Supplemental Material

Epigenetic regulation of lineage commitment in normal and malignant hematopoiesis by histone demethylase KDM2B

Jaclyn Andricovich[1,2], Yan Kai[1,2,3], Weiqun Peng[3], Adlen Foudi[4], and Alexandros Tzatsos[1,2#]

[1]Cancer Epigenetics Laboratory, Department of Anatomy and Regenerative Biology, and

[2]George Washington Cancer Center, George Washington University School of Medicine and Health Sciences, Washington DC 20037, USA

[3]Department of Physics, George Washington University, Washington DC 20037, USA

[4] INSERM UMR-S935, André Lwoff Institute, Paul Brousse Hospital, Villejuif, France

Key words: KDM2B, Cancer, Epigenetics, Hematopoiesis, Leukemia, Histone methylation

INVENTORY OF SUPPLEMENTAL MATERIAL:

SUPPLEMENTAL METHODS

Animal studies. Mice were housed in a pathogen-free animal facility and maintained on a mixed 129SV/C57BL/6 background, unless otherwise indicated. Data presented include both male and female mice at the indicated ages. For the leukemia studies, mice were followed monthly with complete blood counts with differentials on a Heska HemaTrue hematology analyzer (~20µl of blood directly obtained into a capillary from the mouse tail) and blood smears for the presence of immature or abnormal hematopoietic cells. Mice that were moribund were sacrificed for further analysis when they: (a) had WBC over 80,000/µl, (b) developed palpable lymph nodes larger than 1 cm and severe spleenomegaly, (c) had lost >20% of their body weight, (d) showed poor grooming, hunched posture, and decreased activity, or (e) developed severe bleeding and petechiae, indicative of thrombocytopenia, in ears and feet. In the survival studies, we observed that doxycycline administration, to induce the expression of Kdm2b, slowed the progression of Kras driven leukemia and prolonged survival for 2-3 weeks (compare **Figures 7B** and **8B** for Kras only mice).

Generation of conditional Kdm2b knockout mice: To generate the targeting vector for the *Kdm2b* conditional knockout allele, a C57BL/6 BAC clone (Children's Hospital Oakland Research Institute) was used to amplify the short (exons 14-15) and long (exons 20-22) arms and cloned into the pKOII vector which expresses the neomycin gene flanked by *frt* sites (1). Exons 16-19 were similarly amplified and subcloned in the same vector flanked by *loxp* sites. The final construct was linearized with NotI, electroporated into V6.5 ES cells, and positive clones were selected in 400 mg/ml G418. Targeted ES clones were confirmed via Southern blotting and PCR analysis. Correct clones were injected into blastocysts, and chimeric male offspring were then crossed with $Actin^{Flpe}$ mice (Jax 003800, (2)) to excise the *Frt*-flanked PGK-Neo cassette, generating the $Kdm2b^{lox}$ allele (**Figure 1A**). To generate $Kdm2b^{\Delta}$ allele animals were crossed with $Gata1^{Cre}$ (3) mice that cause early embryonic deletion, including germline. Cre mediated excision of exons 16-19 fuses exons 15 and 20 to create an out-of-frame transcript. Successful deletion was confirmed by (a) Western blotting using $Kdm2b^{fl/fl}$ MEFs infected with AdeCre (**Figure 1B**), (b) IHC in $Gata1^{Cre};Kdm2b^{fl/fl}$ embryos (**Figure S1C**), and (c) PCR to detect the $Kdm2b^{\Delta}$ allele in genomic DNA isolated from MEFs and embryos (**Figures 1B** and **S1B**).

Primers to detect the *Kdm2b*^{lox} allele: P1_109141F 5'-TTCCGTCATTATCAGGAGGATCCAGTACCT-‘3 and P2_109320R 5'-ACTTACTTAGCA TCTACCCTGCCGACATTT-‘3, and *Kdm2b*$^{\Delta}$ allele: P3_106621F 5'-GTGTGGTCAACGATGAGCTTCCCAACTGCT-‘3 and P4_109440R 5'-CTGGCCAAACAGGAAATGTCATCAGAGCCG-‘3. Nucleotide numbers refer to NC_000071 (C57BL/6J). To achieve excision in the hematopoietic system we used *Vav1*^{Cre} (4), *Tie2*^{Cre} [B6.Cg-Tg(Tek-cre)12Flv/J; Jax 004128, (5)], and Mx1^{Cre} [C57BL/6J-Tg(Mx1-cre)1Cgn/J; Jax 003556, (6)] mouse strains.

Generation of mice to conditionally express Kdm2b: cDNA of wild-type and demethylation deficient (*Kdm2b*$^{\Delta Jmjc}$) KDM2B were subcloned in the EcoRI sites of pBS31 vector (Open Biosystems, MES4758 gene targeting kit) and were co-electroporated into KH2 murine ES cells with a plasmid expressing Flpe recombinase (pCAGGS-Flpe) to allow Frt/Flpase-mediated single copy site-specific integration into the *Collagen-a1* locus (7). ES cells were screened for integration of the transgene by PCR and correctly targeted clones were injected into blastocysts. Mice generated from these ES cells have doxycycline inducible Kdm2b expression when crossed with strains harboring the *rtTA* transactivator; we used the ROSA26-rtTA^{LSL}-IRES-EGFP [B6.Cg-Gt(ROSA)26Sortm1(rtTA,EGFP)Nagy/J; Jax 005670 (8)] strain that expresses rtTA and GFP in the hematopoietic system when crossed with *Vav1*^{Cre} mice, and Kdm2b upon administration of doxycycline (**Figure 3A**). In all experiments, doxycycline was administered in mice at three weeks of age in the drinkable water at a final concentration of 0.1 mg/ml. Expression of Kdm2b was confirmed by Western blotting and IHC. In some experiments *Kdm2b* transgenic animals were crossed with the LSL-Kras^{G12D} [B6.129-*Kras*^{tm4Tyj}/Nci; strain 01XJ6, (9)] mice which have the Kras^{G12D} allele knocked into the endogenous *Kras* locus, preceded by a *lox-stop-lox* cassette. For transplantation of human leukemia cell lines we used NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (Jax 005557) mice.

Embryo Dissection for AGM. Pregnant dams were sacrificed eleven days post coitus. After removing maternal membranes and yolk sacs, tails were clipped for genotyping. Incisions were made above forelimbs and below hindlimbs, and neural tissue was carefully removed away from the dorsal aorta under a stereoscope. All ventral tissues were dissected away from the aorta, and lateral body walls

were removed. AGMs were placed in a tube of staining buffer, Collagenase Type I was added at a final concentration of 0.25%, and tissue was digested at 37°C for 30 minutes. To ensure single cell suspensions, digestions were reverse pipetted to break up aggregates. Cells were washed twice to remove trace amounts of collagenase and analyzed by flow cytometry.

ChIP-seq. 10^7 cells were cross-linked in 1% formaldehyde for 20 minutes, and then quenched with glycine [final concentration 125 mM] for 5 minutes. Cells were washed with PBS, and then lysed in ice-cold 0.5% Igepal CA-630 in PBS supplemented with protease inhibitors (Thermo Fisher Scientific) for 10 minutes on ice. Nuclei were pelleted, digested with 1 µL micrococcal nuclease (New England Biolabs) for 3 minutes at 37°C to fragment chromatin, lysed with SDS lysis buffer [50 mM Tris-HCl (pH 8.0), 1% SDS, 150 mM NaCl, and 5 mM EDTA], and briefly sonicated. Cells were centrifuged 14,000 x g for 30 min at 4°C, and ten times the volume of dilution buffer [16.7 mM Tris (pH 8.0), 167 mM NaCl, 0.01% SDS, 1.1% Triton X-100, and 1.2 mM EDTA] was added to the soluble fraction. IP was set up on protein A agarose with a rabbit anti-KDM2B antibody (Millipore 09-864), previously validated (10), and binding occurred overnight at 4°C. The IP was then washed twice for 10 minutes with each: low salt buffer [20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1% SDS, 1% Triton X-100, and 2 mM EDTA], high salt buffer [20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 0.1% SDS, 1% Triton X-100, and 2 mM EDTA], LiCl buffer [10 mM Tris-HCl (pH 8.0), 0.25 M LiCl, 1% Igepal CA630, 1% sodium deoxycholate, and 1 mM EDTA], and TE [10 mM Tris pH 7.6, 1 mM EDTA]. DNA-protein cross-links were eluted with elution buffer [0.1 M $NaHCO_3$, 1% SDS] at room temperature for 30 min. Reverse cross-linking was performed overnight at 65°C by increasing the concentration of NaCl. Proteins were digested with proteinase-K for one hour at 42°C and the ChIP was cleaned with PCR Purify Kit (Qiagen). A small aliquot of the ChIP was validated in house prior to sequencing. Library construction and sequencing took place on an Illumina Hi-Seq 2500 instrument.

ChIP-Seq Data Analysis. Quality control and read alignment: FastQC (Version 0.10.1) was used to examine the read quality. Reads were aligned to the reference hg19 genome using Bowtie aligner (Version 1.1.1, (11)), with default parameters. Only uniquely aligned reads were kept. Peak Calling: To identify the bound genomic regions of target proteins, we used MACS package (Version 1.4.2, (12).

Redundant reads, which might be results of PCR artifacts, were filtered before peak calling with a *p*-value threshold $<10^{-5}$. Peak annotation: We used RefSeq transcripts from the UCSC genome browser (https://genome.ucsc.edu/) to annotate KDM2B peaks (enriched binding regions). For each gene, a ± 2kb window spanning the Transcription Start Site (TSS) was used to address whether it overlaps with a peak. If there was an overlap, this gene was considered to be bound by KDM2B. Genomic distribution of ChIP-seq peaks: Each peak was assigned to a category among promoters, gene-bodies, and inter-genic regions. For each gene, promoter region was defined as [TSS ± 2kb], and gene-body as [TSS + 2kb]. If a peak did not fall in those categories, it was considered to bind an inter-genic region. Metagene profiles: From the peaks' genomic distributions, we found that KDM2B preferentially binds TSS (promoter). To obtain an unbiased estimation of binding pattern near TSS loci, we plotted metagene profiles to show average ChIP-seq enrichment signal over TSS in bound genes (13, 14). Raw and processed data have been deposited in the Gene Expression Omnibus repository under accession number GSE70729.

Encode data analysis. Publicly available ChIP-seq data for K562 cells were downloaded from the ENCODE project (https://www.encodeproject.org/, (15)). To allow better comparison of results, raw reads for the selected ENCODE ChIP-seq samples were processed using the same analysis pipeline adopted for our data, including re-alignment of reads over the reference genome. A complete list of the accession numbers is given in a separate excel file. In total, we analyzed 125 datasets for transcription factors and histone modifications as well as 33 control datasets. Peak calling: We employed MACS package (Version 1.4.2) and used the same parameters as previous described. To facilitate the analysis, we compared all 33 control datasets and chose one dataset of comparable size (ENCFF001SAK) to be used as the common input data for MACS. Computation of promoter binding strength: For each factor, we obtained the raw read counts in the promoter region (TSS ± 2kb) of each gene using the peak-filtered reads, and then we calculated the normalized read counts based on the following formula: Normalized Read Count = (Raw Read Count/Library Size of Peak-filtered Reads)* Scaling Factor to generate a genome wide binding matrix. To identify the factors that show a similar binding pattern, first, we identified the 4000 most variable genes, obtained the similarity matrix of all the

factors using Spearman's rank correlation, and then hierarchically clustered them based on their *Spearman's* rank correlation coefficient. To correlate the chromatin status of K562 cells with gene expression, we analyzed the corresponding K562 RNA-seq data (accession number: ENCSR000EYO, 2 replicates, 2x75bp paired end libraries). Briefly, Tophat (v2.0.12) was used to align the raw reads to hg19 genome assembly followed by quantitation and normalization of genes using Cuffnorm from Cufflinks (v2.2.1).

Ingenuity Pathway (IPA) and Gene Set Enrichment (GSEA) analyses. IPA Software (www.ingenuity.com) is a knowledge-based approach for interpreting genome-wide expression profiles through a curated database (Ingenuity Knowledge Database, IKD). The experimental data sets were used to query the IPA and to compose a set of interactive networks taking into consideration canonical pathways, relevant biological interactions, and cellular and disease processes. The significance of the association between the experimental data sets and the IKD concepts was measured in two ways: (a) a ratio of the number of molecules from the data set that map to a particular IKD concept, and (b) a Fisher's exact test was used to estimate the statistical significance of overlap between the two gene sets. To predict alterations in the activity of transcription factors that can explain the observed gene expression changes caused by ablation or overexpression of KDM2B we considered transcription factors with *z*-scores >2 and a *p* overlap value less than 0.01. The regulation z-score represents a statistical metric of the direction of the transcriptional changes (>1.5 fold) caused by KDM2B compared to the changes regulated by the transcription factor of interest in the IKD. The positive or negative sign of *z*-score indicates activation or repression, respectively, of a particular transcription factor. Similarly, GSEA software (http://www.broadinstitute.org/gsea/index.jspis) was used to determine whether an *a priori* defined set of genes in the Molecular Signatures Database (MSigDB) shows statistically significant, concordant differences with changes caused by Kdm2b ablation. Analysis took place with default parameters ("scoring_scheme: weighted", "metric: Signal2Noise", and "nperm:1000").

Immunohistochemistry (IHC). 4 μM sections were deparaffinized in two changes of xylenes, rehydrated sequentially in ethanol, washed in IHC wash buffer [0.5% Triton X-100 in PBS], and rinsed in water. Slides were submerged in buffer A [10 mM sodium citrate pH 6.0], and antigen unmasking was

performed in the Retriever 2100 (Aptum) according to manufacturer's protocol. The slides were allowed to cool, washed three times, incubated with 3% H_2O_2 to block endogenous peroxidase activity, washed three times, and blocked with 10% normal horse serum in IHC wash buffer for one hour. Slides were incubated with primary antibodies for KDM2B (Millipore 09-864) overnight at 4°C. The next day, slides were washed three times, and incubated with 1:500 biotinylated goat anti-rabbit secondary antibody for one hour at room temperature. Specimens were washed three times, treated with ABC reagent (Vector Labs) for 30 minutes, followed by three washes. Finally, staining was developed with the DAB (Di-aminebenzidine) substrate kit (Vector Labs), inactivated in water, counterstained with hematoxylin, dehydrated sequentially ethanol then cleared slides in two changes of xylenes. Slides were photographed with a Carl Zeiss AxionLab.A1 and AxioCam ICc5 and analyzed with the ZEN 2012 software.

Immunofluorescence staining of the AGM area. 4µm paraffin sections were processed as listed in immunohistochemistry section until blocking step. Slides were blocked [5% goat serum/0.5% Triton X-100/PBS] for 1 hour at room temperature. Primary antibodies, rabbit anti-KDM2B (Millipore 09-864) and previously validated (16, 17) rat anti-c-Kit (BD Bioscience 553352), were prepared 1:500 in blocking buffer and slides were incubated overnight at 4°C. The next day, three washes were followed by secondary antibodies AlexaFluor-647 goat anti-rabbit and AlexaFluor-488 goat anti-rat (Cell Signaling) prepared 1:500 in blocking buffer, and incubated overnight at 4°C. Slides were washed three times, and stained with DAPI for 15 minutes. Slides were washed twice, rinsed in PBS, and then rinsed in water prior to mounting with Fluoromount-G (eBioscience). Images were captured using ZEN software on a Zeiss LSM 710 confocal microscope with a Plan-Apochromat 63X/1.40 oil immersion objective and an AxioCam MRc r3.0 camera.

May-Grünwald Giemsa Stain (MGG). Blood smears were dried overnight. Cytospins were performed by pre-clearing funnels with PBS; $1x10^5$ cells were spun at 400 rpm for four minutes, and dried overnight. Slides were stained in May Grünwald solution for two minutes, transferred to a 20X dilution of freshly prepared Giemsa (Sigma) solution for twelve minutes, and washed in two consecutive rinses

of water for 30 seconds each. Slides were air dried, mounted, and imaged with a Zeiss AxionLab.A1 microscope.

Protein Extraction and Western Blotting. Cells were washed in ice-cold PBS and solubilized in lysis buffer [50 mM Tris (pH 7.5), 200 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.5% Na.Deoxycholate, 1 mM EDTA, and supplemented with a mixture of protease and phosphatase inhibitors (Roche)]. The lysates were briefly sonicated and centrifuged for 20 minutes at 14,000xg. The supernatant (soluble whole-cell lysate) was analyzed by Western blotting.

Cell Culture. Human leukemia cell lines were either bought from ATCC (http://atcc.org/) or were a generous gift from Drs. R. Hawley and I. Riz (George Washington University, Washington DC). All cell lines were cultured in IMEM supplemented with 10% (vol/vol) fetal bovine serum, and 1% penicillin and streptomycin. To knockdown KDM2B, we used two previously validated hairpins (TRCN0000118440 and TRCN0000118437 (10)) cloned in the pLKO1 lentiviral system obtained from the RNAi Consortium (Broad Institute, http://www.broadinstitute.org/rnai/trc). The knock down efficiency was >80% as determined by Western blotting (**Figure 6C**) and qRT-PCR (**Figure S5D**) in a BioRad CFX thermal cycler. For cell proliferation assays, $5x10^4$ cells were plated in duplicate in twelve-well plates and passaged at a 1:2 or 1:3 ratio. Cells were counted with a TC20 Cell Counter (BioRad).

SUPPLEMENTAL REFERENCES

1. Bardeesy, N., Sinha, M., Hezel, A.F., Signoretti, S., Hathaway, N.A., Sharpless, N.E., Loda, M., Carrasco, D.R., and DePinho, R.A. 2002. Loss of the Lkb1 tumour suppressor provokes intestinal polyposis but resistance to transformation. *Nature* 419:162-167.

2. Rodriguez, C.I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A.F., and Dymecki, S.M. 2000. High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. *Nat Genet* 25:139-140.

3. Mao, X., Fujiwara, Y., and Orkin, S.H. 1999. Improved reporter strain for monitoring Cre recombinase-mediated DNA excisions in mice. *Proc Natl Acad Sci U S A* 96:5037-5042.

4. Stadtfeld, M., and Graf, T. 2005. Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development* 132:203-213.

5. Koni, P.A., Joshi, S.K., Temann, U.A., Olson, D., Burkly, L., and Flavell, R.A. 2001. Conditional vascular cell adhesion molecule 1 deletion in mice: impaired lymphocyte migration to bone marrow. *J Exp Med* 193:741-754.

6. Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. 1995. Inducible gene targeting in mice. *Science* 269:1427-1429.

7. Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. 2006. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* 44:23-28.

8. Belteki, G., Haigh, J., Kabacs, N., Haigh, K., Sison, K., Costantini, F., Whitsett, J., Quaggin, S.E., and Nagy, A. 2005. Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction. *Nucleic Acids Res* 33:e51.

9. Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D.A. 2001. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev* 15:3243-3248.

10. Tzatsos, A., Paskaleva, P., Ferrari, F., Deshpande, V., Stoykova, S., Contino, G., Wong, K.K., Lan, F., Trojer, P., Park, P.J., et al. 2013. KDM2B promotes pancreatic cancer via Polycomb-dependent and -independent transcriptional programs. *J Clin Invest* 123:727-739.

11. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.

12. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9:R137.

13. Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 42:W187-191.

14. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 28:495-501.

15. ENCODE. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57-74.

16. Yokomizo, T., Yamada-Inagawa, T., Yzaguirre, A.D., Chen, M.J., Speck, N.A., and Dzierzak, E. 2012. Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nat Protoc* 7:421-431.

17. Boisset, J.C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C. 2010. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* 464:116-120.

Figure S1

Andricovich et al

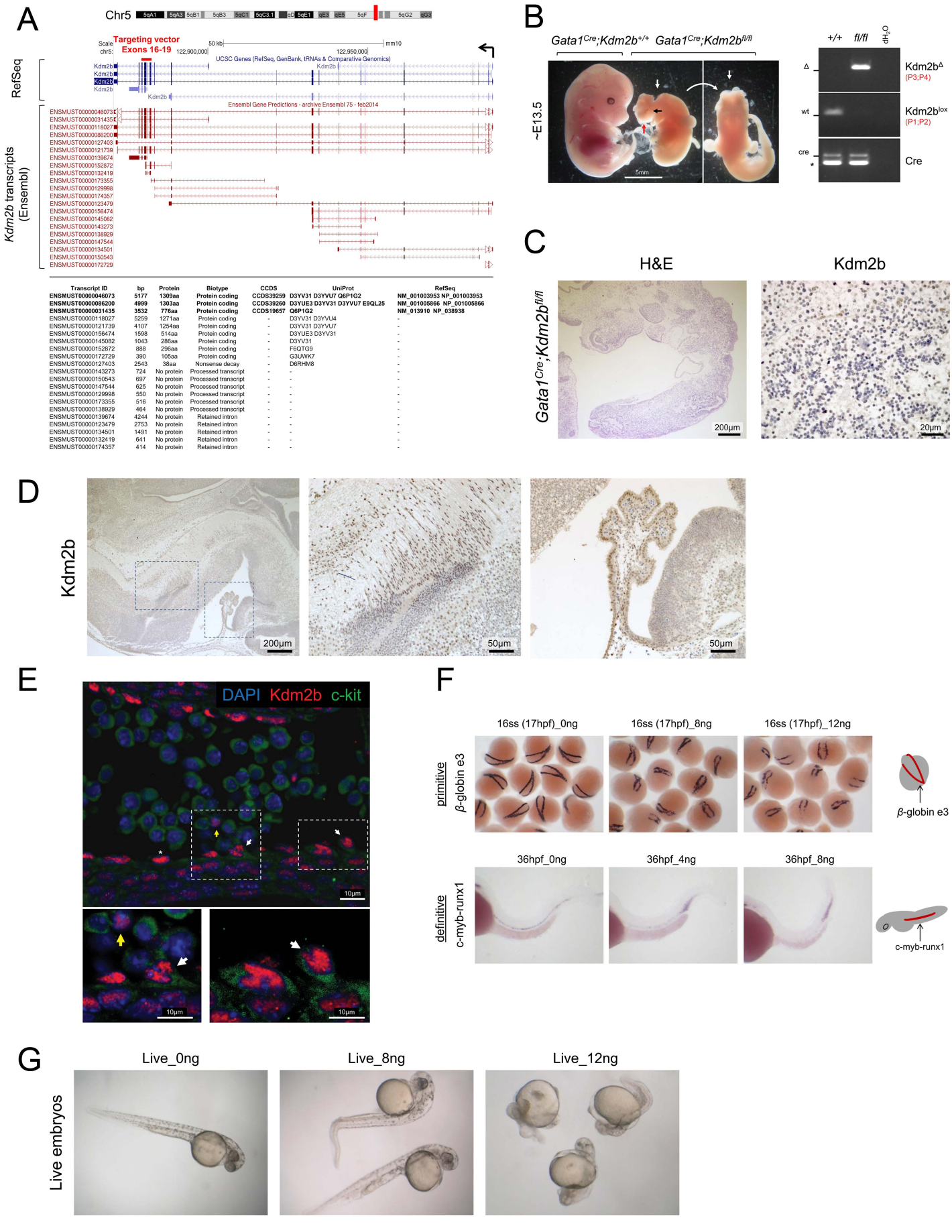


Figure S1

(**A**) The *Kdm2b* genomic locus is at chromosome five and encodes two long isoforms that differ in the use of exon1α, and a short isoform that lacks the Jumonji domain (*RefSeq*). *In silico* analysis through *Ensembl* (http://useast.ensembl.org/index.html) also revealed multiple non-coding transcripts. Bottom: Table with the complete list of identified transcripts and their features. Red line indicates the position of exons 16-19 that were targeted for Cre-mediated recombination.

(**B**) Gross images of wild type and *Kdm2b* null embryos at E13.5. The white arrow shows an open neural tube and excencephaly. Right: PCR with genomic DNA from these embryos confirmed the recombination of the $Kdm2b^{\Delta}$ allele. Scale bar 5mm.

(**C**) H&E and IHC staining for Kdm2b in E10.5 $Gata1^{Cre};Kdm2b^{fl/fl}$ embryos. Note the absence of organogenesis in H&E and the absence of Kdm2b positive cells in knock out embryos.

(**D**) IHC staining for Kdm2b in E13.5 wild type embryos showing broad expression in the developing brain. Insets show Kdm2b positive cells in the medullary hindbrain (middle) and in the developing choroid plexus of the fourth ventricle (right).

(**E**) Immunofluorescent staining of aortas from E10.5 embryos for Kdm2b (red; nuclear) and c-Kit (green; cell surface). White arrow shows budding hemogenic endothelium, yellow arrow shows a circulating “stem like” cell, and white asterisk indicates an aortic endothelial cell. Nuclei were stained with DAPI (blue).

(**F**) Meta-analysis of a previously published RNAi screen in zebrafish (http://zfrhmaps.tch.harvard.edu/pfam/). Embryos injected with morpholino antisense oligonucleotides against *jhdm1b* (*Kdm2b* homolog) exhibited compromised primitive (reduction of embryonic beta globin 3 expression by whole mount ISH) and definitive hematopoiesis (reduction of *cmyb* and *runx1* expression) that was also accompanied by morphological changes.

(**G**) Injection of morpholino antisense oligonucleotides against *jhdm1b* in live embryos caused lethality.

Figure S2

Andricovich et al

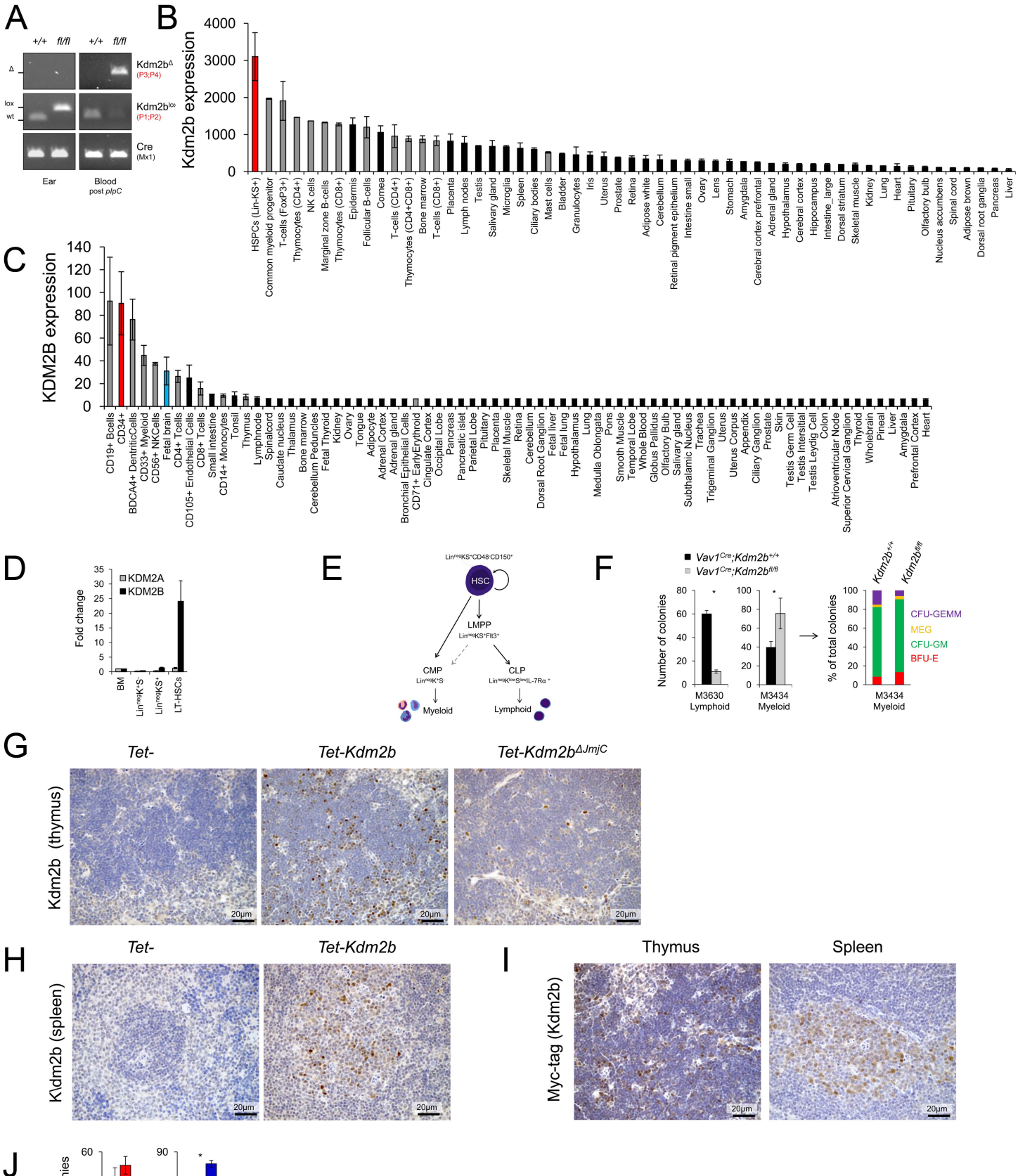


Figure S2

(**A**) PCR showing recombination of the *Kdm2b*$^{\Delta}$ allele in leukocytes, but not ear fibroblasts, of *Mx1*^{Cre};*Kdm2b*$^{fl/fl}$ mice post *pIpC* administration.

(**B-C**) Expression levels of Kdm2b in (**B**) murine (GSE10246; GNF Mouse GeneAtlas V3) and (**C**) human tissues (Human gene atlas, GSE1133) analyzed through the http://www.biogps.org portal. Red colored bars denote HSPCs, gray bars hematopoietic derived cells, black bars non-hematopoietic tissues, and the blue bar fetal human brain.

(**D**) qRT-PCR showing the relative expression of Kdm2b and Kdm2a in the indicated hematopoietic populations sorted by FACS.

(**E**) Schematic of the hematopoietic tree showing distinguishing surface markers of progenitor cells.

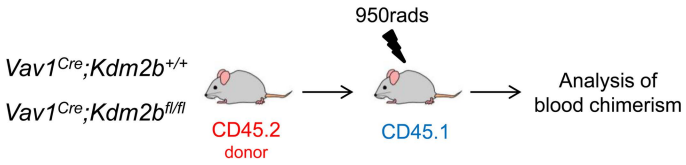
(**F**) Wild-type and *Kdm2b* null BM plated in Methylcellulose M3630 ($5x10^4$ cells) and M3434 ($2x10^4$ cells) to sustain growth of pre-B lymphoid and myeloid progenitors, respectively. The bar graph shows the cumulative number of colonies counted after ten days. *, $p<0.05$. Right: Stacked bar graph shows the percent distribution of several myeloid derived colonies (CFU, colony forming unit; GEMM, Granulocyte Erythroid Megakaryocyte Macrophage; GM, Granulocyte-Macrophage; BFU-E, Blast Forming Unit-Erythroid).

(**G-H**) IHC staining for Kdm2b showing mosaic expression in murine (**G**) thymus and (**H**) spleen of the indicated genotypes post doxycycline administration.

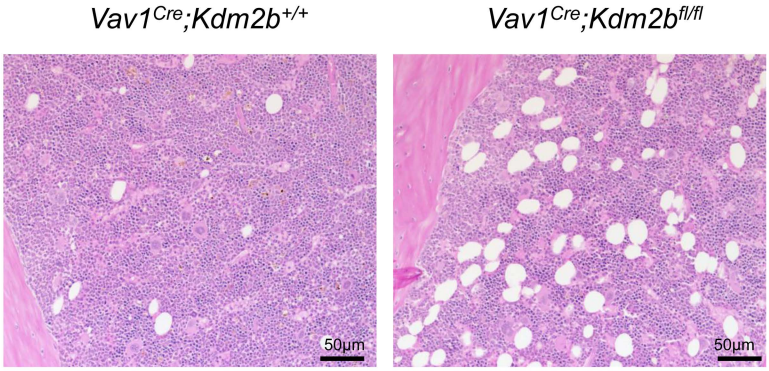
(**I**) The same tissues were also analyzed with an anti-Myc antibody to confirm expression of the exogenous protein.

(**J**) Methylcellulose colony forming assays were performed as in (**F**) in the presence of doxycycline. The bar graph shows the cumulative number of colonies counted after ten days. *, $p<0.05$.

A


950rads
Vav1Cre;Kdm2b+/+
Vav1Cre;Kdm2bfl/fl
CD45.2
donor
CD45.1
Analysis of
blood chimerism

B


Vav1Cre;Kdm2b+/+
Vav1Cre;Kdm2bfl/fl
50μm
50μm

C

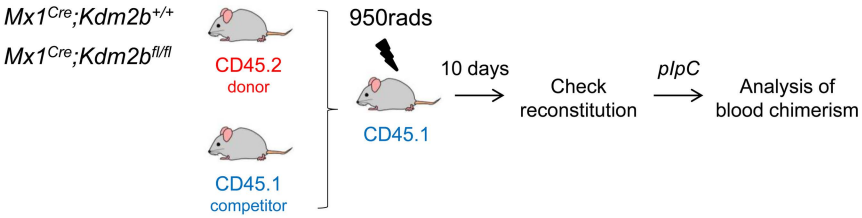

Mx1Cre;Kdm2b+/+
Mx1Cre;Kdm2bfl/fl
CD45.2
donor
CD45.1
competitor
950rads
CD45.1
10 days
Check
reconstitution
pIpC
Analysis of
blood chimerism

Figure S3

(A) Schematic of the non-competitive transplantation.

(**B**) H&E staining of femoral sections showing hypocellularity in mice transplanted *Kdm2b*-null BM.

(**C**) Schematic of the competitive transplantation.

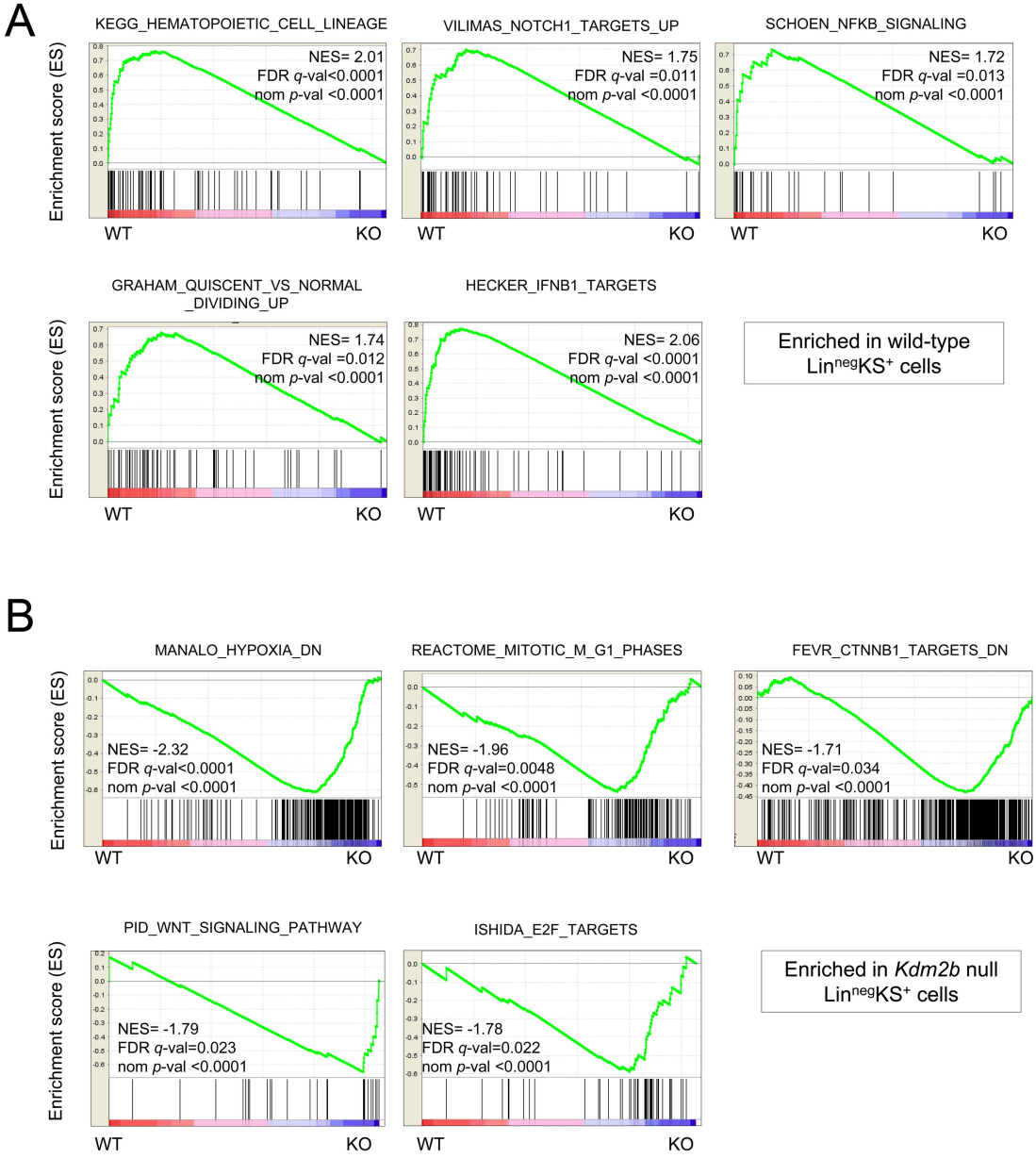

A
KEGG_HEMATOPOIETIC_CELL_LINEAGE
NES= 2.01
FDR q-val<0.0001
nom p-val <0.0001
VILIMAS_NOTCH1_TARGETS_UP
NES= 1.75
FDR q-val =0.011
nom p-val <0.0001
SCHOEN_NFKB_SIGNALING
NES= 1.72
FDR q-val =0.013
nom p-val <0.0001
GRAHAM_QUISCENT_VS_NORMAL_DIVIDING_UP
NES= 1.74
FDR q-val =0.012
nom p-val <0.0001
HECKER_IFNB1_TARGETS
NES= 2.06
FDR q-val <0.0001
nom p-val <0.0001
Enrichment score (ES)
WT
KO
Enriched in wild-type Lin^negKS^+ cells
B
MANALO_HYPOXIA_DN
NES= -2.32
FDR q-val<0.0001
nom p-val <0.0001
REACTOME_MITOTIC_M_G1_PHASES
NES= -1.96
FDR q-val=0.0048
nom p-val <0.0001
FEVR_CTNNB1_TARGETS_DN
NES= -1.71
FDR q-val=0.034
nom p-val <0.0001
PID_WNT_SIGNALING_PATHWAY
NES= -1.79
FDR q-val=0.023
nom p-val <0.0001
ISHIDA_E2F_TARGETS
NES= -1.78
FDR q-val=0.022
nom p-val <0.0001
Enrichment score (ES)
WT
KO
Enriched in Kdm2b null Lin^negKS^+ cells

Figure S4

(**A-B**) Enrichment plots of the indicated gene signatures. *NES*, normalized enrichment score.

Figure S5

Andricovich et al

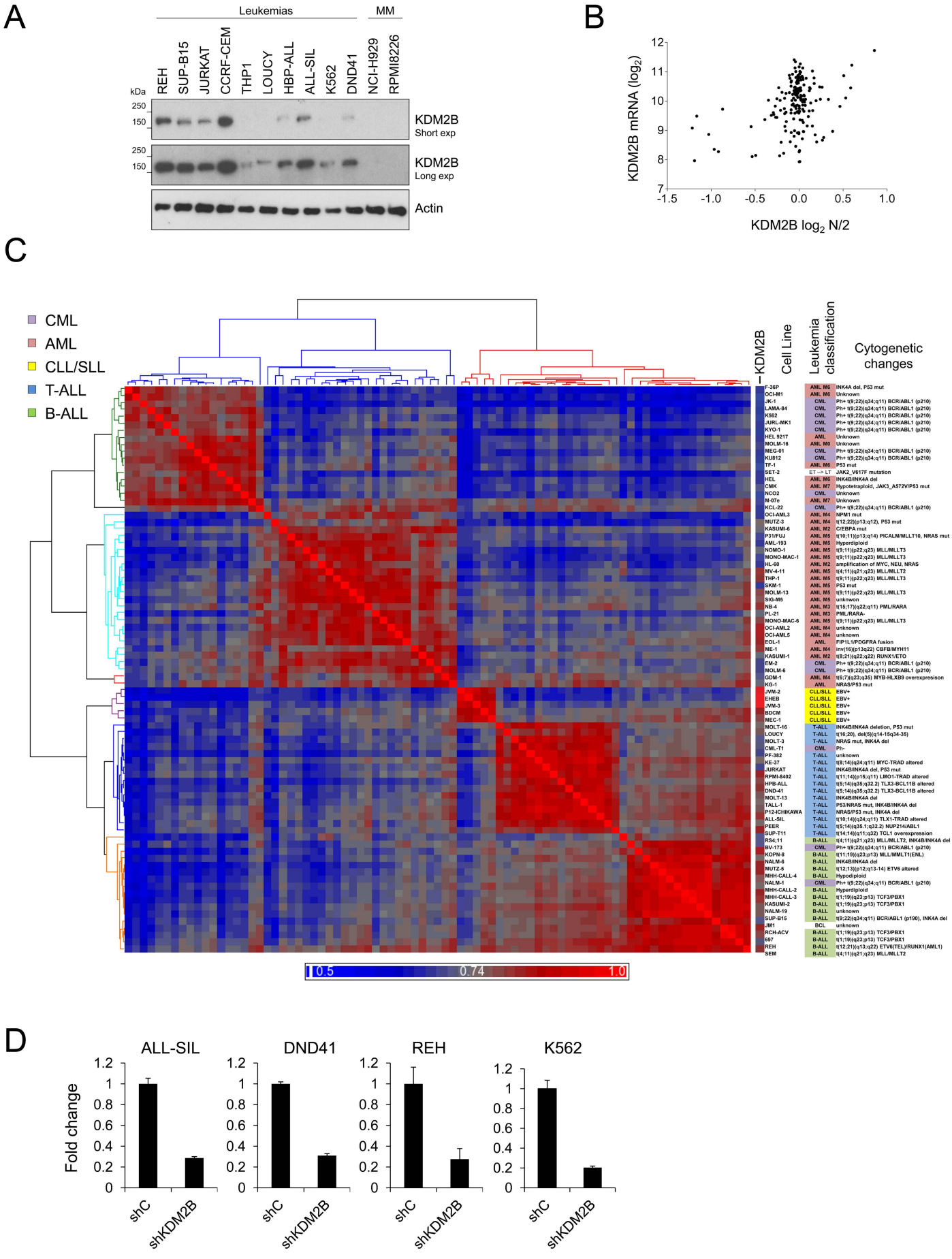


Figure S5

(**A**) Western blot of whole cell lysates showing the levels of KDM2B in the indicated cell lines.

(**B**) Composite plot showing KDM2B expression (y axis) versus copy number changes (x axis) in 159 human cell lines established from hematopoietic malignancies (raw data obtained from the CCLE).

(**C**) Heat map showing the pairwise comparison (*Spearman's* rank) and hierarchical clustering of 81 human leukemia RNA-seq transcriptomes. KDM2B expression, cell line name, leukemia subtype, and cytogenetic abnormalities are shown on the right (information obtained from Drexler HG: Guide to Leukemia-Lymphoma Cell Lines, 2nd Edition. Braunschweig, 2010). The color bar shows the *Spearman's* rank correlation coefficient. BCL, B Cell Lymphoma unspecified; ET, Essential Thrombocythemia; CLL/SIL Chronic Lymphocytic Leukemia-Small Lymphocytic Lymphoma; ALL, Acute Lymphoblastic Leukemia; AML, Acute Myeloid Leukemia; CML, Chronic Myeloid Leukemia; EBV, Epstein-Barr Virus.

(**D**) qRT-PCR shows downregulation of KDM2B transcript in cells treated with shKDM2B lentiviruses.

Figure S6

Andricovich et al

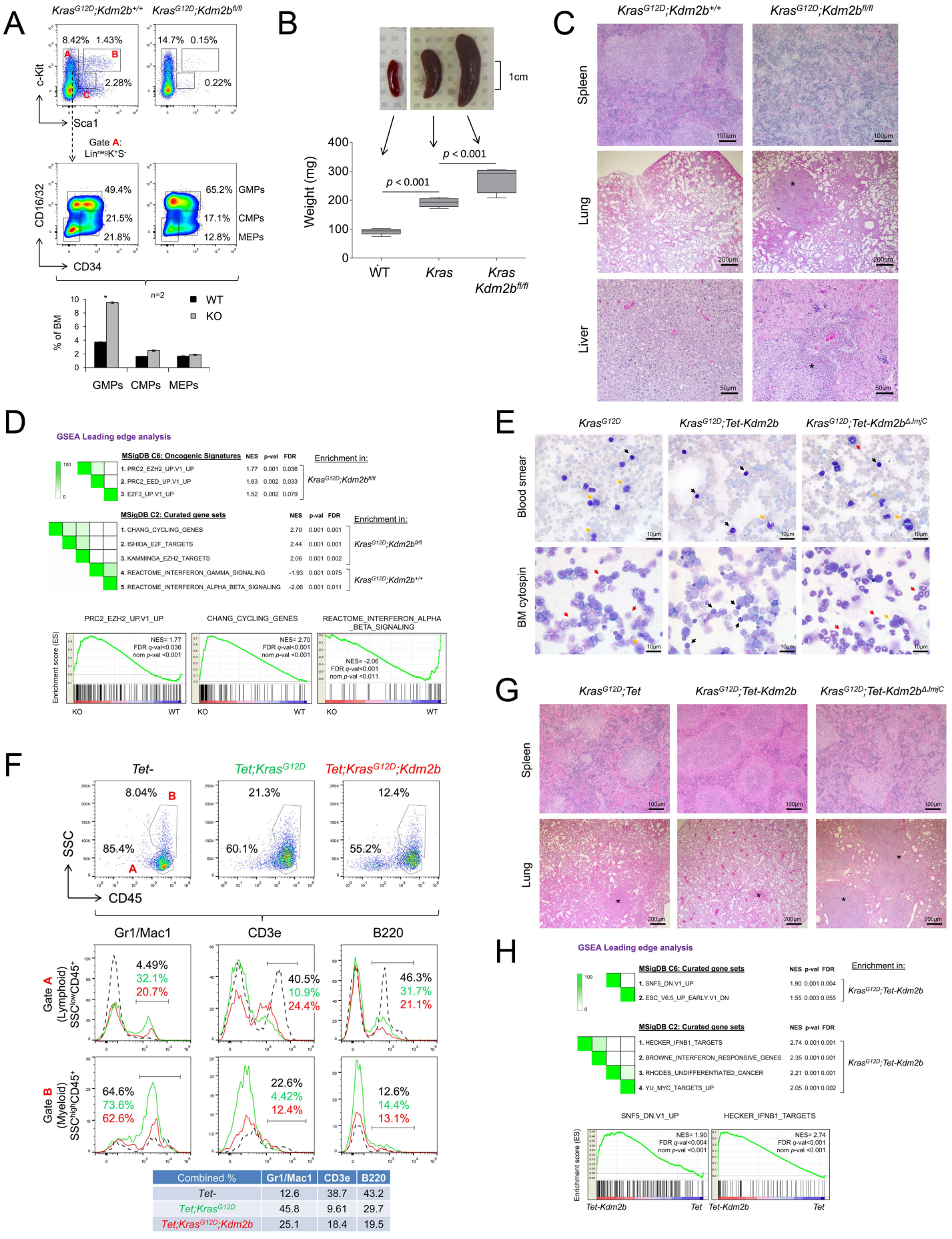


Figure S6

(**A**) Composite analysis showing the frequency (%) of $Lin^{neg}KS+$ (gate B), $Lin^{neg}K^{low}S^{low}$ (gate C), and myeloid (gate A for MEP, CMP, and GMP) progenitors. Bold red type denotes the gates and black numbers denote the percentage of gated cells. Bottom: Flow cytometry plot and bar graph show the frequency (%) of GMPs, CMPs, and MEP in BM. n, number of mice; *, $p<0.05$.

(**B**) Gross pictures of spleens isolated from the indicated geneotypes. Box-whisker graph shows the cumulative weight (mean ± SEM, $n=5$).

(**C**) H&E staining of the indicated tissues from five week old *Vav1^{Cre};Kras^{G12D};Kdm2b$^{+/+}$* and *Vav1^{Cre};Kras^{G12D};Kdm2b$^{fl/fl}$* mice. *, metastatic nodule.

(**D**) GSEA leading edge analysis of *Vav1^{Cre};Kras^{G12D};Kdm2b$^{+/+}$* and *Vav1^{Cre};Kras^{G12D};Kdm2b$^{fl/fl}$* leukemias. The color bar shows the overlap (%) between the individual genes sets. Bottom: enrichment plots of the indicated gene signatures. *NES*, normalized enrichment score.

(**E**) May-Grünwald Giemsa stained blood smears and BM cytospins of eight week old mice of the indicated genotypes. Arrows: Black, lymphocytes; red, immature granulocytes; orange, myeloblasts.

(**F**) Flow cytometric analysis for lymphoid (gate A, $SSC^{low}CD45^{+}$) and myeloid (gate B, $SSC^{high}CD45^{+}$, including blasts) cells in the peripheral blood. Bold red type denotes the gates and black numbers denote the percentage of gated cells. Histograms show the frequency (%) of myeloid ($Gr1^{+}/Mac1^{+}$), T-cells ($CD3e^{+}$), and B-cells ($B220^{+}$) for the indicated genotypes. Bottom: table shows the combined percentage of myeloid, T- and B-cells.

(**G**) H&E staining of the indicated tissues from ten week old mice of the indicated genotypes. *, metastatic nodule.

(**H**) GSEA leading edge analysis of *Vav1^{Cre};Kras^{G12D};Tet* and *Vav1^{Cre};Kras^{G12D};Tet-Kdm2b* leukemias. The color bar shows the overlap (%) between the individual genes sets. Bottom: enrichment plots of the indicated gene signatures. *NES*, normalized enrichment score.

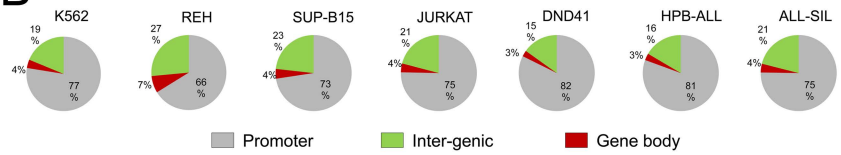
Figure S7

Andricovich et al

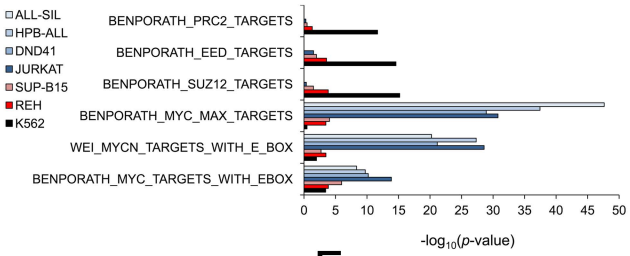
A

	K562	REH	SUP-B15	JURKAT	DND41	HPB-ALL	ALL-SIL
K562	3704	1762	1817	2553	2329	2131	2312
REH		4088	2191	2748	2702	2249	2629
SUP-B15			3847	2867	2867	2637	2808
JURKAT				6789	4639	4475	4726
DND41					6732	4596	4948
HPB-ALL						6453	4657
ALL-SIL							6908

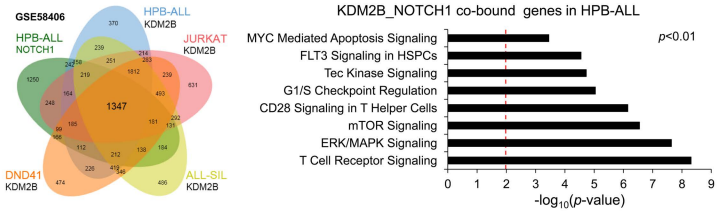
B



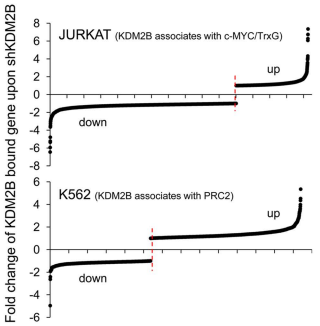
C



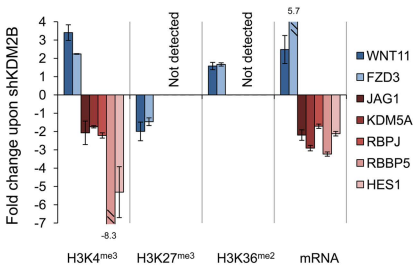
D



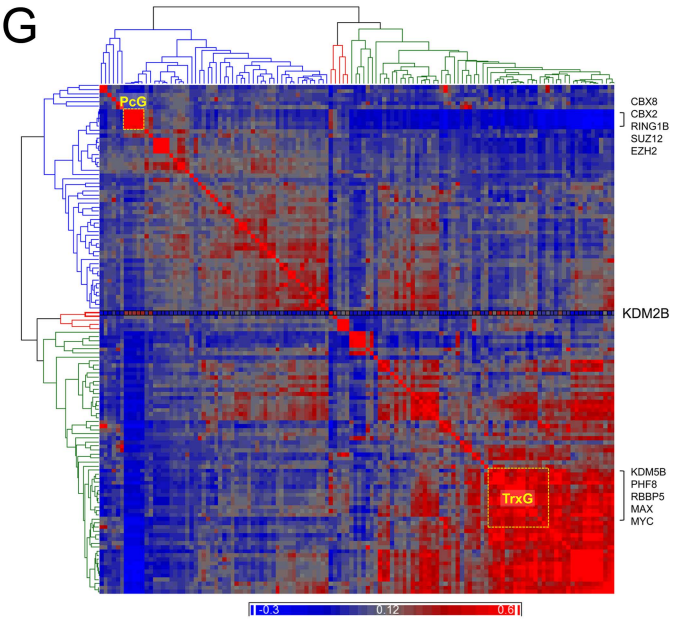
E



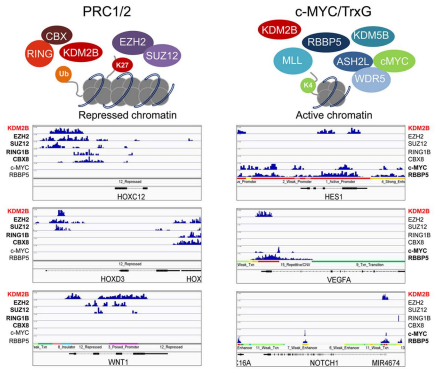
F



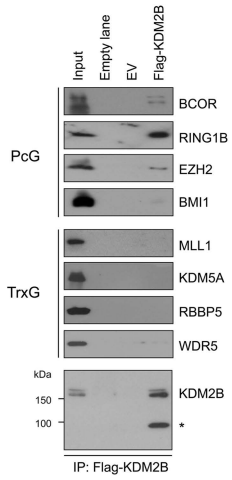
G



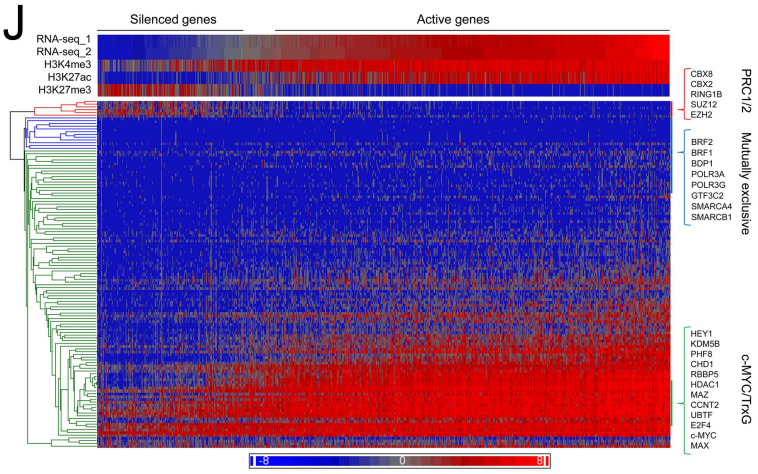
H



I



J



K

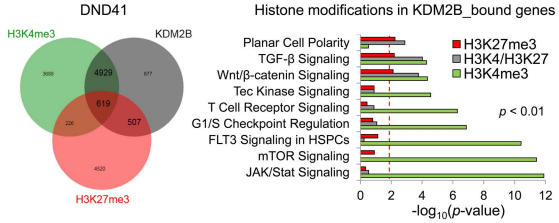


Figure S7

(**A**) Matrix shows the number of common genes bound by KDM2B in the indicated cell lines.

(**B**) Distribution (%) of KDM2B peaks in the indicated cell lines.

(**C**) GREAT (Genomic Regions Enrichment of Annotations Tool, http://bejerano.stanford.edu/great/public/html/) analysis shows the overlap of KDM2B bound genes with previously published profiles of transcription factors and epigenetic regulators contained in The Molecular Signatures Database (MSigDB, http://www.broadinstitute.org/gsea/msigdb/index.jsp).

(**D**) Venn diagram shows the overlap of KDM2B bound genes in the indicated T-ALL cell lines with NOTCH1 bound genes in HPB-ALL (GSE58406). Right: IPA of genes co-bound by KDM2B and NOTCH1 in HPB-ALL. The x axis (log scale) corresponds to the binomial raw *p* values.

(**E**) Effect of KDM2B knockdown on the expression of KDM2B bound genes. Downregulated and upregulated genes are plotted to the left and right of the red dotted line, respectively.

(**F**) ChIP shows the fold change in the enrichment of the indicated histone marks on genes bound by KDM2B in JURKAT cells treated with shKDM2B. Changes in mRNA expression induced by KDM2B knockdown (qRT-PCR) are also shown. Results are mean ± StDev.

(**G**) Heat map showing the genome-wide pairwise comparison and hierarchical clustering (average linkage) of the binding strength of 126 transcription factors and epigenetic regulators at the gene promoters (±2kb window) in K562 cells. The color bar shows the *Spearman's* rank correlation coefficient.

(**H**) Representative screen shots from the IGV (https://www.broadinstitute.org/igv/) genome browser showing co-binding of KDM2B with either PRC1/2 or c-MYC/TrxG complexes in selected targets in K562 cells.

(**I**) HEK293T cells were stably transfected with Flag-KDM2B. Nuclear lysates were immunoprecipitated with an anti-Flag antibody and analyzed by Western blotting for the indicating proteins. *, degradation product from the full length protein.

(**J**) Composite heat map showing stratification of the expression of KDM2B bound genes, the degree of histone methylation, and clustering of the binding strength of 125 factors in K562 cells. The color bar shows the mean-centered normalized read counts in promoter region.

(**K**) Venn diagram shows the presence of H3K4me3 and H3K27me3 modifications in genes bound by KDM2B in DND41 cells. Right: IPA of genes bound by KDM2B in the context of the indicated histone modifications. The x axis (log scale) corresponds to the binomial raw *p* values.