LSD1-mediated repression of *GFI1* super-enhancer plays an essential role in erythroleukemia

Goichi Tatsumi[1,2], Masahiro Kawahara[2]*, Ryusuke Yamamoto[1], Masakatsu Hishizawa[1], Katsuyuki Kito[2], Takayoshi Suzuki[3,4], Akifumi Takaori-Kondo[1] and Akira Andoh[2]

Affiliations

[1] Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University.

[2] Department of Medicine, Shiga University of Medical Science.

[3] The Institute of Scientific and Industrial Research, Osaka University.

[4] CREST, Japan Science and Technology Agency (JST)

Contact information

*Corresponding author: Masahiro Kawahara

E-mail: mkawahar@belle.shiga-med.ac.jp

Address: Seta-Tsukinowa, Otsu, Shiga, 520-2192, Japan

Tel: +81-77-548-2217

Fax: +81-77-548-2219

Running title: *GFI1* super-enhancer repression by LSD1 in leukemia

Conflict of interest

The authors have no conflicts of interest to declare.

Abstract

Super-enhancers (SEs) consist of enhancer clusters with abundant binding of transcription factors (TFs) and cofactors. LSD1 is a histone modifier that eliminates SE activity. However, whether SE suppression by LSD1 is associated with leukemogenesis remains unknown. In erythro-megakaryocyte lineage leukemia cells, activation of the SE of *GFI1* (*GFI1*-SE) is related to induction of myeloid differentiation by LSD1 inhibitors NCD38 and NCD25 and to their anti-leukemia effect. Although functional TF-motifs were concentrated in an evolutionally conserved area, NCD38 barely induced additional TF recruitment. Instead, the transcription cofactors including LSD1, CoREST, HDAC1, and HDAC2 were evicted from *GFI1*-SE. Deletion of *GFI1*-SE impaired induction of myeloid differentiation by NCD38 and NCD25 in erythroleukemia cells. Gene set enrichment analysis revealed that the *GFI1*-SE deletion impaired NCD38-induced programs related to granulocyte differentiation and the CEBPA network, but restored NCD38-suppressed programs related to erythroid development, GATA1 targets, and acute myeloid leukemia (AML) clusters including FAB subtype M6 and AML with myelodysplastic syndrome-related chromosomal abnormalities. Ontologies of genes whose expression changes by NCD38 were cancelled due to the *GFI1*-SE deletion showed enrichment in

AML and neutropenia signatures. Collectively, our data suggest that sustainable repression of *GFI1*-SE by LSD1 is essential for sustenance of erythroleukemia cells.

Introduction

Networks of transcription factors (TFs) play a crucial role in hematopoiesis[1]. TFs function as key regulators for cell fate decisions such as stemness maintenance and terminal differentiation by coordination or cross-antagonism with other TFs. Dysregulation of differentiation is one of the causes of leukemia development[2]. This regulation partially occurs via enhancers near the gene loci of TF genes[3]. Recently, super-enhancers (SEs) have been identified as large clusters of stretched enhancers with abundant binding of TFs and cofactors to confer higher transcriptional activity than typical enhancers[4]. The precise regulation of SEs near the TF gene loci is essential to determine cell identity[4-6]. Moreover, SEs play pivotal roles not only in normal development but also in disease initiation and progression including oncogenesis[7-9]. Recent studies have reported that SEs in the proximity of oncogenes are often aberrantly activated in leukemia cells[10, 11]. However, little is known about whether silencing of SEs is involved in leukemia development or sustenance.

Lysine-specific demethylase 1 (LSD1, also known as KDM1A) is a histone demethylase that catalyzes the demethylation of mono- and di-methylated lysine 4 of histone H3 (H3K4me1/2)[12]. LSD1 is indirectly involved in the deacetylation of acetylated

lysine 27 of histone H3 (H3K27ac) via formation of a complex with RE1 silencing transcription factor corepressor (CoREST, also known as RCOR1) and histone deacetylase1 and 2 (HDAC1 and HDAC2, respectively)[13, 14]. Given that H3K4me1 and H3K27ac are surrogate markers of activated SEs[15, 16], LSD1 is presumed to function as an eraser of activated SEs. Deletion of *Lsd1* causes pancytopenia and increases the H3K27ac level on enhancers of *Lsd1* targets in murine hematopoietic cells[17]. *LSD1* is significantly overexpressed in a number of hematologic malignancies, including acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS)[18, 19]. Moreover, transgenic mice expressing a human *LSD1* isoform exhibit the development of T-cell lymphoblastic leukemia/lymphoma[20]. These reports clearly indicate that alterations in LSD1 expression lead to dysregulation of hematopoiesis and possible leukemia development.

To date, a number of LSD1 inhibitors have been developed, some of which are currently being investigated in clinical trials[21]. These inhibitors commonly induce myeloid differentiation to exert an anti-leukemia effect[18, 22-24]. However, the precise molecular mechanism is not fully understood. We previously developed the LSD1 inhibitors NCD38 and NCD25 and reported that these inhibitors are particularly effective in erythroleukemia and AML transformed from MDS or with MDS-related chromosomal abnormalities[23].

These inhibitors caused myeloid differentiation in such leukemia cells and activated several SEs near TF genes, including *GFI1*, which is a key regulator for neutrophil development[25] and a major partner in the hematopoietic TF network[1]. However, it is yet unclear how the SE localized 14 kb downstream of the *GFI1* gene body[23] (named *GFI1*-SE) is regulated and whether repression of *GFI1*-SE by LSD1 is critical to leukemogenesis in erythroleukemia cells. To clarify these points, we investigated the regulatory mechanism of *GFI1*-SE and the genetic programs influenced by *GFI1*-SE depletion.

Materials and methods

Cells and reagents

HEL, CMK11-5, UT7-EPO and K562 cells were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum and 1% penicillin. Cells were maintained in our laboratory under condition without contamination with mycoplasma after validating the immunophenotype by fluorescence-activated cell sorting (FACS) analysis. NCD38 and NCD25 were synthesized as described previously[26], dissolved in dimethyl sulfoxide (DMSO), and used at 2 μM for all experiments.

Cell proliferation assay

After incubation of cells with 2 μM NCD38, 2 μM NCD25, or DMSO for 2 or 3 days, WST-8 reagent (Nacalai Tesque, Kyoto, Japan) was added. The data was obtained by a microplate reader (Tecan, Männedorf, Switzerland) according to the manufacturer's instructions.

Quantitative PCR

RNA and complementary RNA were synthesized as described in the supplementary

information. Quantitative PCR was performed using LightCycler480 System II (Roche, Basel, Switzerland) and a THUNDERBIRD SYBR qPCR mix (Toyobo, Osaka, Japan). Absolute quantification was calculated using recombinant DNA of targeted amplicons as a standard. Primer information is provided in Supplementary Table 1.

Western blotting

Total cell lysates were prepared as described in the supplementary information. We used anti-GFI1 [sc-8558 or sc-376949; Santa Cruz Biotechnology (SCB), Dallas, TX] and anti-ACTIN (sc-1616; SCB) as primary antibodies, and HRP-conjugated anti-goat (sc-2020; SCB) or anti-mouse (NA931v; GE Healthcare, Little Chalfont, UK) as secondary antibodies.

FACS

We used anti-CD11b-PE-Cy5 (ICRF44; eBioscience, San Diego, CA), and analyzed on FACS Calibur (BD Biosciences, San Jose, CA).

Chromatin immunoprecipitation (ChIP)

ChIP was performed as described previously[27]. For precipitation, anti-H3K27ac (39133; Active Motif, Carlsbad, CA), anti-GFI1 (sc-376949; SCB), anti-GFI1B (sc-28356X; SCB), anti-CEBPA (sc-61X; SCB), anti-TAL1 (ab155195; Abcam, Cambridge, UK), anti-GATA1 (ab11852; Abcam), anti-RUNX1 (ab23980; Abcam), anti- ERG (ab133264; Abcam), anti-LSD1 (ab17721; Abcam), anti-CoREST (ab32631; Abcam), anti-HDAC1 (ab7028; Abcam), anti-HDAC2 (ab7029; Abcam), or control rabbit IgG (sc-2027; SCB), and Dynabeads Protein G (Thermo Fisher Scientific, Waltham, MA) were used. After reverse crosslinking, ChIP DNA was purified with a MinElute PCR Purification kit (Qiagen, Hilden, Germany) and analyzed by quantitative PCR.

Reporter assays

A genomic region including *GFI1*-SE was amplified from HEL genomic DNA and cloned into the pGL3-promoter vector (Promega, Madison, WI). Deletion vectors and mutant vectors were constructed using PCR mutagenesis. Each vector and the Renilla luciferase reporter (phRL-TK, Promega) were cotransfected into HEL cells using Amaxa nucleofector technology (Lonza, Basel, Switzerland). After 48 hours, cells were treated with NCD38, NCD25, or DMSO. After a 24-hour treatment, luciferase activities were measured using the

Dual-Luciferase Reporter Assay System (Promega).

Genome editing

Genome editing was performed using CRISPR-Cas9 as described previously[23]. Briefly, two target-specific oligonucleotides were designed to cleave upstream and downstream sites of *GFI1*-SE using a bioinformatics tool[28, 29] (https://chopchop.rc.fas.harvard.edu/) and were inserted into the CRISPR-Cas9 genome editing vector (pX330-U6-Chimeric_BB-CBh-hSpCas9, plasmid ID: 42230; Addgene, Watertown, MA). Both vectors were transfected into HEL cells using Amaxa nucleofector technology. After limiting dilution and genotype screening, sublines with bi-allelic deletion of *GFI1*-SE (Δ*GFI1*-SE) were established. For control sublines, the non-inserted CRISPR-Cas9 genome editing vector was used.

Microarray experiments and analysis

Total RNA was extracted using an RNeasy Mini Kit (Qiagen). After quality check with an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA), 100 ng total RNA was labeled using a GeneChip WT Terminal Labeling Kit (Affymetrix, Santa Clara, CA) and hybridized to a

GeneChip Human Gene 2.0 ST Array (Affymetrix). Microarray data were analyzed using GeneSpring GX software (Agilent). All raw data were normalized and filtered based on the deviation of the values within a condition (SD < 0.1). Microarray data were submitted to the Gene Expression Omnibus and assigned to GSE128400. Gene set enrichment analysis (GSEA)[30, 31] was performed using GSEA v3.0 software (http://www.broadinstitute.org/gsea/index.jsp). The gene sets are listed in Supplementary Table 2. The gene ontology was analyzed using the Enrichr software[32, 33].

Bioinformatics tools

Our previous ChIP-sequencing data for H3K27ac (accession number: DRA003526) was visualized with the Integrative Genomics Viewer (IGV) software[34, 35] version 2.4 (Broad Institute, Cambridge, MA; https://www.broadinstitute.org/igv). Genomic regions conserved between humans and mice were identified using the ECR browser[36] (http://ecrbrowser.dcode.org/). The genomic sequence spanning the *GFI1*-SE locus was obtained from the UCSC genome browser[37] (http://genome.ucsc.edu/). DNA sequence alignments were performed using the T-coffee server[38] (http://tcoffee.vital-it.ch/apps/tcoffee/do:regular). TF-motifs were analyzed using the

JASPAR database[39] (http://jaspar.genereg.net/).

Statistical analysis

ANOVA followed by Dunnett's test was used for multiple comparison analysis of luciferase reporter activities. All tests were two-sided and a p-value < 0.05 was considered significant.

Results

Relevance of *GFI1*-SE activation to myeloid differentiation and anti-leukemia effect in erythro-megakaryocyte lineage leukemia cells

To investigate the role of *GFI1*-SE in the regulation of GFI1 expression and myeloid differentiation by LSD inhibitors, we compared the changes in H3K27ac levels in the *GFI1*-SE region, *GFI1* transcript levels, a myeloid marker CD11b, an erythroid marker CD235a, cytology, and cell growth following exposure to NCD38 or NCD25 in four erythro-megakaryocyte lineage leukemia cells. The H3K27ac level was elevated up to 25-fold in HEL cells and up to 3-fold in CMK11-5 cells, but not in UT7-EPO cells and K562 cells (Fig. 1a). LSD1 expression was comparable across all cell types and was not changed by exposure to NCD38 and NCD25 (Supplementary Figure 1). In response to the activation of *GFI1*-SE and regardless of the basal H3K27ac values, the *GFI1* transcript level increased in HEL and CMK11-5 cells but not in UT7-EPO and K562 cells (Fig. 1b). Induction of CD11b and neutrophil-like cytology including nuclear segmentation, as well as attenuation of CD235a were also observed only in HEL and CMK11-5 cells (Fig. 1c, Supplementary Figure 2&3). Cell growth was suppressed in HEL and CMK11-5 cells, but barely suppressed in UT7-EPO and K562 cells after 48-hour treatment (Fig. 1d). These data

suggest that the activation of *GFI1*-SE by NCD38 and NCD25 appears to correlate with the induction of myeloid differentiation and subsequent anti-leukemia effect in erythro-megakaryocyte leukemia cells. Although minimal efficacy was observed in UT7-EPO cells after 72-hour treatment, we observed no changes in *GFI1*-SE, *GFI1* expression, and CD11b level, except for attenuation of CD235a (Supplementary Figure 4a-c), suggesting that only this delayed effect is induced in a GFI1-independent manner. We also found *GFI1* upregulation with *GFI1*-SE activation in MDS-L cells transformed from MDS cells[40] but not in THP-1 cells with MLL-AF9 (Supplementary Figure 5), although both were moderately suppressed by NCD38 and NCD25 as previously reported[23]. This data suggests that *GFI1*-SE repression might be essential in MDS overt AML where the border with erythroleukemia has been partially altered[41].

GFI1-SE activation by the coordination of several TFs

Our first aim was to understand which TFs play a significant role in *GFI1*-SE activation by NCD38 or NCD25. Homology search between humans and mice uncovered a highly conserved area (CA) with more than 70% homology in *GFI1*-SE (Fig. 2a). The CA was most strongly activated by NCD38 according to the reconstruction of our previous

ChIP-sequencing data (DRA003526). In murine erythroleukemia cells, treatment with NCD38 or NCD25 elevated the H3K27ac level in the CA, upregulated the level of the *Gfi1* transcript, and induced the granulocyte marker Gr-1 (Supplementary Figure 6). The CA included five CpG sites (CG1-5) (Supplementary Figure 7); CG2-5 were almost fully open across the cell lines we used in the steady state, suggesting that DNA methylation status is not involved in CA activation, while CG1 was open only in CMK11-5 cells, which might relate to the basal H3K27ac level in *GFI1*-SE (Supplementary Figure 8).

Next, to ascertain the impact of the CA on enhancer activity, we cloned the most H3K27ac-enriched area of *GFI1*-SE including the CA into a luciferase vector (wild-type; WT) and generated two types of deletion mutants: a *GFI1*-SE mutant lacking the CA (ΔCA) or the non-CA portion (Δnon-CA). A reporter assay performed after transfection to HEL cells and exposure to NCD38 or NCD25 for 24 hours showed that Δnon-CA induced a high reporter activity that was comparable to the WT, while ΔCA failed to induce reporter activity (Fig. 2b). These data indicate that CA is required to activate *GFI1*-SE.

Within the CA, we identified binding motifs for TFs including GFI1/GFI1B, CEBPA, TAL1, RUNX1, and GATA1 (Supplementary Figure 7). Thus, we generated *GFI1*-SE mutants in which only the binding motifs of each TF were mutated and performed

reporter assays (Fig. 2b). We found that all single TF-motif mutants except the construct with a mutation in the CEBPA upstream-motif (CEBPA #1) exhibited significantly reduced levels of reporter activity compared with the WT after exposure to NCD38 [0.50, 0.61, 0.46, 0.56, and 0.50 of the WT level in constructs with mutations in GFI1/GFI1B-motifs, a CEBPA downstream-motif (#2), a TAL1-motif, a RUNX1-motif, or a GATA1-motif, respectively] and to NCD25 (0.53, 0.66, 0.51, 0.51, and 0.57 of the WT level in the same order as NCD38 exposure, respectively). The GFI1/GFI1B/GATA1/RUNX1/TAL1/CEBPA multiple-motifs mutant (5TFs-mt) achieved the same level of reporter activity reduction as the ΔCA mutant after exposure to NCD38 (0.28 and 0.30 of the WT level in ΔCA and 5TFs-mt, respectively) and NCD25 (0.25 and 0.30 of the WT level in ΔCA and 5TFs-mt, respectively). These data suggest that *GFI1*-SE activation by NCD38 and NCD25 is not controlled by any one TF but rather appears to be regulated by the coordination of various TFs.

Sustainable occupation of TFs and release of the LSD1 complex by NCD38 in *GFI1*-SE

Based on reporter assay results, we hypothesized that the recruitment of several TFs would be required to activate *GFI1*-SE. However, ChIP analysis showed that exposure to NCD38 barely changed the occupancy level of TFs, including GFI1, CEBPA, TAL1, RUNX1, GATA1, and ERG, in *GFI1*-SE (Fig. 3a). Although GFI1B was recruited by more than 2-fold, it is a key transcriptional repressor to silence myeloid-specific genes and consequently switch cell fate towards the erythroid lineage[42]. In addition, the GFI1B-motif mutant did not completely eliminate the reporter activity of the CA. We therefore considered it to be unlikely that GFI1B alone dominantly regulates *GFI1*-SE. In contrast, we and other groups demonstrated previously that LSD1 forms a complex with CoREST, HDAC1, and HDAC2 to silence gene expression and that NCD38 can dissociate LSD1 from GFI1B and RUNX1[27, 43]. ChIP data clearly showed that all members of the LSD1 complex were significantly detached from *GFI1*-SE by NCD38 (Fig. 3b). These findings suggest that *GFI1*-SE can be sufficiently activated by pre-bound TFs after the LSD1 complex has detached from *GFI1*-SE following exposure to NCD38.

Attenuation of NCD38 or NCD25-induced myeloid differentiation by complete loss of *GFI1*-SE in erythroleukemia cells

We next examined whether *GFI1*-SE is essential for LSD1 inhibitor-mediated myeloid differentiation in erythroleukemia cells. Using the CRISPR-Cas9 genome editing system, sublines with complete loss of *GFI1*-SE (Δ*GFI1*-SE) were established from HEL cells (Supplementary Figure 9). ChIP analysis after reinduction of wild-type *GFI1*-SE or each single TF-motif mutant to Δ*GFI1*-SE cells showed that direct recruitment of TFs was reduced in all mutants, indicating that *GFI1*-SE is important as an aggregation and action point of multiple TFs (Supplementary Figure 10). The *GFI1* transcript level following exposure to NCD38 or NCD25 was reduced by approximately 80% in Δ*GFI1*-SE cells (Fig. 4a). This transcriptional reduction caused a decrease in the level of GFI1 protein (Fig. 4b). Finally, induction of CD11b by NCD38 or NCD25 was attenuated in Δ*GFI1*-SE cells (Fig. 4c). We also established sublines with deletion of the CA from murine erythroleukemia cells (ΔCA-MEL, Supplementary Figure 11a). In these sublines, induction of the *Gfi1* transcript and Gr-1 by NCD38 or NCD25 was suppressed (Supplementary Figure 11b-d). These data indicate that *GFI1*-SE is essential for the upregulation of GFI1 and subsequent myeloid differentiation following exposure to NCD38 or NCD25 in erythroleukemia cells.

GFI1-SE is required for the lineage switch from erythroid to myeloid and for the disruption of AML programs upon NCD38 treatment

To investigate the cellular machinery controlled by *GFI1*-SE upon NCD38 treatment in erythroleukemia cells, we conducted gene expression profiling of *GFI1*-SE-non-depleted (control) or Δ*GFI1*-SE HEL cells after treatment with DMSO or NCD38 for 24 hours (Fig. 5a). Subsequently, gene set enrichment analysis (GSEA) was performed to compare DMSO-treated control cells with NCD38-treated control cells to extract gene sets that are influenced by NCD38 treatment, as well as to compare NCD38-treated control cells with NCD38-treated Δ*GFI1*-SE cells to extract gene sets that are influenced by the *GFI1*-SE deletion upon NCD38 treatment. The LSD1 signature that we identified in our previous report[23] was induced by NCD38, whereas this induction was suppressed by the *GFI1*-SE deletion (Supplementary Figure 12), indicating that the gene induction by NCD38 is at least partially mediated by *GFI1*-SE. In terms of differentiation programs, gene sets regulating granulocyte differentiation and the CEBPA network[44, 45] were induced by NCD38 in control cells and suppressed by the *GFI1*-SE deletion upon NCD38 treatment (Fig. 5b). Sets of genes differentially expressed during erythroid development and GATA1 target genes[46, 47] were suppressed by NCD38 but were restored by the *GFI1*-SE deletion (Fig. 5c).

These data indicate that *GFI1*-SE is important in the NCD38-mediated switch from an erythroid to granulocytic lineage by positively regulating the CEBPA-directed machinery and negatively regulating the GATA1-directed machinery. Moreover, we performed GSEA using an officially available gene expression dataset of AML[48] (Fig. 5d and Supplementary Figure 13). Programs characterizing AML cluster 7 and cluster 8 were downregulated by NCD38, whereas this downregulation was significantly impaired by the *GFI1*-SE deletion. According to the original paper containing the dataset[48], the former cluster included all cases registered as FAB subtype M6 and the latter cluster included many cases with MDS-related chromosomal abnormalities including complex karyotypes, deletion of 7q, and monosomy 7. These data suggest that *GFI1*-SE may be important in the exertion of anti-leukemia effects by NCD38 against programs for erythroleukemia or AML with myelodysplasia-related changes.

Finally, we identified 225 genes that were induced more than two-fold by NCD38 when comparing DMSO-treated and NCD38-treated control HEL cells. Among these, only in 17 genes, the NCD38-mediated inductions were attenuated by less than 50% due to the *GFI1*-SE deletion, according to our comparison between DMSO-treated and NCD38-treated Δ*GFI1*-SE HEL cells (Supplementary Table 3). Similarly, we found 57

genes whose expression levels were suppressed by less than 50% by NCD38. Among these, only in 5 genes, the NCD38-mediated suppressions were recovered more than two-fold due to the *GFI1*-SE deletion. According to our previous data set (GSE68348)[23], most of these 22 genes were strongly influenced in CMK11-5 cells following exposure to NCD38. Gene ontology analysis using the Human Phenotype Ontology database revealed that these 22 genes influenced by the *GFI1*-SE deletion upon NCD38 treatment were enriched in signatures related to AML and neutropenia (Fig. 5e). These data suggested that this subset of genes is regulated in a *GFI1*-SE-dependent manner upon NCD38 treatment and may be involved in myeloid differentiation and AML-related programs.

Discussion

In this study, we demonstrated a relationship between myeloid differentiation and activation of *GFI1*-SE by two LSD1 inhibitors, NCD38 and NCD25, in erythro-megakaryocyte lineage leukemia cells. The effect of the LSD1 inhibitors on *GFI1*-SE activity was dependent on the conserved area between humans and mice. Multiple motifs of TFs including RUNX1, GATA1, CEBPA, TAL1, GFI1, ERG, and GFI1B were localized in this CA and mutation of each motif partially impaired the enhancer ability of *GFI1*-SE. The deletion of *GFI1*-SE attenuated the induction of GFI1 and CD11b by NCD38 and NCD25 in erythroleukemia cells. GSEA revealed that NCD38 induced a lineage switch from erythroid to myeloid by activating a granulocytic differentiation program and suppressing an erythroid development program in erythroleukemia cells. However, the *GFI1*-SE deletion clearly impaired this NCD38-induced lineage switch.

The TF network is an essential system for determining cell fate [1]. SEs of TF genes are reported to contain a number of TF-motifs and play an important role as an action point for TFs in controlling the overall TF network[3, 49]. GFI1 is a key regulator for neutrophil development and is reported to be negatively correlated to some TFs including GFI1B, which is a key regulator for erythroid development[1, 25, 42]. Therefore, *GFI1*-SE functions as

a key conductor in the TF network by aggregating the regulation by TFs and controlling GFI1 expression. Silencing of *GFI1*-SE is likely important in directing differentiation to the erythroid lineage.

Our ChIP results revealed that additional recruitment of TFs is not necessary for *GFI1*-SE activation by NCD38. Instead, histone-repressive modifiers including LSD1, CoREST, HDAC1, and HDAC2 were evicted from *GFI1*-SE by NCD38. These modifiers are known to form a complex and associate with several TFs[14, 43, 50, 51]. Our previous comprehensive proteome analysis found that LSD1 interacts with GFI1B and RUNX1, and that NCD38 selectively dissociates LSD1 from GFI1B without disruption of the LSD1-CoREST-HDAC1/2 complex in erythroleukemia cells[27]. This finding has been further supported by another report using a different LSD1 inhibitor[52]. Overall, our data suggest that *GFI1*-SE remains primed by the cooperative binding of TFs but can be sufficiently obscured by recruitment of the LSD1-CoREST-HDAC1/2 complex via TFs such as GFI1B in erythroleukemia cells. Therefore, cross-antagonism between GFI1 and GFI1B may occur by recruitment of the LSD1-CoREST-HDAC1/2 complex rather than by increased binding of opposing TFs. Moreover, a recent study reported that another LSD1 inhibitor was associated with increased chromatin accessibility in an MLL-AF9 mice

leukemia model[53]. Although this study demonstrated that this increased accessibility is mediated by a key TF, PU.1, our data suggest that dissociation of the LSD1-CoREST-HDAC1/2 complex from enhancers may also increase chromatin accessibility at least when using NCD38.

Our results demonstrated that *GFI1*-SE activation affects the anti-leukemia effect of NCD38 and NCD25 in erythro-megakaryocyte lineage leukemia cells. According to our GSEA results, the gene sets characterizing AML clusters including FAB subtype M6 or AML harboring MDS-related chromosomal abnormalities were impaired by NCD38, whereas this impairment was cancelled by the *GFI1*-SE deletion. Gene ontology analysis showed that genes whose expression change by NCD38 was canceled by the *GFI1*-SE deletion were enriched in signatures related to neutropenia and AML. Recent reports have accumulated evidence of a relationship between aberrant SE activation and leukemogenesis[11, 54-56]. Pharmacological suppression of aberrant SE activation has been considered as an option for leukemia treatment. In contrast, a direct association between aberrant repression of SEs and leukemogenesis has been demonstrated in only a few reports[57, 58]. Our data provide a novel finding that sustainable repression of *GFI1*-SE by LSD1 is essential for the sustenance of erythroleukemia cells. Furthermore, as this

repression can be easily reversed by the LSD1 inhibitors NCD38 and NCD25, LSD1-mediated repression of SEs is a promising target for erythroleukemia. However, it remains unclear what causes the differences in the responses of similar erythro-megakaryocyte lineage leukemia cells to NCD38 and NCD25. Although DNA methylation status in the CA was not significantly changed across these cells, a recent paper has demonstrated that a combination of LSD1 inhibitors and a DNA hypomethylating agent, 5-azacytidine, yielded improved efficacy against AML[59]. DNA methylation status surrounding the entire gene body of GFI1 might be involved. In addition, the delayed attenuation of CD235a and minimal growth suppression in UT7-EPO might be caused by activation of other SEs such as *ERG*-SE, as we previously demonstrated that *ERG*-SE activation by NCD38 contributed to the attenuation of CD235a but not CD11b[27]. Further clarification of the epigenetics anomaly network centered on SEs is required to develop more effective epigenetics-targeted therapies in erythroleukemia cells.

LSD1 inhibitors have been tested in several clinical trials targeting AML[21]. Although LSD1 inhibitors experimentally exert an anti-leukemia effect against multiple AML subtypes in mouse models and human cell lines, AML cells frequently exhibit resistance to LSD1 inhibitors[23, 24, 60]. Our data also suggest that the activation status of

GFI1-SE or subsequent *GFI1* upregulation might act as a surrogate marker to predict an anti-leukemia effect, at least in erythroleukemia and megakaryoblastic leukemia cells. Determining whether such SE reactivation correlates with the anti-leukemia effect of LSD1 inhibitors in future clinical studies will help to select patients suitable for LSD1 inhibitor administration. Furthermore, our GSEA results showed that a CEBPA network was induced by NCD38 and negatively regulated by the *GFI1*-SE deletion, while a gene set of GATA1 targets was suppressed by NCD38 and recovered by the *GFI1*-SE deletion, suggesting that the lineage switch from erythroid to myeloid by NCD38 is at least partially executed by machineries directed by CEBPA and GATA1 in erythroleukemia cells. In contrast, a gene set characterizing the AML cluster with CEBPA mutations was prone to enrichment in NCD38-treated cells and was not influenced by the *GFI1*-SE deletion. These data raise the possibility that CEBPA mutations may be an adverse marker of the anti-leukemia effect of NCD38. This possibility is also supported by a previous report in which deletion of *Cebpa* was described to generate resistance of leukemia cells to LSD1 inhibition, despite the use of a different LSD1 inhibitor and a different leukemia mouse model[53]. Further clinical studies are necessary to clarify what type of gene alterations influence the efficacy of LSD1 inhibitors.

In conclusion, activation of *GFI1*-SE by NCD38 is induced by dissociation of the LSD1-CoREST-HDAC1/2 complex from *GFI1*-SE and leads to the induction of programs related to the lineage switch from erythroid to myeloid, as well as impairment of programs related to erythroleukemia and AML with MDS-related chromosomal abnormalities. Sustainable repression of *GFI1*-SE by LSD1 is essential for the sustenance of erythroleukemia cells and therefore may be a surrogate adverse marker to predict the anti-leukemia effect of LSD1 inhibitors.

Acknowledgements

We would like to thank Hideyo Hirai (Department of Transfusion Medicine and Cell Therapy, Kyoto University) for help with vector construction. We also thank the Central Research Laboratory of Shiga University of Medical Science for assistance with quantitative PCR and the AMAXA transfection system. This work was supported by JSPS KAKENHI Grant Number JP16K09846 (MK) and the Takeda Science Foundation (MK).

Author contributions

G.T. performed all experiments, analyzed data, and wrote the paper. M.K. designed and conducted all experiments, analyzed data, provided funding, and wrote the paper. R.Y. and M.H. assisted with several experiments. T.S. provided LSD1 inhibitors. K.K., A.T-K. and A.A supervised the experiments and edited the paper.

References

1. Moignard V, Macaulay IC, Swiers G, Buettner F, Schutte J, Calero-Nieto FJ, *et al.* Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nature cell biology* 2013 Apr; **15**(4)**:** 363-372.

2. Tenen DG. Disruption of differentiation in human cancer: AML shows the way. *Nat Rev Cancer* 2003 Feb; **3**(2)**:** 89-101.

3. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretsky I, Jaitin DA, David E, *et al.* Immunogenetics. Chromatin state dynamics during blood formation. *Science* 2014 Aug 22; **345**(6199)**:** 943-949.

4. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 2013 Apr 11; **153**(2)**:** 307-319.

5. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell* 2013 Mar 7; **49**(5)**:** 825-837.

6. Aranda-Orgilles B, Saldana-Meyer R, Wang E, Trompouki E, Fassl A, Lau S, *et al.* MED12 Regulates HSC-Specific Enhancers Independently of Mediator Kinase Activity to Control Hematopoiesis. *Cell Stem Cell* 2016 Dec 1; **19**(6)**:** 784-799.

7. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, *et al.* Super-enhancers in the control of cell identity and disease. *Cell* 2013 Nov 7; **155**(4)**:** 934-947.

8. Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, *et al.* Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc Natl Acad Sci U S A* 2013 Oct 29; **110**(44)**:** 17921-17926.

9. Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, *et al.* Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol Cell* 2015 Apr 16; **58**(2)**:** 362-370.

10. Herranz D, Ambesi-Impiombato A, Palomero T, Schnell SA, Belver L, Wendorff AA, *et al.* A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat Med* 2014 Oct; **20**(10)**:** 1130-1137.

11. Groschel S, Sanders MA, Hoogenboezem R, de Wit E, Bouwman BAM, Erpelinck C, *et al.* A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* 2014 Apr 10; **157**(2)**:** 369-381.

12. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004 Dec 29; **119**(7)**:** 941-953.

13. Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 2005 Sep 16; **19**(6)**:** 857-864.

14. Lee MG, Wynder C, Cooch N, Shiekhattar R. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* 2005 Sep 15; **437**(7057)**:** 432-435.

15. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America* 2010 Dec 14; **107**(50)**:** 21931-21936.

16. Adam RC, Yang H, Rockowitz S, Larsen SB, Nikolova M, Oristian DS, *et al.* Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage

choice. *Nature* 2015 May 21; **521**(7552)**:** 366-370.

17. Kerenyi MA, Shao Z, Hsu YJ, Guo G, Luc S, O'Brien K, *et al.* Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. *Elife* 2013; **2:** e00633.

18. Harris WJ, Huang X, Lynch JT, Spencer GJ, Hitchin JR, Li Y, *et al.* The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer cell* 2012 Apr 17; **21**(4)**:** 473-487.

19. Niebel D, Kirfel J, Janzen V, Holler T, Majores M, Gutgemann I. Lysine-specific demethylase 1 (LSD1) in hematopoietic and lymphoid neoplasms. *Blood*, vol. 124: United States, 2014, pp 151-152.

20. Wada T, Koyama D, Kikuchi J, Honda H, Furukawa Y. Overexpression of the shortest isoform of histone demethylase LSD1 primes hematopoietic stem cells for malignant transformation. *Blood* 2015 Jun 11; **125**(24)**:** 3731-3746.

21. Magliulo D, Bernardi R, Messina S. Lysine-Specific Demethylase 1A as a Promising Target in Acute Myeloid Leukemia. *Front Oncol* 2018; **8:** 255.

22. Schenk T, Chen WC, Gollner S, Howell L, Jin L, Hebestreit K, *et al.* Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia. *Nat Med* 2012 Mar 11; **18**(4)**:** 605-611.

23. Sugino N, Kawahara M, Tatsumi G, Kanai A, Matsui H, Yamamoto R, *et al.* A novel LSD1 inhibitor NCD38 ameliorates MDS-related leukemia with complex karyotype by attenuating leukemia programs via activating super-enhancers. *Leukemia* 2017 Nov; **31**(11)**:** 2303-2314.

24. Maes T, Mascaro C, Tirapu I, Estiarte A, Ciceri F, Lunardi S, *et al.* ORY-1001, a Potent and Selective Covalent KDM1A Inhibitor, for the Treatment of Acute

Leukemia. *Cancer Cell* 2018 Mar 12; **33**(3)**:** 495-511.e412.

25. Karsunky H, Zeng H, Schmidt T, Zevnik B, Kluge R, Schmid KW, *et al.* Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. *Nat Genet* 2002 Mar; **30**(3)**:** 295-300.

26. Ogasawara D, Itoh Y, Tsumoto H, Kakizawa T, Mino K, Fukuhara K, *et al.* Lysine-specific demethylase 1-selective inactivators: protein-targeted drug delivery mechanism. *Angewandte Chemie* 2013 Aug 12; **52**(33)**:** 8620-8624.

27. Yamamoto R, Kawahara M, Ito S, Satoh J, Tatsumi G, Hishizawa M, *et al.* Selective dissociation between LSD1 and GFI1B by a LSD1 inhibitor NCD38 induces the activation of ERG super-enhancer in erythroleukemia cells. *Oncotarget* 2018 Apr 20; **9**(30)**:** 21007-21021.

28. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res* 2014 Jul; **42**(Web Server issue)**:** W401-407.

29. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res* 2016 Jul 8; **44**(W1)**:** W272-276.

30. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics* 2003 Jul; **34**(3)**:** 267-273.

31. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 2005 Oct 25; **102**(43)**:** 15545-15550.

32. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013 Apr 15; **14:** 128.

33. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016 Jul 08; **44**(W1)**:** W90-97.

34. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, *et al.* Integrative genomics viewer. *Nat Biotechnol*, vol. 29: United States, 2011, pp 24-26.

35. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013 Mar; **14**(2)**:** 178-192.

36. Ovcharenko I, Nobrega MA, Loots GG, Stubbs L. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res* 2004 Jul 1; **32**(Web Server issue)**:** W280-286.

37. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, *et al.* The human genome browser at UCSC. *Genome Res* 2002 Jun; **12**(6)**:** 996-1006.

38. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000 Sep 8; **302**(1)**:** 205-217.

39. Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, *et al.* JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2014 Jan; **42**(Database issue)**:** D142-147.

40. Tohyama K, Tohyama Y, Nakayama T, Ueda T, Nakamura T, Yoshida Y. A novel factor-dependent human myelodysplastic cell line, MDS92, contains haemopoietic

cells of several lineages. *British journal of haematology* 1995 Dec; **91**(4)**:** 795-799.

41. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016 May 19; **127**(20)**:** 2391-2405.

42. Saleque S, Cameron S, Orkin SH. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev* 2002 Feb 01; **16**(3)**:** 301-306.

43. Hu X, Li X, Valverde K, Fu X, Noguchi C, Qiu Y, *et al.* LSD1-mediated epigenetic modification is required for TAL1 function and hematopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* 2009 Jun 23; **106**(25)**:** 10141-10146.

44. Kastner P, Lawrence HJ, Waltzinger C, Ghyselinck NB, Chambon P, Chan S. Positive and negative regulation of granulopoiesis by endogenous RARalpha. *Blood* 2001 Mar 1; **97**(5)**:** 1314-1320.

45. Kamikubo Y, Zhao L, Wunderlich M, Corpora T, Hyde RK, Paul TA, *et al.* Accelerated leukemogenesis by truncated CBF beta-SMMHC defective in high-affinity binding with RUNX1. *Cancer Cell* 2010 May 18; **17**(5)**:** 455-468.

46. da Cunha AF, Brugnerotto AF, Duarte AS, Lanaro C, Costa GG, Saad ST, *et al.* Global gene expression reveals a set of new genes involved in the modification of cells during erythroid differentiation. *Cell Prolif* 2010 Jun; **43**(3)**:** 297-309.

47. Welch JJ, Watts JA, Vakoc CR, Yao Y, Wang H, Hardison RC, *et al.* Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* 2004 Nov 15; **104**(10)**:** 3136-3147.

48. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, *et al.* Prognostically useful gene-expression profiles

in acute myeloid leukemia. *N Engl J Med* 2004 Apr 15; **350**(16)**:** 1617-1628.

49. Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, *et al.* Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* 2015 Dec 17; **163**(7)**:** 1663-1677.

50. Saleque S, Kim J, Rooke HM, Orkin SH. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Molecular cell* 2007 Aug 17; **27**(4)**:** 562-572.

51. Hakimi MA, Bochar DA, Chenoweth J, Lane WS, Mandel G, Shiekhattar R. A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc Natl Acad Sci U S A* 2002 May 28; **99**(11)**:** 7420-7425.

52. Ishikawa Y, Gamo K, Yabuki M, Takagi S, Toyoshima K, Nakayama K, *et al.* A Novel LSD1 Inhibitor T-3775440 Disrupts GFI1B-Containing Complex Leading to Transdifferentiation and Impaired Growth of AML Cells. *Mol Cancer Ther* 2017 Feb; **16**(2)**:** 273-284.

53. Cusan M, Cai SF, Mohammad HP, Krivtsov A, Chramiec A, Loizou E, *et al.* LSD1 inhibition exerts its antileukemic effect by recommissioning PU.1- and C/EBPalpha-dependent enhancers in AML. *Blood* 2018 Apr 12; **131**(15)**:** 1730-1742.

54. Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MG, Qi J, *et al.* Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer cell* 2013 Dec 9; **24**(6)**:** 777-790.

55. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 2013 Apr 11; **153**(2)**:** 320-334.

56. Yamazaki H, Suzuki M, Otsuki A, Shimizu R, Bresnick EH, Engel JD, *et al.* A remote GATA2 hematopoietic enhancer drives leukemogenesis in inv(3)(q21;q26) by activating EVI1 expression. *Cancer cell* 2014 Apr 14; **25**(4)**:** 415-427.

57. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, *et al.* Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet* 2004 Jun; **36**(6)**:** 624-630.

58. Steidl U, Rosenbauer F, Verhaak RG, Gu X, Ebralidze A, Otu HH, *et al.* Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. *Nat Genet* 2006 Nov; **38**(11)**:** 1269-1277.

59. Duy C, Teater M, Garrett-Bakelman FE, Lee TC, Meydan C, Glass JL, *et al.* Rational Targeting of Cooperating Layers of the Epigenome Yields Enhanced Therapeutic Efficacy against AML. *Cancer Discov* 2019 Jul; **9**(7)**:** 872-889.

60. Fiskus W, Sharma S, Shah B, Portier BP, Devaraj SG, Liu K, *et al.* Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. *Leukemia* 2014 Nov; **28**(11)**:** 2155-2164.

Figure Legends

Fig. 1. Correlation of *GFI1*-SE re-activation with myeloid differentiation and anti-leukemia effects in erythroleukemia and megakaryocytic leukemia cells

(a) ChIP-qPCR analysis of H3K27ac in *GFI1*-SE, (b) quantitative PCR analysis of the *GFI1* mRNA level, (c) FACS analysis of CD11b, and (d) a cell proliferation assay, on HEL, CMK11-5, UT7-EPO and K562 cells after 48-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. The Y-axes in (a), (b), and (d) indicate the % input of H3K27ac, the % GAPDH, and the fold change of viable cells measured by MTS assay, respectively. The mean fluorescence intensity (MFI) is presented in (c). FACS experiments were performed independently twice. Other experiments were performed independently three times. The means (± SD) are shown in (a), (b), and (d) and representative data are shown in (c). Mean values are described on the top of each bar in (a) and (b). *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$.

Fig. 2. Role of TFs in *GFI1*-SE activation by NCD38 and NCD25 in erythroleukemia cells

(a) Homology search of *GFI1*-SE between human and mouse. The top portion shows the

reconstructed ChIP-sequencing data of H3K27ac around the *GFI1* locus and *GFI1*-SE using the DRA003526 data set and referencing human genome assembly hg19. The bottom portion shows a comparison of the human genome area presented in the top portion to the murine genome assembly mm10. The conserved area (CA), with homology of more than 70% over 100 base-pairs in *GFI1*-SE, is indicated by a red box. (b) Transcriptional activity analysis of *GFI1*-SE by a reporter assay. The left panel presents schematic diagrams of site-directed deletion or mutagenesis in the *GFI1*-SE region. The center and right panels show reporter activity of HEL transfectants with each indicated vector after a 48-hour treatment with 2 μM NCD38 and 2 μM NCD25, respectively, in comparison to DMSO. Each data set is presented as the relative ratio to the luciferase activity of an empty-vector transfectant treated with DMSO, after normalization of firefly luciferase activity to Renilla luciferase activity. Experiments were performed independently three times and the means (± SD) are shown. ΔCA: mutant lacking CA; Δnon-CA: mutant lacking the non-CA portion; GFI1/GFI1B mt: mutant of two GFI1/GFI1B-motifs; CEBPA #1 or #2 mt: mutant of a mutated CEBPA-motif #1 or #2, respectively; TAL1 mt: mutant of a TAL1-motif; RUNX1 mt: mutant of a RUNX1-motif; GATA1 mt: mutant of a GATA1-motif; 5TFs-mt: mutant of GFI1/GFI1B/CEBPA/TAL1/RUNX1/GATA1 multiple-motifs. The location of all

TF-motifs in *GFI1*-SE is indicated in Supplementary Figure 8. *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$ (ANOVA followed by Dunnett's test).

Fig. 3. Alteration in recruitment of TFs and the LSD1 complex on *GFI1*-SE after exposure to NCD38

(a) ChIP-qPCR analysis for GFI1, GFI1B, CEBPA, TAL1, RUNX1, GATA1, and ERG. (b) ChIP-qPCR analysis for LSD1, CoREST, HDAC1, and HDAC2. Data are presented as the fold change of enrichment calculated by dividing the % input of each TF on *GFI1*-SE in the NCD38 treatment cells by that in the DMSO treatment cells. Experiments were performed independently three times and the means (± SD) are shown.

Fig. 4. Attenuation of myeloid differentiation by NCD38 and NCD25 in Δ*GFI1*-SE erythroleukemia cells

(a) Changes in *GFI1* mRNA levels after 48-hour treatment with 2 μM NCD38 or NCD25. The fold change was calculated in each subline by dividing the *GFI1* mRNA level in NCD38 or NCD25 by that in DMSO after normalization to each *GAPDH* transcript. Experiments were performed independently three times and the means (± SD) are shown.

(b) Western blot analysis for the protein expression level of *GFI1* after 48-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. Actin is used as a control. (c) FACS analysis of CD11b. The histogram shows CD11b expression on the cell surface of each subline after 48-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. The filled histograms indicate DMSO-treated cells. The mean fluorescence intensity (MFI) is shown. The experiments in (b) and (c) were performed independently twice and representative data are shown.

Fig. 5. Involvement of *GFI1*-SE in the lineage switch from erythroid to myeloid programs and in impairment of AML programs upon NCD38 treatment

(a) Scatter plot of the expression profiling. Genes that are changed by more than 2-fold are plotted in log10 raw intensity values and representative gene names including *GFI1* are indicated. (b) GSEA of gene sets involved in regulation of granulocyte differentiation and the CEBPA network. (c) GSEA analysis of gene sets involved in differentially expressed genes during erythroid development and GATA1 targets. (d) GSEA analysis of gene sets of AML cluster 7 and 8 from datasets previously reported by Valk et al[48]. In (a) to (d), the left panels and right panels present results from comparisons between DMSO treatment and

NCD38 treatment in control HEL cells and between control HEL cells and Δ*GFI1*-SE HEL cells upon NCD38 treatment, respectively. The normalized enrichment score (NES) and the false discovery rate (FDR) q-value in GSEA are presented at the bottom of each panel. (e) Gene ontology analysis of genes influenced by *GFI1*-SE upon NCD38 treatment. The left pie charts show genes whose expression was significantly changed by NCD38 in white pies. Among these, genes whose NCD38-mediated inductions were attenuated by less than 50% or whose NCD38-mediated suppressions were recovered more than two-fold due to the *GFI1*-SE deletion are presented in black pies. The list of these genes and the fold changes are presented in Supplementary Table 3. The right panel presents the gene ontology result obtained from these 22 genes with Human Phenotype Ontology provided by the Enrichr software. FC denotes fold change.

Fig. 1

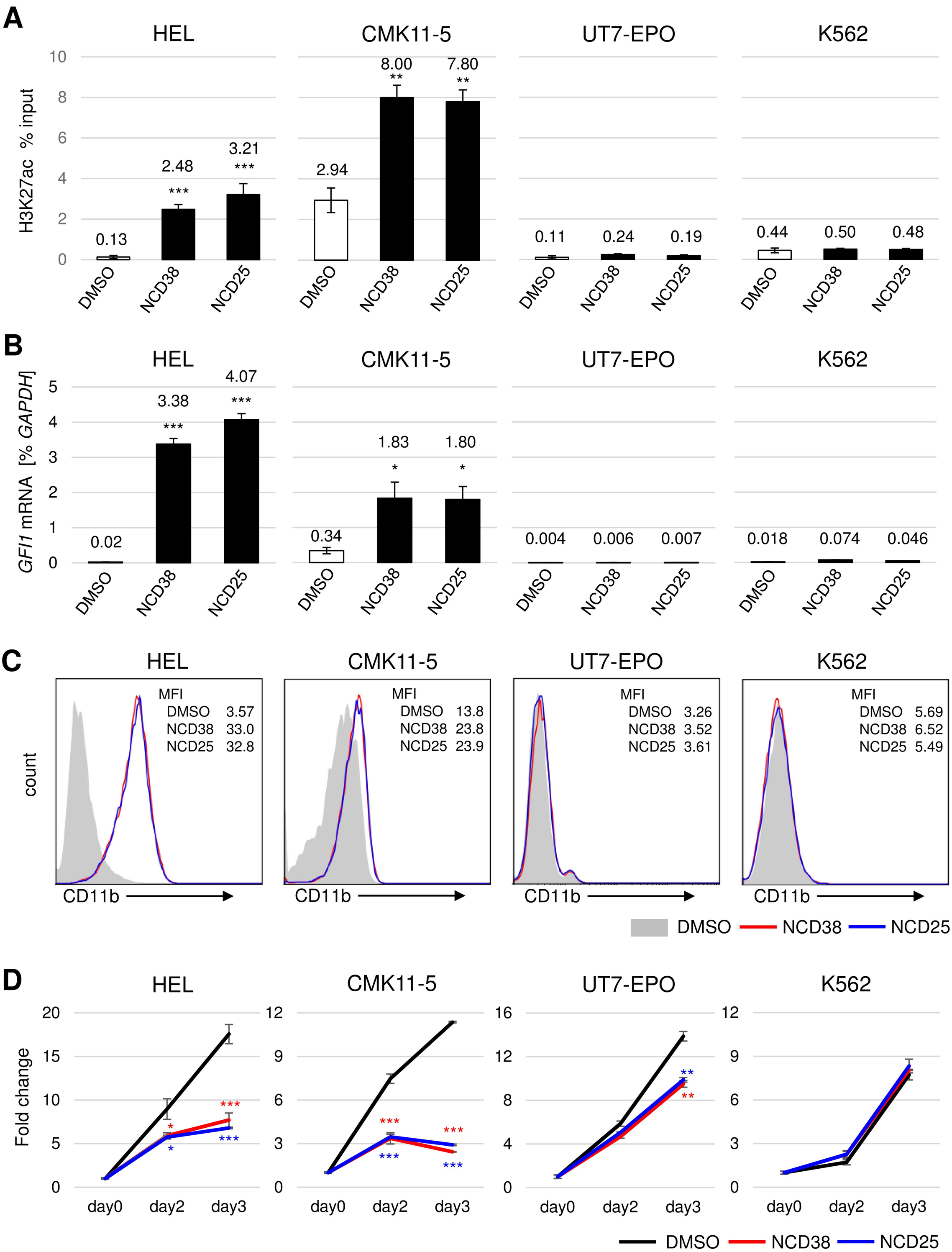
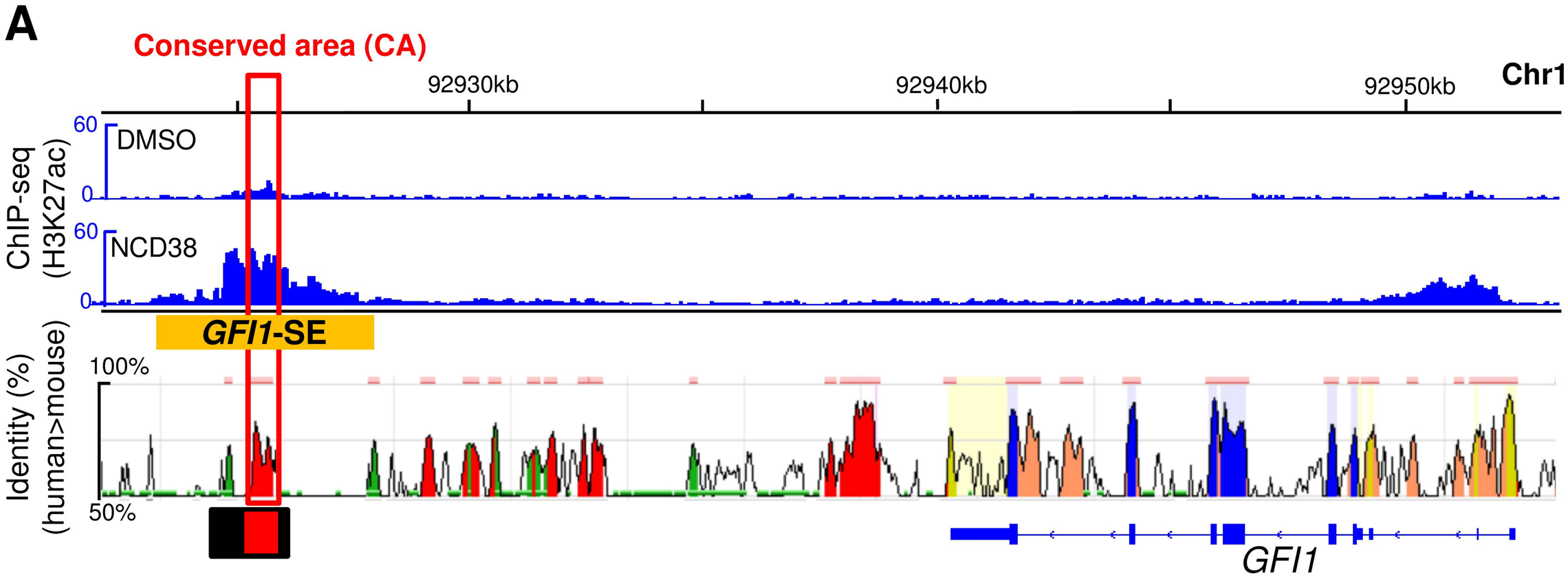

A
HEL
CMK11-5
UT7-EPO
K562
H3K27ac % input
0.13
2.48 ***
3.21 ***
2.94
8.00 **
7.80 **
0.11
0.24
0.19
0.44
0.50
0.48
DMSO
NCD38
NCD25
B
GFI1 mRNA [% GAPDH]
0.02
3.38 ***
4.07 ***
0.34
1.83 *
1.80 *
0.004
0.006
0.007
0.018
0.074
0.046
C
count
CD11b
MFI
DMSO 3.57
NCD38 33.0
NCD25 32.8
DMSO 13.8
NCD38 23.8
NCD25 23.9
DMSO 3.26
NCD38 3.52
NCD25 3.61
DMSO 5.69
NCD38 6.52
NCD25 5.49
D
Fold change
day0
day2
day3

Fig. 2


A
Conserved area (CA)
92930kb
92940kb
92950kb
Chr1
ChIP-seq (H3K27ac)
60
0
DMSO
NCD38
GFI1-SE
Identity (%) (human>mouse)
100%
50%
GFI1

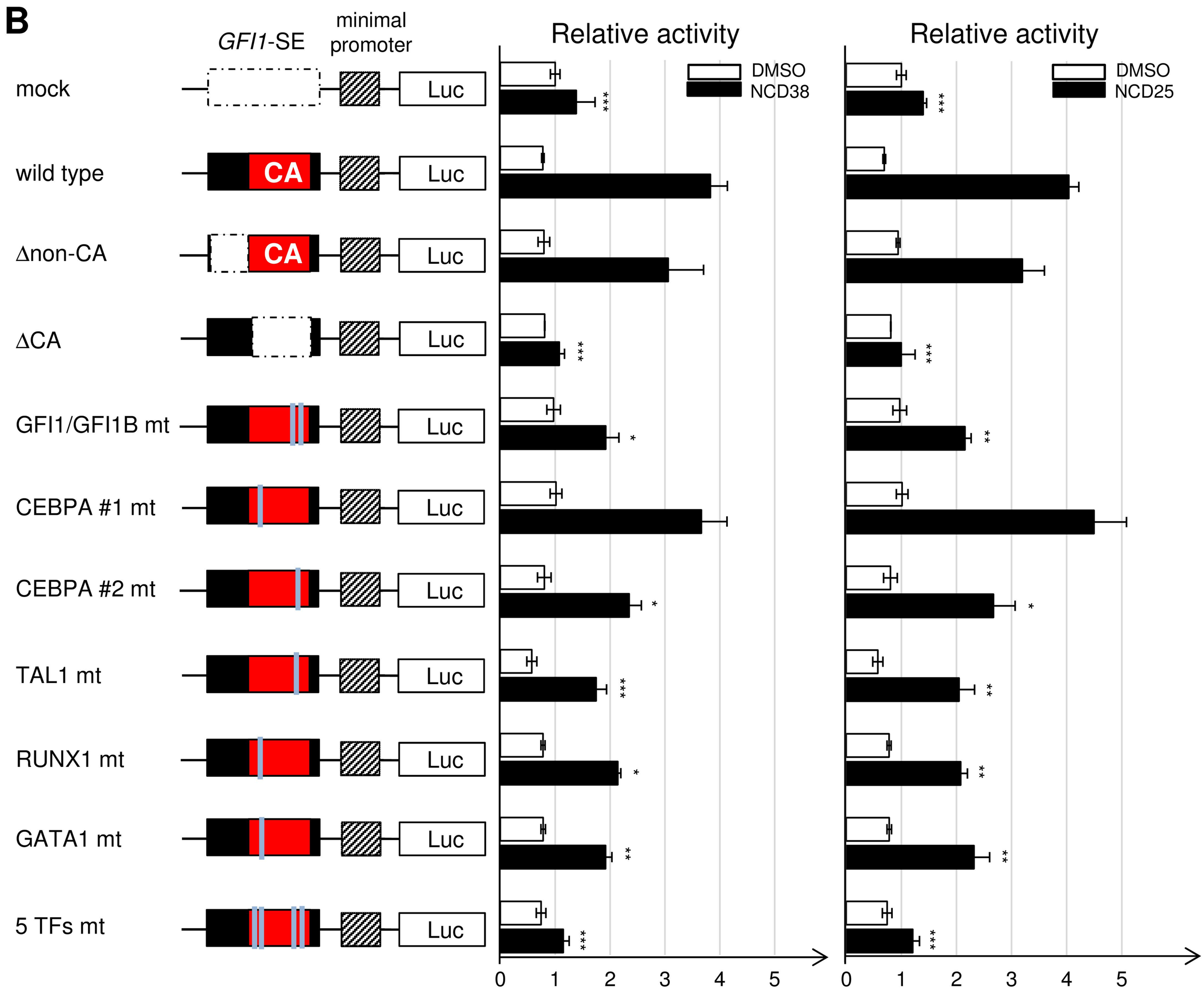
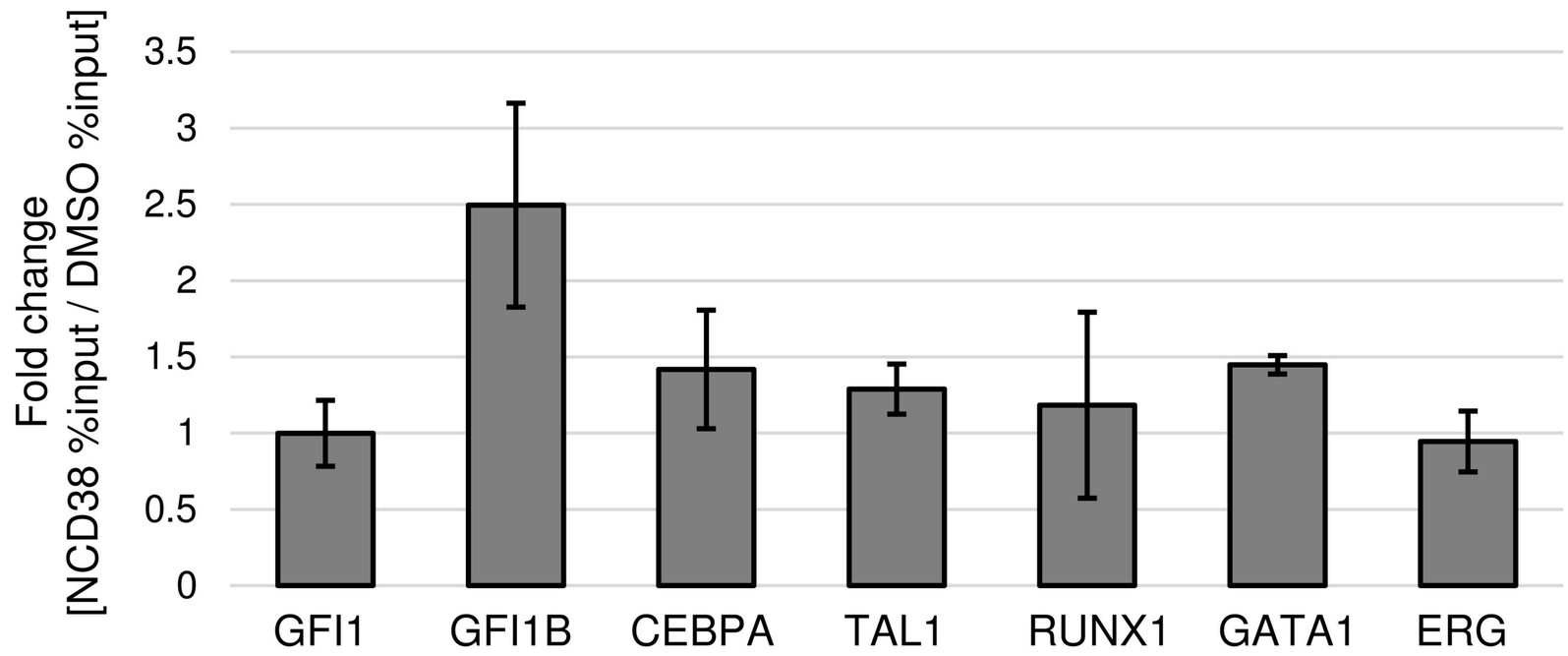

B
GFI1-SE
minimal promoter
Relative activity
DMSO
NCD38
NCD25
mock
wild type
Δnon-CA
ΔCA
GFI1/GFI1B mt
CEBPA #1 mt
CEBPA #2 mt
TAL1 mt
RUNX1 mt
GATA1 mt
5 TFs mt
CA
Luc
0
1
2
3
4
5

Fig. 3

A


Fold change
[NCD38 %input / DMSO %input]
3.5
3
2.5
2
1.5
1
0.5
0
GFI1
GFI1B
CEBPA
TAL1
RUNX1
GATA1
ERG

B

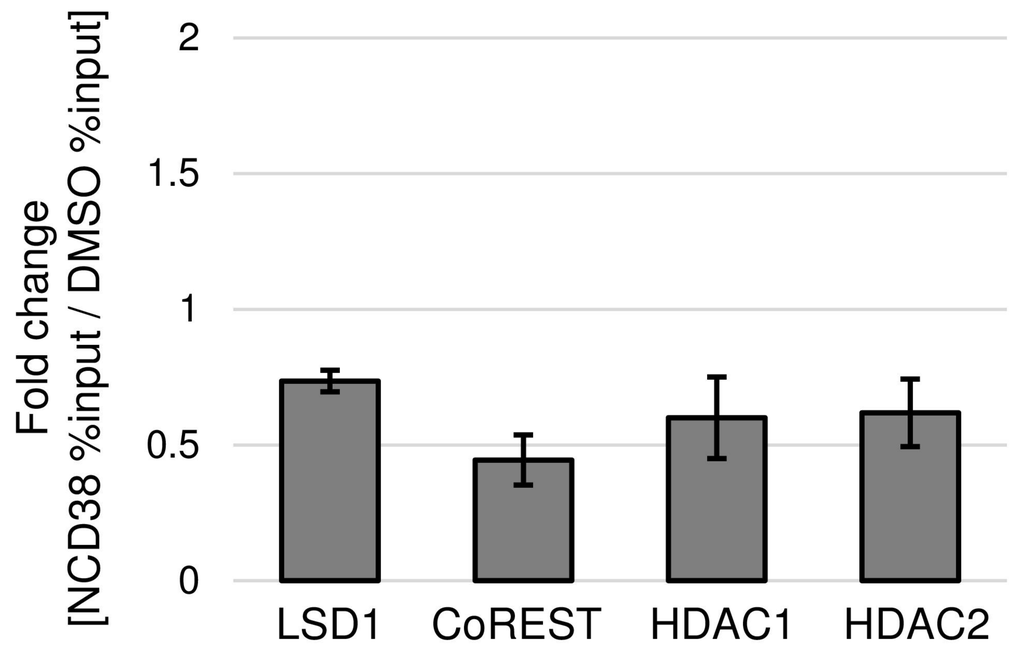
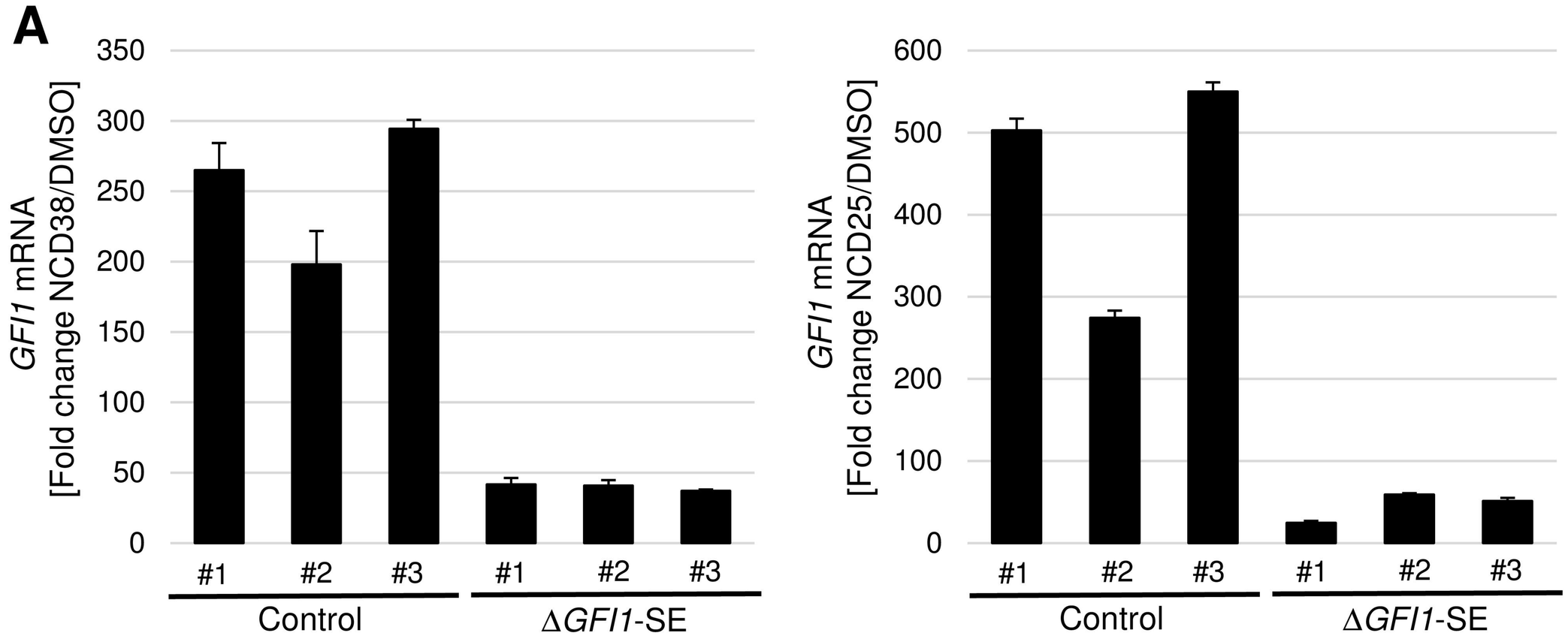
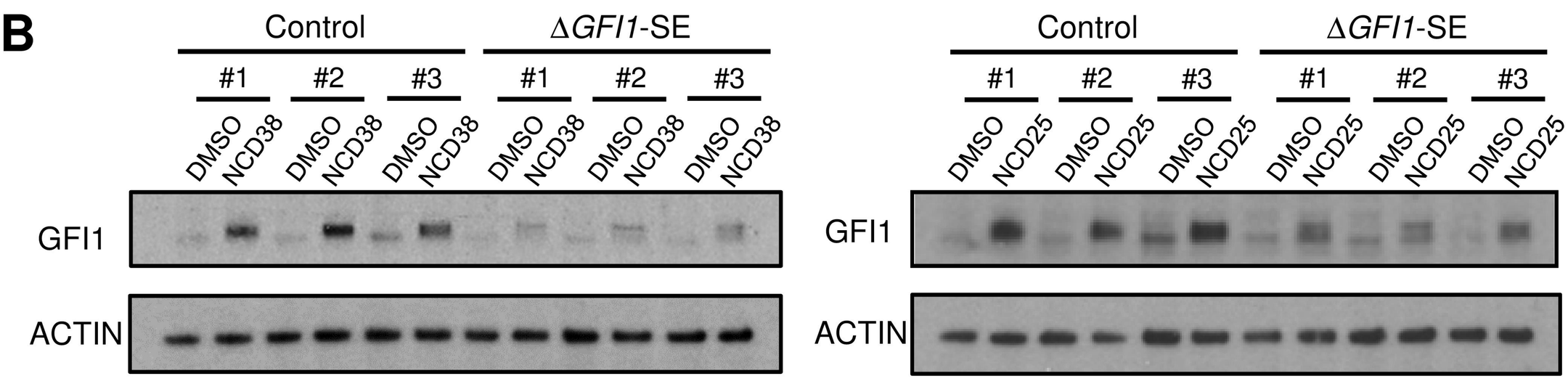

Fold change
[NCD38 %input / DMSO %input]
2
1.5
1
0.5
0
LSD1
CoREST
HDAC1
HDAC2

Fig. 4

A


GFI1 mRNA
[Fold change NCD38/DMSO]
350
300
250
200
150
100
50
0
#1 #2 #3 #1 #2 #3
Control
ΔGFI1-SE
GFI1 mRNA
[Fold change NCD25/DMSO]
600
500
400
300
200
100
0
#1 #2 #3 #1 #2 #3
Control
ΔGFI1-SE

B


Control
ΔGFI1-SE
#1 #2 #3 #1 #2 #3
DMSO NCD38 DMSO NCD38 DMSO NCD38 DMSO NCD38 DMSO NCD38 DMSO NCD38
GFI1
ACTIN
Control
ΔGFI1-SE
#1 #2 #3 #1 #2 #3
DMSO NCD25 DMSO NCD25 DMSO NCD25 DMSO NCD25 DMSO NCD25 DMSO NCD25
GFI1
ACTIN

C

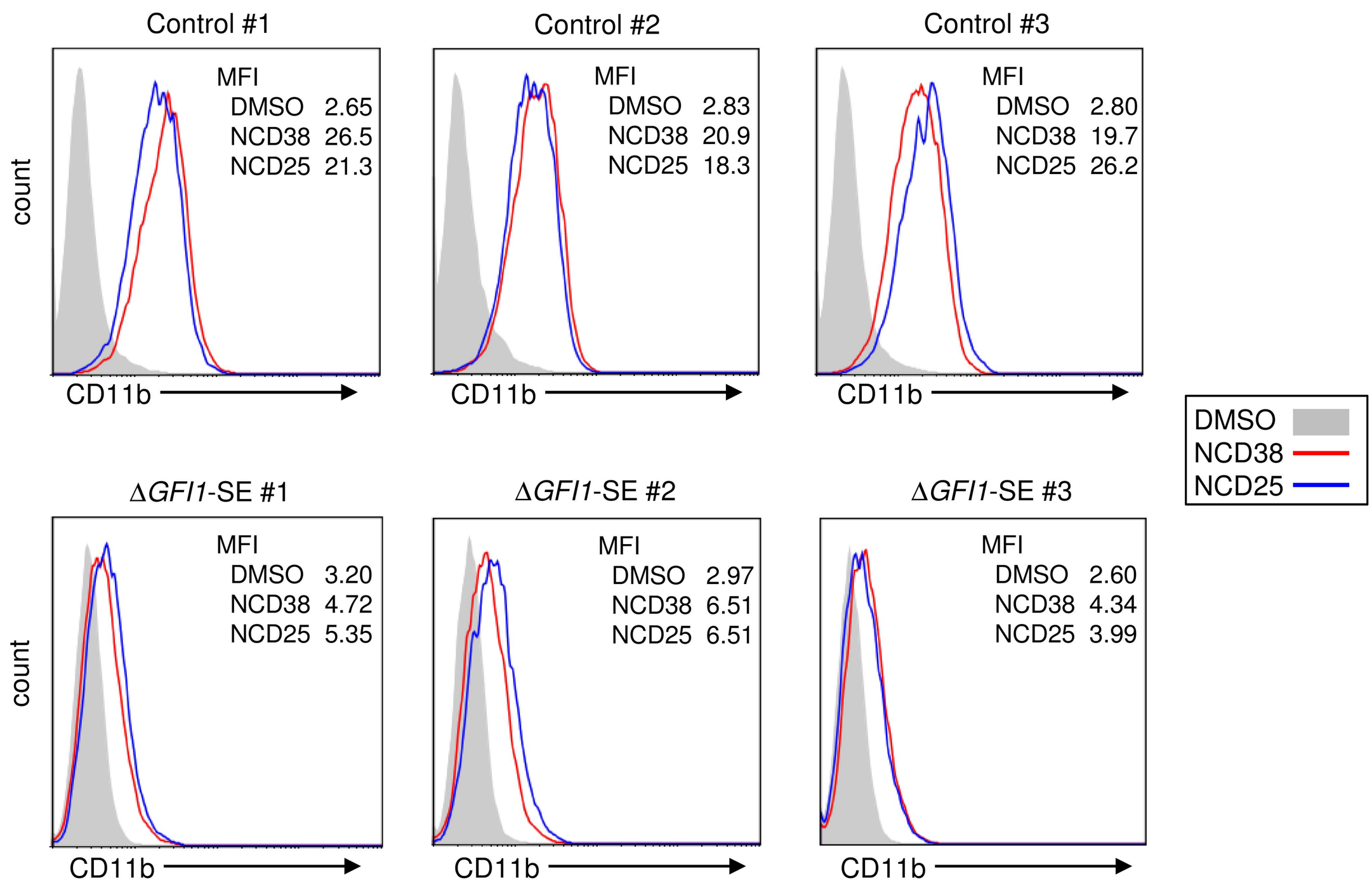
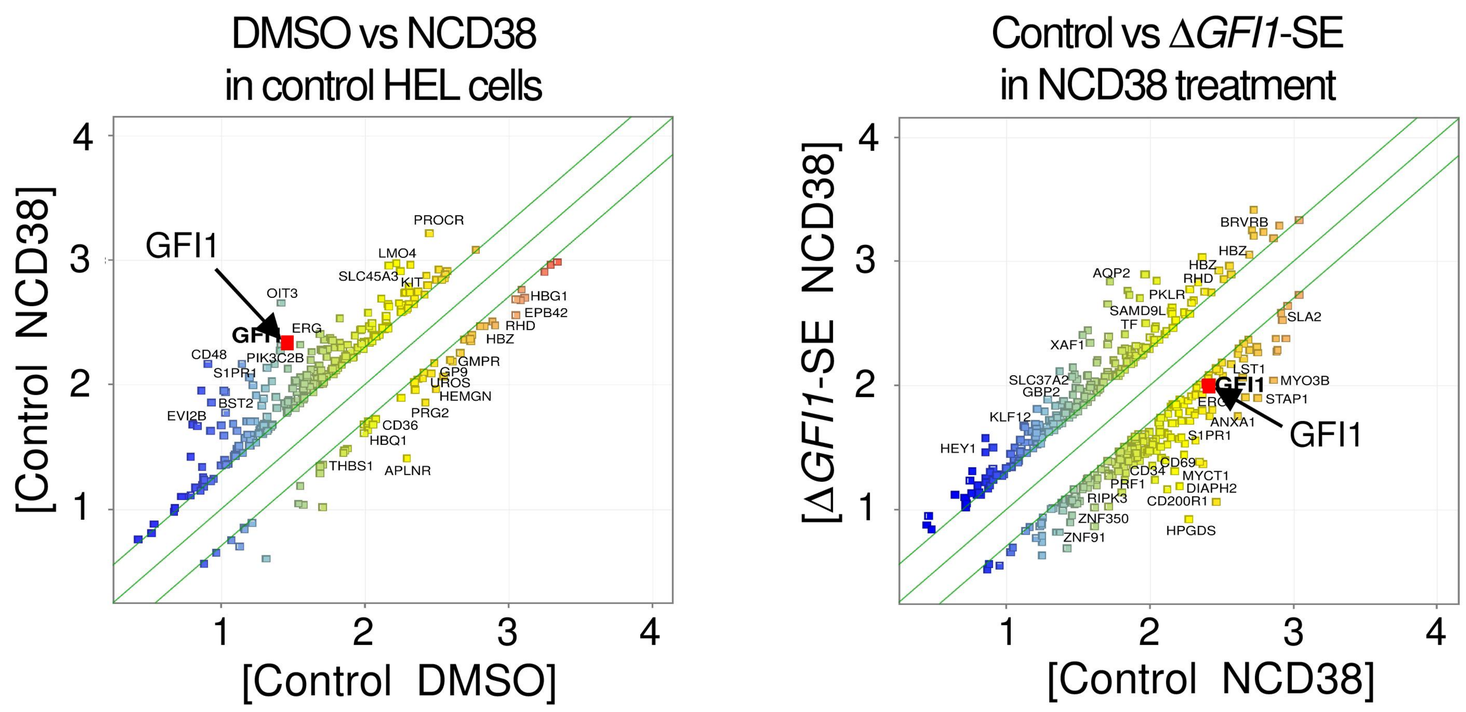
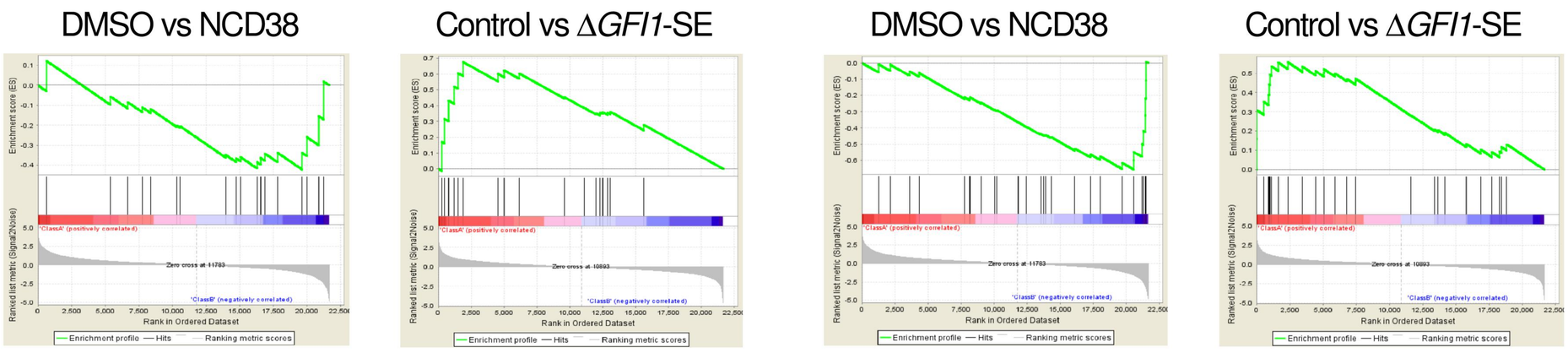

Control #1
MFI
DMSO 2.65
NCD38 26.5
NCD25 21.3
count
CD11b
Control #2
MFI
DMSO 2.83
NCD38 20.9
NCD25 18.3
CD11b
Control #3
MFI
DMSO 2.80
NCD38 19.7
NCD25 26.2
CD11b
DMSO
NCD38
NCD25
ΔGFI1-SE #1
MFI
DMSO 3.20
NCD38 4.72
NCD25 5.35
count
CD11b
ΔGFI1-SE #2
MFI
DMSO 2.97
NCD38 6.51
NCD25 6.51
CD11b
ΔGFI1-SE #3
MFI
DMSO 2.60
NCD38 4.34
NCD25 3.99
CD11b

Fig. 5

A

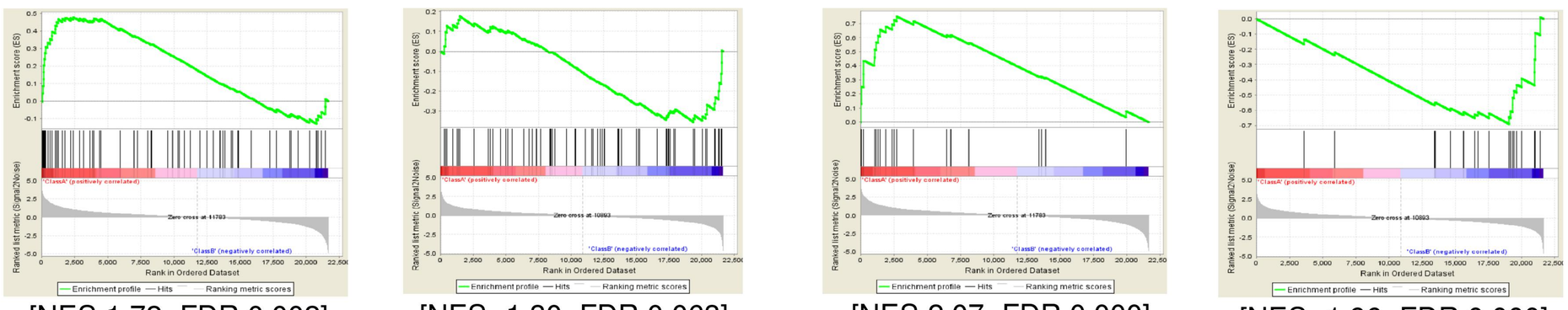


B



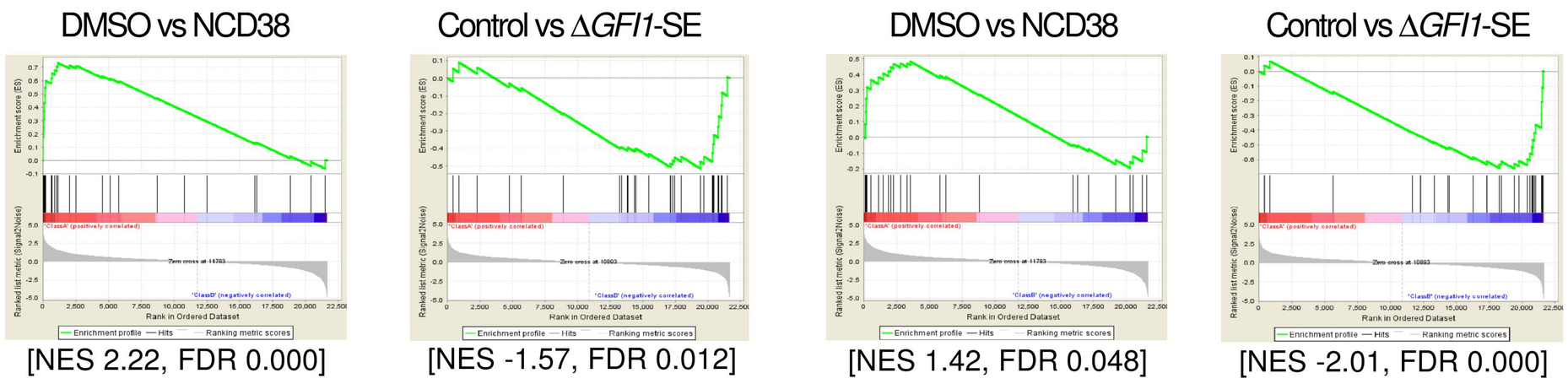
[NES -1.20, FDR 0.095] [NES 1.90, FDR 0.000] [NES -1.98, FDR 0.000] [NES 1.79, FDR 0.002]

C



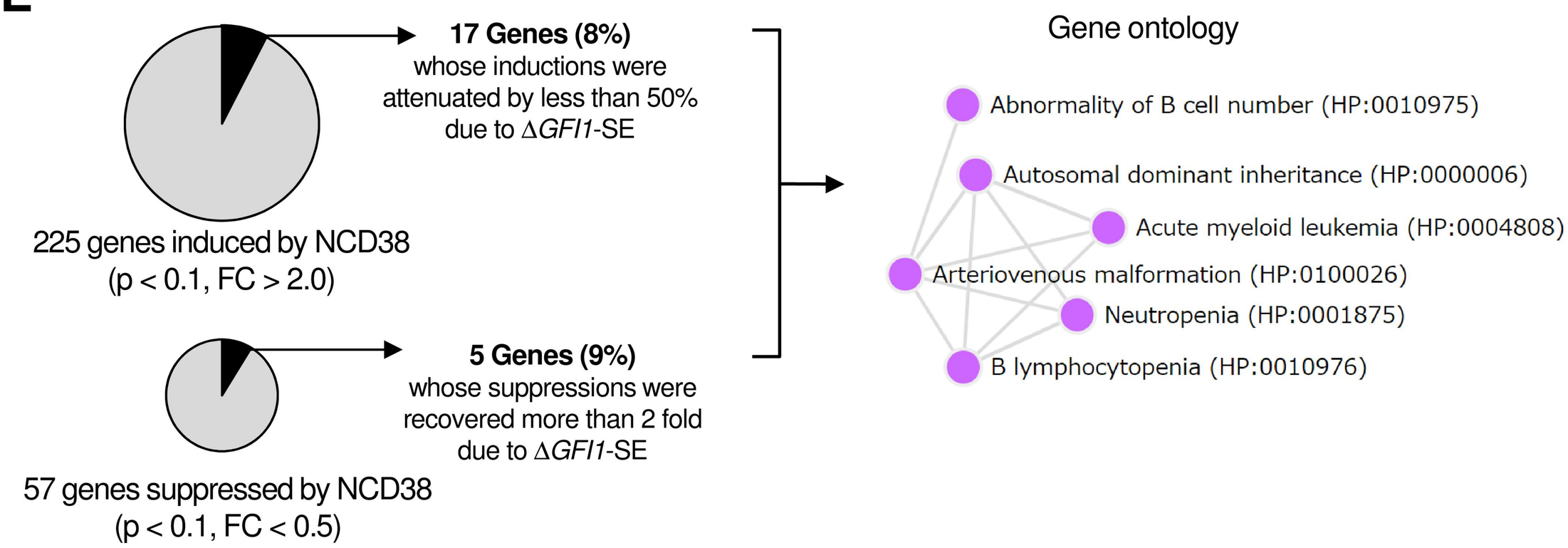
[NES 1.72, FDR 0.002] [NES -1.30, FDR 0.063] [NES 2.07, FDR 0.000] [NES -1.96, FDR 0.000]

D



[NES 2.22, FDR 0.000] [NES -1.57, FDR 0.012] [NES 1.42, FDR 0.048] [NES -2.01, FDR 0.000]

E



LSD1-mediated repression of *GFI1* super-enhancer plays an essential role in erythroleukemia

Goichi Tatsumi[1,2], Masahiro Kawahara[2]*, Ryusuke Yamamoto[1], Masakatsu Hishizawa[1], Katsuyuki Kito[2], Takayoshi Suzuki[3,4], Akifumi Takaori-Kondo[1] and Akira Andoh[2]

Affiliations

[1] Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University.

[2] Department of Medicine, Shiga University of Medical Science.

[3] The Institute of Scientific and Industrial Research, Osaka University.

[4] CREST, Japan Science and Technology Agency (JST)

Contact information

*Corresponding author: Masahiro Kawahara

E-mail: mkawahar@belle.shiga-med.ac.jp

Address: Seta-Tsukinowa, Otsu, Shiga, 520-2192, Japan

Tel: +81-77-548-2217

Fax: +81-77-548-2219

Running title: *GFI1* super-enhancer repression by LSD1 in leukemia

Supplementary Information

includes

Supplementary Materials and Methods

Supplementary Figures 1-13

Supplementary Tables 1-3

Supplementary materials and methods

Cells

THP-1 cells and murine erythroleukemia cells (MEL) were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin. MDS-L cells[1] were cultured in the same medium with 20 ng/ml human interleukin-3 (PeproTech, Rocky Hill, NJ) and 20 μM 2-mercaptoethanol.

Quantitative PCR

RNA was extracted using either TRIzol reagent (Life Technologies, Carlsbad, CA). Complementary RNA was synthesized by Superscript 3 reverse transcriptase (Thermo Fisher Scientific, Waltham, MA). Quantitative PCR was performed using LightCycler480 System II (Roche, Basel, Switzerland) and a THUNDERBIRD SYBR qPCR mix (Toyobo, Osaka, Japan). Absolute numbers were calculated using the recombinant DNA of targeted amplicons for the standard curve. The primer sequences are provided in Supplementary Table 1.

Western blotting

Total cell lysates were extracted in lysis buffer [50 mM Tris-HCl at pH 8.0, 150 mM NaCl, 1% Triton X, 1 mM PMSF, 1 mM EDTA, and protease inhibitor cocktail (Nacalai Tesque, Kyoto, Japan)]. We used primary antibodies anti-LSD1 (C69G12, Cell Signaling Technology, Tokyo, Japan) and anti-ACTIN (sc-1616, Santa Cruz Biotechnology, Dallas, TX). HRP-conjugated anti-rabbit (NA934v, GE Healthcare, Little Chalfont, UK) or anti-goat (sc-2020, Santa Cruz Biotechnology) were used as secondary antibodies.

Cell surface marker assay

For cell surface marker analysis, murine erythroleukemia cells were stained with anti-Ly6G(Gr-1)-PE-Cy7 (1A8; BD Biosciences, San Jose, CA), and human leukemia cells were stained with anti-CD11b-PE-Cy5 (ICRF44; eBioscience, San Diego, CA) and anti-CD235a-PE-Cy5 (GA-R2; BD Biosciences), and then analyzed by FACS Calibur or CantoII (BD Biosciences).

Cytology

Twenty-thousand cells were suspended in phosphate-buffered saline and attached to glass slides by centrifugation at 800 rpm for 4 minutes using Cytospin 4 (Thermo Fisher

Scientific). The glass slides were stained with a traditional Wright-Giemsa staining.

Bisulfite sequencing

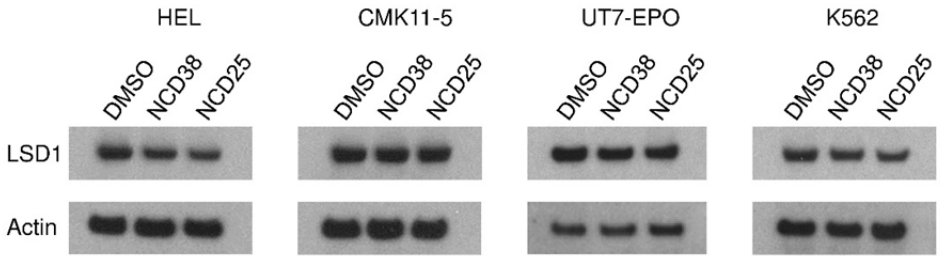
DNA methylation analysis was performed as described previously[2]. Briefly, genomic DNA was isolated from cells using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany), and bisulfite treatment was performed using the MethylEasy Xceed Kit (Human Genetic Signatures, Sydney, Australia). Modified DNA was amplified by PCR and cloned into a pGEM-T Easy Vector System (Promega, Madison, WI). The bisulfite sequencing-specific primers are listed in Supplementary Table 1. The independent colonies were amplified with the Illustra TempliPhi Amplification kit (GE Healthcare) or NucleoSpin Plasmid EasyPure kit (Macherey-Nagel, Düren, Germany), and then sequenced.

Chromatin immunoprecipitation (ChIP)

ChIP was performed as described previously[3]. Briefly, after crosslinking with 0.5% paraformaldehyde for 10 min and quenching with 100 mM glycine for 10 min, cells were lysed and incubated in lysis buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1mM EDTA pH 8.0, 10% glycerol, 0.5% NP-40, 0.25% TritonX-10) on ice for 10 min. After intensive

washes, the pellets were resuspended in shearing buffer (0.1% SDS, 1 mM EDTA pH 8.0, 10 mM Tris-HCl pH 8.0) and sonicated using S220 ultrasonicators (Covaris, Woburn, MA). For precipitation, anti-H3K27ac (39133; Active Motif, Carlsbad, CA), anti-GFI1 (sc-376949; Santa Cruz Biotechnology), anti-GFI1B (sc-28356X; Santa Cruz Biotechnology), anti-CEBPA (sc-61X; Santa Cruz Biotechnology), anti-TAL1 (ab155195; Abcam, Cambridge, UK), anti-GATA1 (ab11852; Abcam), anti-RUNX1 (ab23980; Abcam), anti- ERG (ab133264; Abcam), anti-LSD1 (ab17721; Abcam), anti-CoREST (ab32631; Abcam), anti-HDAC1 (ab7028; Abcam) or anti-HDAC2 (ab7029; Abcam), and Dynabeads Protein G (Thermo Fisher Scientific) were used. After reverse crosslinking, ChIP DNA was purified with a MinElute PCR Purification kit (Qiagen) and analyzed by quantitative PCR.

Supplementary Figure 1

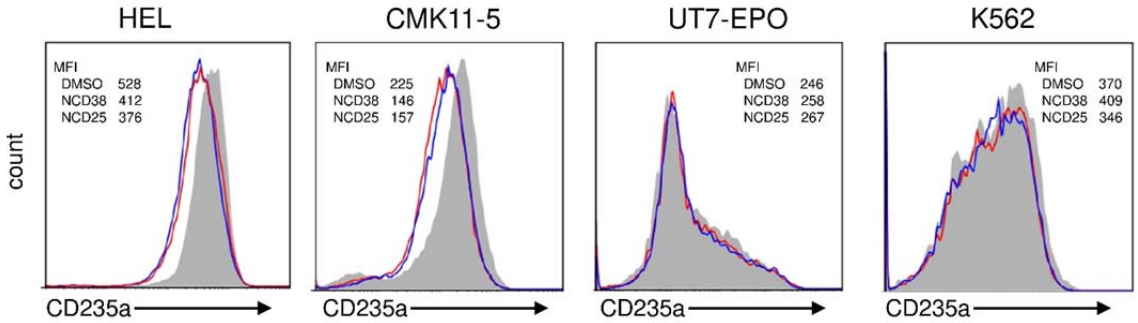


Supplementary Figure 1

LSD1 expression levels in erythro-megakaryocytic leukemia cell lines after exposure to NCD38 or NCD25

Western blot analysis for the protein expression level of LSD1 after 48-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. Actin was used as a control. Experiments were performed independently twice and representative data are shown.

Supplementary Figure 2

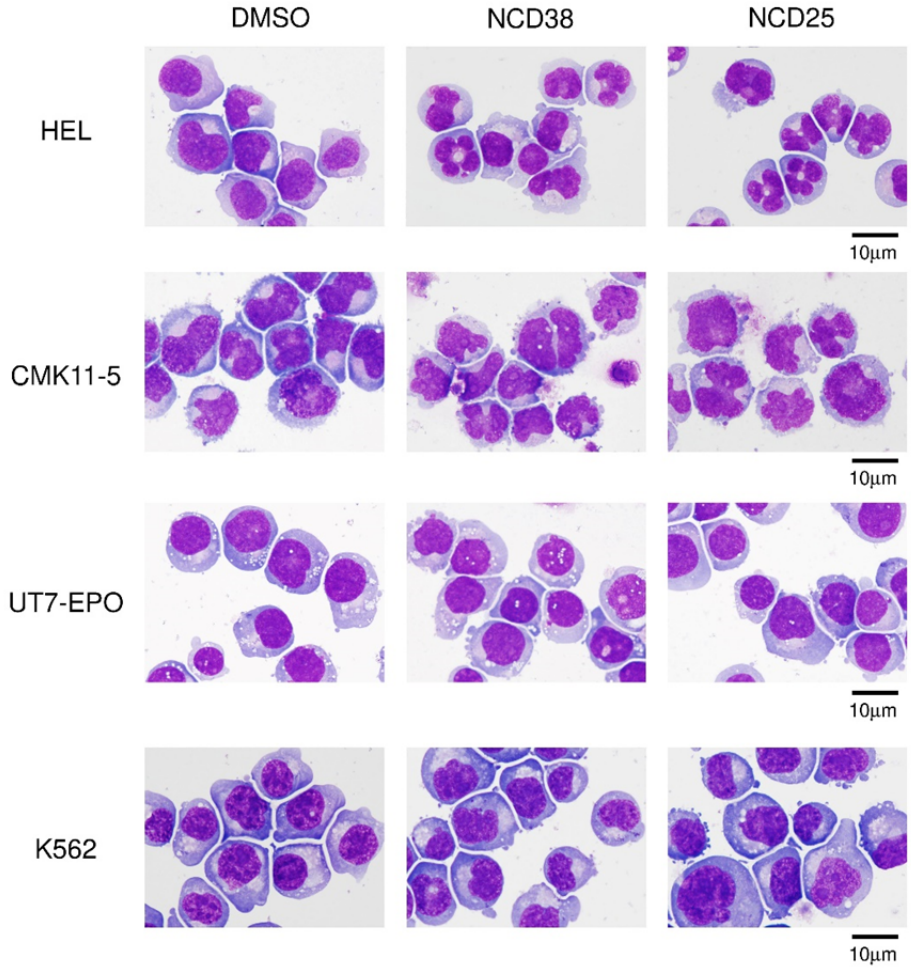


Supplementary Figure 2

Expression change of an erythroid marker, CD235a, in erythro-megakaryocytic leukemia cell lines after exposure to NCD38 or NCD25

FACS analysis of CD235a in HEL, CMK11-5, UT7-EPO and K562 cells after 48-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. The mean fluorescence intensity (MFI) is presented. Experiments were performed independently three times and representative data are shown.

Supplementary Figure 3

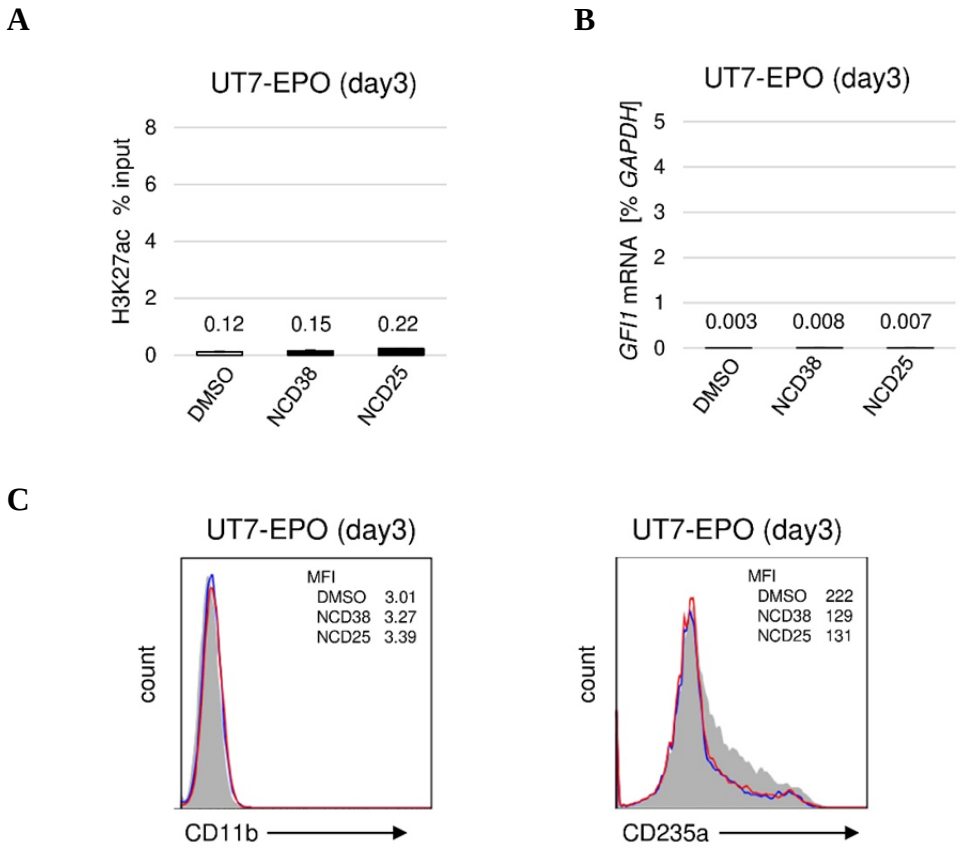


Supplementary Figure 3

Cytological changes in erythro-megakaryocytic leukemia cell lines after exposure to NCD38 or NCD25

Morphology of HEL, CMK11-5, UT7-EPO and K562 cells after 48-hour treatment with 2 μM NCD38, NCD25 or DMSO. Cytospin slides were stained with traditional Wright-Giemsa staining. Representative pictures are shown. Scale bar indicates 10 μm.

Supplementary Figure 4



Supplementary Figure 4

H3K27ac in *GFI1*-SE, GFI1 transcript, and cell surface markers CD11b and CD235a in UT7-EPO cells after 72-hour treatment with NCD25 or NCD38

(a) ChIP-qPCR analysis of H3K27ac in *GFI1*-SE, (b) quantitative PCR analysis of the *GFI1* mRNA level and (c) FACS analysis of CD11b and CD235a in UT7-EPO cells after 72-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. The Y-axes in (a) and (b)

1 indicate the % input of H3K27ac and the % GAPDH, respectively. The mean fluorescence

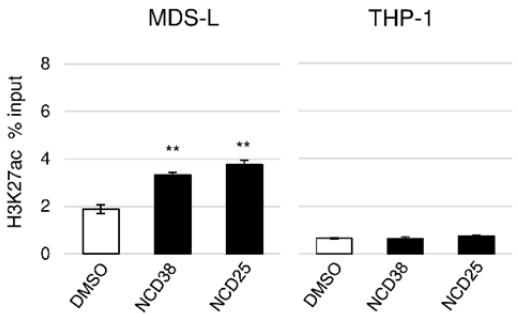
2 intensity (MFI) is presented in (c). All experiments were performed independently three

3 times. The means (±SD) are shown in (a) and (b), and representative data are shown in (c).

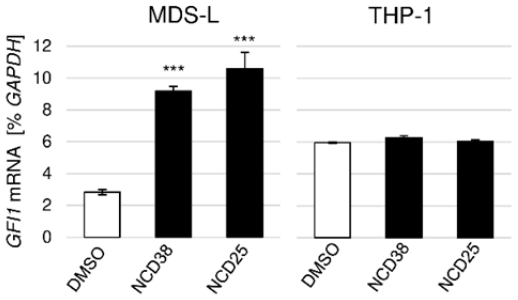
4

Supplementary Figure 5

A



B



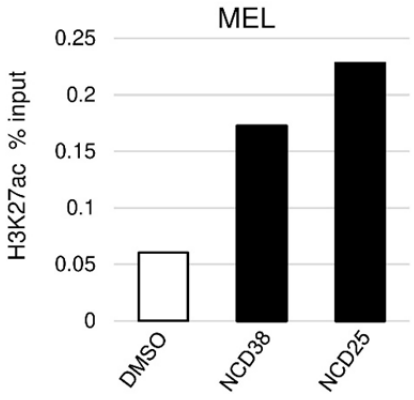
Supplementary Figure 5

***GFI1*-SE activation and *GFI1* expression status in MDS-L and THP-1 cells after exposure to NCD38 or NCD25**

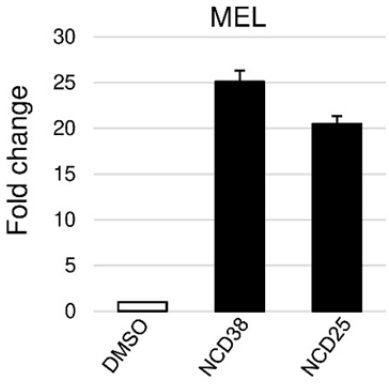
(a) ChIP-qPCR analysis of H3K27ac in *GFI1*-SE and (b) quantitative PCR analysis of the *GFI1* mRNA level in MDS-L and THP-1 cells after 48-hour treatment with 2 μM NCD38, 2 μM NCD25, or DMSO. The Y-axes in (a) and (b) indicate the % input of H3K27ac and the % GAPDH. All experiments were performed independently three times and the means (±SD) are shown. *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$.

Supplementary Figure 6

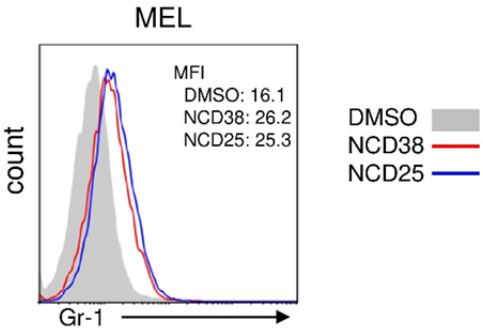
A.



B.



C.



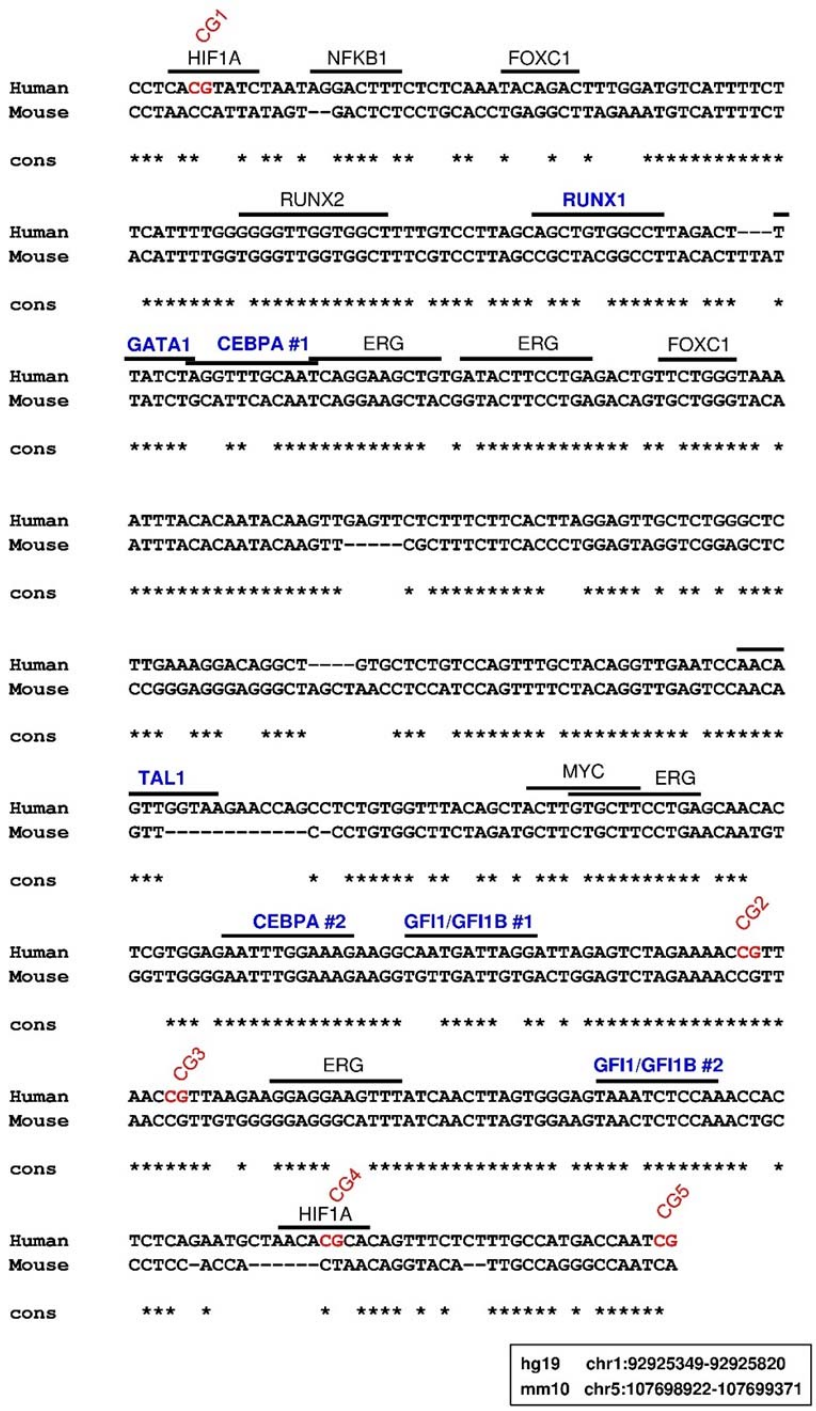
Supplementary Figure 6

Correlation of CA re-activation with myeloid differentiation in murine

erythroleukemia (MEL) cells

(a) ChIP-qPCR analysis of H3K27ac in the CA portion, (b) quantitative PCR analysis of the *Gfi1* mRNA level and (c) FACS analysis of Gr-1 in MEL cells after 48-hour treatment with LSD1 inhibitors including 2 μM NCD38 and 2 μM NCD25, or DMSO as a control. The Y-axes in (a) and (b) indicate % input of H3K27ac and the relative ratio of *Gfi1* mRNA in NCD38 or NCD25 to that in DMSO after normalization by the internal control *Gapdh* mRNA. Mean fluorescence intensity (MFI) is presented in (c). ChIP-qPCR and FACS experiments were performed independently twice. Quantitative PCR experiments were performed independently three times. The means (± SD) are shown in (a) and (b), and representative data are shown in (c).

Supplementary Figure 7

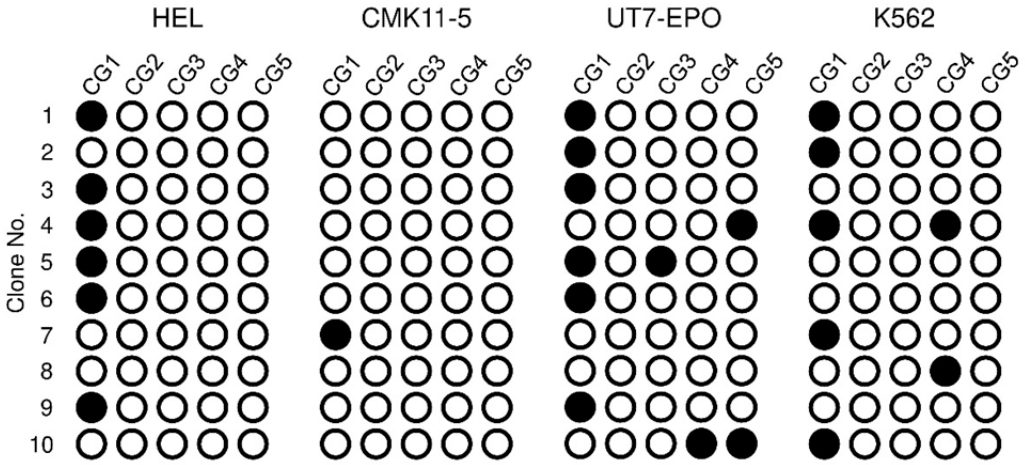


Supplementary Figure 7

TF motifs and CpG sites in the CA of *GFI1*-SE

Nucleotide sequence alignment between human and mouse of the CA of *GFI1*-SE. The lines above the sequences indicate the TF binding motifs that are related to myeloid differentiation. Several TF binding motifs presented in blue letters are the targets of mutagenesis in the luciferase assay (Fig. 2b). Five CpG sites in the CA of *GFI1*-SE are highlighted in red and each CpG site is numbered. The asterisks indicate the conserved nucleotides.

Supplementary Figure 8

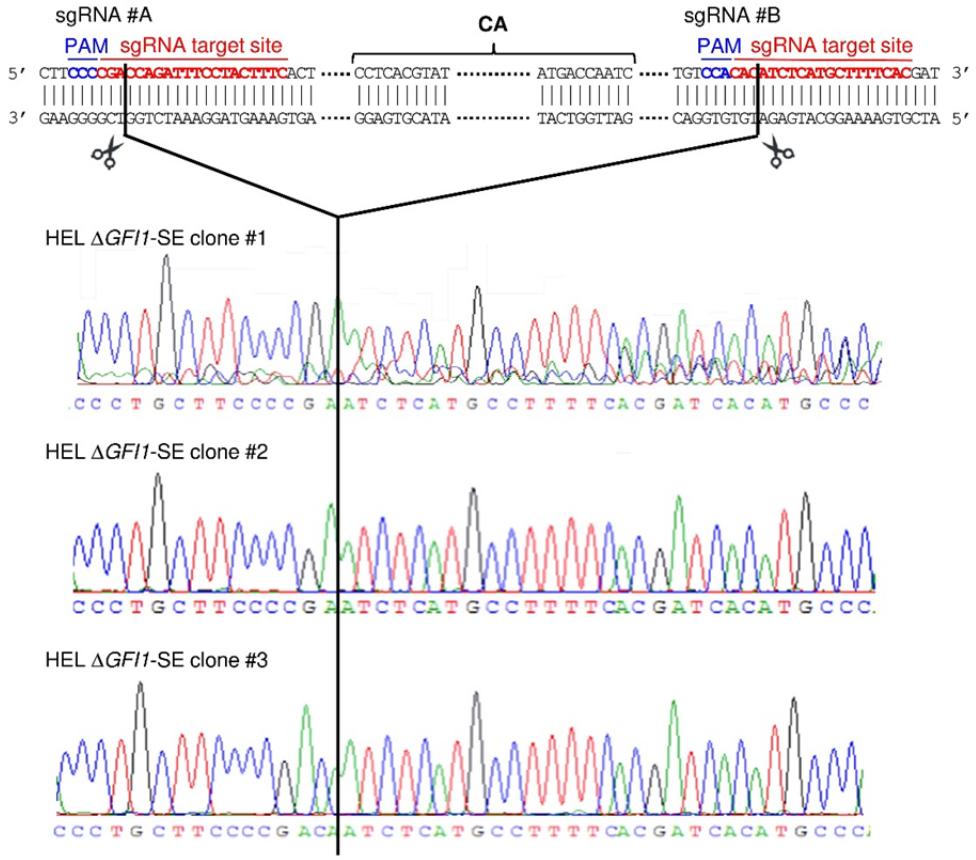


Supplementary Figure 8

DNA methylation status in the CA of *GFI1*-SE in erythro-megakaryocytic leukemia cell lines

CpG methylation status by bisulfite sequencing in the CA of *GFI1*-SE. Transverse rows and vertical lines represent each single clone and each CpG site, respectively. Solid and open circles indicate methylated and unmethylated CG sites, respectively. Five CpG sites in the CA of *GFI1*-SE are indicated in Supplementary Figure 7.

Supplementary Figure 9



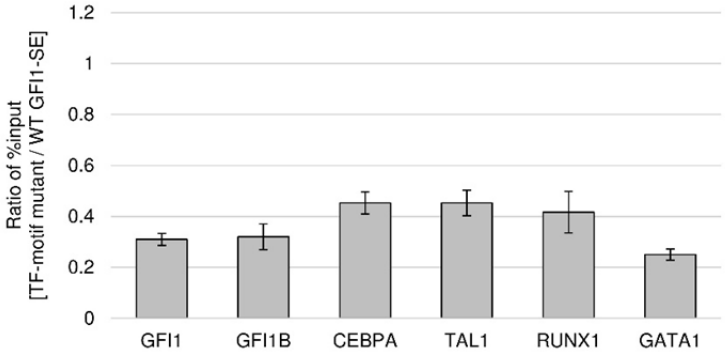
Supplementary Figure 9

Generation of Δ*GFI1*-SE HEL cells by CRISPR-Cas9

Establishment of Δ*GFI1*-SE sublines of HEL cells. The upper part shows a schematic diagram of the establishment of subclones harboring a *GFI1*-SE knockout allele using the CRISPR-Cas9 genome editing system. The vertical lines indicate the predicted Cas9 cleavage between two sgRNA sequences. The lower part shows Sanger-sequencing results

of the targeted region of three Δ*GFI1*-SE sublines (Δ*GFI1*-SE).

Supplementary Figure 10

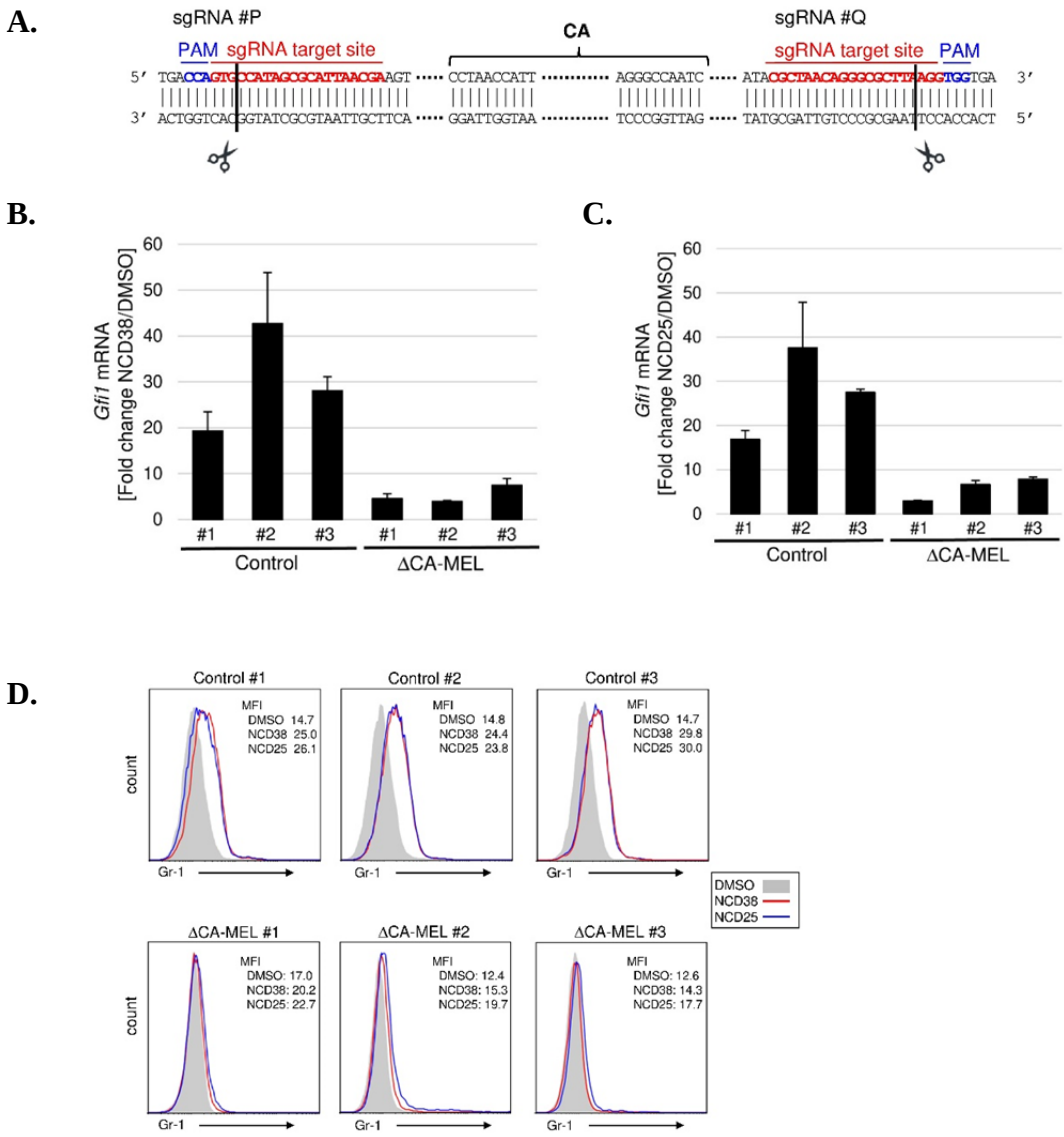


Supplementary Figure 10

Reduction of direct TF recruitment in single TF-motif mutated *GFI1*-SE in Δ*GFI1*-SE cells

ChIP-qPCR analysis for GFI1, GFI1B, CEBPA, TAL1, RUNX1, and GATA1 after reinduction of each single TF-motif mutant vector or wild-type *GFI1*-SE vector into Δ*GFI1*-SE cells. Data are shown as the ratio of % input of indicated TFs on each TF-motif mutated *GFI1*-SE to that in wild-type *GFI1*-SE. Experiments were performed independently three times and the means (± SD) are shown.

Supplementary Figure 11



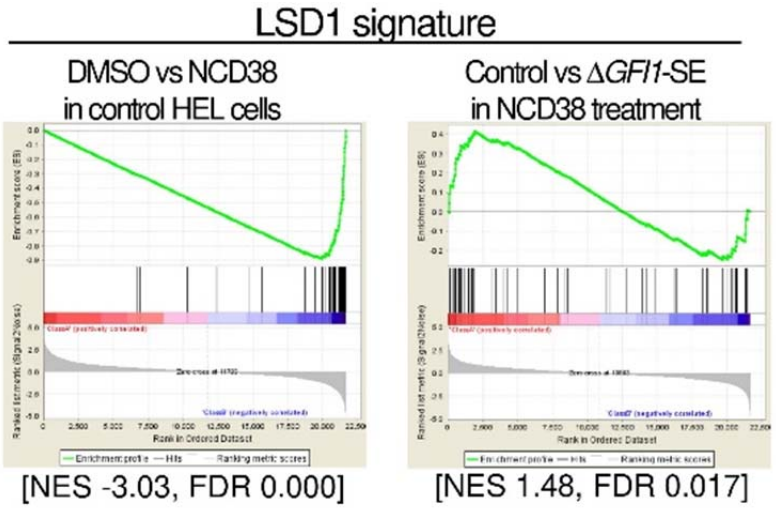
Supplementary Figure 11

Attenuation of myeloid differentiation by NCD38 and NCD25 in ΔCA murine erythroleukemia (MEL) cells

(a) Schematic diagram of the establishment of subclones harboring the ΔCA allele

(MEL-ΔCA) using the CRISPR-Cas9 genome editing system. The vertical lines indicate the predicted Cas9 cleavage between two sgRNA sequences. (b, c) Fold change of *Gfi1* mRNA after 48-hour treatment with 2 μM NCD38 and NCD25. The fold change was calculated in each subline by dividing *Gfi1* mRNA level in NCD38 or NCD25 by that in DMSO. Experiments were performed independently three times and the means (± SD) are shown. (d) FACS analysis of Gr-1 (Ly-6G). Histogram shows Gr-1 expression on the cell surface of each subline after 48-hour treatment with 2 μM NCD38, NCD25 or DMSO. The filled histograms indicate DMSO-treated cells. The mean fluorescence intensity (MFI) is shown. The experiments were performed independently twice and representative data are shown.

Supplementary Figure 12

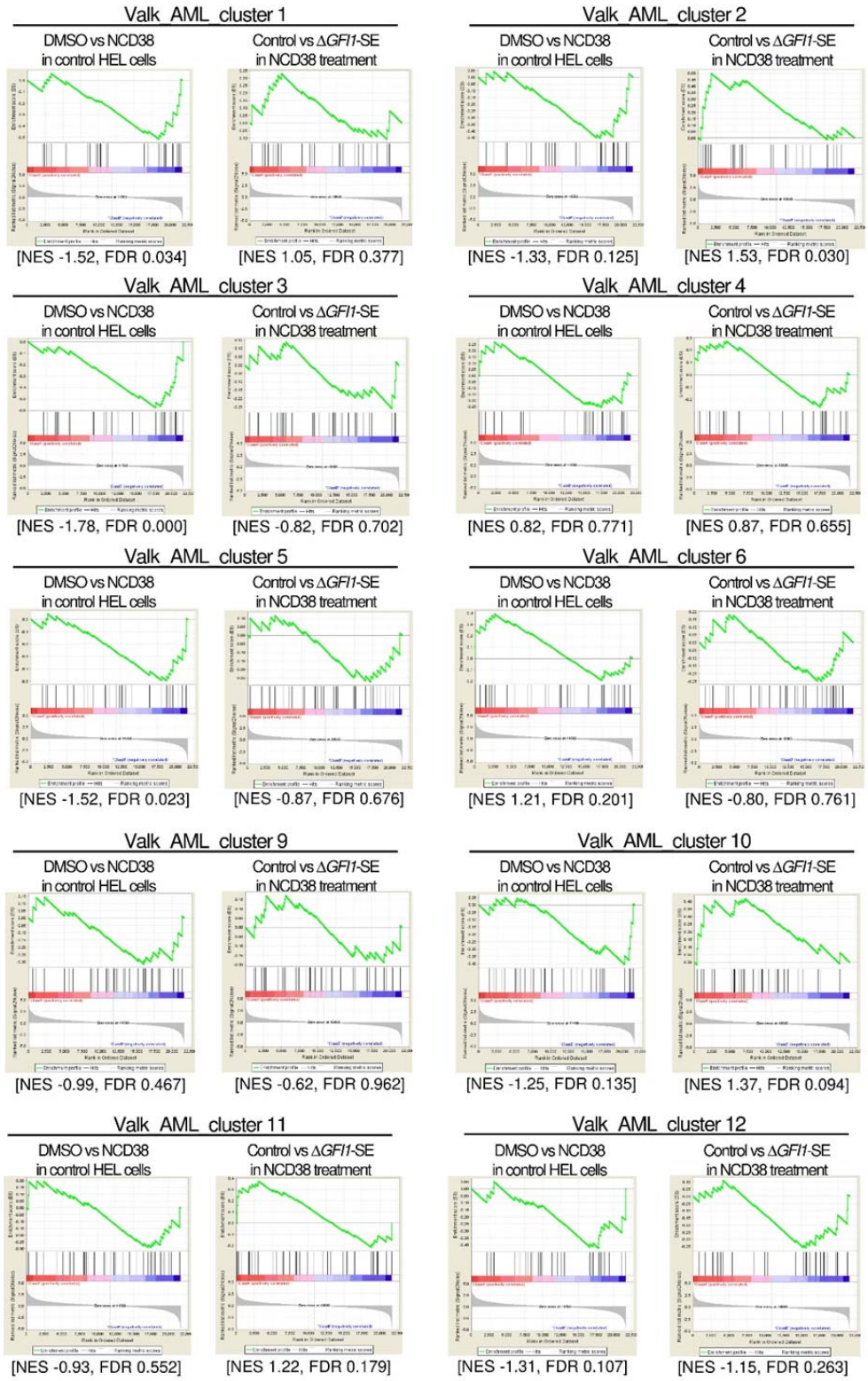


Supplementary Figure 12

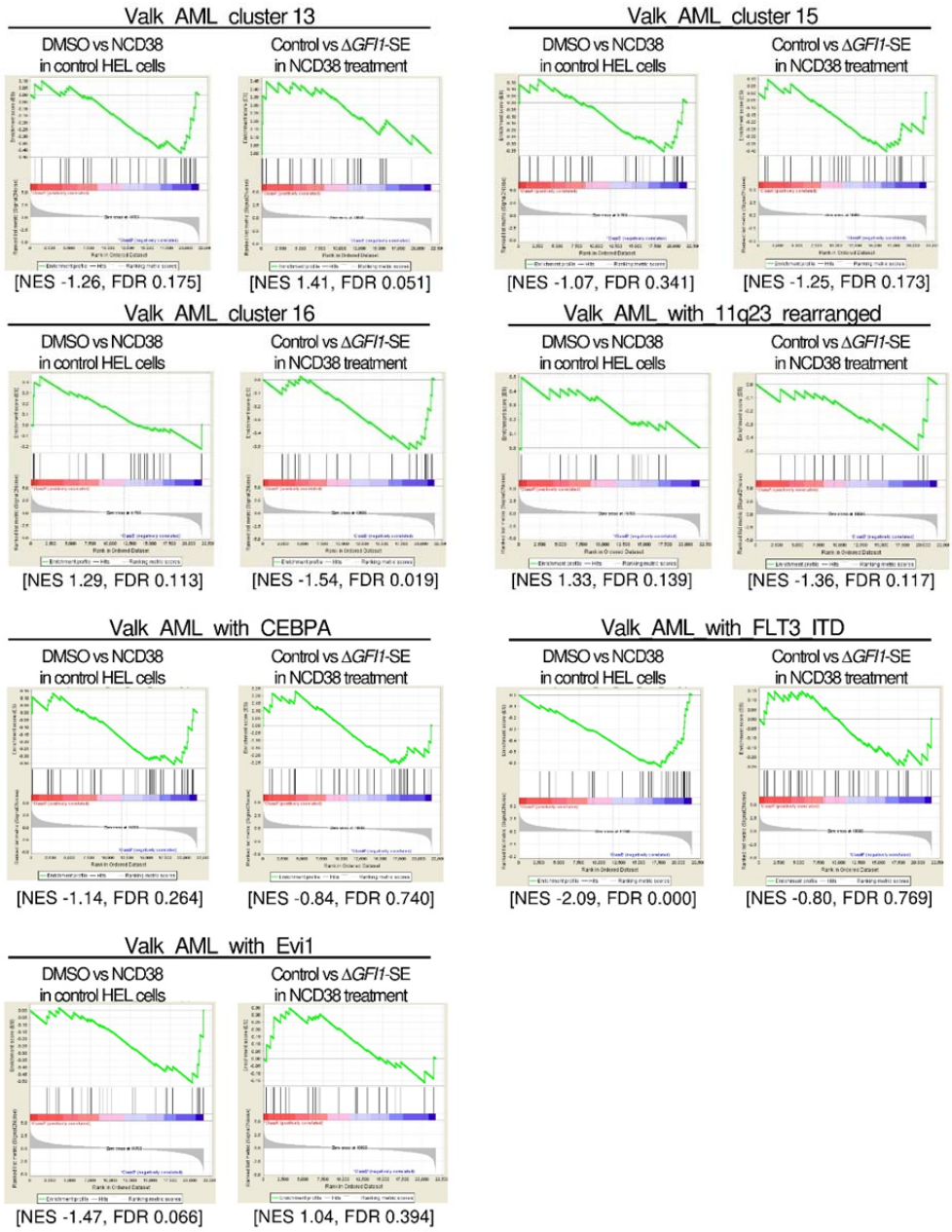
GSEA analysis of the LSD1 signature

The left and right panels present results from comparison between DMSO treatment and NCD38 treatment in control HEL cells and between control HEL cells and Δ*GFI1*-SE HEL cells upon NCD38 treatment, respectively. The normalized enrichment score (NES) and the false discovery rate (FDR) q-value in GSEA are presented at the bottom of each panel. A gene set for the LSD1 signature is listed in Supplementary Table 2.

Supplementary Figure 13



Supplementary Figure 13, continued



Supplementary Figure 13

GSEA analysis of gene sets of AML from datasets previously reported by Valk et al.

In each AML cluster signature, the left and right panels present results from the comparison between DMSO treatment and NCD38 treatment in control HEL cells and between control HEL cells and Δ*GFI1*-SE HEL cells upon NCD38 treatment, respectively. The normalized enrichment score (NES) and the false discovery rate (FDR) q-value in GSEA are presented at the bottom of each panel. Gene sets for each AML cluster signature are listed in Supplementary Table 2.

Supplementary Table 1. Primers and sgRNAs

Primers for RT-PCR

human *GFI1*	Forward	CTCGGAGTTTGAGGACTTCTG
	Reverse	CCGCTCCATGAGTACGGTTTG
human *GAPDH*	Forward	GAAGGTGAAGGTCGGAGTC
	Reverse	GAAGATGGTGATGGGATTTC
murine *Gfi1*	Forward	CTATCCCTGTCAGTACTGTGGC
	Reverse	CTTGAAGCCTGTGTGCTTTCTG
murine *Gapdh*	Forward	AGGTCGGTGTGTGAACGGATTTG
	Reverse	TGTAGACCATGTAGTTGAGGTCA

Primers for ChIP-qPCR

human *GFI1*-SE	Forward	TGTCATTTTCTTCATTTTGGGGG
	Reverse	CCCAGAGCAACTCCTAAGTG
murine CA	Forward	CCAGCCTAACTGTCAGAGGTAAA
	Reverse	TCCGCTCCCCTATTTTCTAAGAG

Primers for bisulfite-specific PCR

First half part of CA	Forward	TAAAGGTGATTTTTGTTTGTTTGAG
	Reverse	ATTCTTACCAACTATTAAATTCAACCTATA
Second half part of CA	Forward	TTTATATAATATAAGTTGAGTTTTTTTT
	Reverse	AATTAATTTCCTTCAACAACCTAAC

sgRNAs for enhancer deletion

human *GFI1*-SE	sgRNA #A	GAAAGTAGGAAATCTGGTCGGGG
	sgRNA #B	GTGAAAAGGCATGAGATGTGTGG
murine CA	sgRNA #P	TCGTTAATGCGCTATGGCACTGG
	sgRNA #Q	CGCTAACAGGGCGCTTAAGGTGG

Genotyping primers for detecting enhancer deletion

human *GFI1*-SE	Forward	CCTGATTCTGTGCCTTCTTCATAC
	Reverse	GAGTGTGGTTTGACTGTGGTATC

murine CA	Forward	GAAAAATCTGGCATGTCTGTCCC
	Reverse	CGGTTTGATTGTTGACACCTGTT

1

2 Primers for generating vectors for luciferase assays

cloning *GFI1*-SE	Forward	TTAGAGGAGGCACTGAAAGCAAG
	Reverse	CATACTTGTAAGCCCAGCTACTTG
deletion of non-CA	Forward	CTTACCTTGCTTTCAGTGCCTCC
	Reverse	ACACCCTCACGTATCTAATAGGAC
deletion of CA	Forward	AGTCCTATTAGATACGTGAGGGTG
	Reverse	GTTGTTGTTGAGATGGAGTCTTGC
GFI1/GF1B motif mutant #1	Forward	ACTCTAATCCTA**C**TCATTGCCTTCTTTC
	Reverse	CTAGAAAACCGTTAAGAAGGAGGAAGTT
GFI1/GF1B motif mutant #2	Forward	AGTGGGAGTAA**C**TCTCCAAACCACTCTC
	Reverse	AAGTTGATAAACTTCCTCCTTCTTAACG
RUNX1 motif mutant	Forward	CTAAGGCCACAGCTGCTAAGGACAAAAG
	Reverse	CTAAGG**GG**ACAGCTGCTAAGGACAAAAG
GATA1 motif mutant	Forward	ACTTTATCTAGGTTTGCAATCAGGAAGC
	Reverse	ACTTTA**C**CTAGGTTTGCAATCAGGAAGC
CEBPA motif mutant #1	Forward	CTTTATCTAGGTTTGC**CC**TCAGGAAGCTGT
	Reverse	TCTAAGGCCACAGCTGCTAAGGACAAAAGC
CEBPA motif mutant #2	Forward	AACACTCGTGGAGAA**GG**TGGAAAGAAGGCA
	Reverse	GCTCAGGAAGCACAAGTAGCTGTAAACCAC
TAL1 motif mutant	Forward	AGGTTGAATCCAA**G**AGTT**T**GTAAGAACCAG
	Reverse	GTAGCAAACTGGACAGAGCACAGCCTGTCC

3 * The mutation in each motif site is underlined.

4

Supplementary Table 2. Gene sets for GSEA analysis

A. REGULATION_OF_GRANULOCYTE_DIFFERENTIATION

ADIPOQ	C1QC	CUL4A	HAX1	HCLS1	IL5	INPP5D
LEF1	MLL5	OGT	PRDM16	RARA	RUNX1	TESC
TRIB1	ZBTB46					

B. KAMIKUBO_MYELOID_CEBPA_NETWORK

ALDH2	ANXA1	CAMP	CASD1	CD177	CEBPA	CSF2RA
CYBB	DGAT2	EMR1	GLRX	HDC	HP	IL18
IRF8	ITGAM	LBP	LCN2	LGALS1	LTF	MMP8
PGD	PRTN3	RAB31	S100A8	S100A9	SERPINB1	SPINT2

C. Differentially expressed genes found during erythroid development

ALAS2	ANK1	ARPC1B	ATP5G1	BDNFOS	BTG1	C11orf17
C19orf48	C19orf6	C1QBP	CA1	CAPG	CCDC114	CCL18
CCL2	CCL5	CCT6A	CD44	CECR1	CSTB	CTSH
CYBA	CYBASC3	CYP27A1	EGR1	ERAF	EYA3	FADS2
FCGRT	FCN1	FKBP5	GDF15	GIPC1	GYPC	H3F3A
HBA1	HBB	HBG1	HBM	HLA□DQA1	HLA□DRA	HMGN1
IGHG1	IL8	ILF3	ITLN1	KCNH2	KHSRP	KIAA1727
LIPA	LOC388588	LOC399761	LOC730200	LXN	LYZ	MGC4677
MMP9	NDUFA3	NOP5/NOP58	NUDT4	PLA2G7	PRG1	PRSS1
PSAP	PSMA2	REXO2	RHAG	RNASE1	RPL22L1	SELENBP1
SLC12A9	SLC25A37	SNHG5	SOD2	STK11	TINP1	TPSAB1
TRIB3	TSPAN17	TYMS	UBE2D3	UQCRQ	VAV2	WDR36

D. WELCH_GATA1_TARGETS

ABCB10	ALAD	ALAS2	ANK1	BACH1	EPB49	FTL

GSTT2	HBZ	HEBP1	HMBS	KLF1	MAFG	MAFK
NFE2	PPOX	SLC4A1	STOM	TFRC	UROD	UROS
ZFPM1						

E. Valk_AML_cluster_7

ANK1	C5ORF4	CLCN3	DNAJC6	EPB41	GAPVD1	GDF15
GYPE	HBBP1	HBZ	KCNH2	KEL	MYL4	OSBP2
PDZK1IP1	RAP1GAP	RHAG	RHD	SELENBP1	SLC25A37	SLC2A1
SLC6A8	SLC6A9	SPTB	TAL1	TNS1	TPM1	TRIM10

F. Valk_AML_cluster_8

ABCG2	ANK1	ARHGEF12	BCAM	C5ORF4	CDC42BPA	CDH1
DCAF11	FECH	GYPA	GYPB	MOSPD1	MXI1	OPTN
OSBP2	PBX1	RHCE	RHD	SELENBP1	SLC6A8	SNCA
TAL1	TMEM158	TNS1	TRAK2	TSPAN5		

G. LSD1 signature

ACVRL1	ARHGEF11	BST2	CALHM2	CCR2	CCSER2	CLEC4A
CNR2	CORO2A	CTSV	EFNA4	ERG	FCGR1B	FCGR2B
FCRLA	FOSL2	GCNT1	GFI1	GFI1B	HLA-DMB	HVCN1
HYLS1	IL18R1	IL18RAP	IL4I1	ISY1-RAB43	KIAA0513	LAIR1
LMO4	LOC388242	LPXN	METTL7A	MILR1	OIT3	OSBPL11
P2RY14	PI16	PIK3C2B	PLD2	PNRC1	PPP1R13B	PROCR
PRR9	RAB43	RASSF4	S1PR1	SELL	SIDT2	SIGLEC17P
SLC35F6	SLC7A8	SNX21	SPTAN1	STAB1	TIGD3	TLR1
TMEM243	TNFRSF10D	TNFSF10	TPM4	TRAF5	ZBTB46	

H. Valk_AML_cluster_1

ATF3	ATP10A	BHLHE41	BLNK	BMI1	COBLL1	DPP4

FHL1	GPR126	IGHM	IRF7	JUN	KCNA5	KIF17
LTBP3	MECOM	MEF2C	MMRN1	NR4A2	PCDH9	PROM1
PRR16	SLC2A3	SLC38A1	SOCS2	SPAG6	SPIB	TRPS1

I. Valk_AML_cluster_2

ARHGAP22	CCL1	CLU	DOCK1	EZR	GLI2	GOLGA8A
GOLGA8B	GPR56	GRB10	GUCY1A3	HBB	HOXA5	HOXB2
IL2RA	JAG1	KCNK5	LAPTM4B	LPIN1	MAP4K4	PDE3B
PIM1	PLA2G4A	PLS1	PTP4A3	SCHIP1	SRSF8	TRIM16
TRPC2						

J. Valk_AML_cluster_3

ADCY2	AGTPBP1	AIM1	ARHGAP22	BLVRA	CCL1	CEBPD
COL4A5	DOCK1	ENPP4	FAM30A	FCGRT	GAS2	GPR56
IL17RA	IL6ST	LAPTM4B	MAP7	NET1	PBXIP1	PDGFD
PIEZO2	PLS1	QPRT	RAP2A	SCN9A	SEPP1	SERPINB8
SH2D1A	SMC4	TNFRSF4	TRIM16	WBP5		

K. Valk_AML_cluster_4

ABCB1	B4GALT6	C18ORF1	CAPN2	CD7	CFD	CSDA
CTNNA1	CYFIP1	DRAM1	FZD6	HPGDS	HPS4	IGFBP7
IKZF2	IL4R	ITGA4	KLF2	LRP10	NDFIP1	P2RX5
PMS2L2	PMS2P3	PRR5L	RAB13	TRAT1	TRDV2	TRIB1
TSPAN7	TUBB6	UGT2B28				

L. Valk_AML_cluster_5

CAMK1	CCR1	CD86	EFHD2	EPB41L3	FAM198B	GNS
HNMT	IFNGR2	KCNQ1	LILRA1	LILRA6	LILRB1	LILRB2
LILRB3	MAFB	NOD2	PILRA	PSAP	PTAFR	RASSF4

SIGLEC7	SIRPA	SIRPB1	SLC15A3	SMPDL3A	STS	TFEB
TLR8	TYMP	UBE2D1	VCAN	VDR		

M. Valk_AML_cluster_6

ADCY2	BST2	CAT	CD74	CEP70	CORO1A	DPPA4
DPYSL3	DSC2	FAM110B	FAM174B	FOXC1	FOXF2	FTO
GPC4	HIGD1A	HLA-DPA1	HLA-DRA	HOXB3	KIAA0930	LTBP1
NT5DC3	PECAM1	PIEZO2	PLXNB1	RGS10	RSL1D1	SLC27A6
SMC4	SNCAIP	TTC27	TUBGCP4	XPA		

N. Valk_AML_cluster_9

AK5	BAHCC1	CBFB	CD1C	CD59	CD81	CHI3L1
CHST12	CLEC10A	CLIP2	CLIP3	COLEC12	DHRS3	EMID1
FAM105A	FAM171A1	FCGR2B	ICAM4	MGLL	MN1	MSLN
MTMR11	MYH11	NDE1	NRP1	NT5E	PAPSS2	PTPRM
RPS6KA2	RUNX3	SPARC	ST18	TGFBI	TPPP3	VSIG4

O. Valk_AML_cluster_10

ARHGEF17	AZU1	BAALC	BCL7A	C3AR1	CD22	CFD
CHRDL1	CRIM1	EEF1A2	F2RL1	FAM30A	FLNB	GNAI1
IGHM	LPHN1	MLLT3	MN1	NPDC1	PAWR	PIK3C2B
PPP1R16B	PRKD2	RBPMS	RNASE2	SETBP1	SMAGP	SNED1
SORBS3	SPON1	SPRY1	SPTBN1	TPM2		

P. Valk_AML_cluster_11

ALDH2	APP	ASS1	CD200	CD34	CIITA	CISH
CLIC4	DAB2	DAPK1	DNM1	DPYSL2	DUSP7	EGFL7
EVL	FYN	GAS2L1	GIMAP5	GIMAP6	H1F0	IFITM1
ITGA6	KCNN4	KLF9	KYNU	LHFPL2	LPAR6	MDFIC

MYO5C	PALM	REEP5	SERPINB9	SLC38A1	SPARC	ST3GAL5
VEGFA						

Q. Valk_AML_cluster_12

AFF2	ALADL1	ARHGAP4	AZI2	CALR	CST7	FGF13
GABRE	GALNT3	HGF	LAMC1	LGALS9	MEG3	MFNG
MST1	MXRA7	NKX3-2	NRIP1	P4HB	PCBP3	PRODH
PTCH1	PTGDS	PTGER1	RAB5B	SIX3	SKAP2	STXBP1
TMEM87A	VCL					

R. Valk_AML_cluster_13

ADCY7	ADRA2C	BAIAP3	C11ORF21	C11ORF9	C15ORF39	CACNA2D2
CAV1	FBLN5	GRK5	HSPG2	IL5RA	ITGB4	KDM4B
LAT2	LCP1	NBL1	NCALD	PNMT	POU4F1	PSD3
RFL	ROBO1	RUNX1T1	SLC25A1	STK32B	THSD7A	TRH
VLDLR	VOPP1					

S. Valk_AML_cluster_15

ARHGEF3	ATN1	BASP1	CD52	CEACAM8	CTNNA1	DRAM1
ECHDC2	EPB41L2	FADS1	GNA12	GRAMD1B	HOXA10	HOXA9
HOXB2	HOXB5	HSPB1	IGF2BP2	IGHM	KLF9	MEIS1
NDFIP1	RUNX1	SEL1L3	SFXN3	SH3TC1	SLC16A1	SUCLG2
TBL1X	TNS3	TUBB6				

T. Valk_AML_cluster_16

ADCY9	AK2	AKR7A2	APOC2	BRE	C1ORF54	C20ORF103
C3ORF14	CACNA2D3	CADM1	CD70	CLSTN2	DACH1	GGA2
ITGA7	KCNE1L	KCNN2	MBNL1	PENK	RET	RPP25
TDRD7	TGM5	TKTL1	TRPM4	UVRAG		

U. Valk_AML_with_11q23_rearranged

APOC2	C1ORF54	C20ORF103	C3ORF14	CACNA2D3	CADM1	CD70
CES1	DACH1	HIF1A	ITGA7	KCNE1L	LOC283683	MSLN
NXT2	P2RY2	PENK	RET	TGM5	TKTL1	TRPM4
UVRAG						

V. Valk_AML_with_CEBPA

ARHGEF3	ATN1	B4GALT6	BASP1	CAMP	CD38	CD7
CEACAM8	CEBPA	CTNNA1	DLC1	GALC	GNA12	HOXA9
HOXB2	HPGDS	HSPB1	IGF2BP2	IGHM	IGLL1	ITM2A
LCN2	LTF	MEST	MMP8	NDFIP1	PGLYRP1	PRR5L
SEL1L3	SFXN3	SLC16A1	SUCLG2	TBL1X	TNS3	TRDV2
TRIB1	TUBB6					

W. Valk_AML_with_FLT3_ITD

ADCY2	APP	BAHCC1	COL4A5	CYSLTR2	ENPP2	GOLGA8A
GOLGA8B	GPR56	HOXA4	HOXA5	HOXA9	HOXB2	HOXB3
HOXB5	HOXB6	IL1RAP	IL2RA	KCNK5	LAPTM4B	LCT
LGALS3BP	LYRM1	MAGED1	MAP1A	MMP2	MRC1	NR6A1
PBX3	PDE4B	PDGFD	PIEZO2	PIM1	QPRT	SEPP1
SMC4	SOCS2	TARP	TRIM16	TRPC2		

X. Valk_AML_with_EVI1

AZU1	CD34	CFD	CRIM1	DMXL2	EEF1A2	F2RL1
FAM30A	GNAI1	IGHM	LPHN1	MN1	NPDC1	PAWR
PIK3C2B	PPP1R16B	PXDN	RBPMS	RNASE2	RNASE3	SEPX1
SMAGP	SORBS3	SPRY1	SPTBN1			

Supplementary Table 3. List of genes influenced by *GFI1*-SE depletion

Genes that were induced more than two-fold by NCD 38 but whose inductions were attenuated by less than 50% via *GFI1*-SE depletion

Gene Symbol	Fold Change [DMSO vs NCD38 in control HEL cells]	Fold Change [Control vs ΔGFI1-SE in NCD38 treatment]	Fold Change [DMSO vs NCD38 in CMK11-5 cells (Dataset: GSE68348)][4]
LOC101928012	12.35	5.50	13.43
PRR9	10.62	4.44	7.00
GFI1	8.05	2.85	9.07
S1PR1	6.27	2.59	7.39
SLC45A3	6.24	3.08	3.17
TEK	5.33	2.28	3.24
TRGJP2	4.37	1.32	2.41
STYK1	3.56	1.56	2.73
ERG	3.54	1.76	4.58
CD200R1	2.69	1.22	7.98
RP11-702L15.4	2.36	1.18	NA
IGFBP4	2.08	0.96	2.47
RNU6-795P	2.07	0.83	NA
MIR296	2.06	0.71	NA
RP11-758M4.4	2.05	0.92	NA
RP11-661A12.4	2.02	0.91	NA
RPL23AP32	2.00	0.93	1.99

Genes that were attenuated by less than 50% by NCD 38 but whose reductions were recovered more than two-fold via *GFI1*-SE depletion

Gene Symbol	Fold Change [DMSO vs NCD38 in control HEL cells]	Fold Change [Control vs ΔGFI1-SE in NCD38 treatment]	Fold Change [DMSO vs NCD38 in CMK11-5 cells (Dataset: GSE68348)][4]
THBS1	0.41	1.07	0.33
CCR4	0.41	0.94	0.91

RHD	0.42	0.84	0.15
FAM19A3	0.46	0.93	0.45
RNA5SP302	0.49	1.03	NA

NA; not available in this data set (GSE68348) due to differences between the platforms used for separate analyses (Affymetrix vs. Agilent).

Supplementary References

1. Tohyama K, Tohyama Y, Nakayama T, Ueda T, Nakamura T, Yoshida Y. A novel factor-dependent human myelodysplastic cell line, MDS92, contains haemopoietic cells of several lineages. *British journal of haematology* 1995 Dec; **91**(4)**:** 795-799.

2. Shimazu Y, Hishizawa M, Hamaguchi M, Nagai Y, Sugino N, Fujii S*, et al.* Hypomethylation of the Treg-Specific Demethylated Region in FOXP3 Is a Hallmark of the Regulatory T-cell Subtype in Adult T-cell Leukemia. *Cancer Immunol Res* 2016 Feb; **4**(2)**:** 136-145.

3. Yamamoto R, Kawahara M, Ito S, Satoh J, Tatsumi G, Hishizawa M*, et al.* Selective dissociation between LSD1 and GFI1B by a LSD1 inhibitor NCD38 induces the activation of ERG super-enhancer in erythroleukemia cells. *Oncotarget* 2018 Apr 20; **9**(30)**:** 21007-21021.

4. Sugino N, Kawahara M, Tatsumi G, Kanai A, Matsui H, Yamamoto R*, et al.* A novel LSD1 inhibitor NCD38 ameliorates MDS-related leukemia with complex karyotype by attenuating leukemia programs via activating super-enhancers. *Leukemia* 2017 Nov; **31**(11)**:** 2303-2314.