding sites (drug binding was determined to be  $\log_2$  fold change  $> \log_2(0.5)$ ). **h**, Representative genome viewer tracks of vehicle (VEH), DB2115 (DB), and CLICK-on-CUT&Tag (purple), showing drug binding at specific PU.1 loci only. **i**, Summary enrichment scores of CLICK-on-CUT&Tag data, represented as the  $\log_2$  fold change of DB2750-pulldown/input. **j**, Highest ranked de novo motif found enriched in drug binding and non-drug-binding PU.1 sites. Class III ETS ( $P = 1 \times 10^{-5838}$ ) and Class I–III ETS ( $P = 1 \times 10^{-766}$ ).

percentage of input cycle threshold score  $\pm$  s.e.m.;  $n = 3$  experimental replicates, \* $P < 0.05$  ( $P = 0.0312$ ) with two-sided Student's  $t$ -test. **g**, Proportion of DB2750 binding sites out of all PU.1 binding sites (drug binding was determined to be  $\log_2$  fold change  $> \log_2(0.5)$ ). **h**, Representative genome viewer tracks of vehicle (VEH), DB2115 (DB), and CLICK-on-CUT&Tag (purple), showing drug binding at specific PU.1 loci only. **i**, Summary enrichment scores of CLICK-on-CUT&Tag data, represented as the  $\log_2$  fold change of DB2750-pulldown/input. **j**, Highest ranked de novo motif found enriched in drug binding and non-drug-binding PU.1 sites. Class III ETS ( $P = 1 \times 10^{-5838}$ ) and Class I–III ETS ( $P = 1 \times 10^{-766}$ ).

proximity to gene bodies (that is, promoter, intron, exon and transcription termination site regions only) and that resulted in respective changes in ATAC status or gene expression (that is, opening and up-regulation for gained PU.1 sites and vice versa) (Fig. 4m,n). This

produced lists of 194 PU.1-lost genes that showed enrichments in homeostatic and cell-cycle-related pathways and signatures of bone marrow and hematopoietic progenitor cells, and 506 PU.1-gained genes that showed enrichment in developmental and maturation pathways



and signatures of differentiated myeloid cells (Supplementary Data 3, Fig. 4o and Extended Data Fig. 6c).

To further investigate the signatures found in TF redistributor exposed cells, we generated a cellular identity map<sup>32</sup> (Extended Data Fig. 6d). This generated ten distinct open chromatin region clusters that associated with human hematopoietic cell types (Fig. 4p). By

applying our PU.1 CUT&Tag data, we identified a unique enrichment of lost PU.1 sites in the stem cell or progenitor cluster (progenitor no. 2), whereas gained PU.1 sites were enriched in ubiquitous regions found in all cell types, including myeloid cell types (common no. 1 and common no. 2). Thus, pharmacological PU.1 redistribution directs cells away from a progenitor or stem cell identity and towards a differentiated

**Fig. 4 | PU.1 repositioning restructures the accessible chromatin landscape and rewires the PU.1-driven transcriptome.** **a**, Schematic of the experimental approach to preparing MOLM13 cells for ATAC and RNA sequencing. **b**, Venn diagram of chromatin accessibility changes following DB2115 treatment (analyzed using Diffbind with an FDR < 0.1,  $n = 2$ ). **c**, De novo motifs found in closing, opening or unchanged accessible chromatin regions (binomial exact test, Homer). **d**, GC content of the central 100 bp of closing, opening and unchanged accessible chromatin. **e**, Annotation of closing, opening or unchanged accessible chromatin. **f**, Summary pie charts of lost, gained and unchanged PU.1 peaks displaying their chromatin accessibility status following treatment. **g**, Representative viewer tracks of genomic loci displaying lost (blue box), gained (red boxes) or unchanged (unmarked) PU.1 binding regions with both PU.1 CUT&Tag (top) and ATAC data (bottom). **h**, Enriched de novo motifs from the following categories of peaks: lost PU.1 and closing, lost PU.1 and unchanged, gained PU.1 and opening and gained PU.1 and unchanged (Homer analysis with background of lost PU.1 and unchanged for lost PU.1 and closing peaks (and vice versa) or background of gained PU.1 and unchanged for gained PU.1 and opening peaks (and vice versa)). **i**, Differentially expressed

genes (DEGs) identified following 20 h of DB2115 treatment (DESeq2,  $\log_2$  fold change  $>/< 0.5$  and FDR  $< 0.1$ ,  $n = 3$ ). **j,k**, Gene set enrichment analysis (GSEA; MSigDB) of upstream TF pathways (**j**) and chemical and genetic perturbations (**k**). **l**, Volcano plots of gained, lost and unchanged PU.1 peak-associated gene expression. Accompanying pie charts illustrate the proportions of DEGs that are up-regulated and down-regulated. **m,n**, Volcano plots of the transcriptional changes of gained (**m**) and lost (**n**) PU.1 peak-associated genes filtered on opening or closing chromatin, and in promoter–intron–exon regions only. Accompanying pie charts illustrate the proportion of DEGs that are up-regulated and down-regulated. **o**, The promoter–intronic–exonic DB2115-target genes that were lost and closing (194) or gained and opening (506) were analyzed for cell identity from the human gene atlas (Enrichr). Average expression,  $P$  value (Fisher's exact test) and odds ratio (OR) are shown. RPKM, reads per kilobase per million mapped reads. **p**,  $K$ -means clustering of binarized human cell ATAC peaks<sup>32</sup> uncovered ten clusters associating with differing cell identities. **q**, Using this  $K$ -means cluster as a reference, the 'identity' of PU.1 CUT&Tag lost, gained and unchanged peaks was determined; only clusters showing enrichment were labeled.

state (Fig. 4q). Taken together with the transcriptomic data, these findings indicate that DB2115 directly displaces PU.1 from progenitor or stem cell-essential gene regulatory regions, leading to PU.1 relocation at pro-differentiation regions.

### Temporally resolved consequences of redistributed PU.1

Although PU.1 has been described as possessing 'non-classical' pioneering abilities<sup>29</sup>, to our knowledge, the ability of PU.1 to pioneer chromatin under steady-state endogenous conditions has not been investigated. To understand the kinetics of PU.1 redistribution, we examined PU.1 localization after 1 h, 4 h and 12 h of DB2115 exposure. A total of 30,683 PU.1 binding regions were identified, with 11,487 sites changing in occupancy. A substantial proportion of these changes occurred after just 1 h (10%) or 4 h (22%) of incubation (Fig. 5a). These early changes predominantly consisted of PU.1 losses (Fig. 5a, right panels), but equilibrated to equivalent gains and losses at the later time points. Examination of the relationship of drug affinity or GC content over the PU.1 redistribution time course identified that the earliest losses exhibited the lowest GC content and highest CLICK enrichment score (Extended Data Fig. 7a,b, left panels), reflective of ordered and affinity-driven PU.1 displacement. Early PU.1 gains also exhibited lower GC content and higher CLICK enrichment, suggesting that PU.1 is ushered to the most GC-rich, drug-free PU.1 binding sites over time (Extended Data Fig. 7a,b, right panels). These findings suggest that PU.1 is rapidly displaced from chromatin by DB2115 but then requires a 'searching' period to locate and bind to alternative lower-affinity chromatin sites.

To understand the dynamics of chromatin accessibility changes in relation to PU.1 redistribution, we performed an identical DB2115 time course with ATAC-seq. Approximately 78,271 sites exhibited changing chromatin status, with a substantial proportion occurring

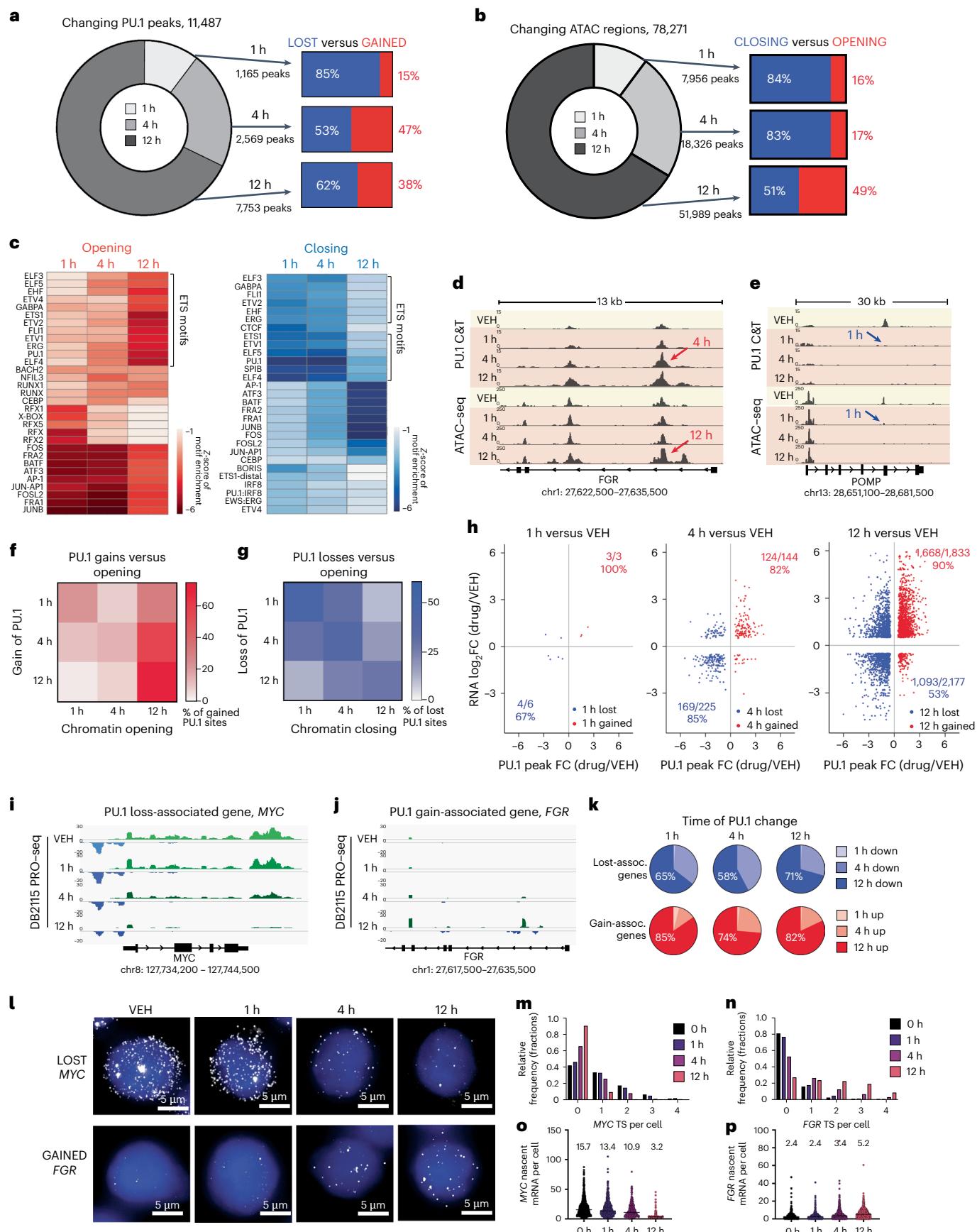
after just 1 h (10%) and 4 h (23%) (Fig. 5b). Interestingly, both the 1 h and 4 h changes predominated with closing events, but this equilibrated to equal opening and closing events by 12 h. Homer motif analysis revealed that the early closing regions and late opening regions were most highly enriched for ETS motifs (Fig. 5c).

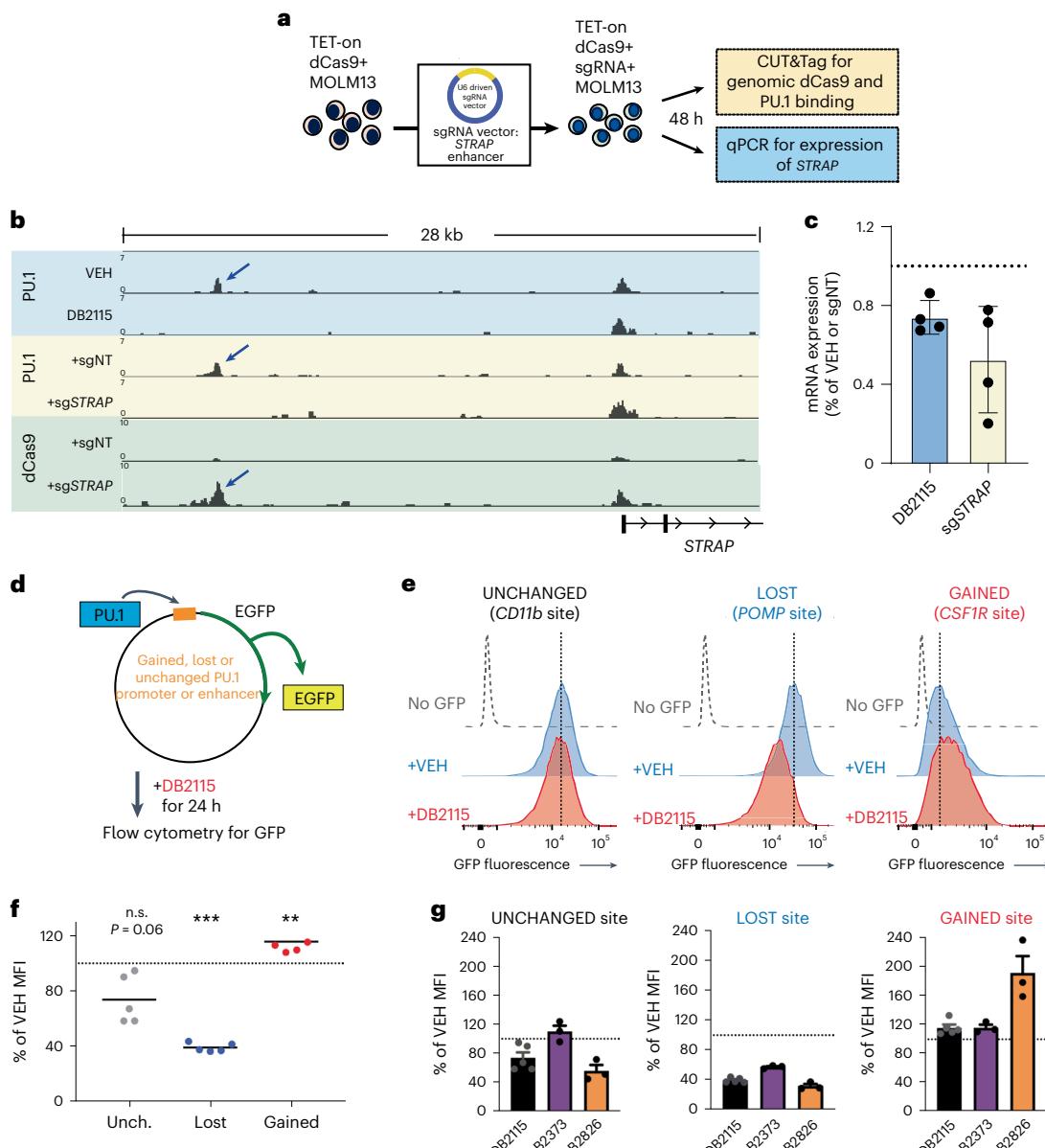
Focusing specifically on chromatin accessibility changes at PU.1 redistributed sites highlighted that many gained sites displayed PU.1 binding before chromatin opening (*FGR* and *CSF1R*; Fig. 5d and Extended Data Fig. 7c), whereas lost sites predominantly exhibited simultaneous closing (*RUNX1* and *POMP*; Fig. 5e and Extended Data Fig. 7c). Global examination of these redistribution versus chromatin accessibility events revealed that PU.1 gains were preceding the opening of the chromatin (Extended Data Fig. 7d). This effect was most noticeable when examining PU.1-gained sites at 4 h, with the majority opening at the later 12 h time point (Fig. 5f). Conversely, the relationship between loss of PU.1 and closing of chromatin appeared more synchronous (Extended Data Fig. 7e), with predominant simultaneous losses and closing (Fig. 5g).

To determine the effect of pharmacologically mediated PU.1 redistribution on nascent transcription, we performed precision nuclear run-on with sequencing (PRO-seq). Overall, the number of differentially transcribed genes increased from 185 to 2,358 to 8,387 at 1 h, 4 h and 12 h of DB2115 exposure, respectively (Extended Data Fig. 7f). Focusing on genes associated with changes in PU.1 binding, gained sites typically led to increases in transcription at all time points and vice versa for lost sites (Fig. 5h), consistent with our RNA sequencing data (Fig. 4m and Extended Data Fig. 7g). Redistributed PU.1-associated nascent transcripts changed at later time points than the corresponding changes in PU.1 binding or chromatin accessibility, with the majority of nascent RNA transcription fluctuations occurring at 12 h (Fig. 5i,j) and Extended Data Fig. 7h; for example, *MYC* and *FGR* loci in Fig. 5i,j).

**Fig. 5 | Temporally resolved pioneering of chromatin accessibility and nascent transcription by redistributed endogenous PU.1.** MOLM13 cells were treated with 5  $\mu$ M DB2115 for 1 h, 4 h or 12 h and CUT&Tag, ATAC-seq and PRO-seq were performed. **a**, Time of PU.1 peak changes and corresponding gains and losses at each time point. **b**, Time of chromatin accessibility changes and corresponding opening and closing at each time point. **c**, Known motif enrichment z-scores from closing and opening regions at 1 h, 4 h and 12 h (Homer analyses). **d,e**, Representative viewer tracks of genomic loci displaying both PU.1 CUT&Tag (top tracks) and ATAC-seq (bottom tracks) over the time course. Arrows indicate time of first detection of gain (**d**) or loss (**e**) of PU.1 or opening (**d**) and closing (**e**) chromatin. **f,g**, Heatmap depicting the time of chromatin opening for PU.1-gained sites (**f**) or depicting the time of chromatin closing for PU.1-lost sites over the time course (**g**). Sites were filtered to remove gained and lost PU.1 sites without changes in chromatin accessibility. **h**, Comparison of  $\log_2$  fold change of gene expression from PRO-seq data versus associated

PU.1 peak  $\log_2$  fold change from PU.1 CUT&Tag conducted over the time course. The proportions of reduced DEGs out of all lost PU.1-associated DEGs (blue font) and increased DEGs out of gained PU.1-associated DEGs (red font) are shown in each dot plot. **i,j**, Representative IGV tracks of positive and negative sense PRO-seq data displaying loss (**i**) or gain (**j**) of transcription over the time course. **k**, Cumulative pie charts depicting time of gene up-regulation or down-regulation from PRO-seq data grouped by time (1 h, 4 h or 12 h) of associated PU.1 loss or gain. **l**, Representative smFISH images from MOLM13 cells for the lost and gained PU.1-associated genes *MYC* (upper panels), and *FGR* (lower panels) over the DB2115 time course. *MYC* and *FGR* transcripts are in white pseudo-color; DNA is in blue pseudo-color. **m,n**, Transcription site (TS) burst frequency per cell for *MYC* (**m**) and *FGR* (**n**) over the time course. PU.1 CUT&Tag and ATAC sequencing peaks were created using Diffbind,  $n = 2$  and FDR  $< 0.1$ . PRO-seq DEGs were called using the NRSA pipeline;  $n = 2$  and  $P_{adj} < 0.05$ .





**Fig. 6 | Both genetic and pharmacological PU.1 redistribution activates gene expression at known and newly identified alternate PU.1 target sites.** **a**, Experimental schematic for generation and evaluation of dCas9 and sgRNA-expressing MOLM13 cells. **b**, Representative viewer tracks of genomic loci displaying PU.1 CUT&Tag of vehicle or DB2115-treated cells (top two tracks), sgNT-expressing or sgSTRAP-expressing cells (middle two tracks) and dCas9 CUT&Tag of sgNT-expressing and sgSTRAP-expressing cells (bottom two tracks). Arrows highlight the PU.1 or dCas9 binding at the STRAP enhancer. sgNT, non-targeting control sgRNA. **c**, Relative mRNA expression of STRAP by qPCR for DB2115-treated cells (for 24 h, compared to vehicle) and sgSTRAP dCas9+ cells (48 h of doxycycline, compared to sgNT dCas9+);  $n = 4$  experimental

replicates displaying mean  $\pm$  s.e.m. **d**, Experimental schematic for the design and evaluation of the native PU.1-driven eGFP reporter assay in MOLM13 cells. **e,f**, Representative eGFP fluorescence histograms (e) and summary mean fluorescence intensity (MFI) data for the three eGFP reporter-transduced MOLM13 cells (unchanged-CD11b site, lost-POMP site and gained-CSF1R site; see Extended Data Fig. 8b) following 24 h treatment with 5  $\mu$ M DB2115 or vehicle (f);  $n = 4-5$  experimental replicates, two-sided, paired Student's  $t$ -tests, \*\*\* $P < 0.001$  ( $P = 0.0096$ ), \*\* $P < 0.0005$  ( $P = 0.0005$ ). **g**, Mean change in eGFP MFI for the three reporter MOLM13 cell lines following 24 h treatment with 5  $\mu$ M DB2115 (black), DB2373 (purple) or DB2826 (orange) compared to vehicle  $\pm$  s.e.m.,  $n = 3-5$  experimental replicates.

To examine the transcriptional dynamics of PU.1 redistribution at single-molecule resolution, we performed targeted single-molecule RNA-fluorescence *in situ* hybridization (smFISH) for genes associated with loss (*MYC*), gain (*FGR*) or unchanged (*SPI1*) PU.1 binding (Fig. 5l and Extended Data Fig. 7i). For *MYC*, we observed significant reductions in transcription site burst frequency (Fig. 5m) and total nascent transcripts per cell (Fig. 5o). By contrast, for *FGR*, we observed significant increases in transcription site burst frequency (Fig. 5n) and total nascent transcripts per cell (Fig. 5p). *SPI1* nascent transcripts and transcription site burst frequency did not change at any time point

analyzed (Extended Data Fig. 7j,k). Overall, the dynamics of pharmacologically mediated PU.1 redistribution provide evidence for both the pioneering ability of endogenous PU.1 within native chromatin and, in turn, the ability of DB2115 to redirect transcriptional activity.

#### PU.1 redistribution activates genes at alternate PU.1 target sites

To identify direct regulation of PU.1 at candidate target sites, we used a genetic system to imitate pharmacological PU.1 displacement, whereby a catalytically dead Cas9 (dCas9) occupies *cis*-regulatory regions to

block TF binding<sup>33</sup>. We designed sgRNA against a DB2115-sensitive PU.1 binding site in the enhancer of the *STRAP* gene (Extended Data Fig. 8a) (Fig. 6a). Characteristic pharmacological displacement of PU.1 by DB2115 in MOLM13 cells was recapitulated in sg*STRAP*+ dCas9 cells (Fig. 6b, top and middle panels). A dCas9 peak was evident at the identical location as PU.1 loss, indicating dCas9-mediated PU.1 displacement. Quantification of *STRAP*mRNA confirmed a DB2115-like reduction in expression in sg*STRAP*+ dCas9 cells, thereby confirming that PU.1 displacement from *cis*-regulatory sites has direct effects on gene expression (Fig. 6c).

To further investigate the effects of PU.1 redistribution on target gene expression, we designed several genetic eGFP reporter systems in which eGFP fluorescence is driven by an endogenous PU.1 binding site (Fig. 6d). The PU.1 binding sites selected were from a gained (*CSF1R* alternative promoter), lost (*POMP* intergenic enhancer) and an unchanged (*CD11b* promoter) site (Extended Data Fig. 8b). Upon stable transduction of these reporters, we found that the baseline level of eGFP was different for each reporter but was reflective of our CUT&Tag results (Fig. 6e and Extended Data Fig. 8c). Upon DB2115 exposure, the lost-site reporter showed a robust decrease in eGFP fluorescence whereas the gained-site reporter showed a significant increase in fluorescence (Fig. 6f). Additionally, two second-generation compounds, DB2373 and DB2826, demonstrated comparable changes to eGFP expression in the three reporters (Fig. 6g). These data establish that pharmacological PU.1 redistribution has both positive and negative effects on gene expression.

### Myeloid differentiation is driven by PU.1 repositioning

Many myeloid surface receptor transcripts were directly up-regulated by pharmacologically induced PU.1 redistribution (Supplementary Data 3). To understand whether PU.1 redistribution can enhance protein expression or function, we investigated two myeloid and immune surface receptors, CSF1R (known PU.1 target) and IL-4R (non-canonical PU.1 target). Encouragingly, surface expression of CSF1R drastically increased from 0% to 30% (Fig. 7a) and IL-4R increased by 15% following DB2115 exposure (Fig. 7b).

To investigate the functional effects of pharmacological induction of CSF1R or IL-4R expression, we induced receptor expression with DB2115 before assessing ligand stimulation responses via phospho-protein flow cytometry (Fig. 7c). Vehicle-pretreated MOLM13 cells did not respond to CSF1 stimulation at any dose, whereas DB2115-pretreatment led to dose-dependent increases in pS6 levels (Fig. 7d). Similarly, IL-4 stimulations led to enhanced pSTAT6 activity in drug-invoked over vehicle-invoked MOLM13 cells (Fig. 7e). Next, we grew drug-invoked MOLM13 cells in methylcellulose containing either CSF1 or IL-4 for 1 week (Fig. 7f). DB2115-pretreatment led to reduced frequency and size of MOLM13 colonies compared to vehicle

(Fig. 7f). Addition of CSF1 or IL-4 increased the number of colonies in the DB2115-invoked group (Fig. 7g); however, the surviving cells displayed high expression of the myeloid marker CD15, indicating a commitment to differentiation (Extended Data Fig. 9a–e).

Next, we sought to investigate whether MOLM13 cells would differentiate following longer-term TF redistributor exposure. After a 5 day exposure to low-dose DB2115 in liquid culture, MOLM13 cells displayed myeloid differentiation, gaining macrophage-like morphology (Fig. 7h) and increased surface expression of myeloid markers CD11b and CD86, but not the stem cell marker CD34 (Fig. 7i,j). Consistently, THP1 cells displayed a similar myeloid differentiation phenotype under TF redistributor exposure, with macrophagic morphology (Extended Data Fig. 9f) and changes to myeloid surface marker expression (Extended Data Fig. 9g,h).

To test whether pharmacological repositioning of PU.1 can drive myeloid differentiation in primary cells, we examined the effect of TF redistributor exposure on seven samples from patients with AML or MDS that were grown in methylcellulose for 8–13 days. Cell growth and colony formation were severely hampered by DB2115 in a dose-dependent manner in all seven primary AML or MDS samples (Fig. 7l,m). Flow cytometry showed that the remaining viable DB2115-treated cells exhibited increases in myeloid surface markers, including CD11b, CD14, CD15 and CD86 (Fig. 7n and Extended Data Fig. 10a–c). Overall, this provides proof of concept that TF redistributors can be wielded to redeploy PU.1 to drive a differentiation program and induce cellular cytokine sensitivity.

### Discussion

In this study, we document evidence for TF redistribution caused by genome-wide, small-molecule blockade of *cis*-regulatory PU.1 binding sites (Fig. 8a). Consequentially, redistribution of PU.1 leads to robust and rapid changes in chromatin accessibility and rewires PU.1 transcriptional networks to drive increased expression of alternative PU.1 gene targets, which encourages myeloid differentiation of AML cells (Fig. 8b). Therefore, use of ‘TF redistributors’ such as DB2115, DB2313, DB2373 and DB2826 can provide unique insight into the behavior and dynamics of TFs while maintaining the stoichiometry of endogenous protein, and offers a molecular strategy to abrogate the aberrant transcriptional circuits in cancer and other diseases.

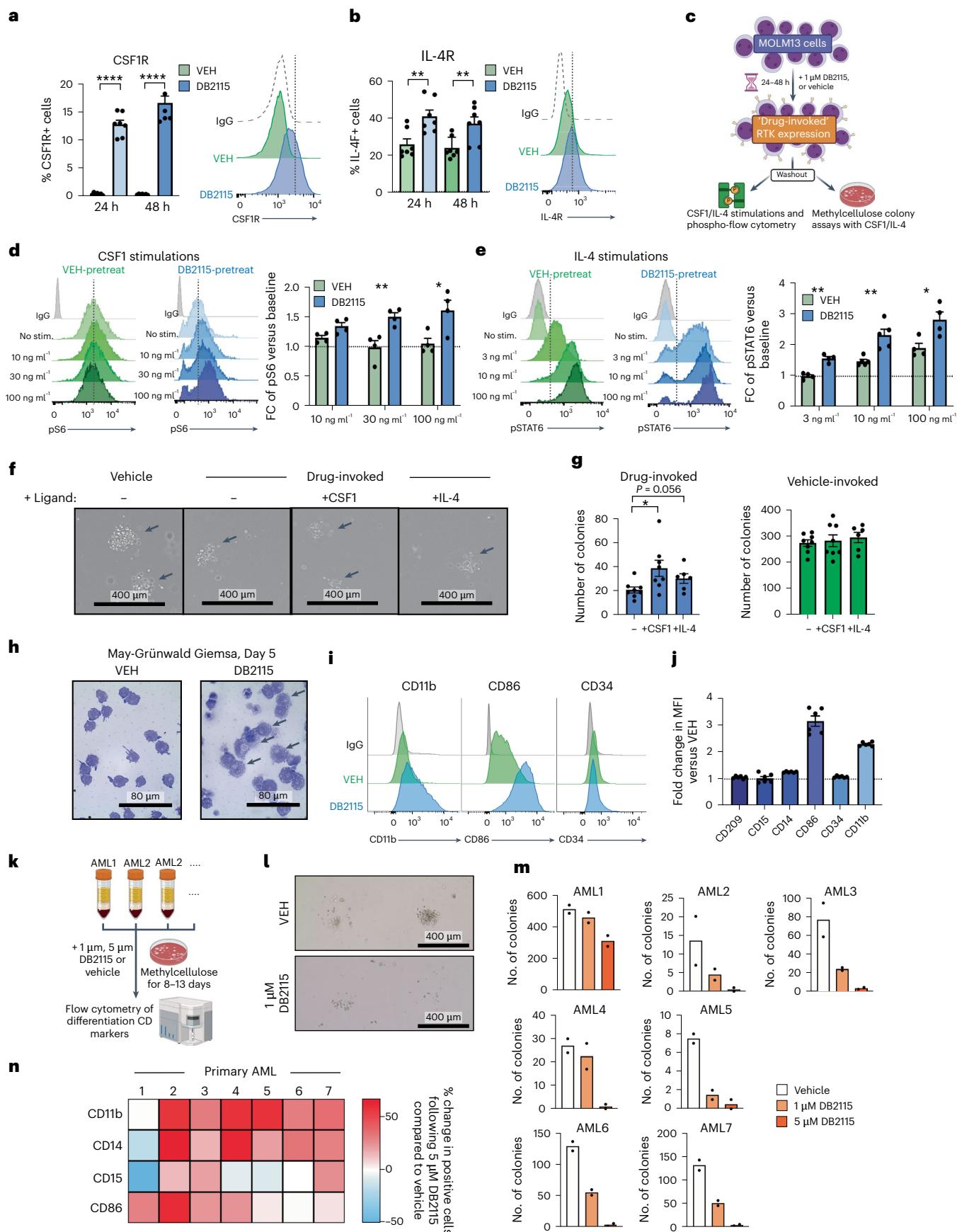
We identified that A-rich surrounding PU.1 binding motifs were most sensitive to inhibition by the described class of agents. One possible explanation for this could be that specific A-rich binding co-factors are aiding or stabilizing PU.1 binding at these sites. This could include the reported A-rich binding and PU.1 interacting TFs SATB1 (ref. 4), IRF4/8 (refs. 34,35) or C/EBP family factors<sup>36</sup>. It is possible that the anti-leukemic effects of the compounds are linked to the blockade of co-factor–PU.1 complexes forming on the DNA<sup>37,38</sup>. Thus, further

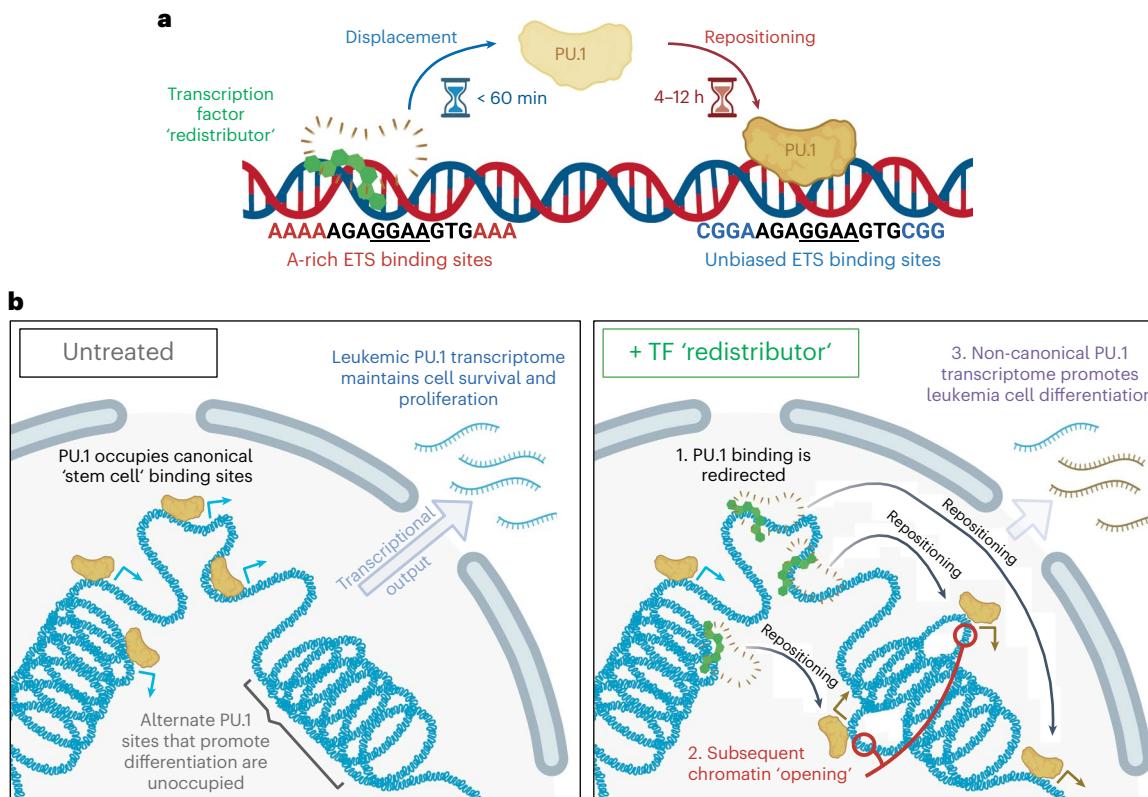
### Fig. 7 | Pharmacological TF redistribution induces myeloid lineage receptor responsiveness and promotes differentiation of leukemic cells.

**a,b**, Proportion of CSF1R+ (a) and IL-4R+ MOLM13 cells (b) after 24 h or 48 h of treatment with vehicle or 1 μM DB2115, and a representative histogram of staining intensity after 48 h; n = 7 experimental replicates displaying mean ± s.e.m. \*\*P < 0.01 (P = 0.0057 (24 h IL-4)) and P = 0.0096 (48 h IL-4)) and \*\*\*\*P < 0.0001. **c**, Schematic of approach to invoke surface receptor expression on MOLM13 cells and assess response to ligands (CSF1 and IL-4). **d**, Fold change of pS6 MFI versus baseline for CSF1 stimulations of drug-invoked MOLM13 cells, including representative histograms; n = 3 experimental replicates displaying mean ± s.e.m.; \*P < 0.05 (P = 0.0234) and \*\*P < 0.01 (P = 0.005). **e**, Fold change of pSTAT6 MFI versus baseline for IL-4 stimulations of drug-invoked MOLM13 cells, including representative histograms; n = 3 experimental replicates displaying mean ± s.e.m.; \*P < 0.05 (P = 0.0164) and \*\*P < 0.01 (P = 0.006 (3 ng ml<sup>-1</sup>) and P = 0.0025 (10 ng ml<sup>-1</sup>)). **f**, Representative day 7 colony assay images of drug-invoked or vehicle-invoked MOLM13 cells grown in methylcellulose containing no growth factors, +100 ng ml<sup>-1</sup>CSF1 or +100 ng ml<sup>-1</sup>IL-4. **g**, Summary D7 colony

numbers from DB2115-pretreated or vehicle-pretreated MOLM13 cells in the absence or presence of ligands, n = 3–4 experimental replicates displaying mean ± s.e.m., \*P < 0.05 (P = 0.0248). **h**, May–Grünwald Giemsa cytopsin image of MOLM13 cells treated with vehicle or 1 μM DB2115 for 5 days. The experiment was repeated independently three times with similar results. **i,j**, Representative histograms (i) and fold change in MFI versus vehicle for cell surface markers from 5-day treated cells, n = 7 experimental replicates displaying mean ± s.e.m. **j,k**, Experimental schema illustrating the evaluation of primary AML samples following vehicle or DB2115 exposure in methylcellulose colony assays.

**l**, Representative light microscope images of AML colonies (sample 6) after vehicle or 5 μM DB2115 treatment for 8 days. Example image is from one experiment with two technical replicates with similar results. **m**, Colony counts from samples from the seven patients with AML exposed to vehicle, 1 μM or 5 μM DB2115 for 8–13 days, n = 2 technical replicates per sample. **n**, Summary heatmap of the percentage change in CD marker expression in primary AML cells following 5 μM DB2115 compared to vehicle. All statistical tests displayed were unpaired, two-sided Student’s t-tests.





**Fig. 8 | The molecular mechanism of action and cellular consequences of pharmacological PU.1 redistribution.** **a**, TF 'redistributors' (for example, DB2115, DB2313, DB2373 or DB2826) directly and rapidly displace PU.1 from canonical adenine-rich ETS binding sites and subsequently redistribute it to second-tier, sequence-unbiased ETS binding sites. **b**, Under steady-state conditions, canonical PU.1 binding and the ensuing PU.1-driven transcriptome

is essential for leukemia cell survival; however, this is perturbed through the administration of TF redistributors. PU.1 is repositioned to alternate binding sites, redirecting its pioneer activity leading to subsequent chromatin opening and a rewiring of the PU.1-driven transcriptome, ultimately driving myeloid differentiation.

investigation is required to uncover which co-factors are essential for stabilizing PU.1 binding to adenine-biased sites and whether the co-factors themselves are sensitive to TF redistributors, thus contributing to the observed anti-leukemic effects of the drug class<sup>9</sup>.

As a direct result of PU.1 displacement, PU.1 was redeployed to alternative binding sites throughout the genome. It is noteworthy that the motifs derived from the localization data were in line with *in vitro* PU.1 affinity assays<sup>39,40</sup>, with GC-rich, low-affinity ETS sites only becoming populated with PU.1 once the high-affinity, A-rich ETS sites were blocked. Additionally, our temporal investigations into PU.1 redistribution provided evidence for the pioneering ability of endogenous PU.1, which supports the 'non-classical' pioneering function of PU.1 and dependence on other remodeling factors to modify the chromatin landscape<sup>1,41,42</sup>.

Interestingly, some PU.1 binding loci were unchanged following pharmacological binding site blockade despite these loci containing a typical PU.1 binding motif. There are a number of possibilities to explain this finding, including that PU.1 co-factors or nearby chromatin binding proteins may stabilize the PU.1-DNA interaction and prevent pharmacological displacement (this could include any of the known PU.1 interactors; for example, HOX, IRF or C/EBP family factors)<sup>35,38,43,44</sup> or that the drug is unequally distributed throughout the chromatin landscape. It has been reported that pharmacological compounds can be hindered in their ability to freely diffuse throughout the dense and fractal nuclear compartments<sup>45,46</sup>. Additional research will be required to interrogate the underlying forces driving sensitivity and insensitivity to PU.1 displacement.

TF redistribution is a biological phenomenon that can be driven by changes in the chromatin context and/or the protein milieu<sup>4,7,47</sup>.

This study has highlighted an interesting form of TF redistribution: direct pharmacological TF displacement and redistribution, which not only has therapeutic potential but also allows for rapid and tunable investigations into the fast biology and temporal chromatin dynamics of endogenous TFs. Pharmacological redistribution of TF localization could be applied to investigate the global behaviors of many TFs, including those that have been documented to be displaced by DNA binding site inhibitors (for example EVI1, PIT-1 and HER2 (refs. 48–50)) or those unexplored disease regulatory factors which rely upon AT-rich minor-groove contacts (for example, FOXA1 in breast cancer<sup>51,52</sup>).

Although still in its infancy, the potential for TF redistributors to be used for therapeutic application is highly attractive and theoretically plausible. Pharmacological targeting of transcriptional dysregulation found in diseases such as MDS and AML has proven enigmatic, with most approaches limited to targeting protein–protein interactions in the chromatin rather than TF–DNA interactions<sup>24,53–55</sup>. Initial pre-clinical investigations in mouse models of AML have demonstrated that these tool TF redistributors display on-target efficacy for eliminating leukemia<sup>9</sup>, warranting further development of next-generation versions of these agents pursuing superior pharmacokinetics. Additionally, it will be important to determine how best to use the molecular PU.1 redistribution phenomenon therapeutically and investigate whether a combination with other drugs can synergize to better control the aberrant PU.1 transcriptional network. Thus, with further investigation, TF redistributors offer an exciting potential therapeutic class of small molecules that could have far-reaching effects for the treatment of leukemias and other aberrant TF-mediated diseases.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-024-01911-7>.

## References

1. Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 (2018).
2. Perry, J. A., Seong, B. K. A. & Stegmaier, K. Biology and therapy of dominant fusion oncproteins involving transcription factor and chromatin regulators in sarcomas. *Annu. Rev. Cancer Biol.* **3**, 299–321 (2019).
3. Henley, M. J. & Koehler, A. N. Advances in targeting ‘undruggable’ transcription factors with small molecules. *Nat. Rev. Drug Discov.* **20**, 669–688 (2021).
4. Hosokawa, H. et al. Transcription Factor PU.1 represses and activates gene expression in early Tcells by redirecting partner transcription factor binding. *Immunity* **48**, 1119–1134.e7 (2018).
5. Zhang, P. et al. PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood* **96**, 2641–2648 (2000).
6. Rekhtman, N., Radparvar, F., Evans, T. & Skoultchi, A. I. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev.* **13**, 1398–1411 (1999).
7. McKenzie, M. D. et al. Interconversion between tumorigenic and differentiated states in acute myeloid leukemia. *Cell Stem Cell* **25**, 258–272.e9 (2019).
8. Munde, M. et al. Structure-dependent inhibition of the ETS-family transcription factor PU.1 by novel heterocyclic diamidines. *Nucleic Acids Res.* **42**, 1379–1390 (2014).
9. Antony-Debré, I. et al. Pharmacological inhibition of the transcription factor PU.1 in leukemia. *J. Clin. Invest.* **127**, 4297–4313 (2017).
10. Békés, M., Langley, D. R. & Crews, C. M. PROTAC targeted protein degraders: the past is prologue. *Nat. Rev. Drug Discov.* **21**, 181–200 (2022).
11. Ungerbäck, J. et al. Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). *Genome Res.* **28**, 1508–1519 (2018).
12. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–1577 (1994).
13. Stephens, D. C. et al. Pharmacologic efficacy of PU.1 inhibition by heterocyclic dicitations: a mechanistic analysis. *Nucleic Acids Res.* **44**, 4005–4013 (2016).
14. Ogbonna, E. N. et al. Drug design and DNA structural research inspired by the Neidle laboratory: DNA minor groove binding and transcription factor inhibition by thiophene diamidines. *Bioorg. Med. Chem.* **68**, 116861 (2022).
15. Sive, J. I. et al. Genome-scale definition of the transcriptional programme associated with compromised PU.1 activity in acute myeloid leukaemia. *Leukemia* **30**, 14–23 (2016).
16. Will, B. et al. Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. *Nat. Med.* **21**, 1172–1181 (2015).
17. Steidl, U. et al. A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. *J. Clin. Invest.* **117**, 2611–2620 (2007).
18. Steidl, U. et al. Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. *Nat. Genet.* **38**, 1269–1277 (2006).
19. Bonadies, N., Pabst, T. & Mueller, B. U. Heterozygous deletion of the PU.1 locus in human AML. *Blood* **115**, 331–334 (2010).
20. Mueller, B. U. et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood* **100**, 998–1007 (2002).
21. Viny, A. D. et al. Cohesin members Stag1 and Stag2 display distinct roles in chromatin accessibility and topological control of HSC self-renewal and differentiation. *Cell Stem Cell* **25**, 682–696.e8 (2019).
22. Mizuki, M. et al. Suppression of myeloid transcription factors and induction of STAT response genes by AML-specific Flt3 mutations. *Blood* **101**, 3164–3173 (2003).
23. Vangala, R. K. et al. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood* **101**, 270–277 (2003).
24. Kühn, M. W. M. et al. Targeting chromatin regulators inhibits leukemogenic gene expression in NPM1 mutant leukemia. *Cancer Discov.* **6**, 1166 (2016).
25. Gu, X. et al. Leukemogenic nucleophosmin mutation disrupts the transcription factor hub that regulates granulomonocytic fates. *J. Clin. Invest.* **128**, 4260–4279 (2018).
26. Ralvenius, W. T. et al. A novel molecular class that recruits HDAC/MECP2 complexes to PU.1 motifs reduces neuroinflammation. *J. Exp. Med.* **220**, e20222105 (2023).
27. Munde, M., Poon, G. M. K. & Wilson, W. D. Probing the electrostatics and pharmacological modulation of sequence-specific binding by the DNA-binding domain of the ETS family transcription factor PU.1: a binding affinity and kinetics investigation. *J. Mol. Biol.* **425**, 1655–1669 (2013).
28. Wontakal, S. N. et al. A large gene network in immature erythroid cells is controlled by the myeloid and B cell transcriptional regulator PU.1. *PLoS Genet.* **7**, e1001392 (2011).
29. Minderjahn, J. et al. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat. Commun.* **11**, 402 (2020).
30. Pham, T. H. et al. Mechanisms of in vivo binding site selection of the hematopoietic master transcription factor PU.1. *Nucleic Acids Res.* **41**, 6391–6402 (2013).
31. Liu, Y. et al. Heterocyclic diamidine Interactions at AT base pairs in the DNA minor groove: effects of heterocycle differences, DNA AT sequence and length. *J. Phys. Chem. B* **112**, 11809 (2008).
32. Corces, M. R. et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* **48**, 1193–1203 (2016).
33. Shariati, S. A. et al. Reversible disruption of specific transcription factor-DNA interactions using CRISPR/Cas9. *Mol. Cell* **74**, 622–633.e4 (2019).
34. Escalante, C. R. et al. Crystal structure of PU.1/IRF-4/DNA ternary complex. *Mol. Cell* **10**, 1097–1105 (2002).
35. Kanno, Y., Levi, B. Z., Tamura, T. & Ozato, K. Immune cell-specific amplification of interferon signaling by the IRF-4/8-PU.1 complex. *J. Interferon Cytokine Res.* **25**, 770–779 (2005).
36. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
37. Konstantinos Tzelepis, A. et al. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep.* **17**, 1193–1205 (2016).
38. Cao, Z. et al. ZMYND8-regulated IRF8 transcription axis is an acute myeloid leukemia dependency. *Mol. Cell* **81**, 3604 (2021).
39. Poon, G. M. K. & Macgregor, R. B. Base coupling in sequence-specific site recognition by the ETS domain of murine PU.1. *J. Mol. Biol.* **328**, 805–819 (2003).
40. Poon, G. M. K. Sequence discrimination by DNA-binding domain of ETS family transcription factor PU.1 is linked to specific hydration of protein-DNA interface. *J. Biol. Chem.* **287**, 18297–18307 (2012).

41. Ghisletti, S. et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32**, 317–328 (2010).
42. Barozzi, I. et al. Co-regulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol. Cell* **54**, 844 (2014).
43. Al Sadoun, H., Burgess, M., Hentges, K. E. & Mace, K. A. Enforced expression of Hoxa3 inhibits classical and promotes alternative activation of macrophages in vitro and in vivo. *J. Immunol.* **197**, 872–884 (2016).
44. Hohaus, S. et al. PU.1(Spi-1) and C/EBP $\alpha$  regulate expression of the granulocyte-macrophage colony-stimulating factor receptor  $\alpha$  gene. *Mol. Cell. Biol.* **15**, 5830–5845 (1995).
45. Bancaud, A. et al. Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J.* **28**, 3785–3798 (2009).
46. Wang, T. et al. Chemical-induced phase transition and global conformational reorganization of chromatin. *Nat. Commun.* **14**, 5556 (2023).
47. Bell, C. C. et al. Targeting enhancer switching overcomes non-genetic drug resistance in acute myeloid leukaemia. *Nat. Commun.* **10**, 2723 (2019).
48. Zhang, Y. et al. Targeting a DNA binding motif of the EVI1 protein by a pyrrole-imidazole polyamide. *Biochemistry* **50**, 10431–10441 (2011).
49. Chiang, S. Y. et al. Targeting the Ets binding site of the HER2/neu promoter with pyrrole-imidazole polyamides. *J. Biol. Chem.* **275**, 24246–24254 (2000).
50. Peixoto, P. et al. Direct inhibition of the DNA-binding activity of POU transcription factors Pit-1 and Brn-3 by selective binding of a phenyl-furan-benzimidazole dication. *Nucleic Acids Res.* **36**, 3341–3353 (2008).
51. Cirillo, L. A. & Zaret, K. S. Specific Interactions of the wing domains of FOXA1 transcription factor with DNA. *J. Mol. Biol.* **366**, 720 (2007).
52. Fu, X. et al. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. *Proc. Natl Acad. Sci. USA* **113**, E6600–E6609 (2016).
53. Lu, R. & Wang, G. G. Pharmacologic targeting of chromatin modulators as therapeutics of acute myeloid leukemia. *Front. Oncol.* **7**, 241 (2017).
54. Xu, B. et al. Selective inhibition of EZH2 and EZH1 enzymatic activity by a small molecule suppresses MLL-rearranged leukemia. *Blood* **125**, 346–357 (2015).
55. Zhang, S., Liu, M., Yao, Y., Yu, B. & Liu, H. Targeting LSD1 for acute myeloid leukemia (AML) treatment. *Pharmacol. Res.* **164**, 1043–6618 (2021).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024

## Methods

### Study approval

Human samples were collected after informed consent was obtained from patients and upon approval of the Institutional Review Board of Albert Einstein College of Medicine (protocol 2005–536).

### Small molecules

The synthesis of DB2115 (ref. 56) and DB2313 (ref. 9) has been previously reported; the synthesis of DB2373, DB2826 and DB2750 is provided in the Supplementary Information<sup>57,58</sup>. Following synthesis, all compounds were dissolved in 10% dimethylsulfoxide and water at a stock concentration of 5 mM and stored at –20 °C.

### Cell culture and drug treatment

THP1, MOLM13, MV411, JURKAT and HL60 cells were originally purchased from ATCC. MOLM13, THP1, JURKAT and MV411 were cultured in RPMI 1640 medium supplemented with 10% heat-inactivated FBS and 1% penicillin–streptomycin. HL60 were cultured in IMDM medium supplemented with 20% heat-inactivated FBS and 1% penicillin–streptomycin. All cell lines were maintained in an incubator at 37 °C and 5% CO<sub>2</sub>.

Cells were treated with 5 µM of DB2115, DB2313, DB2373 or DB2750 at a starting concentration of one million cells per ml for the specified time.

### CUT&Tag

CUT&Tag was performed as previously reported<sup>59</sup> but with a few technical alterations. In brief, 5 × 10<sup>5</sup> cells per cell line were collected and, to prevent the osmotic displacement of PU.1 (or other factors) from chromatin during the high-salt downstream protocol<sup>60</sup>, lightly fixed with 2% formaldehyde for 2 min. The cells were bound to Concanavalin A-coated beads (Bangs Laboratories) and incubated with the primary antibodies (anti-PU.1, Santa Cruz, sc-352; anti-RUNX1, Cell Signaling, 4334S; anti-ELF1, Proteintech, 22565-1-AP; anti-FLI1, Invitrogen, PA5-29597; anti-GAPBA, Invitrogen, PA5-27735; and anti-GATA3, Cell Signaling, 5852T) or IgG control (Santa Cruz, sc-3888; Extended Data Fig. 1e) at 4 °C overnight. Samples were then incubated with a secondary antibody (guinea pig anti-rabbit; Antibodies Online, ABIN101961) followed by a pre-loaded pA-Tn5 adaptor complex (generated in-house). Tagmentation buffer with magnesium was used to induce fragmentation. DNA was extracted by phenol–chloroform–isoamyl alcohol and amplified with NEBNext HiFi 2×PCR Master Mix and universal i5 and barcoded i7 primers<sup>61</sup> for 13 cycles. AMPure XP beads (A63880) were used for post-PCR clean-up of libraries. Libraries were subjected to 35-bp paired-end sequencing using an Illumina NextSeq 500 platform on high output mode at the Einstein Epigenomics Core. Fastq files were generated using Picard tools (v.2.17.1), with adaptor trimming by TrimGalore (v.0.3.7) and QC assessment using FASTQC (v.0.11.4).

### CLICK-on-CUT&Tag

To determine binding preferences for DB2115 within the genome, we examined an alkyne-linker-tagged DB2115 (called DB2750) for in vitro DNA binding ability. Preparation of drug-coated magnetic beads was performed as previously described<sup>62</sup>, whereby a 25 µM solution of DB2750 in H<sub>2</sub>O was incubated with 150 µM Biotin Azide (Click Chemistry tools), 100 µM E301 (Sigma-Aldrich) and 4 mM CuSO<sub>4</sub> for 1 h at 4 °C before adding streptavidin-coupled Dynabeads (MyOne Streptavidin T1, Invitrogen) and incubating for 1 h. Coated beads were washed and incubated overnight with either a 25 mM equal mix of the two synthesized 68-bp dsDNA fragments (AT-rich and AT-poor; see Supplementary Table 1b) or DNA libraries from PU.1 CUT&Tag of vehicle-treated MOLM13 cells. Beads were washed three times with 25 mM Tris-HCl, 1% Triton-X, 0.5% SDS, 150 mM NaCl and 5 mM EDTA wash buffer before being eluted in 50 µl of 0.1 M NaHCO<sub>3</sub>, 0.1% SDS solution. QiaQuick PCR Purification Kit (Qiagen) was used to purify

fragments. Synthesized fragment purified samples were quantified by quantitative PCR with reverse transcription (RT-qPCR) (see Supplementary Table 1b for AT-rich and AT-poor primers). CUT&Tag pull-down samples were amplified with NEBNext HiFi 2× PCR Master Mix and universal i5 and barcoded i7 primers<sup>61</sup> for five cycles before being sequenced with Illumina NextSeq 500 35-bp PE sequencing (Einstein Epigenomics Core). Bioinformatic processing was performed identically to CUT&Tag.

### Classical ChIP

ChIP was performed similarly to the previously described protocol<sup>9</sup>. In brief, MOLM13 cells were treated with 5 µM DB2115 for 12 h before cells were crosslinked in 0.5% formaldehyde for 7 min at room temperature (25 °C). Samples were sonicated for 12 min before being incubated with 5 µg of PU.1 antibody (Santa Cruz, sc-352) or rabbit IgG (Santa Cruz, sc-3888). Also included was Activ-motif *Drosophila* normalization control DNA and antibody (cat. nos. 53083 and 61686). Following pulldown with Dynabeads Protein A (Thermo Fisher Scientific), DNA fragments were purified using QIAquick PCR Purification Kit (Qiagen). Library preparation was carried out using the NEB Ultra II DNA Prep Kit following the manufacturer's recommendations. Next-generation sequencing was performed with Illumina NextSeq 500, 75-bp paired-end reads on high output at Einstein Epigenomics Core. Fastq files were generated using Picard tools (v.2.17.1), with adaptor trimming by TrimGalore (v.0.3.7) and QC assessment using FASTQC (v.0.11.4).

### Cell titer blue proliferation assay

To assess the IC<sub>50</sub> of the small molecules in MOLM13 cells, we used the resazurin cell viability assay (Cell Titer Blue, Promega) following the manufacturer's guidelines. In brief, cells were plated at a density of 1 × 10<sup>3</sup> per 100 µl and incubated with vehicle or compounds for 48 h at a range of concentrations (0.01–50 µM) before 20 µl of Cell Titer Blue reagent was added and fluorescence (560<sub>EX</sub>/590<sub>EM</sub>) was measured using a FLUOstar Omega instrument (BMG Labtech).

### Omni-ATAC-sequencing

Omni-ATAC-seq was performed as previously described<sup>63</sup>. In brief, 50,000 MOLM13 cells treated with either 5 µM DB2115 or vehicle were subjected to nuclei isolation with ATAC resuspension buffer (containing 0.1% NP-40, 0.1% Tween20 and 0.01% digitonin) before transposition was performed using Tn5 transposase in TD buffer (Illumina Tagment DNA kit). Transposed DNA was purified using Minelute PCR purification (Qiagen) before DNA was amplified by PCR with custom PCR primers<sup>64</sup> and cycle number was determined using a KAPA Library Quantification Kit (KAPA Biosystems). Libraries were sequenced on an Illumina NextSeq 500, with 75-bp paired-end reads on high output at Einstein Epigenomics Core. Fastq files were generated using Picard tools (v.2.17.1), with adaptor trimming by TrimGalore (v.0.3.7) and QC assessment using FASTQC (v.0.11.4).

### GFP reporter assays

Reporter constructs were designed around the promoter–enhancer cassette as previously described<sup>8,9</sup>. In brief, a minimal promoter (TATA box) was placed upstream from an open reading frame encoding destabilized D<sub>2</sub>EGFP. The original enhancer element was a five-repeat tandem of the λB motif of the murine Igλ2-4 enhancer. For this study, the λB motifs were replaced by POMP, CSF1R-alt or NS sequences (Supplementary Table 1c). PU.1 binding site underlined between NdeI and SdAI/PstI cleavage sites upstream of the TATA box. For the myeloid CD11b promoter, the entire regulatory region, including the TATA box, was replaced by the –412 to +98 sequence of the 5' flanking region of the human CD11b gene (GenBank accession no. M84477.1)<sup>64</sup>. The cassettes plus open reading frames were inserted between the NdeI/EcoRI sites of pLJM1–EGFP, a puromycin-selectable lentiviral vector (Addgene, plasmid no. 19319)<sup>65</sup>. The inserts replaced the

CMV-EGFP construct in the vector and the resultant plasmids were denoted with the suffix pLJM1. All insertions were sequenced-verified by Sanger sequencing in the forward direction with a hU6-F primer (5'-GAGGGCCTATTCATGATT-3') and/or in the reverse direction with a standard EGFP-NR primer.

Next, 10 µg of pLJM1 vectors were transfected to 293T cells spread in a 100 mm culture dish with helpers (10 µg psPAX2 and 4 µg VSV-G) using lipofectamine 2000 reagent (Invitrogen). Virus-containing supernatant was collected and filtered at 48 h from transfection and was transduced into MOLM13 cells by centrifuging at 1,000g, 37 °C, 1 h with Transdux reagent (System Biosciences). Following 3 days of puromycin selection, pLJM1 MOLM13 cell lines were treated for 24 h with 5 µM DB2115 (or other drugs) before analysis on an LSRII flow cytometer was performed (Becton–Dickinson). Forward scatter and side scatter gated, DAPI-negative cells were quantified for GFP fluorescence using FlowJo software.

### CRISPRd: dCas9 blocking studies

Lentiviral vectors for constitutive TET3G expression driven by a CMV promoter (hygromycin selection marker), and TRE3G-inducible dCas9 (blasticidin selection marker) were designed and produced by VectorBuilder and transduced using Transdux (System Biosciences) into MOLM13 cells. Successful transduction was selected with 100 µg ml<sup>-1</sup> hygromycin and 10 µg ml<sup>-1</sup> blasticidin treatment for 1 week. Expression of dCas9 was confirmed following 48 h of 1 µg ml<sup>-1</sup> doxycycline by western blot; however, partial or leaky dCas9 expression was evident in untreated cells. A constitutive sgRNA-expressing lentiviral vector against *STRAP* or a non-targeting control were synthesized by Cellecta and subsequently transduced into TET3G-dCas9-expressing MOLM13 cells. Selection was performed for 3 days post transduction with puromycin before cells were treated with 1 µg ml<sup>-1</sup> doxycycline for 48 h and collected for CUT&Tag or RT-PCR analysis.

### Western blotting

Whole-cell extracts were prepared using lysis buffer (150 mM NaCl, 50 mM Tris-HCl, 5 mM EDTA, 1% NP-40, 1% phosphatase inhibitor cocktail, 1× Protease Inhibitor Cocktail, 1 mM phenylmethyl sulfonyl fluoride (PMSF), 10% glycerol). SDS-PAGE was performed with equal protein per sample. PVDF membranes were used for western blot transfer, and immunoblotting using primary antibodies (1:5,000 for actin, Sigma-Aldrich, A2066; 1:1,000 for PU.1, Santa Cruz, sc-352; 1:500 Cas9, Cell Signaling, 14697S) followed by HRP-conjugated secondary antibodies (1:5,000). Imaging of western blots was performed using chemiluminescent ECL substrate on a LI-COR Odyssey Fc imager.

### CLICK-chemistry immunofluorescence staining

MOLM13 cells were attached to lysine-coated coverslips following 6 h of 5 µM DB2750 or vehicle treatment and subjected to processing as directed by a Click-iT EdU imaging Kit (Invitrogen) but with a substitution of DB2750 for the Click-Edu reagent. In brief, cells were fixed (3.7% formaldehyde) and permeabilized (0.5% Triton-X) before being stained with a Click-iT reaction mix containing Alexa Fluor 488 Azide. Prolong gold containing DAPI (Thermo Fisher) was added to stain DNA before imaging on an Olympus BX83 microscope with an X-Cite 120 PC lamp (EXFO) and an ORCA-R2 digital charge-coupled device camera (Hamamatsu) with a ×100 objective.

### RT-qPCR and mRNA sequencing

RNA was extracted from 5 µM DB2115-treated cells using an RNeasy Mini Kit (Qiagen) and RT-qPCR was performed using iScript cDNA Synthesis Kit (Bio-Rad). qPCR reactions were performed on a ViiA7 instrument using Power SYBR Green PCR Master Mix (Thermo Fisher). Expression levels were normalized to *GAPDH*, and the primers used are listed in Supplementary Table 1a. For mRNA sequencing, library preparation and transcriptome sequencing were performed by Novogene.

### Primary AML samples

Adult AML or MDS bone marrow and peripheral blood samples were obtained after written informed consent and following Albert Einstein College of Medicine Institutional Review Board approval (2005–536). Characteristics of the primary MDS and AML samples used for experiments are listed in Supplementary Table 1f.

For CUT&Tag experiments, primary AML cells were subjected to CD34<sup>+</sup> column enrichment (Miltenyi Biotech) before being viably frozen. Upon thawing, one to two million CD34<sup>+</sup> enriched cells were thawed and incubated in 1 ml StemSpan SFEMII medium plus 1× CC100 growth cocktail (StemCell Tech) and TPO (50 ng ml<sup>-1</sup>, R&D Systems) in the presence of 5 µM DB2115 or vehicle for 12 h before viable cells were collected for PU.1 CUT&Tag.

For colony assays, 50–100,000 unfractionated mononuclear AML cells were plated per ml of Methocult H4435 enriched medium (StemCell Tech) along with 1 µM or 5 µM of DB2115 or vehicle. After 8–13 days, colonies were counted and cells were collected for flow cytometry.

### Surface receptor induction colony assays

For MOLM13 colony assays, cells were first ‘invoked’ for receptor expression with a 1 µM dose of DB2115 or vehicle for 24 h in 10% FBS, RPMI medium. After a washout, 3,000 cells were immediately plated per ml of methylcellulose in HSC002SF (R&D Systems) containing added recombinant IL-4 (100 ng ml<sup>-1</sup>, R&D Systems) or CSF1 (100 ng ml<sup>-1</sup>, R&D Systems) and scored after 7 days.

Light microscopy images to detail colony morphology were obtained using an EVOS FL Auto microscope (Life Technologies) with an objective at ×4 or ×10 magnification.

**Cytomorphology.** Cell morphology was assessed after cytopsin of 100,000 cells onto a glass slide (5 min at 500 rpm) and May–Grünwald Giemsa staining, according to standard protocols. Images were obtained using an EVOS FL Auto microscope (Life Technologies) with an objective at ×10 or ×50 magnification.

### Flow cytometry

Antibody staining was performed for 30 min before flow cytometry analyses, using the following Biolegend antibodies: CD14 FITC (no. 325603), CD15 APCCy7 (no. 323048), CD86 PEcy7 (no. 374210), CD34 PE (no. 343506), CD209 AF647 (no. 330112), CD11b PerCP Cy5.5 (no. 101228), CSF1R APC (no. 347306), IL-4 RBV421 (no. 355014) and DAPI for alive cell gating. Flow cytometry analysis was performed on a BD LSRII containing a yellow laser running FACSDiva 8 software. Flow cytometry data was analyzed using FlowJo (v.10.6.1). Log<sub>10</sub> scales are used for plots showing flow cytometry fluorescence intensity data.

### Phospho-flow cytometry

MOLM13 cells were ‘invoked’ for surface receptor tyrosine kinase (RTK) expression with a 1 µM DB2115 treatment for 24 h, followed by a washout and incubation with media-only for 24 h. Cells were then incubated for 1 h in serum-free IMDM at 37 °C before a 5 min stimulation with recombinant mouse huCSF-1 (10–100 ng ml<sup>-1</sup>, R&D Systems) or huLL-4 (3–100 ng ml<sup>-1</sup>, R&D Systems). After fixation with 1% PFA and permeabilization with ice-cold acetone, intracellular staining for phospho-S6 ribosomal protein (608604, Biolegend) and pSTAT6 (686012, Biolegend) before flow cytometry analysis.

### PRO-seq

Initially, 30 million MOLM13 cells were treated with 5 µM DB2115 or vehicle and collected at the indicated time points for nuclei extraction. Cells were washed with ice-cold PBS and lysed with cell lysis buffer (10 mM Tris-Cl pH 7.4, 300 mM sucrose, 3 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 0.5% NP-40, 5 mM dithiothreitol, 1 mM PMSF, EDTA-free protease cocktail inhibitor tablet). Using Dounce homogenization, nuclei were pelleted by centrifugation and washed with nuclei storage buffer (50 mM Tris-Cl

pH 8.3, 40% glycerol, 5 mM MgCl<sub>2</sub>, 5 mM dithiothreitol, 0.1 mM EDTA, 1 mM PMSF, EDTA-free protease cocktail inhibitor tablet). After counting, pelleted nuclei were resuspended in storage buffer and stored at -80 °C.

PRO-seq was performed in two biological replicates as previously described using approximately 20 million nuclei per run on with GTP, ATP, UTP and biotin-11-CTP (PerkinElmer) using 0.5% Sarkosyl (Fisher Scientific) to prevent transcription initiation<sup>66–68</sup>. RNA was reversed-transcribed and amplified to make the cDNA library for sequencing by the Vanderbilt University Medical Center (VANTAGE Genome Sciences Shared Resource on an Illumina NovaSeq 6000 (PE-150, 50 million reads). Following adaptor trimming with Cutadapt (v.1.18), these sequences were aligned and mapped using bowtie2 (v.2.5.1). Samtools (v.1.9) was used for the file format conversion before using the Nascent RNA Sequencing Analysis (NRSA)<sup>67</sup> pipeline to determine the gene body changes.

### smFISH

To design mRNA-specific probes for sequential smFISH, a full-length transcript of each gene was used as input for PaintSHOP (<https://paintshop.io>)<sup>69</sup> to retrieve 10–29 primary targeting sequences (23–39 bp). Sequences were screened for off-target activity using NCBI Blast (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Selected sequences were then concatenated on both 5' and 3' ends with flanking 20-mer sequences (RO1, ATACTGGAGCGACGCGTGAT; RO2, GTTTGAAGATTG-GACCTGGA; or RO4, CTAAGGTACCTAATTGCCTAG), generating a final 'primary probe' (Supplementary Table 1d,e). SmFISH immunofluorescence staining procedure and analysis were performed as described previously<sup>70</sup>. In brief, MOLM13 cells treated with 5 μMDB2115 for 0 h, 1 h, 4 h and 12 h were fixed and permeabilized before primary and secondary hybridization reactions were performed. Samples were mounted in Prolong Diamond Antifade reagent plus DAPI (Invitrogen). Images were acquired using an oil immersion ×100 objective on a Leica Thunder fluorescence microscope. For data analysis, single-molecule mRNA and transcription site detection was performed using FISH-quant<sup>71</sup> by 3D Gaussian fitting of thresholded spots implemented in MATLAB R2024a.

### Data processing

CUT&Tag-generated fastq files were mapped to the human genome (hg38) using bowtie2 (v.2.2.3) with options –end-to-end –very-sensitive –no-mixed –no-discordant –phred33 –I 10 –X 700. Normalized bed-graph files were generated by using bedtools genomecov with a normalization factor of 1,000,000/o. total human reads. Peak calling was performed using bdgpeakcall from MACS2 (v.2.1.0) with options –I 100 and –c 2. Bigwig files were generated using bedGraphToBigWig for visualization in IGV\_2.4.15. Motif analysis and peak annotation was performed using the Homer package (v.4.11.1)<sup>36</sup>.

Differential peak analysis was performed in RStudio with either the Diffbind package (v.3.12.0) using DESEQ2 analyses<sup>72</sup> if  $n > 2$  or a custom script we termed 'GoodpeaksScript' (<https://github.com/steidl-lab/rePU.1sitioning>) if  $n = 1$ . In Diffbind, the FDR cutoff was set to 0.1 for determining significantly changed peaks, and technical variability between replicates is displayed in PCA and heatmap plots (Extended Data Fig. 1a,b). Additionally, Diffbind analysis was performed on a randomization of vehicle and DB2115 pairs, with no differential peaks identified (FDR < 0.1), providing confidence in the redistribution phenomenon (Extended Data Fig. 1c,d). For GoodpeakAnalysis, three stringent filters were used for the differential peak analysis of the average peak intensity: minimum intensity of >7.5, minimum fold change of >4 and minimum summit of >3.

For CLICK-on-CUT&Tag, the average MACS2 peak scores across three replicates were calculated and compared to input vehicle-treated PU.1 CUT&Tag peak scores to generate a log<sub>2</sub> fold change CLICK score. Peaks with CLICK scores above 0.5 were considered enriched for drug binding, whereas scores below 0.5 were considered non-drug-binding.

Classical ChIP-generated fastq files were mapped to the human or *Drosophila melanogaster* genome (hg38 or dm3) using bowtie2 (v.2.2.3). Duplicates were then removed with Picard, and normalization was performed by subsampling .bam files by a *Drosophila* reads ratio of vehicle over DB2115. The MACS2 package and callpeaks function was used to identify peaks in normalized files, and bigwig files were generated for Integrative Genomics Viewer (IGV) visualization. Differential peak analysis was performed in RStudio with 'GoodpeakAnalysis'.

For RNA sequencing datasets, quality control was performed based on error distribution along the length of reads, GC distribution, N content, base quality and adaptor content. Reads were mapped to the hg38 transcriptome using STAR aligner<sup>73</sup>. Raw counts were subsequently normalized and analyzed for differential expression in R using the Bioconductor package DESeq2 (ref. 74). An enrichment score was generated using the negative logarithm of the adjusted *P* value multiplied by the sign of the fold change for each gene and input into Fast Gene Set Enrichment Analysis (FGSEA, Bioconductor<sup>75,76</sup>). Pre-ranked gene lists were queried against standard c1–8 and hallmark MSigDB gene lists (v.7.4, Broad Institute<sup>77</sup>). Additional gene list enrichments were conducted with Enrichr<sup>78–80</sup>.

For the *k*-means ATAC cell identity mapping, raw ATAC sequencing counts were obtained for healthy donor cell populations<sup>32</sup>. The raw counts were normalized using variance stabilizing transformation (DESeq2)<sup>74</sup>, and each locus was then binarized into open or closed chromatin regions using mean normalized count followed by *k*-means clustering of the binarized data (Extended Data Fig. 5). The loci in each cluster was overlapped with the PU.1 CUT&Tag data, and the significance of the overlap was calculated using hypergeometric distribution *P* value.

### Statistics and reproducibility

Statistical tests were performed in RStudio (2023.12.1+402) or GraphPad Prism (v.9.5). No statistical method was used to predetermine sample size. No data were excluded from the analyses. The investigators were not blinded to allocation during experiments and outcome assessment. The experiments were not randomized.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All datasets are available online in the Gene Expression Omnibus under SuperSeries GSE267389. The hg38 human genome dataset was used for alignment and analysis and is available at [https://www.ncbi.nlm.nih.gov/datasets/genome/GCF\\_000001405.26](https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.26). Source data are provided with this paper.

### Code availability

Custom code used to generate results in this study is available via GitHub at <https://github.com/steidl-lab/rePU.1sitioning> or Zenodo<sup>81</sup>.

### References

56. Liu, Y. et al. Designed compounds for recognition of 10 base pairs of DNA with two at binding sites. *J. Am. Chem. Soc.* **134**, 5290–5299 (2012).
57. Casitas, A., Canta, M., Solà, M., Costas, M. & Ribas, X. Nucleophilic aryl fluorination and aryl halide exchange mediated by a Cu<sup>I</sup>/Cu<sup>II</sup> catalytic cycle. *J. Am. Chem. Soc.* **133**, 19386–19392 (2011).
58. Ates-Alagöz, Z. et al. Synthesis and potent antimicrobial activities of some novel retinoidal monocationic benzimidazoles. *Arch. Pharm. (Weinheim)* **339**, 74–80 (2006).
59. Kaya-Okur, H. S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).

60. Xhani, S., Esaki, S., Huang, K., Erlitzki, N. & Poon, G. M. K. Distinct roles for interfacial hydration in site-specific DNA recognition by ETS-family transcription factors. *J. Phys. Chem. B* **121**, 2748 (2017).
61. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
62. Tyler, D. S. et al. Click chemistry enables preclinical evaluation of targeted epigenetic therapies. *Science* **356**, 1397–1401 (2017).
63. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
64. Pahl, H., Rosmarin, A. & Tenen, D. Characterization of the myeloid-specific CD11b promoter. *Blood* **79**, 865–870 (1992).
65. Sancak, Y. et al. The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* **320**, 1496–1501 (2008).
66. Zhang, S. et al. PAX3-FOXO1 coordinates enhancer architecture, eRNA transcription, and RNA polymerase pause release at select gene targets. *Mol. Cell* **82**, 4428–4442.e7 (2022).
67. Wang, J. et al. Nascent RNA sequencing analysis provides insights into enhancer-mediated gene regulation. *BMC Genomics* **19**, 633 (2018).
68. Zhao, Y. et al. High-resolution mapping of RNA polymerases identifies mechanisms of sensitivity and resistance to BET inhibitors in t(8;21) AML. *Cell Rep.* **16**, 2003–2016 (2016).
69. Hershberg, E. A. et al. PaintSHOP enables the interactive design of transcriptome- and genome-scale oligonucleotide FISH experiments. *Nat. Methods* **18**, 937 (2021).
70. Wheat, J. C. et al. Single-molecule imaging of transcription dynamics in somatic stem cells. *Nature* **583**, 431–436 (2020).
71. Imbert, A. et al. FISH-quant v2: a scalable and modular tool for smFISH image analysis. *RNA* **28**, 786–795 (2022).
72. Stark, R. & Brown, G. DiffBind: differential binding analysis of ChIP-Seq peak data. *Bioconductor* <https://bioconductor.org/packages/release/bioc/html/DiffBind.html> (2021).
73. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
74. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
75. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at <https://doi.org/10.1101/060012> (2021).
76. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
77. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
78. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
79. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
80. Xie, Z. et al. Gene set knowledge discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).
81. Stauber, J. steidl-lab/rePU1sitioning: rerePU1sitioning. Zenodo <https://zenodo.org/records/13313910> (2024).

## Acknowledgements

We thank A. Skoultschi, W. D. Wilson, B. Will and L. LaFave for insightful discussion and manuscript review, as well as S. Heaton for providing us with the recombinant pA-Tn5 used in the CUT&Tag experiments. We thank the Einstein Epigenomics and Flow Cytometry Cores, as well as the Stem Cell Isolation and Xenotransplantation Core Facility (funded by NYSTEM grant no. C029154) of the Gottesman Institute for Stem Cell Research and Regenerative Medicine for expert services and support. This work was supported by National Institutes of Health grants R35CA253127 (to U.S.), HL155178 (to G.M.K.P.) and a National Science Foundation grant MCB2028902 (to G.M.K.P.). S.J.T. was supported by a Young Investigator award from the Edward P. Evans Foundation, a Leukemia & Lymphoma Society Fellowship, a NYSCF Druckenmiller Fellowship and the Einstein Training Program in Stem Cell Research from the Empire State Stem Cell Fund through NYSDOH Contract C3029266. R.K. was supported by a Leukemia & Lymphoma Society Fellowship. S.S. was supported by grants from the National Cancer Institute (K00CA223044) and the Leukemia & Lymphoma Society (3439-25). This work was supported by Jane A. and Myles P. Dempsey. U.S. holds the Edward P. Evans Endowed Professorship in Myelodysplastic Syndromes at Albert Einstein College of Medicine. The Endowed Professorship was supported by a grant from the Edward P. Evans Foundation. Some panels in Figs. 1, 4, 7 and 8 were created with [Biorender.com](https://biorender.com).

## Author contributions

U.S. and S.J.T. conceived the study. S.J.T., U.S., O.B., G.M.K.P., J.C.W., S.S., G.T., J.C., K.R.S., M.G. and A.V. developed the methodology. A.A.F., A.K. and D.W.B. performed chemical synthesis. S.J.T., O.B. and G.M.K.P. performed the investigation. S.J.T., B.A.B., J.S., E.S., J.C. and R.K. conducted the formal data analysis. S.J.T. and U.S. wrote and edited the paper. S.J.T., J.S., R.K. and E.S. were involved with data visualization. U.S. supervised the project and acquired funding.

## Competing interests

U.S. has received research funding from GlaxoSmithKline, Bayer Healthcare, Aileron Therapeutics and Novartis; has received compensation for consultancy services and for serving on scientific advisory boards from GlaxoSmithKline, Bayer Healthcare, Celgene, Aileron Therapeutics, Stelexis Therapeutics and Pieris Pharmaceuticals; and has equity ownership in and is serving on the board of directors of Stelexis Therapeutics. The other authors declare no competing interests.

## Additional information

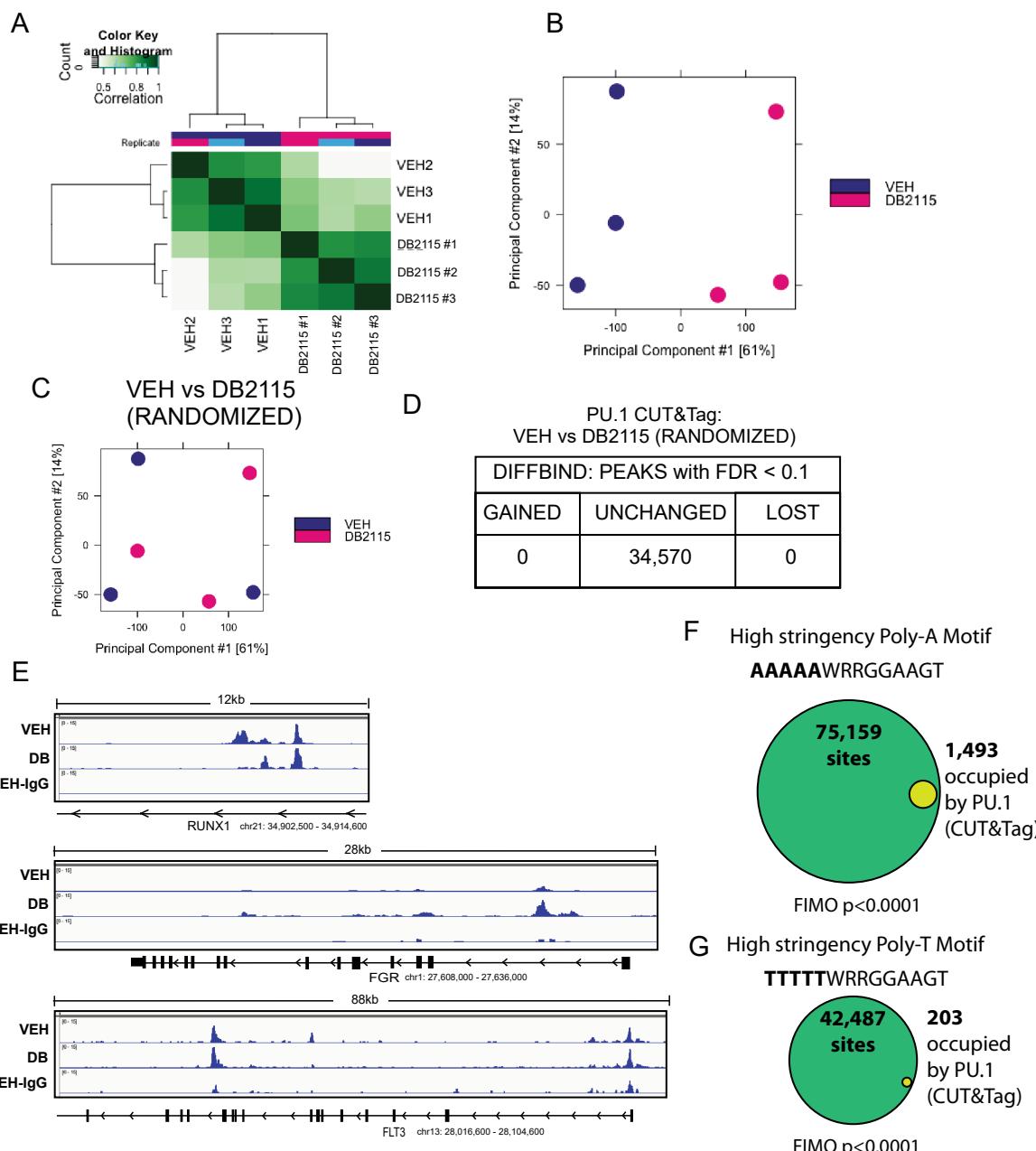
**Extended data** is available for this paper at <https://doi.org/10.1038/s41588-024-01911-7>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41588-024-01911-7>.

**Correspondence and requests for materials** should be addressed to Ulrich Steidl.

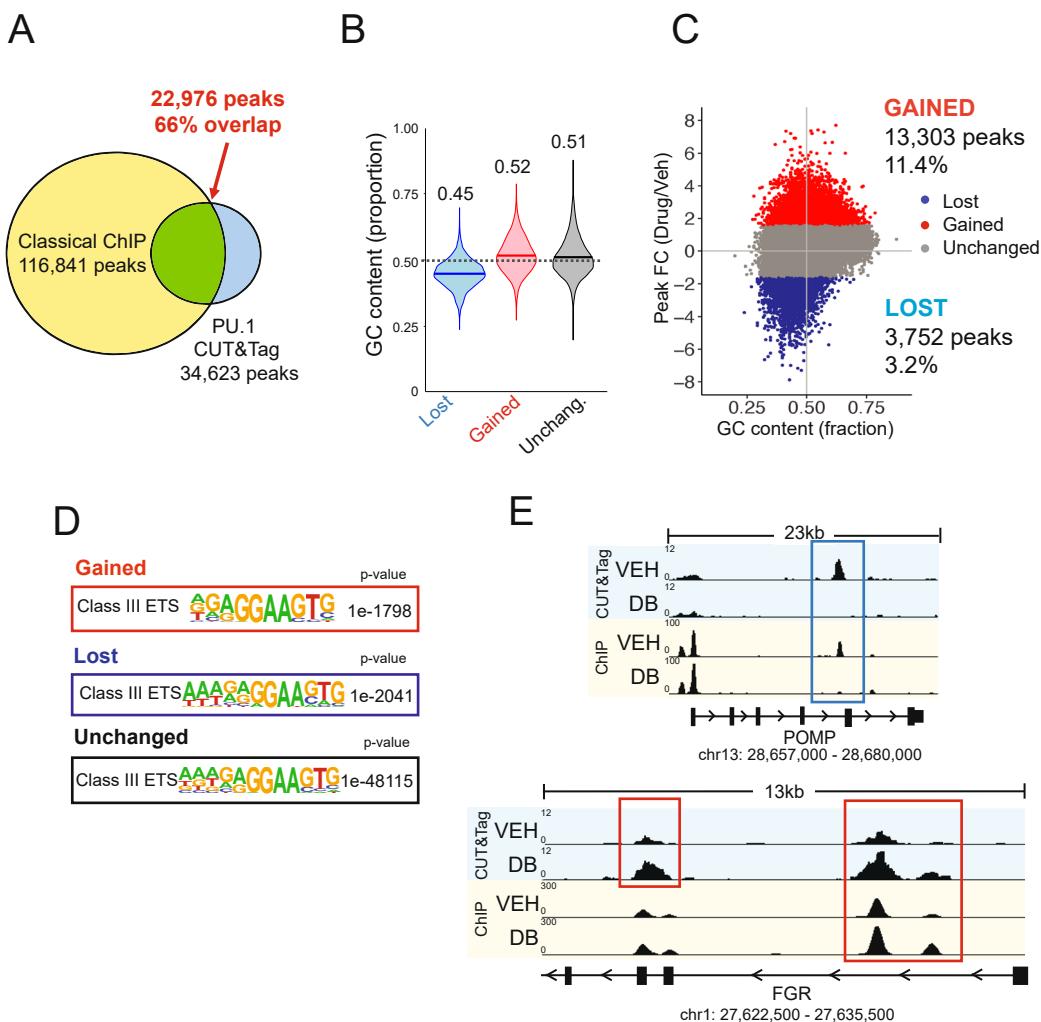
**Peer review information** *Nature Genetics* thanks Jason Carroll and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).



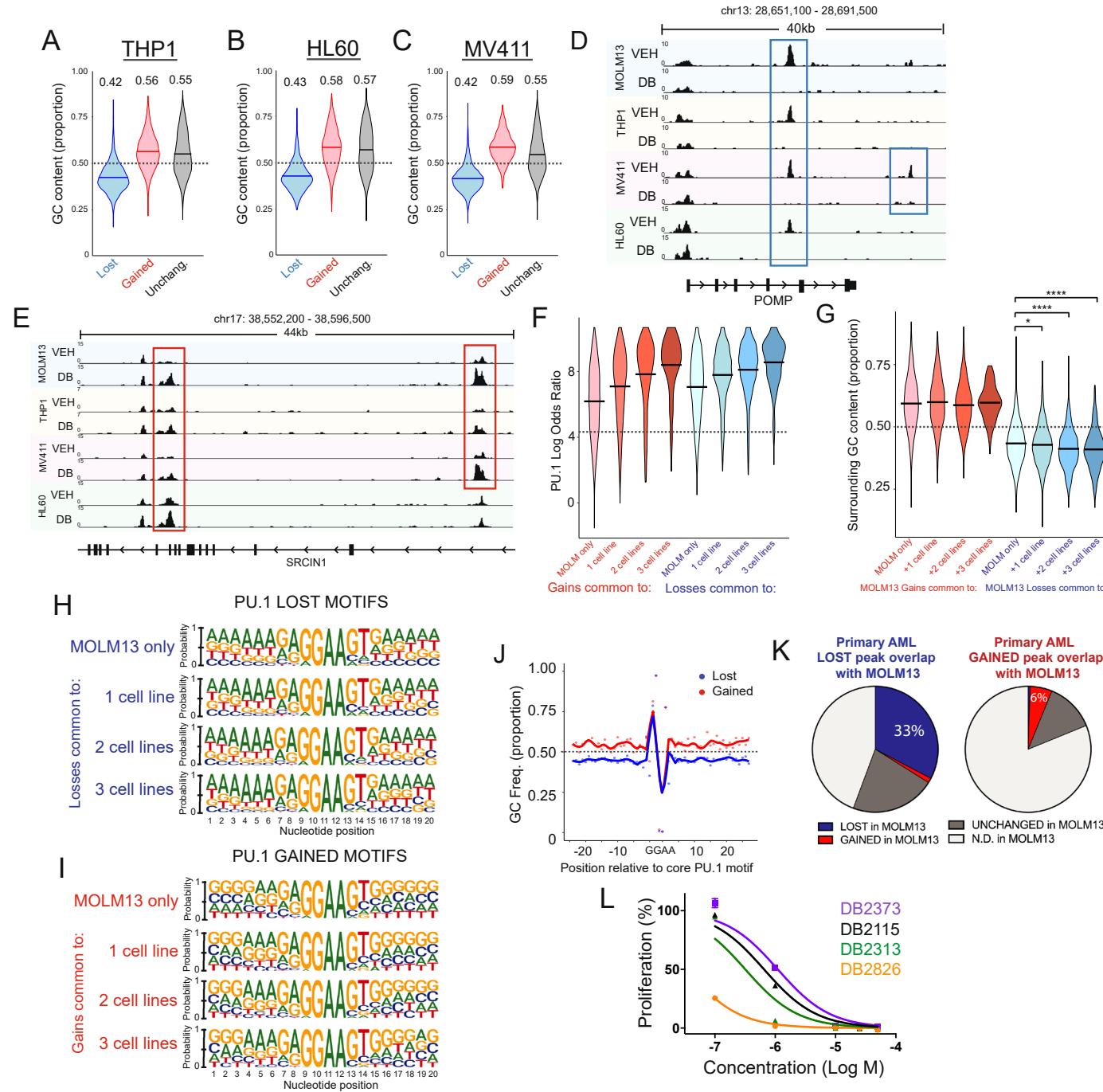
**Extended Data Fig. 1 | PU.1 redistribution in MOLM13.** (a) Correlation heatmap of the differentially bound PU.1 peaks from Diffbind analysis of MOLM13 VEH vs DB2115 treated samples. (b) PCA plot showing association between replicates of VEH vs DB2115 treated MOLM13 differential peaks from Diffbind analysis. (c) Randomization of the VEHvsDB2115 treatment pairs to determine Diffbind peak calling robustness. PCA plot of differential peaks from randomized treatment pairs and (d) numbers of significantly changed peaks, FDR < 0.1, are

shown. (e) Representative viewer tracks of genomic loci displaying lost, gained and unchanged PU.1 binding from Fig. 1g, plus an additional track displaying minimal reads detected from IgG CUT&Tag. The FIMO tool (MEME suite, p-value cut-off of  $p < 0.0001$ ) was used to identify (f) poly-A upstream and (g) poly-T upstream PU.1 motifs (AAAAAWRRGGAAAGT and TTTTTWRRGGAAAGT respectively) in the entire human (hg38) genome. Also shown is the overlap between these genomic sites and total PU.1 CUT&Tag sites.



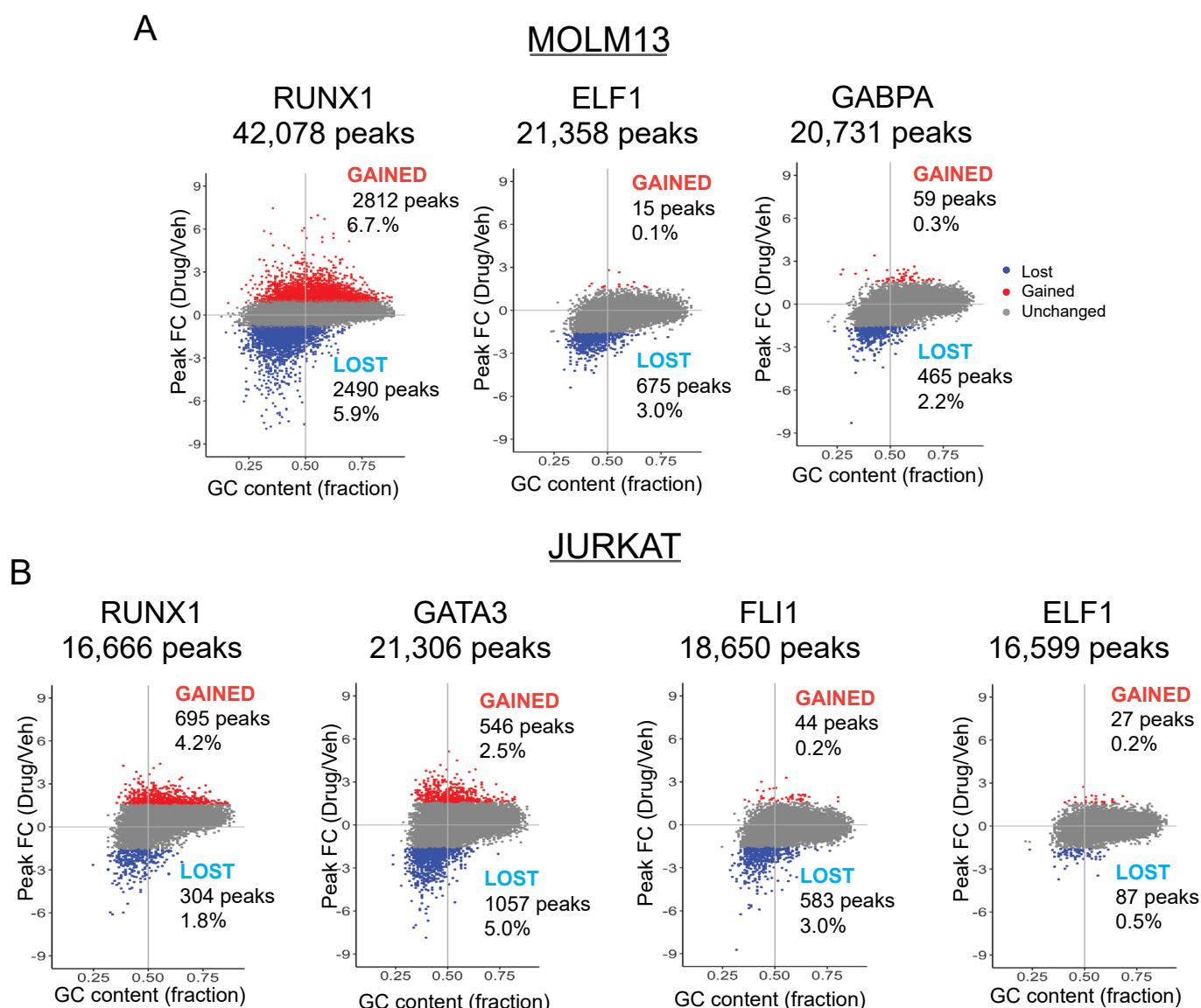
**Extended Data Fig. 2 | Classical ChIP examination of PU.1 redistribution in MOLM13.** (a) Summary of the overlap of MACS2 called peaks between classical PU.1 ChIP ( $n = 1$ ) and PU.1 CUT&Tag ( $n = 3$ ). (b) The proportion of GC content in the central 100 bp of lost, gained and unchanged classical ChIP PU.1 peaks. (c) Comparison of log2fold change in classical ChIP PU.1 peak score (DB2115-treated/Vehicle) with GC content of central 100 bp of all peaks

(colored according to peak groups). (d) The top *de novo* motif identified from homer analysis of each group of classical PU.1 peaks. (e) Representative viewer tracks of genomic loci displaying PU.1 CUT&Tag (top tracks) and classical PU.1 ChIP (bottom tracks). Highlighted are lost (blue boxes), gained (red boxes) and unchanged (unmarked) PU.1 binding instances.


**Extended Data Fig. 3 | Multi-cell line analysis of PU.1 redistribution.**

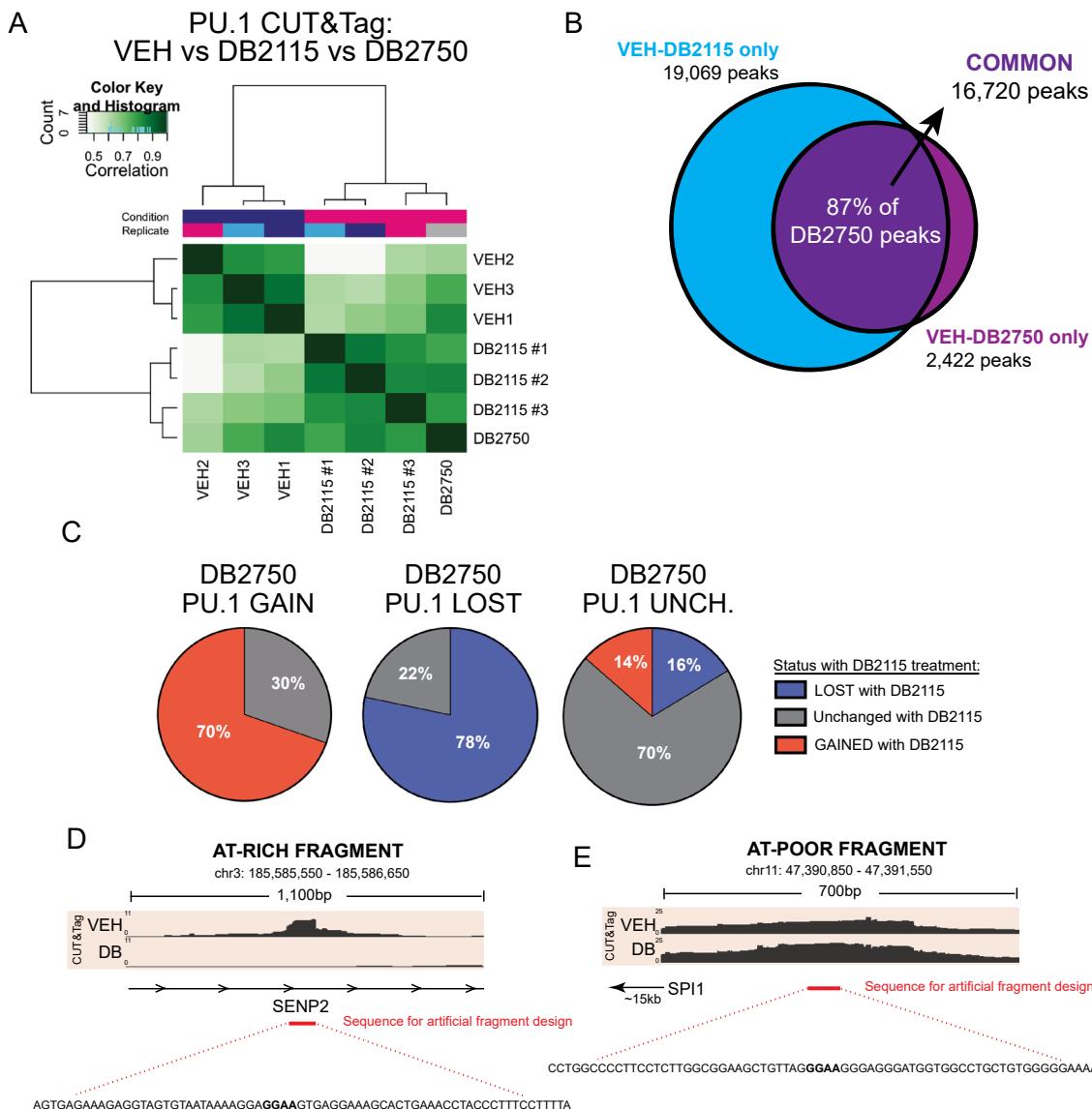
(a-c) GC content of the central 100 bp of lost, gained and unchanged PU.1 peaks from the three cell lines THP1, HL60 and MV411 following 12 hr of 5  $\mu$ M DB215. (d & e) Representative viewer tracks of genomic loci displaying lost (blue boxes), gained (red boxes) and unchanged (unmarked) PU.1 binding for all four cell lines. (f) Log odds ratio score for the PU.1 consensus sequence (Pham et al., 2013) in the 8 categories of commonly lost/gained peaks identified from MOLM13, THP1, HL60 and MV411 cell lines. (g) GC content of the central 100 bp of PU.1 gained and lost peaks which were classified according to their degree of commonality across the 4 cell lines (MOLM only, common to 1, 2 or 3 other

cell lines). One-way ANOVA with Tukey multiple comparisons was performed, \* $p < 0.05$ ( $p = 0.0445$ [MOLM only vs +1 cell line]), \*\*\* $p < 0.0001$ . (h) Central PU.1 motif identified in the 4 categories of commonly lost PU.1 peaks or (i) 4 categories of commonly gained peaks. (j) GC-content position frequency matrix of lost and gained pooled primary AML sample peaks, which have been centered on a short consensus ETS motif (GAGGAAGT) and examined  $\pm 25$  bp. (k) Similarity between primary AML PU.1 peak changes and MOLM13 data. (l) Dose-response curve for MOLM13 viability (Cell titer blue assay) for the four diamidines compounds, DB2115, DB2373, DB2313 and DB2826,  $n = 3$  experimental replicates per drug displaying mean  $\pm$  SEM.



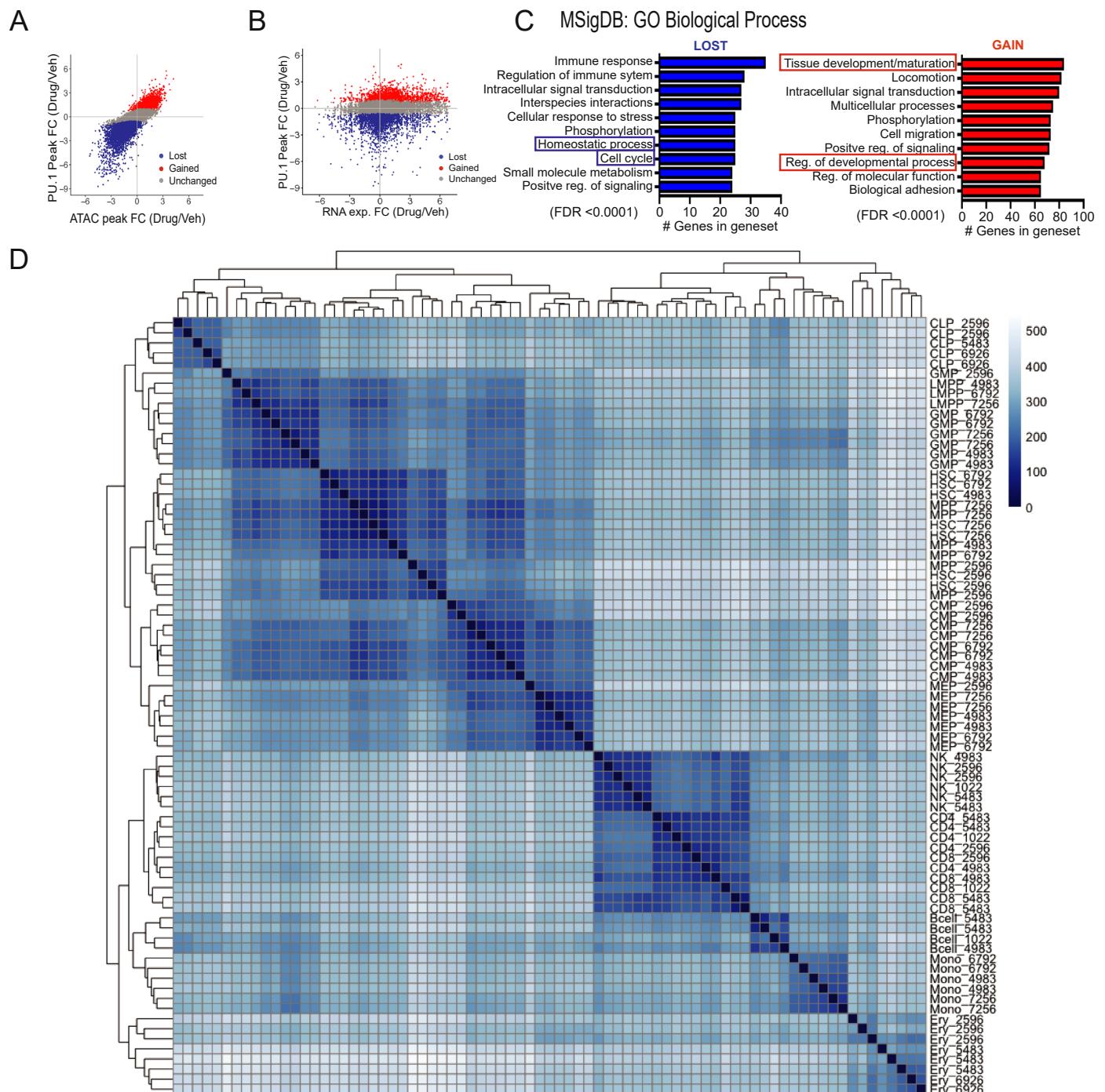
**Extended Data Fig. 4 | Other transcription factor redistribution after DB2115 treatment.** (a) Comparison of log2fold change of RUNX1, ELF1 and GABPA peak score (12 hr of 5  $\mu$ M DB2115-treated/Vehicle) versus GC content of the central 100 bp in the PU.1-null cell line, JURKAT. n = 1 and differential analyses was conducted via Goodpeaks script.

log2fold change of RUNX1, GATA3, ELF1 and GABPA peak score (12 hr of 5  $\mu$ M DB2115-treated/Vehicle) versus GC content of the central 100 bp in the PU.1-null cell line, JURKAT. n = 1 and differential analyses was conducted via Goodpeaks script.



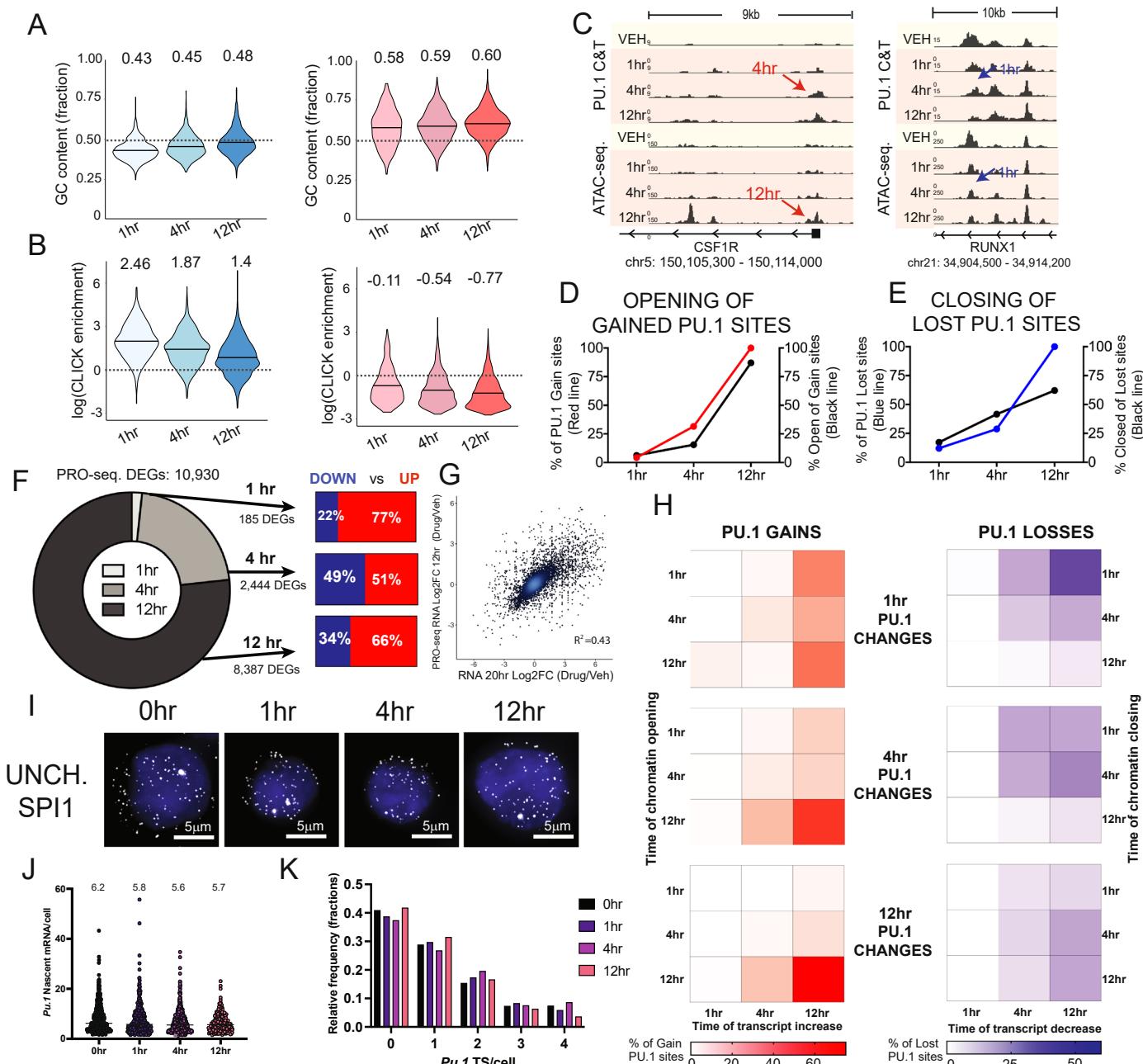
**Extended Data Fig. 5 | Characterization of linker-tagged DB2115 (known as DB2750).** (a) Correlation heatmap of the differentially bound PU.1 peaks from Diffbind analysis of MOLM13 VEH vs DB2115 vs DB2750 treated samples, showing grouping of DB2750 with DB2115 samples. (b) Overlap of all PU.1 peaks from DB2750 versus DB2115 CUT&Tag. (c) DB2750 PU.1 peaks grouped as gain,

loss or unchanged and examined for the peak status in the DB2115 peak dataset. Representative viewer tracks of PU.1 CUT&Tag from vehicle and DB2115 treated MOLM13 cells showing the (d) AT-rich/SENP2 and (e) AT-poor/SPI1 upstream sequences used to synthesize artificial fragments used in experiments from Fig. 3e.



**Extended Data Fig. 6 | Additional transcriptomic and chromatin accessibility characterization following DB2115 exposure.** (a) Comparison of log<sub>2</sub>fold change in PU.1 peak score (DB2115-treated/Vehicle) with log<sub>2</sub>fold change of ATAC peak score (colored according to PU.1 CUT&Tag groups). (b) Comparison of log<sub>2</sub>fold change in PU.1 peak score (DB2115-treated/Vehicle) with log<sub>2</sub>fold change of RNA expression of associated genes (colored according to PU.1 peak

**(c)** A high confidence list of promoter/intronic/exonic DB2115-target genes being either lost/closing (194) or gained/opening (506) were analyzed for enrichment of pathways from GO biological processes using the molecular signature database. **(d)** K-means clustering of normalized and binarized raw count ATAC data from Corces et al., 2016; confirming correct groupings of cellular identity, used for Fig. 4p,q.

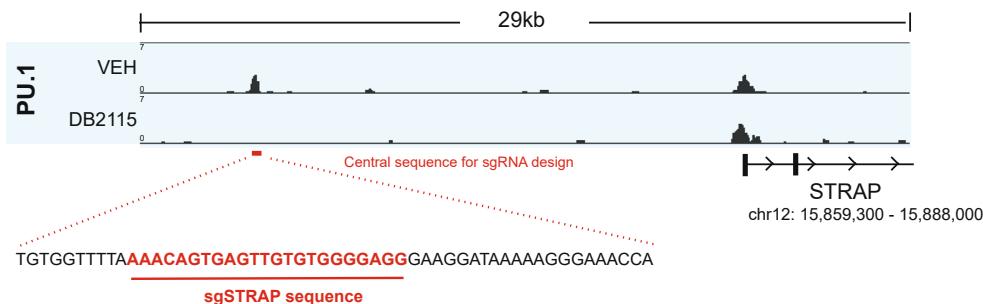


Extended Data Fig. 7 | Additional PU.1 redistribution time course data.

(a) Proportion of GC content or (b) CLICK enrichment score of lost and gained PU.1 peaks (left & right panels respectively) for each timepoint of DB2115 treatment (1, 4 and 12 hr). (c) Representative viewer tracks of genomic loci displaying both PU.1 CUT&Tag (top tracks) and ATAC sequencing (bottom tracks) over the time course. Arrows indicate time of first detection of gain/loss of PU.1, or opening/closing chromatin. (d) Occurrence of first detectable PU.1 gained sites over time (as a % of total gained sites, red) compared to the occurrence of detectable open chromatin at these same PU.1 gained sites over time (black). (e) Occurrence of first detectable PU.1 loss over time (as a % of total lost sites, blue) compared to the occurrence of detectable open chromatin at these same PU.1 lost sites over time (black). (f) The time at which DEGs (NRSA pipeline

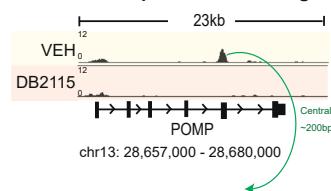
with a p-adj. cut-off <0.1) are first detected and the corresponding proportion of increases and decreases at 1, 4, and 12 hr. (g) Comparison of 20 hr RNA-seq log2FC and 12 hr PRO-seq log2FC expression values from DB2115 treated MOLM13 cells. (h) Heatmaps depicting the time of chromatin opening versus the time of nascent transcript increase of PU.1 gained sites from 1, 4 or 12 hrs, (left panels) or depicting the time of chromatin closing versus the time of nascent transcript decrease for PU.1 lost sites from 1, 4, or 12 hrs. (i) Representative smFISH images from MOLM13 cells for the unchanged PU.1-associated gene, *SPI1*, over the DB2115 time course. *SPI1* transcripts are in white pseudo-color, DNA is in blue pseudo-color. (j) Total nascent transcript counts for *SPI1* mRNA per cell and (k) frequency of transcription burst sites per cell for *SPI1* over the DB2115 time course.

## A STRAP enhancer guide design



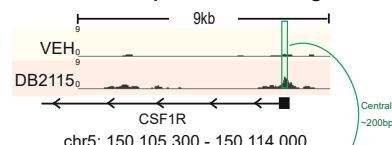
B

## **LOST site reporter insert design**



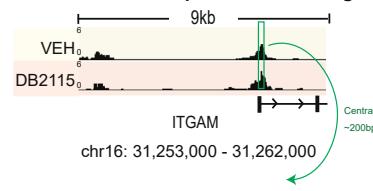
TACTATTTGATGACATTCTTGGCTTGTCAGATTAACAGCGCTGCCCTCCCTGCTTGTGATCCTTAGACTGTACTCTGATCTAACCTAGACCCCTGCCTGCTGGTTTACCTGACTGATCTTGTGTTTGTCTCCCTCCCTCCACCTCTTTTTTTAAGTTGTTGACATCTAGAGAAATTTTCCACATAAAAGATTAAATGCTCTGATAGATGTTGAAATTTATATGTTGCTGGATGATTGATGGGGGAATGTTGAAACTGCCCTTCTGAGATGCTGGATGTT

## GAINED site reporter insert design



chr5: 150,105,300 - 150,114,000

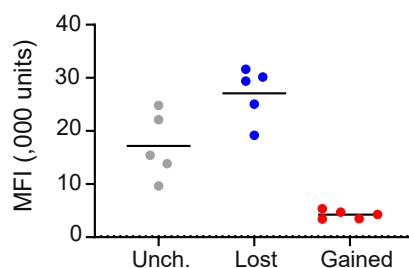
**UNCHANGED site reporter insert design**



IT GAM

GCGAGATCTTGTAGAGACAGGGTCTCTATGTTGCCAGGCTGG  
TTCAACACTCCAGCGCTAACGACTCTCTCCGTCCTGGCCCTCCAAAGTCG  
TGCCTAACAGCGCTGACCCTACGCCCCTGGCCCTATTATGGTAAACA  
CGAATTCAGGAGGCAGGGCTAAGTCATTCTGTTGATATGCTTGCGGC  
ACCAAGAAAACAAGATGGTGGCAAATCAGGACCTTTGGATAGTGTATTGACT  
TTGAAGATTGGTGGCAGGAACCTGGGAGGAAGGGTGGCAAGCTGGTGG  
CAGTCCTGGGCGGAAGACCAAGCCAGGGCTATGTCCTACTGAGCTCGGCC  
TCCTCTTGAATCTCTGATAGACTCTGCCTCCTACTCTCCCTTCTGCCCTC  
TTCTGCTGACGGCTTAT

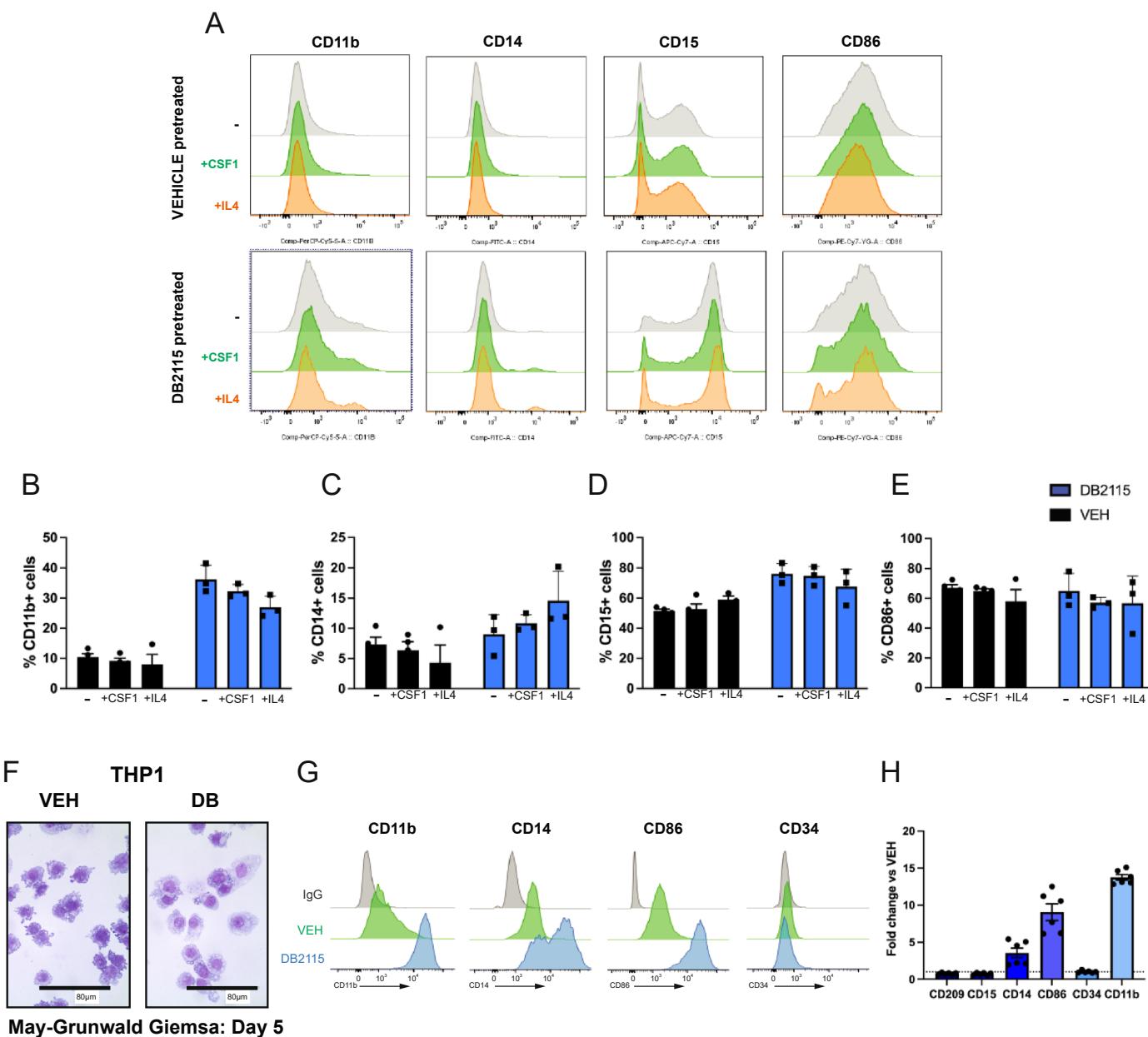
C



## Extended Data Fig. 8 | dCAS9 and GFP reporter assay details.

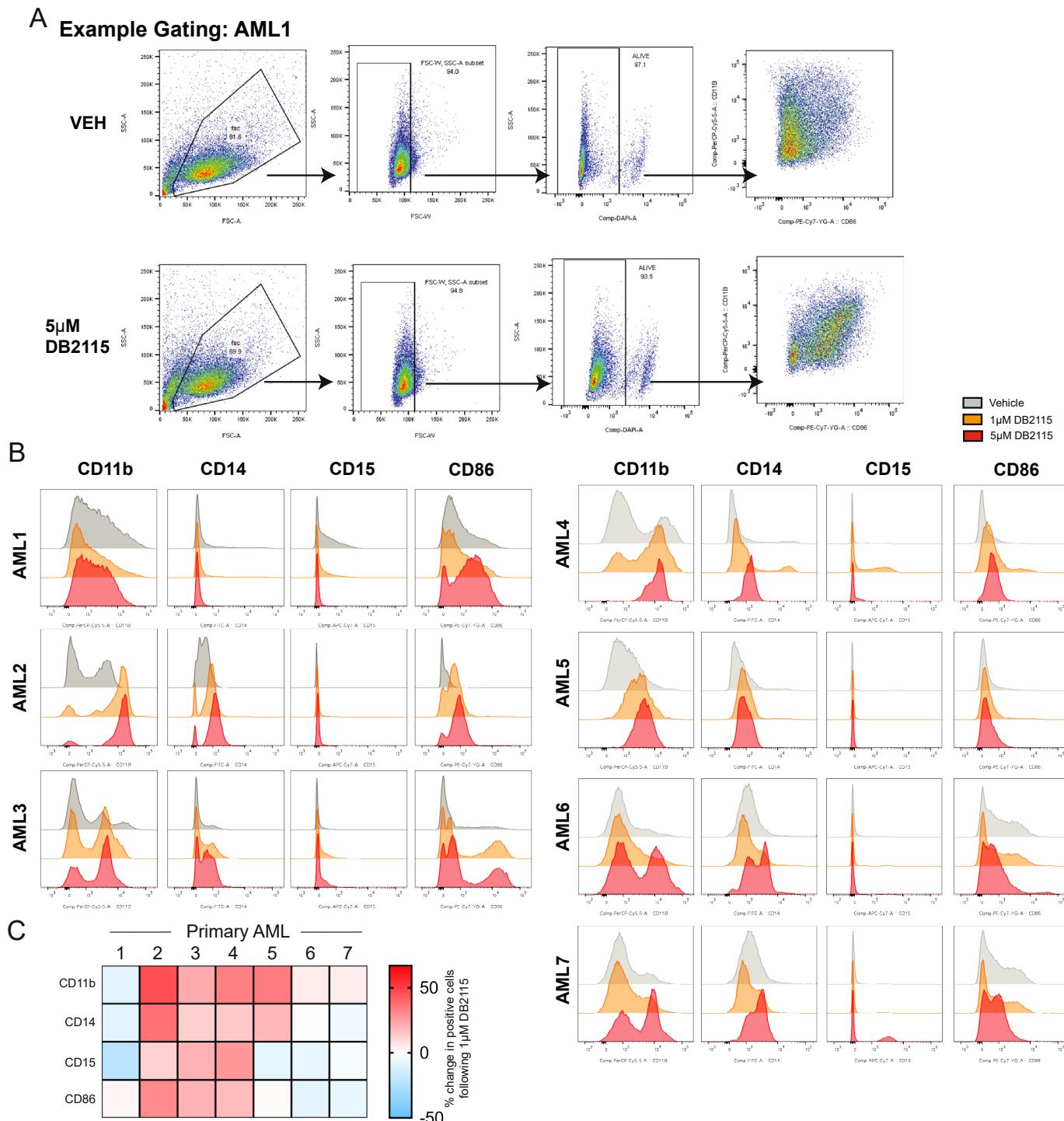
**(a)** Representative viewer tracks of PU.1 CUT&Tag from vehicle and DB2115 treated MOLM13 cells showing the STRAP locus. Identified is the PU.1 binding site where the sgRNA was designed to enable dCas9 targeting (sequence is highlighted in red). **(b)** Representative viewer tracks of PU.1 CUT&Tag from

vehicle and DB2115 treated MOLM13 cells showing the lost POMP, gained CSF1R (alternate promoter) and unchanged CD11b sites. Identified in green is the PU.1 peak and corresponding sequence which was used as the enhancer for eGFP expression in the lentiviral vector. (c) Raw baseline MFI value of the unchanged, lost and gained eGFP reporter cells without drug treatment, n = 5.



**Extended Data Fig. 9 | Additional data from drug-induced differentiated cell lines.** (a) Representative histograms and (b–e) average percentage positive for CD15, CD14, CD86 and CD11b in drug-invoked (or vehicle) MOLM13 cells following 7 days of culture in methylcellulose, n = 3 experimental replicate displaying mean ± SEM. (f) May-Grünwald Giemsa cytopsin image of THP1

cells treated with vehicle or 1 μM DB2115 for 5 days. Experiment was repeated independently 3 times with similar results. (g) Representative histograms and (h) fold change in MFI vs vehicle for CD209, CD15, CD14, CD86, CD34 and CD11b in 5 day treated cells, n = 3 experimental replicates displaying mean ± SEM.



**Extended Data Fig. 10 | Raw flow cytometry data for primary AML samples.** (a) Representative gating schema for primary AML samples grown in methylcellulose with vehicle, 1 μM or 5 μM DB2115 for 8–13 days. (b) Histograms displaying expression of CD15, CD14, CD86 and CD11b in primary AML samples

following treatment with DB2115 or vehicle after 8–13 days. (c) Summary heatmap displaying changes in expression of CD15, CD14, CD86 and CD11b after 8–13 days of 1 μM DB2115 vs vehicle.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection Data collection for Flow cytometry occurred with FACSDiva 8 software on a BD LSRII, and data was analyzed using FlowJo v10.6.1.

Data analysis CUT&Tag generated fastq files were mapped to the human genome (hg38) using bowtie2 (Version 2.2.3) with options: --end-to-end --very-sensitive --no-mixed --no-discordant --phred33 -I 10 -X 700. Normalized bedgraph files were generated by using bedtools genomecov with a normalization factor of 1,000,000/No. total human reads. Peak calling was performed using bdgpeakcall from MACS2 (Version 2.1.0) with options -I 100 and -c 2. Bigwig files were generated using bedGraphToBigWig for visualization in IGV\_2.4.15. Motif analysis and peak annotation was performed using the Homer package (v4.11.1). Differential peak analysis was performed in R studio with either the Diffbind package (v3.12.0) utilizing DESEQ2 (v1.30.1) if the n>2, or a custom script we termed "GoodpeaksScript" (<https://github.com/steidl-lab/rePU.1sitioning>) if n=1. Classical ChIP generated fastq files were mapped to the human or drosophila genome (hg38 or dm3) using bowtie2 (Version 2.2.3). Duplicates were then removed with Picard (v2.20.1) and normalization was performed by subsampling .bam files by a drosophila reads ratio of vehicle over DB2115. The MACS2 package (v2.1.0) and callpeaks function was used to identify peaks in normalized files, and bigwig files were generated for IGV\_2.4.15 visualization. For RNA sequencing datasets, quality control was performed on the basis of error distribution along the length of reads, GC distribution, N content, base quality, and adapter content. Reads were mapped to the hg38 transcriptome using STAR aligner (v2.7.7a). Raw counts were subsequently normalized and analyzed for differential expression in R using the Bioconductor package DESeq2 (v1.30.1). An enrichment score was generated using the negative logarithm of the adjusted p-value multiplied by the sign of the fold-change for each gene and input into Fast Gene Set Enrichment Analysis (FGSEA v1.19.2). Pre-ranked gene lists were queried against standard c1-8 and hallmark MSigDB gene lists (v7.4, Broad Institute). Additional gene list enrichments were conducted with Enrichr v3.0. For Pro-seq data analysis, adaptor trimming was performed with Cutadapt (v1.18), then the sequences were aligned and mapped using

bowtie2 (v2.5.1). Samtools (v1.9) was used for the file format conversion before using the Nascent RNA Sequencing Analysis (NRSA) pipeline to determine the gene body changes.

Statistical tests were performed in R studio (2023.12.1+402) or GraphPad Prism (v9.5).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Raw and processed sequencing data (CUT&Tag, ChIP, ATAC, RNA seq & PRO-seq) has been uploaded to the GEO online database under Super Series GSE267389 publicly available July 30th 2024). <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267389>

Hg38 genomic data is available from [https://www.ncbi.nlm.nih.gov/datasets/genome/GCF\\_000001405.26/](https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.26/).

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Adult AML/MDS BM, and PB samples were obtained for proof-of-concept studies, but no analyses were made based on gender or sex. The gender of patient samples were listed in S. Table 1F.

Reporting on race, ethnicity, or other socially relevant groupings

Not applicable (only small proof-of-concept sample size was used in this study)

Population characteristics

The covariate-relevant patient characteristics include age, gender and cytogenetics of the leukemia and have been reported in S.Table 1F)

Recruitment

Patient samples with MDS or AML were obtained after written informed consent, from Montefiore Medical Center / Albert Einstein Cancer Center. These samples were obtained from a tissue repository, they were deidentified, and not specifically collected for this study.

Ethics oversight

Adult AML/MDS BM, and PB samples were obtained for proof-of-concept studies after written informed consent and following Albert Einstein College of Medicine institutional review board approval (CCI 2008-842 and 2006-536).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample sizes were not based on formal power calculations to detect pre-specified effect sizes for this proof-of-concept study. Sample size was based on availability of primary AML samples, or reproducibility from at least 2 replicates of Ch IP/CUT& Tag sequencing.

Data exclusions

No data was excluded. All data that passed quality control parameters using standard sequencing and analysis algorithms (described in detail in the methods) were included in the analysis.

Replication

All attempts of replication (biological and technical) were successful and have been presented in the manuscript. Multiple approaches were employed to validate the observations/phenomenon described in this paper, with both CUT&Tag and ChIP seq. in multiple different cellular systems showing the same phenomenon of PU.1 redistribution. In summary, all MOLM13 experiments (CUT&Tag, Flow cytometry, Differentiation) were conducted with 2-5 experimental replicates. Primary sample studies were conducted with 2 technical replicates over 7 different biological samples (from different patients). Other cell line and drug CUT&Tag studies were typically from a n=1, but were providing consistent observations over the different contexts/pharmacologica used.

Randomization

This is not relevant for the present study (no group allocation was performed).

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

## Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

CD14 FITC (#32S603) Biolegend, 1:100  
 CD15 APCCy7 (#323048) Biolegend, 1:100  
 CD86 PE Cy7 (#374210) Biolegend, 1:100  
 CD34 PE (#343506) Biolegend, 1:100  
 CD209 AF647 (#330112) Biolegend, 1:100  
 CD11b PerCP Cy5.5 (#101228) Biolegend, 1:100  
 CSF1R APC (#347306) Biolegend, 1:100  
 IL4R BV421 (#355014) Biolegend, 1:100  
 PU.1 (Santa Cruz, sc-352), 1:50 for CnT, 1:1000 for WB  
 rabbit IgG (Santa Cruz; sc-3888), 1:100  
 Cas9(Cell signaling, 14697S), 1:50  
 Actin (Sigma-Aldrich, A2066), 1:5000 for WB  
 GABPA (Invitrogen, PA5-27735), 1:50  
 FLI1 (Invitrogen, PA5-295977), 1:50  
 RUNX1 (Cell signaling, 4334S), 1:50  
 GATA3 (Cell signaling, 5852T), 1:50  
 ELF1 (Proteintech, 22565-1-AP), 1:50  
 Guinea Pig anti-Rabbit secondary (Antibodies Online, ABIN101961) 1:100  
 HRP-conjugated anti-rabbit secondary (Cell signaling, 7074S), 1:5000  
 pSTAT6 AF647 (Biolegend, 686012), 1:100  
 phospho-S6 ribosomal protein PE (Biolegend, 608604), 1:100

### Validation

Standard FACS antibodies obtained from widely used commercial providers were used in this study. All FACS antibodies have been validated by the company through quality control in-house Flow cytometry, and also by comparisons with isotype control antibodies within our laboratory. Western Blot and CUT&Tag antibodies are all commercially available and have been documented by the company to be specific and valid (via Western blot, immunofluorescence or flow cytometry).

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

### Cell line source(s)

MOLM13-ATCC  
 TH PI -ATCC  
 HL60-ATCC  
 MV411-ATCC  
 JURKAT - ATCC  
 293T - ATCC

### Authentication

ATCC authenticated these cell lines at purchase, which had been authenticated by short tandem repeat (STR) profiling. They were not independently authenticated upon receipt.

### Mycoplasma contamination

All cell lines are confirmed negative for Mycoplasma contamination

### Commonly misidentified lines (See [ICLAC](#) register)

None used

## Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A

## ChIP-seq

### Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

*May remain private before publication.*

Raw and processed sequencing data (CUT&Tag, ChIP, ATAC, RNA seq & PRO-seq) has been uploaded to the GEO online database under Super Series GSE267389 publicly available July 31st 2024).

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267389>

Files in database submission

GSM6262719  
 MOLM13, VEH, 12hr [MOLM\_VEH1]  
 GSM6262720  
 MOLM13, VEH, 12hr [MOLM\_VEH2]  
 GSM6262721  
 MOLM13, VEH, 12hr [MOLM\_VEH3]  
 GSM6262722  
 MOLM13, 5uM DB2115, 12hr [MOLM\_DB1]  
 GSM6262723  
 MOLM13, 5uM DB2115, 12hr [MOLM\_DB2]  
 GSM6262724  
 MOLM13, 5uM DB2115, 12hr [MOLM\_DB3]  
 GSM6262725  
 MOLM13, VEH, 12hr [MOLM\_IgGV1]  
 GSM6262726  
 MOLM13, 5uM DB2115, 12hr [MOLM\_IgGDB1]  
 GSM6262727  
 THP1, VEH, 12hr [THP1\_VEH1]  
 GSM6262728  
 THP1, VEH, 12hr [THP1\_VEH2]  
 GSM6262729  
 THP1, 5uM DB2115, 12hr [THP1\_DB1]  
 GSM6262730  
 THP1, 5uM DB2115, 12hr [THP1\_DB2]  
 GSM6262731  
 HL60, VEH, 12hr [HL60\_VEH1]  
 GSM6262732  
 HL60, VEH, 12hr [HL60\_VEH2]  
 GSM6262733  
 HL60, 5uM DB2115, 12hr [HL60\_DB1]  
 GSM6262734  
 HL60, 5uM DB2115, 12hr [HL60\_DB2]  
 GSM6262735  
 MV411, VEH, 12hr [MV411\_VEH1]  
 GSM6262736  
 MV411, VEH, 12hr [MV411\_VEH2]  
 GSM6262737  
 MV411, 5uM DB2115, 12hr [MV411\_DB1]  
 GSM6262738  
 MV411, 5uM DB2115, 12hr [MV411\_DB2]  
 GSM6262739  
 MV411, VEH, 12hr [MV411\_IgG]  
 GSM6262740  
 THP1, VEH, 12hr [THP1\_IgG]

GSM6262741  
HL60, VEH, 12hr [HL60\_IgG]  
GSM6262742  
MOLM13, 5uM DB2373, 12hr [MOLM\_DB2373]  
GSM6262743  
MOLM13, 5uM DB2313, 12hr [MOLM\_DB2313]  
GSM6262744  
MOLM13, 5uM DB2826, 12hr [MOLM\_DB2826]  
GSM6262745  
MOLM13, 5uM DB2750, 12hr [MOLM\_DB2750]  
GSM6262746  
MOLM13, 200nM Daunorubicin, 12hr [MOLM\_Daun]  
GSM6262747  
MOLM13, 400nM Cytarabine, 12hr [MOLM\_AraC]  
GSM6262748  
MOLM13, 200nM Daun, 12hr [MOLM\_Daun\_IgG]  
GSM6262749  
MOLM13, VEH, 12hr [MOLM\_VEH\_DaunAraC]  
GSM6262750  
MOLM13, 5uM DB2115, 1hr [MOLM\_1hr\_DB\_1]  
GSM6262751  
MOLM13, 5uM DB2115, 1hr [MOLM\_1hr\_DB\_2]  
GSM6262752  
MOLM13, 5uM DB2115, 4hr [MOLM\_4hr\_DB\_1]  
GSM6262753  
MOLM13, 5uM DB2115, 4hr [MOLM\_4hr\_DB\_2]  
GSM6262754  
MOLM13, 5uM DB2115, 12hr [MOLM\_12hr\_DB]  
GSM6262755  
MOLM13, VEH, 12hr [MOLM\_VEH\_timecourse]  
GSM6262756  
MOLM13, CRISPRd sgNon-targeting [sgNT\_PU1\_1]  
GSM6262757  
MOLM13, CRISPRd sgNon-targeting [sgNT\_PU1\_2]  
GSM6262758  
MOLM13, CRISPRd sgNon-targeting [sgNT\_dCas9]  
GSM6262759  
MOLM13, CRISPRd targeting STRAP upstream region [sgSTRAP\_dCas9]  
GSM6262760  
MOLM13, CRISPRd targeting STRAP upstream region [sgSTRAP\_PU1\_1]  
GSM6262761  
MOLM13, CRISPRd targeting STRAP upstream region [sgSTRAP\_PU1\_2]  
GSM6262762  
MOLM13, CLICKonCUT&TAG with DB2750 [CLICK1]  
GSM6262763  
MOLM13, CLICKonCUT&TAG with DB2750 [CLICK2]  
GSM6262764  
MOLM13, CLICKonCUT&TAG with DB2750 [CLICK3]  
GSM8264925  
MOLM13, VEH, ATAC-seq Rep1  
GSM8264926  
MOLM13, VEH, ATAC-seq Rep2  
GSM8264927  
MOLM13, VEH, ATAC-seq Rep3  
GSM8264928  
MOLM13, 1hr DB2115, ATAC-seq Rep1  
GSM8264929  
MOLM13, 1hr DB2115, ATAC-seq Rep2  
GSM8264930  
MOLM13, 4hr DB2115, ATAC-seq Rep1  
GSM8264931  
MOLM13, 4hr DB2115, ATAC-seq Rep2  
GSM8264932  
MOLM13, 12hr DB2115, ATAC-seq Rep1  
GSM8264933  
MOLM13, 12hr DB2115, ATAC-seq Rep2  
GSM8264934  
MOLM13, VEH, 12hr, ChIP-seq  
GSM8264935  
MOLM13, DB2115, 12hr, ChIP-seq  
GSM8264936  
MOLM13, VEH, 12hr, RUNX1, Rep1  
GSM8264937  
MOLM13, VEH, 12hr, RUNX1, Rep2  
GSM8264938  
MOLM13, 5uM DB2115, 12hr, RUNX1 Rep1

GSM8264939  
 MOLM13, 5uM DB2115, 12hr, RUNX1, Rep2  
 GSM8264940  
 JURKAT, VEH, 12hr, RUNX1  
 GSM8264941  
 JURKAT, 5uM DB2115, 12hr, RUNX1  
 GSM8264942  
 JURKAT, VEH, 12hr, ELF1  
 GSM8264943  
 JURKAT, 5uM DB2115, 12hr, ELF1  
 GSM8264944  
 JURKAT, VEH, 12hr, GATA3  
 GSM8264945  
 JURKAT, 5uM DB2115, 12hr, GATA3  
 GSM8264946  
 JURKAT, VEH, 12hr, FLI1  
 GSM8264947  
 JURKAT, 5uM DB2115, 12hr, FLI1  
 GSM8264948  
 MOLM13, VEH, 12hr, ELF1  
 GSM8264949  
 MOLM13, 5uM DB2115, 12hr, ELF1  
 GSM8264950  
 MOLM13, VEH, 12hr, GABPA  
 GSM8264951  
 MOLM13, 5uM DB2115, 12hr, GABPA  
 GSM8264952  
 AML sample1, VEH, 12hr  
 GSM8264953  
 AML sample1, 5uM DB2115, 12hr  
 GSM8264954  
 AML sample2, VEH, 12hr  
 GSM8264955  
 AML sample2, 5uM DB2115, 12hr  
 GSM8264956  
 MOLM13, VEH, 12hr PRO-seq rep1  
 GSM8264957  
 MOLM13, VEH, 12hr PRO-seq rep2  
 GSM8264958  
 MOLM13, 5uM DB2115, 1hr PRO-seq rep1  
 GSM8264959  
 MOLM13, 5uM DB2115, 1hr PRO-seq rep2  
 GSM8264960  
 MOLM13, 5uM DB2115, 4hr PRO-seq rep1  
 GSM8264961  
 MOLM13, 5uM DB2115, 4hr PRO-seq rep2  
 GSM8264962  
 MOLM13, 5uM DB2115, 12hr PRO-seq rep1  
 GSM8264963  
 MOLM13, 5uM DB2115, 12hr PRO-seq rep2  
 GSM8264964  
 MOLM13, VEH, 20hr, RNA-seq rep1  
 GSM8264965  
 MOLM13, VEH, 20hr, RNA-seq rep2  
 GSM8264966  
 MOLM13, VEH, 20hr, RNA-seq rep3  
 GSM8264967  
 MOLM13, DB2115, 20hr, RNA-seq rep1  
 GSM8264968  
 MOLM13, DB2115, 20hr, RNA-seq rep2  
 GSM8264969  
 MOLM13, DB2115, 20hr, RNA-seq rep3

Genome browser session  
 (e.g. [UCSC](#))

No longer applicable

## Methodology

### Replicates

The central datasets were repeated multiple times for consistency, minimum of 2 independent Chip or CUT&Tag replicates (MOLM13, THP1, HL60, MV411, dCas9-STRAP, RUNX1, MOLM13 timecourse). For exploratory or primary datasets with low cell material available, n=1 (Daunorubicin, Ara-C, DB2373, DB2826, DB2313, GATA3, FLII, GABPA, ELF1, primary patient samples (multiple)).

### Sequencing depth

Sequencing for CUT&Tag was conducted with Paired-end sequencing (35-75bp reads) with an average coverage of 10-15million reads

## Sequencing depth

per sample. This is well above the recommended sequencing depth of~5million for CUT&Tag.

## Antibodies

Antibodies used for ChIP and CUT&Tag:  
 PU.I (Santa Cruz, sc-352, 1:50)  
 rabbit IgG (Santa Cruz; sc-3888, 1:100)  
 Cas9(Cell signaling, 14697S, 1:50)  
 GABPA (Invitrogen, PA5-27735, 1:50)  
 FLII (Invitrogen, PA5-295977, 1:50)  
 RUNX1 (Cell signaling, 4334S, 1:50)  
 GATA3 (Cell signaling, 5852T, 1:50)  
 ELF1 (Proteintech, 22565-1-AP, 1:50)

## Peak calling parameters

CUT&Tag generated fastq files were mapped to the human genome (hg38) using bowtie2 (Version 2.2.3) with options: --end-to-end --very-sensitive --no-mixed --no-discordant --phred33 -I 10-X 700. Normalized bedgraph files were generated by using bedtools genomecov with a normalization factor of 1,000,000/No. total human reads. Peak calling was performed using bdgpeakcall from MACS2 (Version 2.1.0) with options -n100 and -c 2. Bigwig files were generated using bedGraphToBigWig for visualization in IGV 2.4.15. Motif analysis and peak annotation was performed using the Homer package.  
 For CLICKonCUT& Tag, the average MACS2 peak scores across 3 replicates were calculated and compared to input vehicle-treated PU.I CUT& Tag peak scores to generate a log2fold change CLICK score. Peaks with CLICK scores above 0.5 were considered enriched for drug binding, whereas scores below 0.5 were considered non-drug binding.  
 Classical ChIP generated fastq files were mapped to the human or drosophila genome (hg38 or dm3) using bowtie2 (Version 2.2.3). Duplicates were then removed with Picard and normalization was performed by subsampling .bam files by a drosophila reads ratio of vehicle over DB2115. The MACS2 package and call peaks function was used to identify peaks in normalized files, and bigwig files were generated for IGV visualization. Differential peak analysis was performed in R studio with "GoodpeakAnalysis".

## Data quality

Initial Fastq files were confirmed passable via multi-FASTQC analyses.  
 After peak calling peaks were confirmed for significant changes using Diffbind package, in particular: Differential peak analysis was performed in R studio with either the Diffbind package utilizing DESEQ2 analyses if the n>2, or a custom script we termed "GoodpeaksScript" (<https://github.com/steidl-lab/rePU.lsitioning>). if n=1. In Diffbind, FDR cutoff was set to 0.1 for determining significantly changed peaks. For GoodpeakAnalysis, three stringent filters were used for the differential peak analysis of the average peak intensity; min. intensity> 7.5, min. fold change>4, and min. summit >3.

## Software

We used the following software for ChIP-seq data analyses: Bowtie2 (Version 2.2.3), Picard (v2.20.1), bedtools genomecov, MACS2 (bdgpeakcall, Version 2.1.0), bedGraphToBigWig, IGV\_2.4.15., Homer v4.11.1, Diffbind package utilizing DESEQ2 v3.12.0, STAR aligner (2.7.7a), FGSEA (1.19.2), Bioconductor, DESeq2 v1.30.1.  
 Custom scripts from R studio were used and are described at: <https://github.com/steidl-lab/rePU.lsitioning>

## Flow Cytometry

## Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

## Sample preparation

Antibody staining was performed for 30 min prior to flow cytometry analyses, employing the following Biologend antibodies; CD14 FITC (#325603), CD15 APCCy7 (#323048), CD86 PECy7 (#374210), CD34 PE (#343506), CD209 AF647 (#330112), CD11b PerCP Cy5.5 (#101228), CSF1R APC (#347306), IL4R BV421 (#355014) and DAPI for alive cell gating.

## Instrument

Flow cytometry analysis was performed on a BD LSRII containing a yellow laser (BD LSR II-Y).

## Software

Flow cytometric data was acquired using FACSDiva software v 8.0 and analyzed using FlowJo Software (version 10.6.1)

## Cell population abundance

Sorting was not performed

## Gating strategy

The gating strategy used followed established markers and schemes for the identification of viable cultured AML cells or cell lines. Gating scheme for primary AML cells is displayed in Supplemental Figure 10. Negativity for any marker was defined as the threshold defined by staining with the respective isotype control antibodies.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.