

## HMG20B stabilizes association of LSD1 with GFI1 on chromatin to confer transcription repression and leukaemia cell differentiation block

Alba Maiques-Diaz<sup>1</sup>, Luciano Nicosia<sup>1</sup>, Naseer J Basma<sup>1</sup>, Isabel Romero-Camarero<sup>1</sup>, Francesco Camera<sup>1</sup>, Gary J Spencer<sup>1</sup>, Fabio MR Amaral<sup>1</sup>, Fabrizio Simeoni<sup>1</sup>, Bettina Wingelhofer<sup>1</sup>, Andrew J.K. Williamson<sup>2</sup>, Andrew Pierce<sup>2,3</sup>, Anthony D Whetton<sup>2,4</sup>, Tim CP Somervaille<sup>1</sup>.

<sup>1</sup>Leukaemia Biology Laboratory, Cancer Research UK Manchester Institute, The University of Manchester, Manchester Cancer Research Centre Building, 555 Wilmslow Road, Manchester, M20 4GJ, United Kingdom

<sup>2</sup>Stem Cell and Leukaemia Proteomics Laboratory, Manchester Academic Health Science Centre, The University of Manchester, Wolfson Molecular Imaging Centre, 27 Palatine Road, Manchester, M20 3LJ, United Kingdom

<sup>3</sup>School of Medical and Health Sciences, College of Human Sciences, Fron Heulog Bangor University, Bangor, LL57 2TH, United Kingdom

<sup>4</sup>School of Veterinary Medicine and School of Biosciences and Medicine, University of Surrey, VSM Building, University of Surrey, Guildford, GU2 7AL, United Kingdom

### Abstract

Pharmacologic inhibition of LSD1 induces molecular and morphologic differentiation of blast cells in acute myeloid leukemia (AML) patients harboring *MLL* gene translocations. In addition to its demethylase activity, LSD1 has a critical scaffolding function at genomic sites occupied by the SNAG domain transcription repressor GFI1. Importantly, inhibitors block both enzymatic and scaffolding activities, in the latter case by disrupting the protein:protein interaction of GFI1 with LSD1. To explore the wider consequences of LSD1 inhibition on the LSD1 protein complex we applied mass spectrometry technologies. We discovered that the interaction of the HMG-box protein HMG20B with LSD1 was also disrupted by LSD1 inhibition. Downstream investigations revealed that HMG20B is co-located on chromatin with GFI1 and LSD1 genome-wide; the strongest HMG20B binding co-locates with the strongest GFI1 and LSD1 binding. Functional assays demonstrated that HMG20B depletion induces leukemia cell differentiation

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: <https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms>

Lead contact & corresponding author: Tim C P Somervaille - [tim.somervaille@cruk.manchester.ac.uk](mailto:tim.somervaille@cruk.manchester.ac.uk).

#### Author contributions

AM-D, LN, NB, FC, IR-C, BW, GS and FS performed experiments; AM-D, FA, FC and TS performed and interpreted bioinformatics analyses; AP, AJKW and ADW designed, performed and analyzed the SILAC proteomic assays; AM-D and TS designed the study and wrote the manuscript; all authors reviewed the final version.

#### Competing interests

All contributing authors declare that there is no conflict of interest with regard to the data presented in this study.

and further revealed that HMG20B is required for the transcription repressor activity of GFI1 through stabilizing LSD1 on chromatin at GFI1 binding sites. Interaction of HMG20B with LSD1 is through its coiled-coil domain. Thus, HMG20B is a critical component of the GFI1:LSD1 transcription repressor complex which contributes to leukemia cell differentiation block.

## Introduction

Epigenetic and transcription factor complexes are emerging therapeutic targets across a range of human malignancies. The histone demethylase LSD1 (Lysine Specific Demethylase 1, also known as KDM1A) has been the focus of recent interest in view of its high level expression in acute myeloid leukemia (AML) and a range of solid tumors, as well as encouraging pre-clinical and early phase clinical studies using candidate inhibitors [1–4]. LSD1 forms a corepressor complex with RCOR1 (CoREST), histone deacetylase (HDAC1/2) and other components [5] and has enzymatic capacity to demethylate monomethyl and dimethyl lysine 4 of histone H3 (H3K4) in a flavin adenine dinucleotide (FAD) dependent manner [6]. In addition, LSD1 has a critical scaffolding function. It binds the N-terminal sequence of SNAG domain transcription factors such as GFI1 and SNAIL through a peptide binding cleft formed by the amine oxidase domain [7]; this interaction is essential for the transcription repressor function of these DNA binding proteins. SNAG domain sequences are structural mimics of the N-terminal tail of Histone H3 [7] and when LSD1's binding cleft is occupied there is no access for the histone tail to the catalytic activity of LSD1. In leukemia cells, recruitment of LSD1 to chromatin genome-wide is overwhelmingly at sites of GFI1 binding [8] and release of LSD1 from GFI1-occupied sites by disruption of protein:protein interaction is required for LSD1 inhibitor-induced leukemia cell differentiation [8]. Pharmacologic inhibition of LSD1 has also proven to be effective in pre-clinical models of SNAG transcription factor-driven neoplasia [8–14].

GFI1 is important in normal and leukemic hematopoiesis where it is required for the functional integrity of hematopoietic stem cells, lymphoid development and generation of neutrophils [15–17]. Mutations in *GFI1* cause severe congenital neutropenia [18] and *GFI1* expression levels have been variably linked to prognosis in AML according to disease subtype [19]. The interaction between GFI1 and LSD1 can be disrupted by a single SNAG domain point mutation (P2A) which inactivates GFI1's function as a transcriptional repressor [20]. Despite important prior insights, the role of many of the proteins found in the LSD1 complex, and how they interact at sites of GFI1 binding to confer transcription repression, remains to be elucidated [21].

In this study our goal was to identify LSD1 protein binding partners that facilitate the stability of LSD1's interaction with GFI1 on chromatin. For this purpose, we used *MLL*-rearranged leukemia cell models which are dependent on the physical interaction of LSD1 with GFI1 to maintain their proliferative, undifferentiated cellular state [4,8].

## Materials and Methods

### Reagents and antibodies

Puromycin (#P8833), blasticidin (#15205) and doxycycline (#24390-14-5) were from Sigma (St Louis, MO). Trans-N-[(2-methoxy-pyridin-3-yl)methyl]-2-phenylcyclopropan-1-amine (OG86) was synthesized in house [2]. Details of other reagents, antibodies and biochemical methods are in the Supplementary Information.

### Human cells and cell cultures

Use of human tissue was in compliance with the ethical and legal framework of the UK's Human Tissue Act, 2004. Primary human AML samples were from Manchester Cancer Research Centre's Tissue Biobank (instituted with the approval of the South Manchester Research Ethics Committee; 18/NW/0092). Their use was authorized following ethical review by the Tissue Biobank's scientific sub-committee and with the informed consent of the donor. THP1 cells were purchased from DMSZ (Braunschweig, Germany), verified by STR analysis and confirmed to be mycoplasma-free. Details of cell culture methods are in the Supplementary Information.

### Expression constructs and lentiviral vectors

Details of vectors and cloning strategies are given in the Supplementary Information. Lentiviral supernatants were prepared, and human and murine cells were infected, as described previously [8].

### Mass spectrometry, ChIPseq, RNA-seq and data analysis

Mass spectrometry protocols, ChIPseq and RNA-seq protocols, and data analysis methods were as previously reported [8] and are as described in detail in the Supplementary Information.

## Results

### Pharmacologic inhibition of LSD1 disrupts its interaction with the DNA-binding cofactor HMG20B

To identify proteins displaced from physical interaction with LSD1 following treatment of AML cells with OG86 (Oryzon Genomics compound #86; trans-N-((2-methoxy-pyridin-3-yl)methyl)-2-phenylcyclopropan-1-amine) which is a representative tranylcypropane-derivative inhibitor of LSD1, we used both Stable Isotope Labeling by Amino acids in cell Culture (SILAC) and Immunoprecipitation-Mass Spectrometry (IP-MS) as relative quantitation proteomics approaches (Figure 1A) [2,8]. THP1 AML cells were selected because they exhibit a t(9;11) *MLL* gene rearrangement and respond to LSD1 inhibition in a similar manner to primary patient *MLL*-translocated AML cells, with differentiation and loss of clonogenic activity [4,8]. For SILAC, cells were grown in heavy or light media for seven passages, treated with OG86 250nM or DMSO vehicle respectively for 48 hours and then protein lysates from both conditions were combined and subjected to endogenous LSD1 immunoprecipitation (Figure 1A; Table S1). In parallel, we performed LSD1 MS-IP using nuclear-enriched protein lysates from THP1 AML cells exposed to either

OG86 250nM or DMSO, also for 48 hours (Figure 1A; Table S2). Both IP experiments identified proteins previously reported in multiple studies to be found in complex with LSD1 [1]. In keeping with our prior findings [8] we found no evidence that OG86 reduced the interaction of LSD1 with complex components RCOR1/2/3 and HDAC1/2. Likewise, the same was the case for PHF21A and zinc finger proteins ZMYM2/3, ZNF217, ZNF516 and ZEB2. However, in both experiments we observed reduced interaction with LSD1 of the high-mobility group protein HMG20B (also known as BRAF35) which is known to interact with LSD1 and with DNA in a non-sequence specific manner [22,23], as well as the coiled-coil domain protein GSE1. The ratio of RREB1 binding was reduced in the SILAC experiment but to a lesser extent in the MS-IP experiment. We focused our attention on HMG20B, given its role in repressing erythroid differentiation [24] and the reported roles of its paralog HMG20A in SNAG domain transcription factor repressive complexes associated with SNAI1 [25], INSM1 [26], and GFI1B [27] in different cellular contexts.

To confirm the reduced interaction of HMG20B with LSD1 following LSD1 inhibition, we performed HMG20B MS-IP using nuclear-enriched protein lysates from THP1 AML cells exposed to either OG86 250nM or DMSO for 48 hours (Figure 1B, Table S3) and reciprocal IP-western analysis for both HMG20B and LSD1 (Figure 1C). HMG20B IP-MS confirmed an interaction of HMG20B with LSD1, RCOR1/3 and GSE1 (Figure 1B, Table S3) although not zinc finger proteins ZMYM2/3, ZNF217, ZNF516 and ZEB2. Both MS and IP-western approaches confirmed that LSD1 inhibition reduced the physical association of HMG20B with LSD1 and RCOR1. We confirmed that the interaction of HMG20B with LSD1 was direct. In vitro transcribed and translated HMG20B interacted with, and was efficiently depleted from solution by, purified GST-LSD1 (Figure 1D). Thus, we concluded that an exemplar tranylcypromine-derivative inhibitor of LSD1 destabilizes the physical association of LSD1/RCOR1 with its associated complex component HMG20B.

### Depletion of HMG20B induces leukemia cell differentiation

To explore the functional role of HMG20B, we induced *HMG20B* knockdown (KD) in THP1 AML cells and in five primary patient AML samples with *MLL* gene rearrangements. In THP1 cells *HMG20B* KD resulted in up regulation of the myeloid differentiation marker CD11b and reduced clonogenic activity in semi-solid culture (Figures 1E-G). We made similar observations in primary patient *MLL* leukemia cells with up regulation of differentiation markers CD11b and CD86, reduced clonogenic activity, increased maturation of cells to mature macrophages (Figures 1H-L; Table S4) and a G1 cell cycle arrest without significant enhancement of apoptosis (Figures S1A-B). These phenotypic observations were very similar to those we reported following pharmacological inhibition of LSD1 in cell line and patient cells [8], in keeping with the concept that HMG20B has an important role in the LSD1 complex.

### Close association of HMG20B on chromatin with LSD1 and GFI1

We previously reported that LSD1 near exclusively co-locates on chromatin with the transcription factor GFI1 in THP1 AML cells, with the strength of LSD1 chromatin immunoprecipitation (ChIP) signal correlating strongly and positively with GFI1 ChIP signal [8]. To evaluate the co-localization of endogenous HMG20B in relation to

endogenous LSD1 and GFI1 genome-wide, we performed CHIP followed by next generation sequencing (ChIPseq) for the three proteins. After excluding blacklisted genomic regions prone to artefact and making use of stringent threshold criteria (called peaks had pileup value  $\geq 50$  and fold enrichment over input  $\geq 5$ ), MACS2 [28] identified 18,385 HMG20B peaks, 22,600 LSD1 peaks and 12,221 GFI1 peaks. As we previously reported for LSD1 and GFI1 [8], whether all peaks were considered, or the strongest 20.7% (i.e. those with pileup values  $\geq 110$ ;  $n=3,797$ ), HMG20B binding was distributed predominantly over intronic and intergenic regions versus promoter regions (Figure 2A), consistent with putative roles at enhancers. Genome-wide, HMG20B peaks were co-located with GFI1 and LSD1 binding peaks, and stronger HMG20B binding correlated positively and significantly with stronger GFI1 and LSD1 binding (Figures 2B-E). Indeed 91.0% and 94.8% respectively of strong HMG20B peaks overlapped with either a GFI1 or an LSD1 binding peak (Figure 2F) (i.e. absolute summit of peaks separated by  $\leq 100$  base pairs). In keeping with the close physical association of HMG20B with both GFI1 and LSD1 on chromatin, MEME-ChIP [29] confirmed that genomic sequences surrounding the absolute summits of the strongest 20.7% of HMG20B binding peaks were strongly enriched for the GFI1 consensus binding motif (Figure 2G).

Interestingly, there were notable differences in the distribution of GFI1 motifs according to whether the absolute summits of strong GFI1 peaks, strong LSD1 peaks or strong HMG20B peaks were considered (Figure 2G). Relative to the absolute summit of strong GFI1 peaks (i.e., the strongest 20.3%, with pileup values  $\geq 135$ ;  $n=2,479$ ), the enriched region of GFI1 transcription factor binding motifs was narrow (63 base pairs), with the apex of the motif probability graph coinciding with the absolute summit of GFI1 ChIPseq peaks (Figure 2G, bottom panel). By contrast, relative to the absolute summit of strong HMG20B or strong LSD1 ChIPseq peaks (i.e. the strongest 20.5%, with pileup values  $\geq 118$ ;  $n=4,641$ ), the enriched region of GFI1 transcription factor binding motifs was broader (95 or 123 base pairs respectively) with, in addition, the apex of the GFI1 motif probability curve offset from the absolute summit of HMG20B or LSD1 ChIPseq peaks by a modal value of 8-10 base pairs on each side (Figure 2G, upper and middle panels). These motif distribution plots are consistent with the GFI1/LSD1/HMG20B repressor complex being recruited to chromatin through sequence specific binding of GFI1, with HMG20B serving an accessory, stabilizing function through binding of its HMG-box to nearby bent, kinked or unwound non-B-type DNA structures [30].

Next we determined base pair distances between the summits of strong GFI1 peaks and co-located strong HMG20B and strong LSD1 peaks. To maximize the signal:noise ratio we focused on the strongest 526 GFI1 peaks (i.e. those with pileup values  $\geq 300$ ) where there were also overlapping strong HMG20B and LSD1 peaks, all with absolute summits within 100 base pairs of one other. The base pair distances from HMG20B peak summits to GFI1 peak summits were strongly correlated with those from LSD1 peak summits to GFI1 peak summits (Figure 3A). In contrast there was little correlation between HMG20B $\rightarrow$ LSD1 versus HMG20B $\rightarrow$ GFI1 base pair distances (Figure S2A), nor HMG20B $\rightarrow$ LSD1 versus LSD1 $\rightarrow$ GFI1 base pair distances (Figure S2B). The mean base pair distance between summits of HMG20B and LSD1 peaks was significantly lower than that between either HMG20B and GFI1 peaks, or LSD1 and GFI1 peaks (Figure 3B). There was no significant

difference in base pair distances between summits of HMG20B and GFI1 peaks versus distances between LSD1 and GFI1 peaks (Figures 3B-D). These data suggest that HMG20B and LSD1 co-occupy chromatin to one side or the other of GFI1.

### **HMG20B interacts with LSD1 through its coiled-coil domain**

To further evaluate the interaction between HMG20B, LSD1 and GFI1, we performed nuclear immunoprecipitation experiments in the presence of benzonase to digest nucleic acid. While immunoprecipitation of both HMG20B and GFI1 pulled down LSD1, IP of HMG20B did not pull down GFI1 and vice versa (Figure 3E), suggesting that stable association of HMG20B with GFI1 on chromatin requires intervening DNA. To determine which domain of HMG20B recruits LSD1 to chromatin we generated and expressed HMG20B deletion mutants (Figure 3F). While full length HMG20B pulled down LSD1, deletion mutants lacking the coiled-coil domain ( C or CC) did not (Figures 3G-H). In contrast, mutants lacking the DNA binding HMG-box readily interacted with LSD1 (Figures 3G-H), despite the low-level expression of the full N-terminal deletion mutant following doxycycline induction, likely due to reduced protein stability (Figure 3G). HMG20B and GFI1 therefore bind DNA close to one another on chromatin through, respectively, the HMG box of HMG20B and the zinc finger domain of GFI1. LSD1 interacts through two points of contact: the coiled-coil domain of HMG20B and the N-terminal SNAG domain of GFI1 [7] (Figure 3I).

### **Functional recruitment of LSD1 to chromatin by a GFI1 DNA binding domain:HMG20B fusion protein**

Assembly of the GFI1:LSD1 transcription repressor complex on chromatin is, like other complexes, likely to be dynamic with variable chromatin scanning and residence times of individual proteins determining stability of the whole. We hypothesized that by directly linking HMG20B to the DNA-binding domain of GFI1 (GFI1-ZNF) the resulting GFI1-ZNF-HMG20B fusion protein would serve as a stable platform for recruitment of LSD1 by circumventing a requirement for co-localization of separate full length HMG20B and GFI1 proteins at the same site. Making use of conditional constructs (Figure 4A), immunoprecipitation experiments confirmed that as expected GFI1.ZNF-HMG20B interacted with both LSD1 and RCOR1 (Figure 4B).

We previously demonstrated that LSD1 inhibitor-induced up regulation of a myeloid differentiation program in THP1 AML cells is dependent upon drug induced physical separation of LSD1 from GFI1 rather than inhibition of the enzymatic activity of LSD1, and that expression of a GFI1.ZNF-LSD1 fusion protein which is resistant to drug-induced physical separation is sufficient to confer cellular resistance to LSD1 inhibitors [8]. Our experiments here show that AML cells exposed to OG86 250nM or DMSO vehicle for 48 hours undergo the expected up regulation of myeloid differentiation markers CD11b and CD86, and that this is unaffected by induced expression of either the DNA binding domain of GFI1 (GFI1-ZNF) or HMG20B alone. In contrast, induced expression of the GFI1.ZNF-HMG20B fusion protein completely blocked CD11b and CD86 upregulation (Figure 4C-E). To confirm this was due to stabilization of LSD1 on chromatin by the induced fusion protein, we performed ChIPseq. While ChIP signal for Flag-HMG20B and LSD1 decreased

following exposure of cells to OG86 250nM, in the presence of the GFI1-ZNF-HMG20B fusion LSD1 ChIP signal was largely retained (Figures 4F,G). The interaction of LSD1 with the GFI1.ZNF-HMG20B fusion protein, and thus its recruitment to chromatin, was through the HMG20B coiled-coil domain (CC) because the CC GFI1.ZNF-HMG20B fusion protein mutant failed to interact and was unable to block the upregulation of CD11b or CD86 (Figure 4C,D).

Thus, even in the absence of the SNAG domain sequence of GFI1, a fusion of the DNA binding domain of GFI1 with HMG20B was sufficient to stably recruit LSD1 to chromatin in the presence of an LSD1 inhibitor to maintain transcription repression.

### **HMG20B stabilizes LSD1:GFI1 interaction on chromatin preventing enhancer activation**

Ahead of experiments to determine the functional importance of HMG20B to the LSD1:GFI1 interaction, we performed characterization of the chromatin and DNA accessibility surrounding HMG20B binding peaks. Consistent with the transcription repressor function of the GFI1/LSD1/HMG20B complex, among the strongest 30% of HMG20B ChIPseq peaks (corresponding to a threshold pileup value of 90), there was an inverse quantitative relationship between HMG20B peak strength and surrounding H3K27Ac ChIP and ATACseq signal (Figures 5A-C). There was no relationship between HMG20B peak strength and surrounding H3K4Me1 ChIP signal (Figure S3A). This suggests that chromatin binding of GFI1/LSD1/HMG20B below an HMG20B pileup value of ~90 had limited function, at least as far as direct or indirect repression of enhancer acetylation and chromatin accessibility is concerned.

To evaluate the consequences of HMG20B depletion upon the interaction between GFI1 and LSD1 on chromatin, as well as on surrounding histone acetylation, ChIPseq was performed for HMG20B, GFI1, LSD1 and H3K27Ac following in control and *HMG20B* KD cells. Fold change in normalized H3K27Ac ChIP signal surrounding each HMG20B peak was determined and then the mean fold change in H3K27Ac ChIP signal for each of the ten cohorts of HMG20B peaks ranked in order of peak strength (n=1838 peaks per cohort). For the strongest three cohorts (corresponding to the strongest 30% of HMG20B ChIPseq peaks) there was a positive relationship between HMG20B peak strength and fold-change increase in the surrounding ( $\pm 1$ kb from absolute peak summit) H3K27Ac ChIP signal following *HMG20B* KD (Figure 5D). As a control for this analysis we calculated the fold change in H3K27Ac ChIP signal surrounding each of 24,479 binding peaks identified in THP1 AML cells for the transcription factor CEBPA following *HMG20B* KD (Figure S3B, C). While there was a positive relationship between CEBPA peak strength and the level of surrounding histone acetylation, there was no relationship between CEBPA peak strength and fold change in H3K27Ac ChIP signal following *HMG20B* KD (Figure S3B-C) demonstrating the specificity of the observed changes surrounding HMG20B binding peaks.

Next, to evaluate the consequences for GFI1 and LSD1 chromatin binding at HMG20B binding sites following HMG20B depletion, we performed a similar analysis and calculated the ratio of GFI1:LSD1 ChIP signal surrounding each HMG20B binding site ( $\pm 300$ bps from absolute peak summit) before and after *HMG20B* KD. Once more, above a threshold level of HMG20B peak strength (in this case the strongest four cohorts of HMG20B peaks,

corresponding to the strongest 40% of HMG20B peaks (pileup value > 78)) we observed a significant reduction in the mean ratio of LSD1 ChIP signal to GFI1 ChIP signal (Figures 5E, F). This is further illustrated by the density plots shown in Figure 5G. *HMG20B* KD led to localized reduction of LSD1 ChIP signal (normalized to assume constant GFI1 ChIP signal following KD), with the mean absolute reduction correlating strongly across HMG20B peak cohorts with the mean absolute reduction in HMG20B ChIP signal (Figure 5H). Taken together, these data indicate that on a genome-wide basis HMG20B serves to stabilize the physical interaction on chromatin of LSD1 with GFI1.

### **HMG20B is required for GFI1 repressor activity in normal and leukemic myeloid cells**

To evaluate the consequences of *HMG20B* transcript depletion on the transcriptome, we performed *HMG20B* KD. Considering the 8,733 expressed protein coding genes (i.e., those with expression levels of 0.25 fragments per kilobase per million mapped reads (FPKM) in at least one of the six samples), 942 were differentially expressed 72 hours after initiation of KD (Figure 6A): 342 genes were upregulated and 600 downregulated (Figure 6A & Table S5). We confirmed by qPCR that both *HMG20B* KD constructs induced increased expression of selected genes (*SPPI1*, *BTG2*, *CD84* and *CD209*) identified as up regulated by RNAseq (Figure S4). We made use of gene set enrichment analysis (GSEA) to confirm that *HMG20B* KD induced expression of a leukemia cell differentiation program. Using the Molecular Signatures Database Hallmark Gene Set collection (each of which conveys a specific biological state or process and displays coherent expression) [31] we noted that immune-associated gene sets which are highly expressed in terminally differentiated monocytes and macrophages (e.g. INFLAMMATORY\_RESPONSE, INTERFERON\_GAMMA\_RESPONSE) were strongly enriched in *HMG20B* KD versus control cells. Likewise, those characteristic of cycling, metabolically active cells (e.g. E2F\_TARGETS, MYC\_TARGETS\_V2) were depleted (Figure 6B). In keeping with induction of a leukemia cell differentiation program we also noted that *HMG20B* KD resulted in strong depletion of a gene set which promotes leukemia stem cell maintenance in murine MLL-AF9 AML cells [32, 33]. There was also strong up regulation of genes whose expression is inversely correlated with leukemia stem cell activity (Figure 6C). Most notably, comparison of the transcriptional changes induced by *HMG20B* KD with those observed following LSD1 inhibition in THP1 AML cells with OG86 [8] revealed a highly significant overlap (Figure 6D). These transcriptional analyses provide further support to the concept that HMG20B is an essential cofactor involved in the repressive activity of GFI1, through stabilization of the interaction of GFI1 with LSD1 on chromatin.

To explore the link between HMG20B peak strength and change in gene expression following *HMG20B* KD, we mapped each of the 18,385 HMG20B binding peaks to the nearest gene using Genomic Regions Enrichment of Annotations Tool (GREAT) [34] and grouped them into 10 separate cohorts according to HMG20B peak strength. We considered peaks located at promoters or the 5'UTR of genes ("promoter peaks"), and those bound at intergenic or intronic regions ("enhancer peaks"), separately. Within the limitations of this approach – enhancers do not necessarily control expression of the nearest gene - we nevertheless observed that, among upregulated genes, the strongest enrichment was observed



for genes mapping to the strongest decile of promoter-bound peaks and enhancer-bound peaks (Figure 6E, F). In keeping with the observed changes in histone acetylation and the change in the GFI1:LSD1 ChIP signal ratio (Figures 5D, E), there was also among upregulated genes significant enrichment of those mapping to HMG20B promoter peaks in the second and third decile (corresponding to a threshold pileup value of 90), and down to the fifth decile for the enhancer cohorts (down to a threshold pileup value of 70).

To provide additional evidence that HMG20B regulates myeloid lineage differentiation through stabilizing the interaction of LSD1 with GFI1 on chromatin, we performed KD experiments in murine KIT<sup>+</sup> hematopoietic stem and progenitor cells (HSPCs) and cultured them in semisolid conditions favoring terminal differentiation. GFI1 is required for terminal granulocytic differentiation and so we hypothesized that *HMG20B* depletion would reduce the proportion of granulocytes in culture after seven days. As expected, control cells generated both granulocytes and macrophages while *Gfi1* KD substantially decreased granulocytic differentiation both by morphology and flow cytometry. Likewise, *Hmg20b* depletion also resulted in production of significantly fewer mature granulocytes (Figure 7A-E). The ratio of immunophenotypic mature granulocytes to mature macrophages was also significantly lower following either *Gfi1* or *Hmg20b* depletion compared with control conditions (Figure 7F). These observations are consistent with the concept that HMG20B serves a critical accessory role in the function of the transcription repressor GFI1.

## Discussion

Our studies reveal the coincident genome-wide binding of the HMG-box protein HMG20B with LSD1 and GFI1 in myeloid leukemia cells, and the functional relevance of the interaction of the coiled-coil domain of HMG20B with LSD1 in stabilizing the interaction of LSD1 with GFI1. Depletion of HMG20B was sufficient to trigger a leukemia cell differentiation programme similar to that observed seen following treatment of leukemia cells with the LSD1 inhibitor OG86; and also to block granulocytic differentiation of normal murine hematopoietic stem and progenitor cells similar to that seen following *Gfi1* depletion.

While the location of the GFI1 transcription repressor complex is dictated by sequence-specific binding of the GFI1 zinc finger domain to its DNA consensus motif, the residency time of the assembled complex on chromatin, and thus the extent and duration of local transcription repression, is likely influenced by accrued co-location of additional co-factors. Our data suggest that HMG20B, which binds DNA independent of sequence via its HMG-box domain [30], may be one such factor. Other points of contact of the LSD1 complex with nucleosomal DNA include through the SANT2 domain of RCOR1 [35] and possibly through positively charged residues in the amino oxidase domain of LSD1 [36]. The role of other proteins identified as components of the LSD1 complex, such as ZMYM2, ZNF516, ZNF217, PHF21A and RREB1, in stabilizing the whole and interacting with DNA remains largely unclear. It is also not clear whether there are important differences in function of LSD1 complex proteins according to whether the complex is tethered at repressed sites of SNAG domain transcription factor binding or, for example, at active OCT4-bound enhancers as is seen in embryonic stem cells [37].

Related to this, although recruitment of LSD1 is a prerequisite for GFI1-mediated transcription repression [20], a detailed understanding of how this results in transcription repression is lacking. It has been argued that recruitment of histone tail deacetylase, H3K4 demethylase and H3K9 methyltransferase enzymatic activities by GFI1 may be sufficient to induce gene silencing [21]. However, (i) we identified neither EHMT1 or EHMT2 in our LSD1 or HMG20B pulldown experiments, (ii) the enzymatic activity of LSD1 is not required for transcription repression at sites of GFI1 binding [8] and (iii) expression of a GFI1 zinc finger-HDAC1 fusion protein only partially rescued GFI1-mediated transcription repression following treatment of AML cells with an LSD1 inhibitor [8]. It is possible that additional, as yet undetermined, mechanisms are of importance in GFI1-mediated transcription repression, such as phase separation and enhancement of RNA Pol II promoter pausing.

The role of HMG20B, and its paralog HMG20A, as components of the LSD1-repressor complex have been recently highlighted in other cellular contexts. In Merkel cell carcinoma, an aggressive neuroendocrine carcinoma of the skin, LSD1 inhibition derepresses key regulators of the neuronal lineage via the disruption of LSD1-CoREST complex integrity, including induction of HMG20B degradation [38]. Moreover, the dual depletion of HMG20B and HMG20A blocked GFI1B-mediated erythroid differentiation in the human erythroleukemia cell line K562 [27]. These data suggest that the involvement of HMG20B in stabilizing LSD1/CoREST may be a general mechanism. The human THP1 AML cell line used in our study showed little expression of the HMG20A paralog (data not shown), explaining why we did not identify HMG20A as an LSD1 co-partner. Interestingly, it has been proposed that HMG20A and HMG20B are mutually exclusive subunits of the complex [39], and in the developing central nervous system they display complementary expression patterns and functions [39]. Further studies are required to clarify the role of HMG20A in the CoREST complex in the hematopoietic system.

Our study adds to emerging data which further clarify the mechanisms of action of tranylcypromine-derivative LSD1 inhibitors such as iadademstat and bomedemstat which are currently being evaluated for efficacy alone or in combination in diverse malignancies such as AML, myelofibrosis and essential thrombocythemia [4, 40]. The potential of LSD1 inhibition to sensitize AML cells to other differentiation-inducing agents, like ATRA or BET protein inhibitors, has been postulated [3, 41–42]. Interestingly, GFI1 may play a central role in some of these synergistic therapeutic effects. For instance, GFI1 expression is high in t(8;21) driven AML and is required for the growth of this AML subtype [43], which may explain the high sensitivity these cells show to LSD1 inhibition [44]. GFI1 has also been proposed to be functionally important in acute promyelocytic leukemia (APL). Its expression is directly regulated by PML/RAR $\alpha$  and it coordinates with the fusion protein to maintain APL, in such a way that the PML/RAR $\alpha$ -induced differentiation block can be relieved upon GFI1 knockdown [45]. Accordingly, LSD1 inhibition sensitizes AML cell lines to retinoic acid treatment through the blockage of LSD1:GFI1 interaction [46]. Finally, it is interesting to note that treatment of NB4 promyelocytic leukaemia cells with LSD1 inhibitors leads to reduced association of the coiled-coil protein GSE1 with LSD1 [47], as we also noted (Tables S1-3). This appears to be as a result of impaired protein translation, and GSE1 downregulation is required for LSD1 inhibitor-induced differentiation [47].

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This work was supported by Cancer Research UK grant numbers C5759/A20971 and C5759/A27412 (AM-D, LN, FC, IR-C, GS, BW, FA, FS, TS); a Kay Kendall Leukaemia Fund Junior Research Fellowship (KKL1185) (to IR-C); Blood Cancer UK grant numbers 13005 and 19007 (to AP, AJKW, ADW); an educational grant from Imago Biosciences (to NB); and Leukaemia UK grant number 2020/JGF/005 (to BW). We are grateful to Jeff Barry, Wolfgang Breitwieser, Gillian Newton, Duncan Smith, John Weightman for technical support; and to Andrew Porter for pre-submission review of the manuscript.

## Sources of support

Cancer Research UK grant numbers C5759/A20971 and C5759/A27412; Kay Kendall Leukaemia Fund Junior Research Fellowship KKL1185; Blood Cancer UK grant numbers 13005 and 19007; Leukaemia UK grant number 2020/JGF/005.

## Data availability

High throughput sequencing data files are available at Gene Expression Omnibus with accession number GSE192975. Mass spectrometry proteomics data are available via ProteomeXchange with identifier PXD030581.

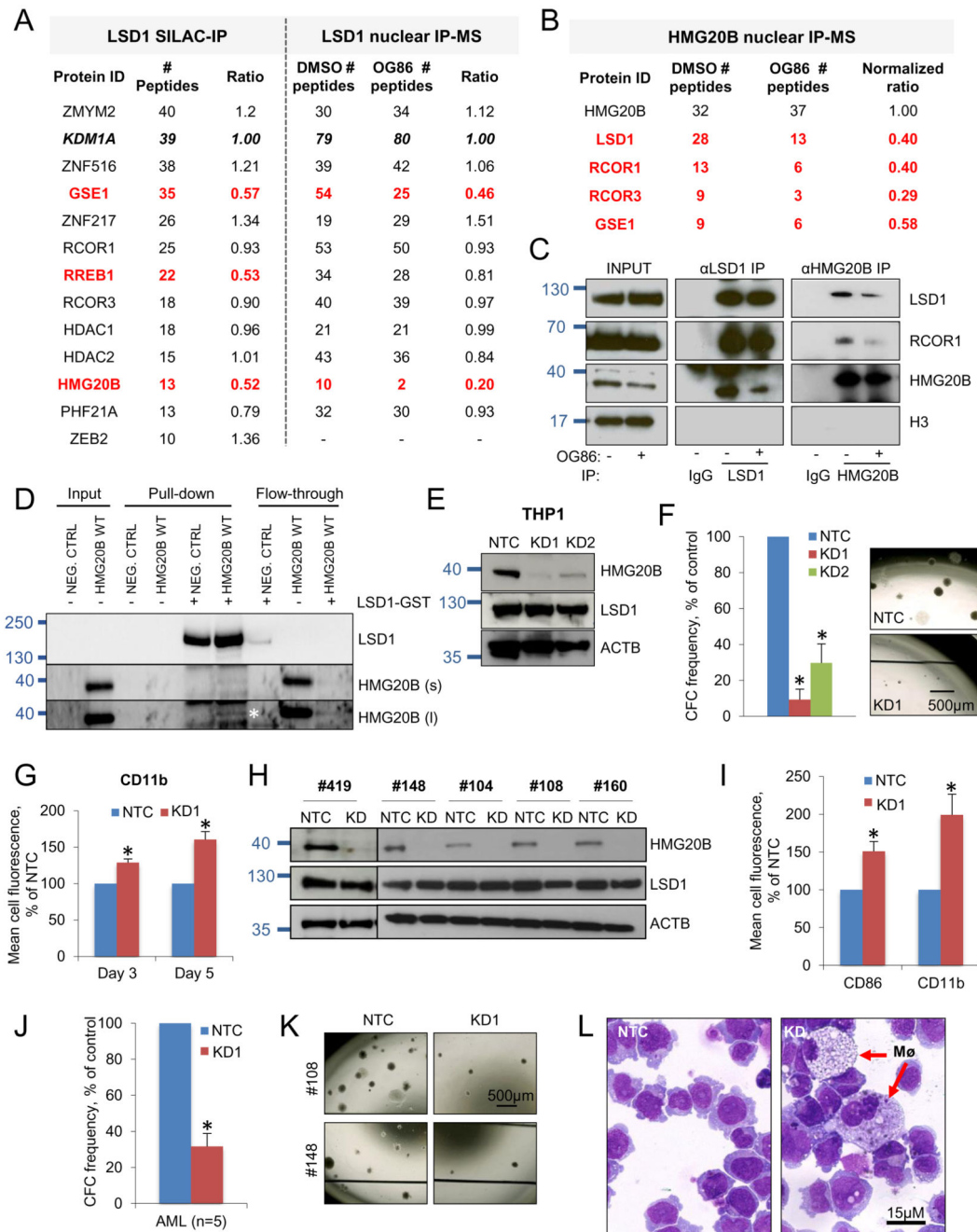
## References

1. Maiques-Diaz A, Somerville TC. LSD1: biologic roles and therapeutic targeting. *Epigenomics*. 2016; 8 (8) 1103–16. [PubMed: 27479862]
2. 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; 21 (4) 473–487. [PubMed: 22464800]
3. Schenk T, Chen WC, Göllner 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; 18 (4) 605–11. [PubMed: 22406747]
4. Salamero O, Montesinos P, Willekens C, Pérez-Simón JA, Pigneux A, Récher C, et al. First-in-Human Phase I Study of Iadademstat (ORY-1001): A First-in-Class Lysine-Specific Histone Demethylase 1A Inhibitor, in Relapsed or Refractory Acute Myeloid Leukemia. *Journal of Clinical Oncology*. 2020; 38 (36) 4260–4273. [PubMed: 33052756]
5. You A, Tong JK, Grozinger CM, Schreiber SL. CoREST is an integral component of the CoREST-human histone deacetylase complex. *Proc Natl Acad Sci USA*. 2001; 98 (4) 1454–8. [PubMed: 11171972]
6. Shi YY, Lan F, Matson C, Mulligan P, Whetstone JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 2004; 119 (7) 941–953. [PubMed: 15620353]
7. 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. *Mol Cell*. 2007; 27 (4) 562–572. [PubMed: 17707228]
8. Maiques-Diaz A, Spencer GJ, Lynch JT, Ciceri F, Williams EL, Amaral FMR, et al. Enhancer Activation by Pharmacologic Displacement of LSD1 from GFI1 Induces Differentiation in Acute Myeloid Leukemia. *Cell Rep*. 2018; 22 (13) 3641–3659. [PubMed: 29590629]
9. Lee C, Rudneva VA, Erkek S, Zapatka M, Chau LQ, Tacheva-Grigorova SK, et al. Lsd1 as a therapeutic target in Gfi1-activated medulloblastoma. *Nat Commun*. 2019; 10 (1) 332. [PubMed: 30659187]

10. Barth J, Abou-El-Ardat K, Dalic D, Kurre N, Maier AM, Mohr S, et al. LSD1 inhibition by tranlycypromine derivatives interferes with GFI1-mediated repression of PU.1 target genes and induces differentiation in AML. *Leukemia*. 2019; 33 (6) 1411–1426. [PubMed: 30679800]
11. Carmichael CL, Wang J, Nguyen T, Kolawole O, Benyoucef A, De Mazière C, et al. The EMT modulator SNAI1 contributes to AML pathogenesis via its interaction with LSD1. *Blood*. 2020; 136 (8) 957–973. [PubMed: 32369597]
12. 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 impairment growth of AML cells. *Mol Cancer Ther*. 2017; 16: 273–284. [PubMed: 27903753]
13. Takagi S, Ishikawa Y, Mizutani A, Iwasaki S, Matsumoto S, Kamada Y, et al. LSD1 Inhibitor T-3775440 Inhibits SCLC Cell Proliferation by Disrupting LSD1 Interactions with SNAG Domain Proteins INSM1 and GFI1B. *Cancer Res*. 2017; 77 (17) 4652–4662. [PubMed: 28667074]
14. 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; 9 (30) 21007–21021. [PubMed: 29765516]
15. Hock H, Hamblen MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y, et al. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature*. 2004; 431: 1002–7. [PubMed: 15457180]
16. Hock H, Hamblen MJ, Rooke HM, Traver D, Bronson RT, Cameron S, Orkin SH. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity*. 2003; 18: 109–120. [PubMed: 12530980]
17. 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; 30: 295–300. [PubMed: 11810106]
18. Person RE, Li F-Q, Duan Z, Benson KF, Wechsler J, Papadaki HA, et al. Mutations in protooncogene GFI1 cause human neutropenia and target ELA2. *Nat Genet*. 2003; 34: 308–312. [PubMed: 12778173]
19. van Bergen MGJM, van der Reijden BA. Targeting the GFI1/1B-CoREST Complex in Acute Myeloid Leukemia. *Front Oncol*. 2019; 9: 1027. [PubMed: 31649884]
20. Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tschlis PN. The Gfi-1 protooncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. *Mol Cell Biol*. 1996; 16: 6263–6272. [PubMed: 8887656]
21. Möröy T, Khandanpour C. Role of GFI1 in Epigenetic Regulation of MDS and AML Pathogenesis: Mechanisms and Therapeutic Implications. *Front Oncol*. 2019; 9: 824. [PubMed: 31508375]
22. Sumoy L, Carim L, Escarceller M, Nadal M, Gratacòs M, Pujana MA, et al. HMG20A and HMG20B map to human chromosomes 15q24 and 19p13.3 and constitute a distinct class of HMG-box genes with ubiquitous expression. *Cytogenet Cell Genet*. 2000; 88 (1-2) 62–7. [PubMed: 10773667]
23. 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; 99 (11) 7420–5. [PubMed: 12032298]
24. Esteghamat F, van Dijk TB, Braun H, Dekker S, van der Linden R, Hou J, et al. The DNA binding factor Hmg20b is a repressor of erythroid differentiation. *Haematologica*. 2011; 96 (9) 1252–60. [PubMed: 21606163]
25. Rivero S, Ceballos-Chávez M, Bhattacharya SS, Reyes JC. HMG20A is required for SNAI1-mediated epithelial to mesenchymal transition. *Oncogene*. 2015; 34 (41) 5264–76. [PubMed: 25639869]
26. Welcker JE, Hernandez-Miranda LR, Paul FE, Jia S, Ivanov A, Selbach M, Birchmeier C. Insm1 controls development of pituitary endocrine cells and requires a SNAG domain for function and for recruitment of histone-modifying factors. *Development*. 2013; 140 (24) 4947–58. [PubMed: 24227653]
27. McClellan D, Casey MJ, Bareyan D, Lucente H, Ours C, Velinder M, et al. Growth Factor Independence 1B-Mediated Transcriptional Repression and Lineage Allocation Require Lysine-

- Specific Demethylase 1-Dependent Recruitment of the BHC Complex. *Mol Cell Biol.* 2019; 39 (13) e00020-19 [PubMed: 30988160]
28. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008; 9 (9) R137. [PubMed: 18798982]
  29. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics.* 2011; 27 (12) 1696–7. [PubMed: 21486936]
  30. Marmorstein LY, Kinev AV, Chan GK, Bochar DA, Beniya H, Epstein JA, et al. A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression. *Cell.* 2001; 104 (2) 247–57. [PubMed: 11207365]
  31. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 2011; 27 (12) 1739–40. [PubMed: 21546393]
  32. Somerville TC, Matheny CJ, Spencer GJ, Iwasaki M, Rinn JL, Witten DM, et al. Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. *Cell Stem Cell.* 2009; 4 (2) 129–40. [PubMed: 19200802]
  33. Deb G, Wingelhofer B, Amaral FMR, Maiques-Diaz A, Chadwick JA, Spencer GJ, et al. Pre-clinical activity of combined LSD1 and mTORC1 inhibition in MLL-translocated acute myeloid leukaemia. *Leukemia.* 2020; 34 (5) 1266–1277. [PubMed: 31780813]
  34. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol.* 2010; 28 (5) 495–501. [PubMed: 20436461]
  35. Yang M, Gocke CB, Luo X, Borek D, Tomchick DR, Machius M, et al. Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. *Mol Cell.* 2006; 23 (3) 377–87. [PubMed: 16885027]
  36. Song Y, Dagil L, Fairall L, Robertson N, Wu M, Ragan TJ, et al. Mechanism of Crosstalk between the LSD1 Demethylase and HDAC1 Deacetylase in the CoREST Complex. *Cell Rep.* 2020; 30 (8) 2699–2711. [PubMed: 32101746]
  37. Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT, et al. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature.* 2012; 482 (7384) 221–5. [PubMed: 22297846]
  38. Leindecker L, Jung PS, Krecioch I, Neumann T, Schleiffer A, Mechtler K, et al. LSD1 inhibition induces differentiation and cell death in Merkel cell carcinoma. *EMBO Mol Med.* 2020; 12 (11) e12525 [PubMed: 33026191]
  39. Wynder C, Hakimi MA, Epstein JA, Shilatifard A, Shiekhattar R. Recruitment of MLL by HMG-domain protein iBRAF promotes neural differentiation. *Nat Cell Biol.* 2005; 7 (11) 1113–7. [PubMed: 16227968]
  40. Pettit K, Gerds AT, Yacoub A, Watts JM, Tartaczuch M, Bradley TJ, et al. A Phase 2a Study of the LSD1 Inhibitor Img-7289 (bomedemstat) for the Treatment of Myelofibrosis. *Blood.* 2019; 134 (Supplement\_1) 556.
  41. Schenk T, Chen WC, Göllner 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; 18 (4) 605–11. [PubMed: 22406747]
  42. Wass M, Göllner S, Besenbeck B, Schlenk RF, Mundmann P, Göthert JR, et al. A proof of concept phase I/II pilot trial of LSD1 inhibition by tranlycypromine combined with ATRA in refractory/relapsed AML patients not eligible for intensive therapy. *Leukemia.* 2021; 35 (3) 701–711. [PubMed: 32561840]
  43. Fiskus W, Mill CP, Nabet B, Perera D, Birdwell C, Manshoury T, et al. Superior efficacy of co-targeting GFII/KDM1A and BRD4 against AML and post-MPN secondary AML cells. *Blood Cancer J.* 2021; 11 (5) 98. [PubMed: 34016956]
  44. Marneth AE, Botezatu L, Hönes JM, Israël JCM, Schütte J, Vassen L, et al. GFII1 is required for RUNX1/ETO positive acute myeloid leukemia. *Haematologica.* 2018; 103 (9) e395–e399. [PubMed: 29674496]
  45. McGrath JP, Williamson KE, Balasubramanian S, Odate S, Arora S, Hatton C, et al. Pharmacological Inhibition of the Histone Lysine Demethylase KDM1A Suppresses the Growth

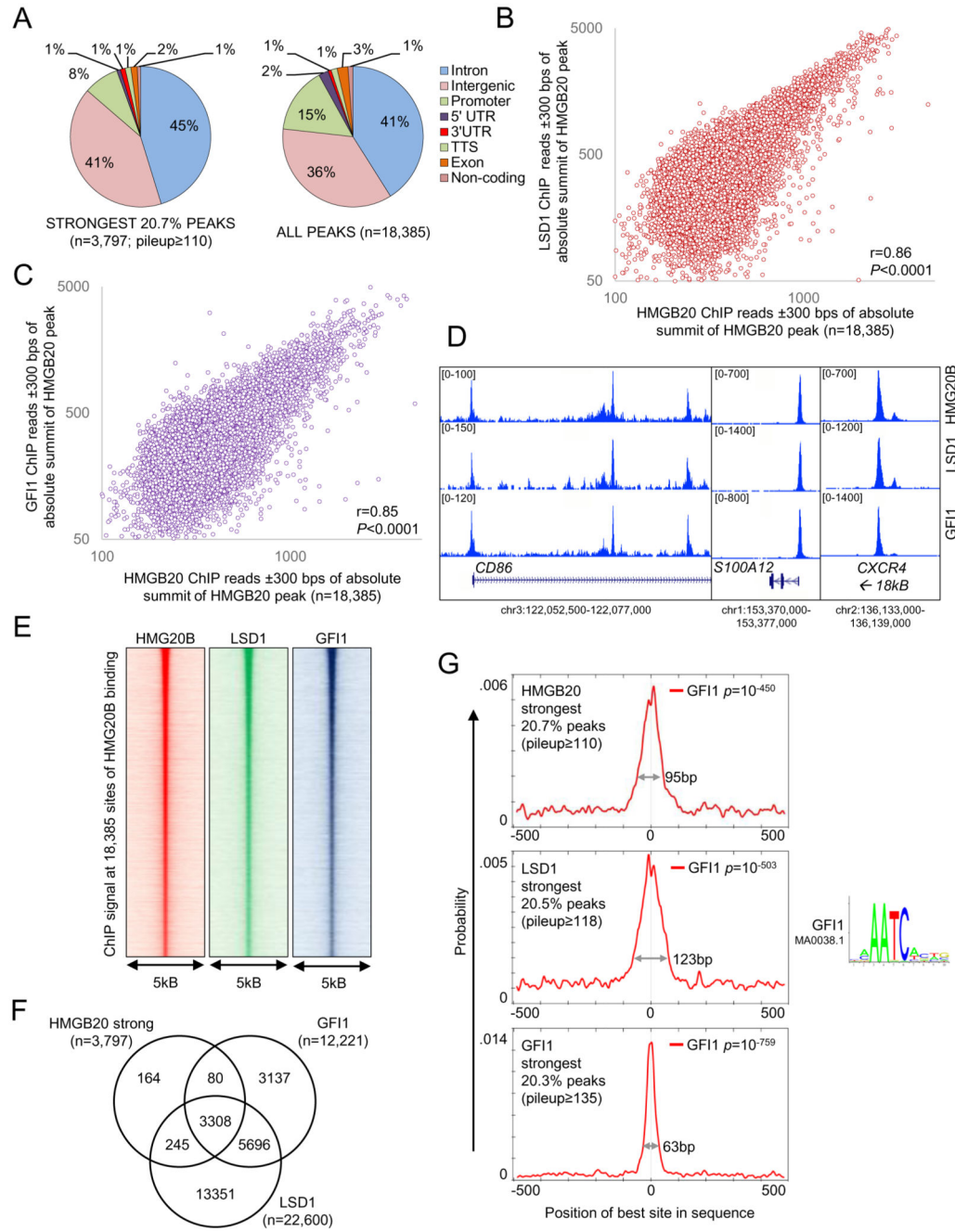
- of Multiple Acute Myeloid Leukemia Subtypes. *Cancer Res.* 2016; 76 (7) 1975–88. [PubMed: 26837761]
46. Tayari MM, Santos HGD, Kwon D, Bradley TJ, Thomassen A, Chen C, et al. Clinical Responsiveness to All-trans Retinoic Acid Is Potentiated by LSD1 Inhibition and Associated with a Quiescent Transcriptome in Myeloid Malignancies. *Clin Cancer Res.* 2021; 27 (7) 1893–1903. [PubMed: 33495312]
47. Ravasio R, Ceccacci E, Nicosia L, Hosseini A, Rossi PL, Barozzi I, et al. Targeting the scaffolding role of LSD1 (KDM1A) poises acute myeloid leukemia cells for retinoic acid-induced differentiation. *Sci Adv.* 2020; 6 (15) eaax2746 [PubMed: 32284990]
48. Nicosia L, Boffo FL, Ceccacci E, Conforti F, Pallavicini I, Bedin F, et al. Pharmacological inhibition of LSD1 triggers myeloid differentiation by targeting GSE1 oncogenic functions in AML. *Oncogene.* 2021; 41 (6) 878–894. [PubMed: 34862459]



**Figure 1. Pharmacological inhibition of LSD1 destabilizes its interaction with HMG20B** (A-C) THP1 AML cells were treated with 250nM OG86 or DMSO vehicle for 48 hr. Proteins known to form part of an LSD1-containing complex and identified by mass spectrometry techniques following (A) LSD1 or (B) HMG20B immunoprecipitation (IP) are shown. Proteins in red exhibit reduced interaction in the presence of OG86. (C) IP westerns of nuclear lysates using the indicated antibodies. (D) IP western shows interaction (white asterisk) of in vitro synthesized HMG20B with purified GST-LSD1, and efficient depletion of HMG20B from flow through. s = short exposure; l = long exposure. (E-G)

THP1 AML cells were infected with lentiviral vectors targeting *HMG20B* for KD, or a non-targeting control (NTC), with puromycin drug resistance as the selectable marker. (E) Western blots show expression of the indicated proteins three days after infection. (F) Mean+SEM colony-forming cell (CFC) frequencies relative to controls after ten days in semi-solid culture (n=3); right panels show representative colonies. (G) Mean+SEM CD11b cell fluorescence relative to control as determined by flow cytometry three and five days after initiation of knockdown (n=2). (H-L) Five independent primary *MLL*-rearranged AML cells were infected with lentiviral vectors targeting *HMG20B* for KD (KD1 construct used), or a NTC, with puromycin drug resistance as the selectable marker (# refers to Biobank identifier). (H) Western blots show the indicated proteins three days after infection. (I) Mean+SEM CD11b and CD86 cell fluorescence relative to control five days after initiation of knockdown (n=5 separate patient samples). (J) Mean+SEM CFC frequencies after ten days in semi-solid culture (n=5 separate patient samples). (K) Representative images of colonies. (L) Representative images of cytopsin preparations. MØ, macrophage. \* indicates  $P<0.05$  by unpaired t-test or, in (F), by one-way ANOVA with a Tukey post hoc test.

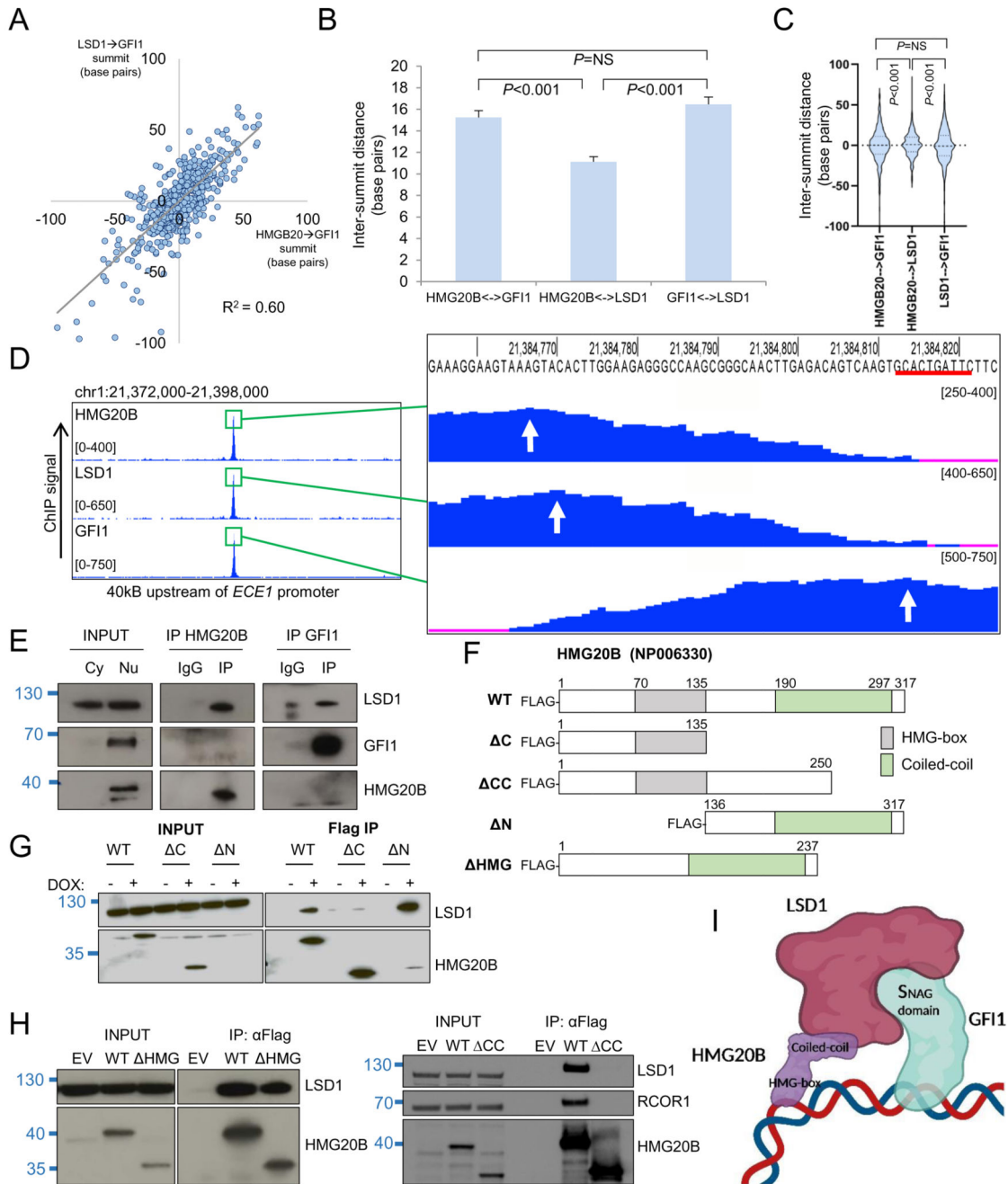




**Figure 2. HMG20B cooccupies chromatin with GFI1 and LSD1**

(A) Pie charts show genome annotations for the strongest (left) or all (right) HMG20B peaks. UTR, untranslated region; TTS, transcription termination sequence. (B-C) Dot plots show correlation between HMG20B ChIP signal and (B) LSD1 ChIP signal or (C) GFI1 ChIP signal. (D) Exemplar ChIPseq tracks. (E) Heatmaps show HMG20B, LSD1 and GFI1 ChIP signal +/- 2.5kb surrounding absolute summit of all HMG20B peaks ranked by peak strength. (F) Venn diagram shows intersection of strong HMG20B peaks with LSD1 and

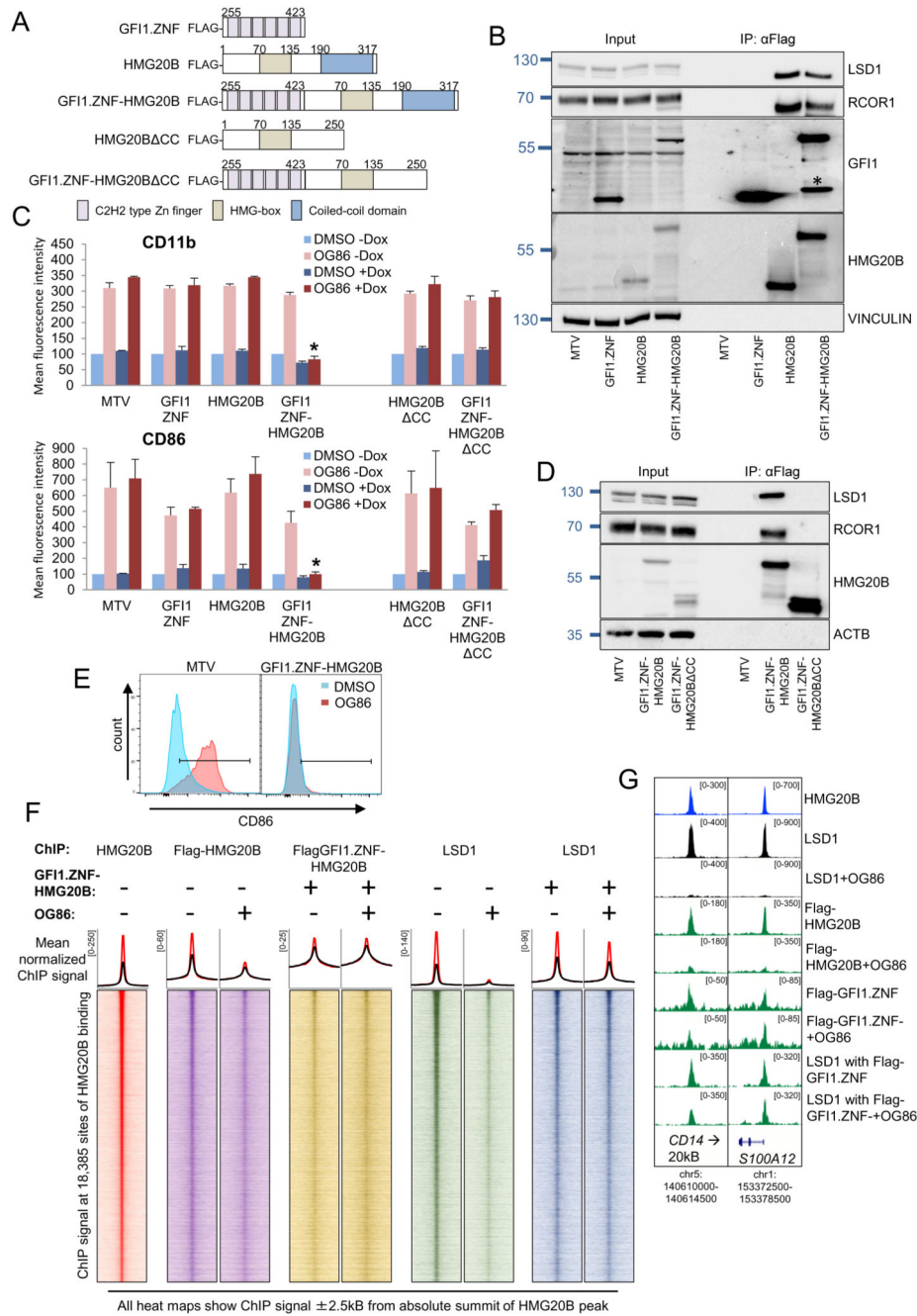
GFI1 binding peaks. (G) MEME-ChIP GFI1 motif enrichment plots surrounding absolute summits of strong HMG20B, LSD1 or GFI1 binding peaks.



**Figure 3. Co-localization pattern of HMG20B, LSD1 and GFI1 on chromatin**

(A) Dot plot show base pair distances between the absolute summits of HMG20B and GFI1 binding peaks (with pileup value = 300) versus base pair distances between the absolute summits of LSD1 and GFI1 binding peaks. (B) Bar graph shows mean+SEM intersummit base pair distances between the indicated factor binding peaks. (C) Violin plots show distribution, median (thick dotted line) and interquartile range (light dotted lines) for intersummit base pair distances between the indicated factor binding peaks. (D) Exemplar ChIPseq tracks. Red line indicates GFI1 binding motif. White arrows indicate called peak

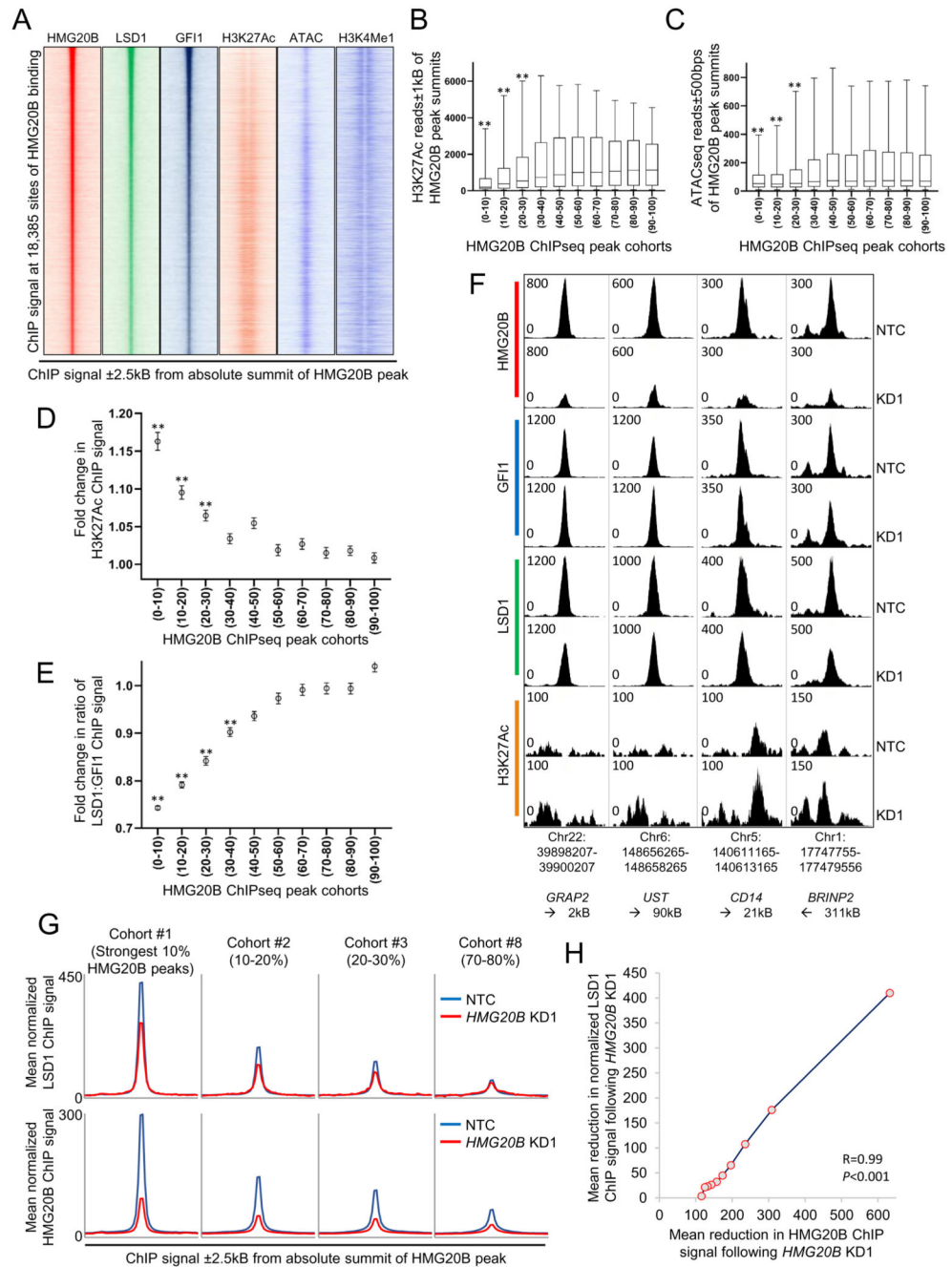
summits. (E) Western shows nuclear immunoprecipitation for the indicated proteins in THP1 AML cells. Cy, cytoplasmic; Nu, nuclear. (F) HMG20B constructs. WT = wild type. (G-H) Westerns shows doxycycline-induced (DOX) expression of HMG20B constructs and  $\alpha$ Flag immunoprecipitations in THP1 AML cells. (I) Proposed model of chromatin binding (created with BioRender). For (B) and (C), one way ANOVA with Tukey post hoc test was used to determine statistical significance.



**Figure 4. Functional recruitment of LSD1 to chromatin by a GFII DNA binding domain:HMG20B fusion protein**

(A) Construct maps. (B) Western shows doxycycline-induced expression of the indicated constructs and αFlag IPs in THP1 AML cells. \* indicates degradation product. (C) THP1 AML cells were infected with lentiviruses expressing the indicated doxycycline-inducible (Dox) GFII fusion or control constructs and then treated with 250nM OG86 or DMSO vehicle. Bar graphs indicate means+SEM fluorescence intensity for the indicated markers, as determined by flow cytometry, in the indicated conditions 24 hrs later (n=3). \* indicates  $P < 0.05$  for the indicated condition versus all other OG86+Dox conditions by one way

ANOVA and a Tukey post hoc test. (D) Western blot shows Dox-induced expression of the indicated constructs and  $\alpha$ Flag IPs in THP1 AML cells. (E) Representative flow cytometry plots. (F) Heatmaps show ChIP signal for the indicated proteins and conditions. Line graphs above show mean normalized ChIP signal for each heatmap surrounding all (black line) or the strongest 20% (red line) of HMG20B peaks. (G) Exemplar ChIPseq peaks tracks.

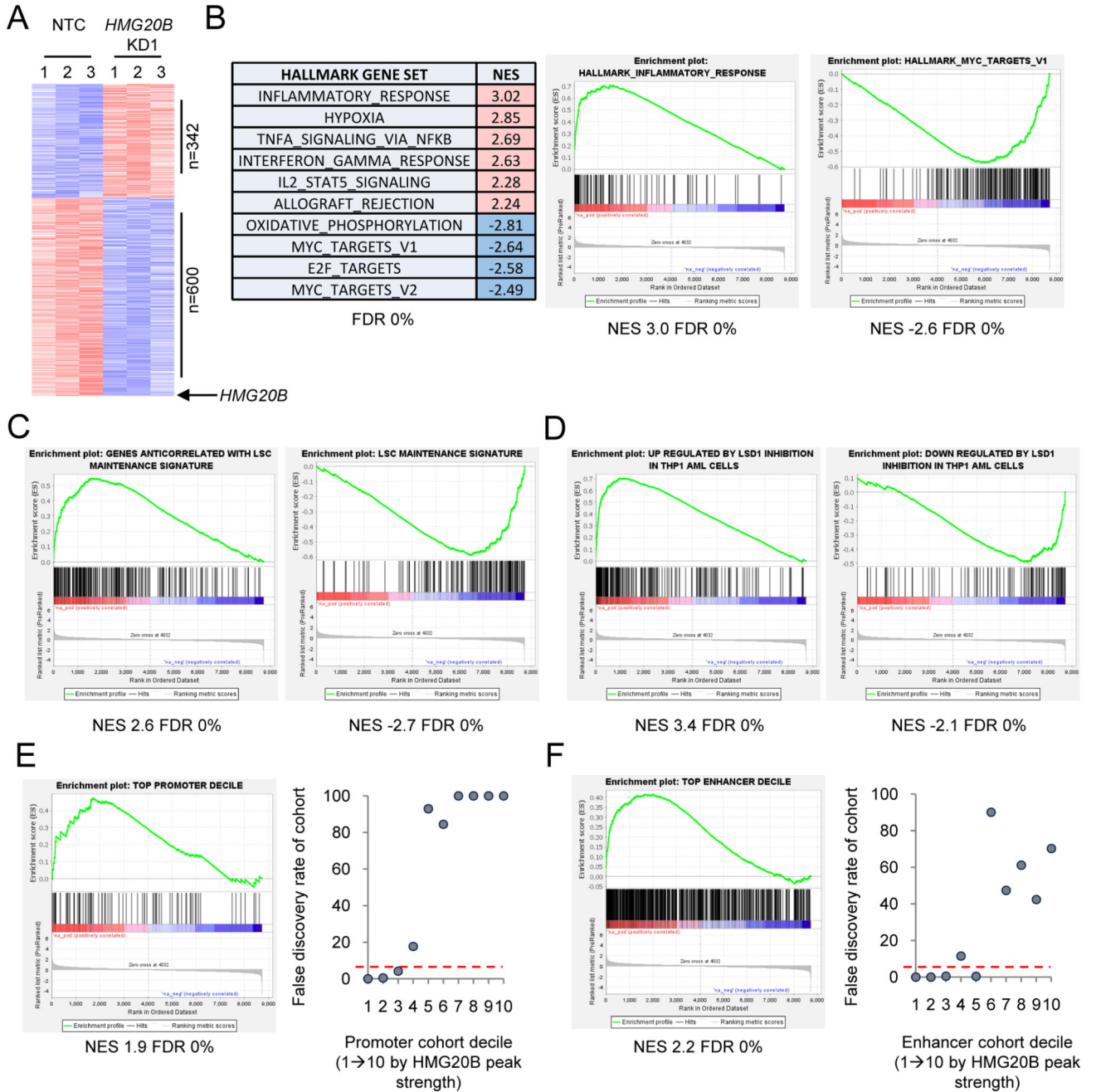


**Figure 5. HGM20B stabilizes the interaction of LSD1 with GFI1 on chromatin**

(A) Heatmaps show ChIP signal for the indicated proteins, histone modifications and chromatin accessibility. (B-C) HMG20B ChIPseq peaks were grouped into ten cohorts according to peak strength. Boxplots show median, 25th and 75th centile values (box), and 5th and 95th centile values (whiskers) for (B) H3K27Ac ChIP signal or (C) ATACseq signal surrounding the indicated cohort of HMG20B binding peaks. (D-F) THP1 AML cells were infected with lentiviral vectors targeting *HMG20B* for KD, or a non-targeting control (NTC). 72 hrs later ChIPseq was performed on puromycin resistant cells. HMG20B ChIPseq

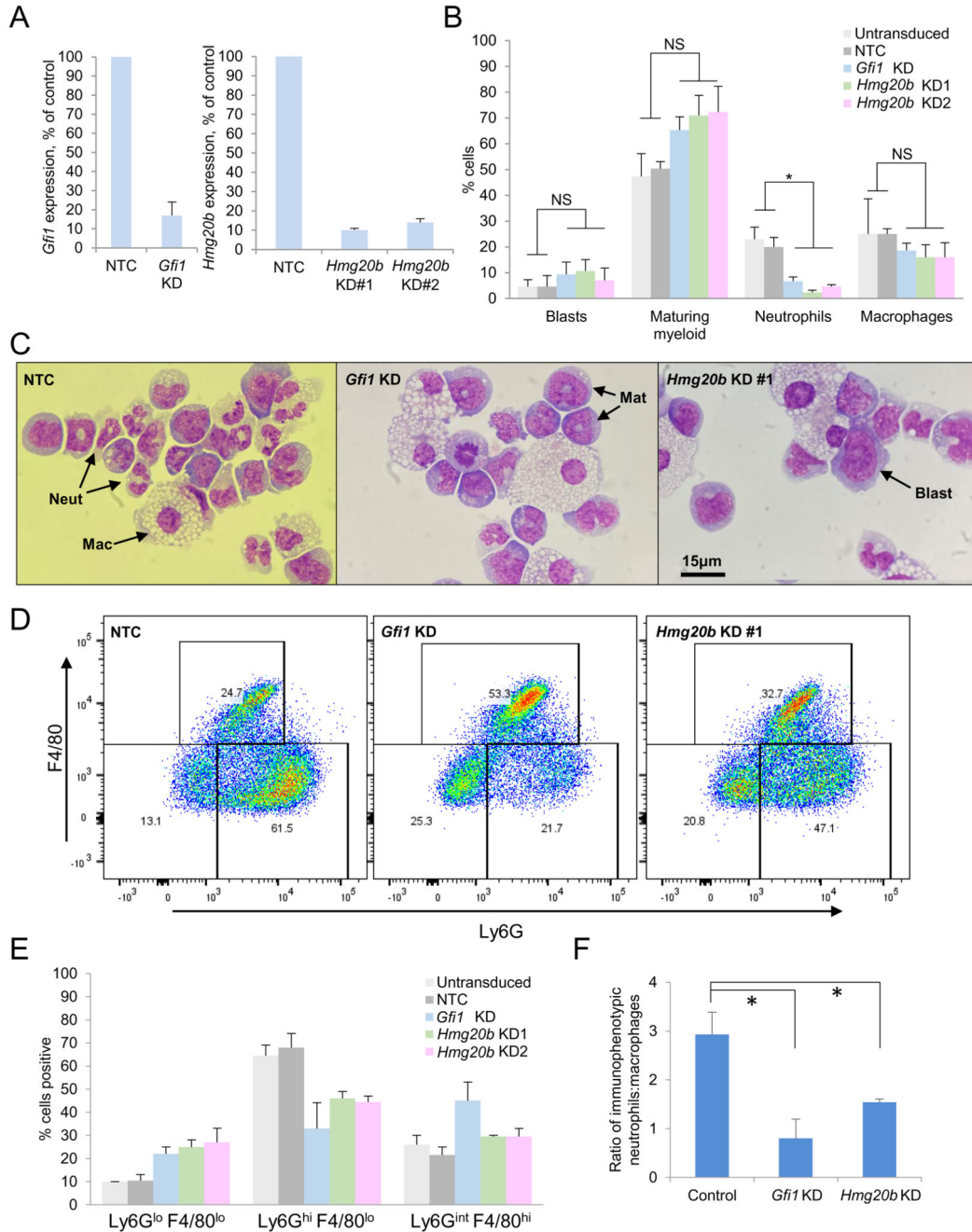
peaks were grouped into ten cohorts according to peak strength. Graphs show mean $\pm$ SEM fold change in (D) H3K27ac ChIP signal or (E) normalized ratio of GFI1:LSD1 ChIP signal surrounding the indicated cohort of HMG20B binding peaks. For (B-E) \*\* indicates  $P < 0.01$  for comparison of each of the top three cohorts versus each of the bottom five cohorts by one way ANOVA and a Tukey post hoc test. (F) Exemplar ChIPseq tracks showing normalized ChIP signal. (G) Density plots show ChIP signal for (upper row) LSD1 (normalized to assume equal GFI1 ChIP signal in NTC and KD conditions) and (lower row) HMG20B in the indicated conditions and for the indicated exemplar cohorts. (H) Graph shows mean absolute loss of LSD1 ChIP signal versus mean loss of HMG20B ChIP signal ( $\pm 300$ bps from HMG20B peak summit) for the ten HMG20B cohorts after *HMG20B* KD.





**Figure 6. Gene expression changes induced by *HMG20B* knockdown**  
(A-F) THP1 AML cells were infected with lentiviral vectors targeting *HMG20B* for KD, or a non-targeting control (NTC) and selected for puromycin resistance. (A) Heatmap shows differentially expressed genes 72 hrs after initiation of KD. (B) Summary table and representative GSEA plots. (C-D) GSEA plots. (E-F) *HMG20B* binding peaks were grouped into ten cohorts according to peak strength and mapped to nearby genes to generate gene sets for use in GSEA. Graphs (right panels) show false discovery rates by cohort for genes

bound by HMG20B at (E) promoters and (F) putative enhancers. GSEA plots are shown for the top cohorts (left panels).



**Figure 7. HMG20B knockdown in murine hematopoietic stem and progenitor cells**  
 (A-C) Murine bone marrow KIT<sup>+</sup> cells were infected with a lentiviral vector targeting *Hmg20b* or *Gfi1* for KD, or a non-targeting control (NTC), and selected with puromycin for 48 hrs prior to semisolid culture. (A) KD efficiency for the indicated genes at 72 hrs, as determined by qPCR. (B) Bar chart shows mean+SEM percentage of the indicated cell types in cytopsin preparations after seven days in semisolid culture (n=3). \* indicates  $P < 0.05$  for the indicated comparisons by one way ANOVA with a Tukey post hoc test. (C) Exemplar cytopsin preparations. Mat, maturing; Mac, macrophage; Neut, neutrophil. (D) Exemplar

flow cytometry plots. (E) Bar chart shows mean+SEM percentage of cells with the indicated immunophenotype after seven days in semisolid culture (n=2); and (F) mean+SEM ratio of immunophenotypic granulocytes (Ly6G<sup>hi</sup> F4/80<sup>lo</sup>) to macrophages (Ly6G<sup>int</sup> F4/80<sup>hi</sup>). Control = untransduced or NTC (n=4); *Hmg20b* KD = KD #1 and #2 (n=4); *Gfi1* KD (n=2). \* indicates  $P < 0.05$  for the indicated comparisons by one way ANOVA with a Tukey post hoc test.