nature chemical biology

SUPPLEMENTARY INFORMATION

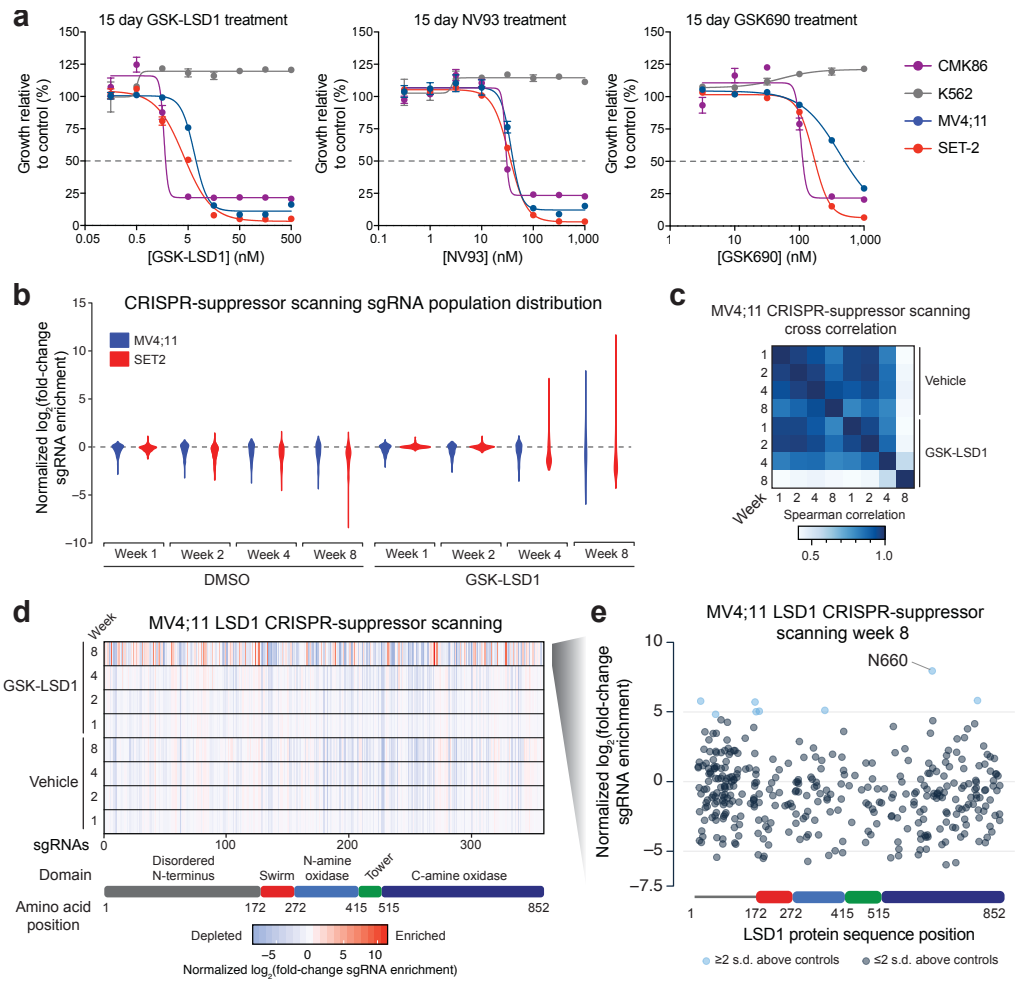
https://doi.org/10.1038/s41589-019-0263-0

In the format provided by the authors and unedited.

CRISPR-suppressor scanning reveals a nonenzymatic role of LSD1 in AML

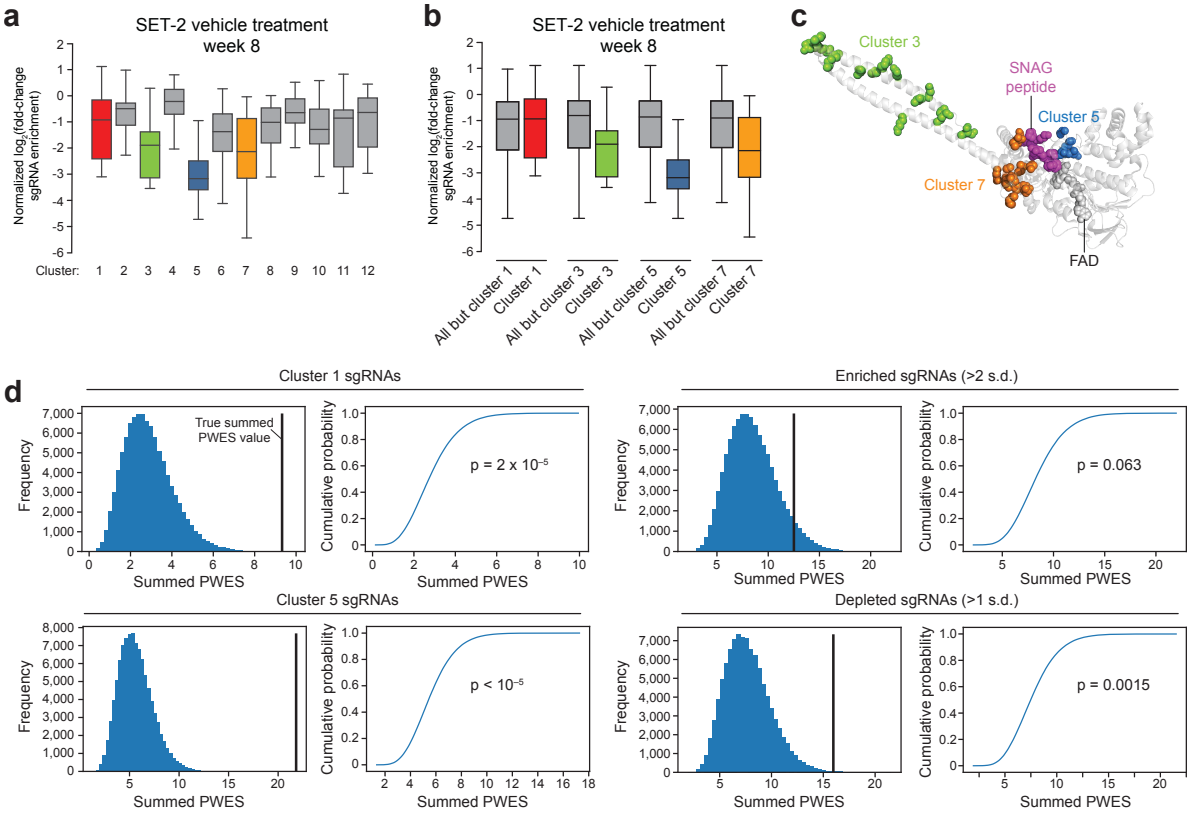
Michael E. Vinyard[1,7], Cindy Su[1,7], Allison P. Siegenfeld[1,7], Amanda L. Waterbury[1,7], Allyson M. Freedy[1,7], Pallavi M. Gosavi[1,7], Yongho Park[1], Eugene E. Kwan[1], Benjamin D. Senzer[1], John G. Doench[2], Daniel E. Bauer[3,4,5], Luca Pinello[2,6] and Brian B. Liau[1,2*]

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. [2]Broad Institute of Harvard and MIT, Cambridge, MA, USA. [3]Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA. [4]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [5]Department of Pediatrics, Harvard Medical School and Harvard Stem Cell Institute, Harvard University, Boston, MA, USA. [6]Molecular Pathology Unit and Cancer Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. [7]These authors contributed equally: Michael E. Vinyard, Cindy Su, Allison P. Siegenfeld, Amanda L. Waterbury, Allyson M. Freedy, Pallavi M. Gosavi. *e-mail: liau@chemistry.harvard.edu



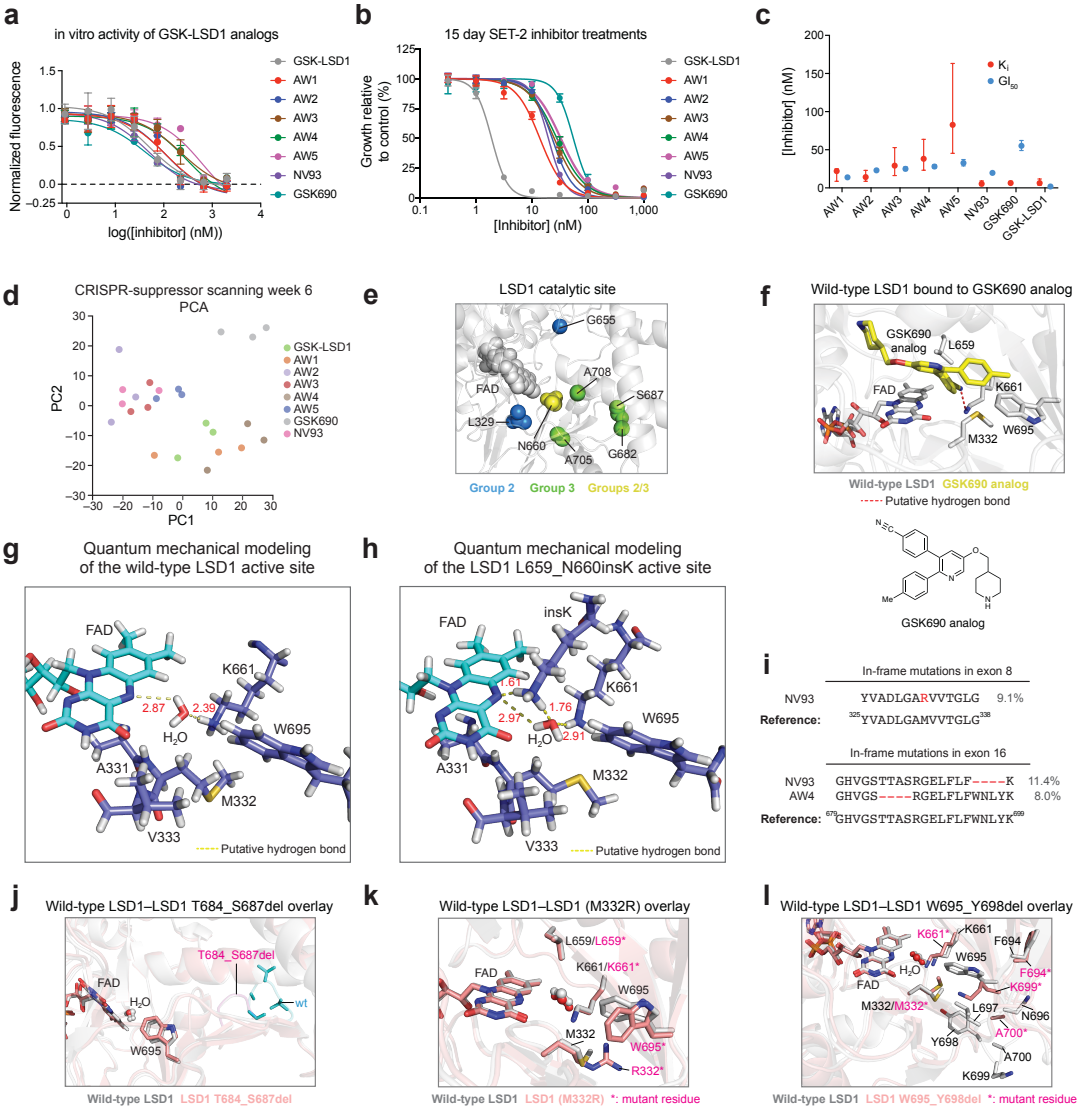
Supplementary Figure 1 | AML cell lines are sensitive to LSD1 inhibitors and LSD1 CRISPR-suppressor scanning in MV4;11

a. Dose-response curves for cell lines treated with GSK-LSD1, NV93, and GSK690 are shown. Data represent mean values ± s.e. across three replicates. One of two independent replicates is shown.
b. Violin plots showing $\log_2$(fold-change sgRNA enrichment) in SET-2 and MV4;11 versus week 0 normalized against functionally neutral genome-targeting control sgRNAs.
c. Heat map showing cross correlation of overall sgRNA enrichment at different time points during the CRISPR-suppressor scanning screen in MV4;11. Data represent mean values across three replicate transductions.
d. Heat map depicting $\log_2$(fold-change sgRNA enrichment) in MV4;11 at the time points specified versus week 0 normalized against functionally neutral genome-targeting control sgRNAs. The sgRNAs are arrayed on the *x* axis by the *LSD1* CDS. Color represents mean values across three replicate transductions.
e. Scatter plot showing sgRNA enrichment in MV4;11 for GSK-LSD1 treatment at week 8 versus week 0 normalized against functionally neutral genome-targeting control sgRNAs. Data represents mean values across three replicate transductions.



Supplementary Figure 2 | CRISPR-suppressor scanning reveals functionally important regions of LSD1 through negatively selected clusters

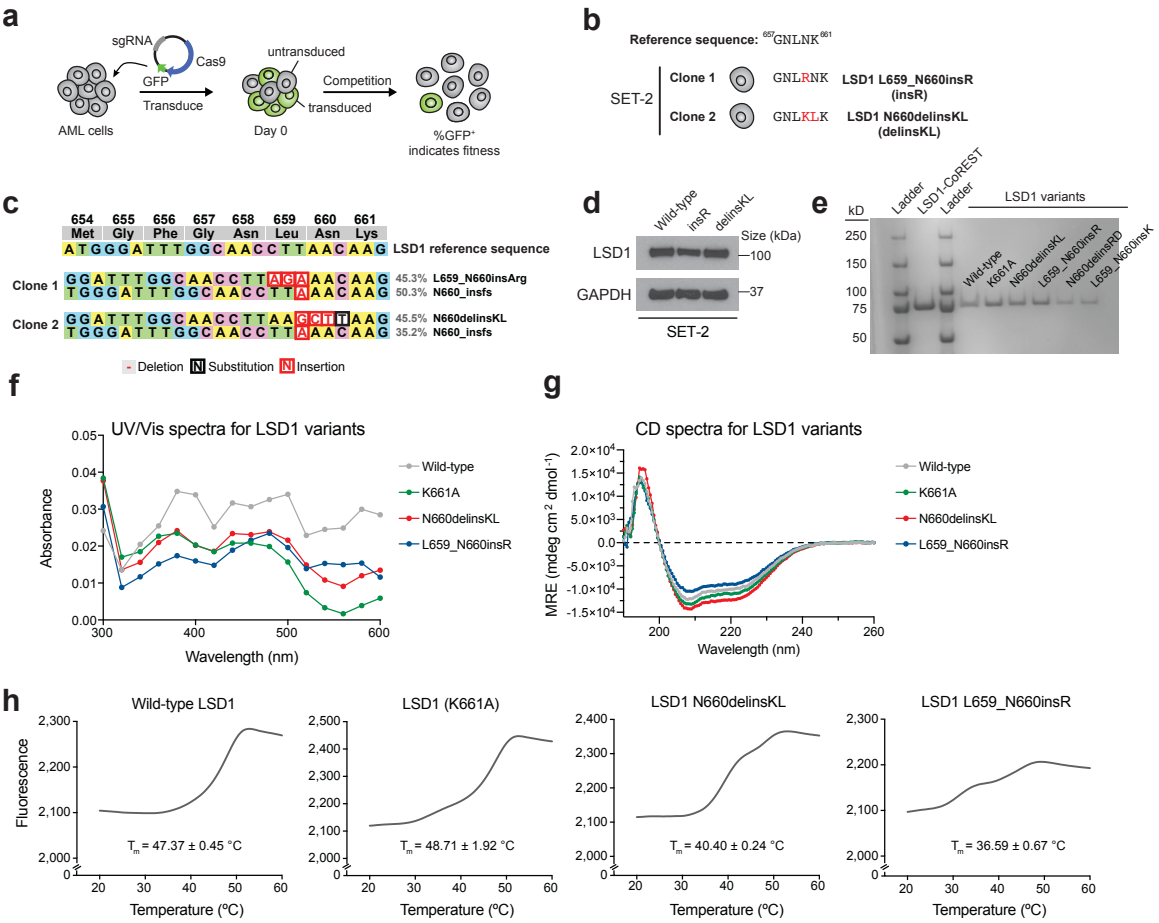
a-b. Box plot showing normalized log_2(fold-change sgRNA enrichment) at week 8 versus week 0 for clusters identified by hierarchical clustering in SET-2 under vehicle treatment. In the plot, bars represent the median, the box represents IQR, and the whiskers represent 1.5 x IQR across three replicate transductions.
c. Structural view of LSD1 depicting the location of clusters 3, 5, and 7 (PDB: 2Y48).
d. Histograms and corresponding cumulative distribution functions (CDF) showing the simulated summed PWES values for the following categories of sgRNAs: cluster 1, cluster 5, sgRNAs with enrichment values >2 s.d. above the mean, and sgRNAs with enrichment values >1 s.d. below the mean. The empirical *P* value, determined to be 1 – CDF(True summed PWES value), is displayed on the graph.



Supplementary Figure 3 | Characterization of CRISPR-suppressor scanning

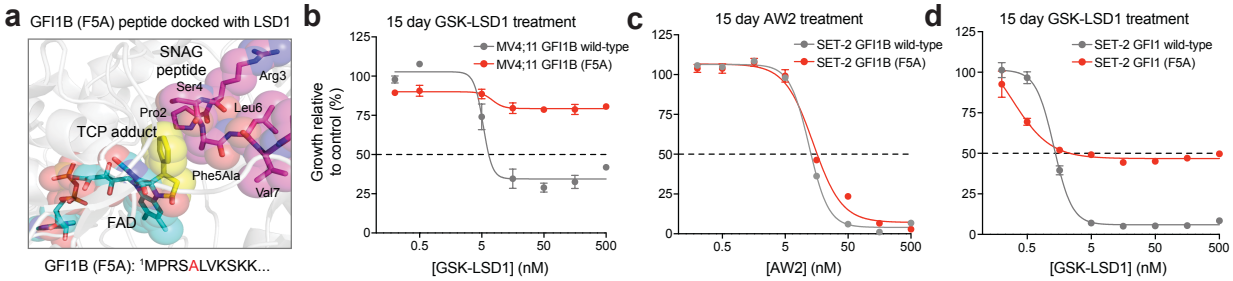
a. Dose-response curves for in vitro LSD1 inhibition by GSK-LSD1, GSK690, NV93, and AW1-AW5 are shown. Data represent mean values ± s.e. across three technical replicates. One of two independent replicates is shown.
b. Dose-response curves for SET-2 cell lines treated with GSK-LSD1, GSK690, NV93, and AW1-AW5 are shown. Data represent mean values ± s.e. across three technical replicates. One of two independent replicates is shown.
c. Scatter plot showing the K_i and GI_{50} values for each inhibitor in **Supplementary Fig. 3a–3b**. Data represents mean values ± transformed s.e. of log(K_i) and log(GI_{50}) respectively. One of two independent replicates is shown.
d. Scatter plot showing projection of compound sgRNA enrichment profiles at week 6 onto PC space where every replicate transduction (n = 3) is considered separately.
e. Structural view depicting the LSD1 active site where residues targeted by sgRNAs enriched by compounds in Group 2 and Group 3. sgN660 was enriched by all compounds.
f. Structural view depicting the wild-type LSD1 active site bound to a GSK690 analog, which is shown due to the lack of structural data of GSK690 or NV93 bound to LSD1. The hydrogen bond between K661, a key residue involved in the catalytic activity of LSD1, and the 4-cyanophenyl ring of the GSK690 analog would be disrupted upon mutation to K661E (PDB: 5YJB).
g-h. Semi-empirical quantum mechanical modeling of wild-type LSD1 and LSD1 L659_N660insK active sites are shown. The insK, through competing interactions, disrupts the water-mediated hydrogen bonding network between the N5 atom of FAD and K661 in LSD1 L659_N660insK.
i. Schematic showing the genotypes of the most abundant mutations identified in (top) exon 8 for NV93 treatment and (bottom) exon 16 for NV93 and AW4 treatment at week 6. Percentages indicate the allele frequency.
j. Homology model of LSD1 T684_S687del overlaid with wild-type LSD1 is shown to highlight the deletion of the wild-type loop compared to the truncated mutant loop (PDB: 2HKO).

k. Homology model of LSD1 (M332R) overlaid with wild-type LSD1 suggests that M332R may remove a potential hydrophobic interaction between M332 and LSD1 inhibitors (PDB: 2HKO).
l. Homology model of LSD1 W695_Y698del overlaid with wild-type LSD1 is shown. This deletion removes W695, a key residue in the hydrophobic substrate-binding cavity involved in GSK690 analog binding (PDB: 2HKO).



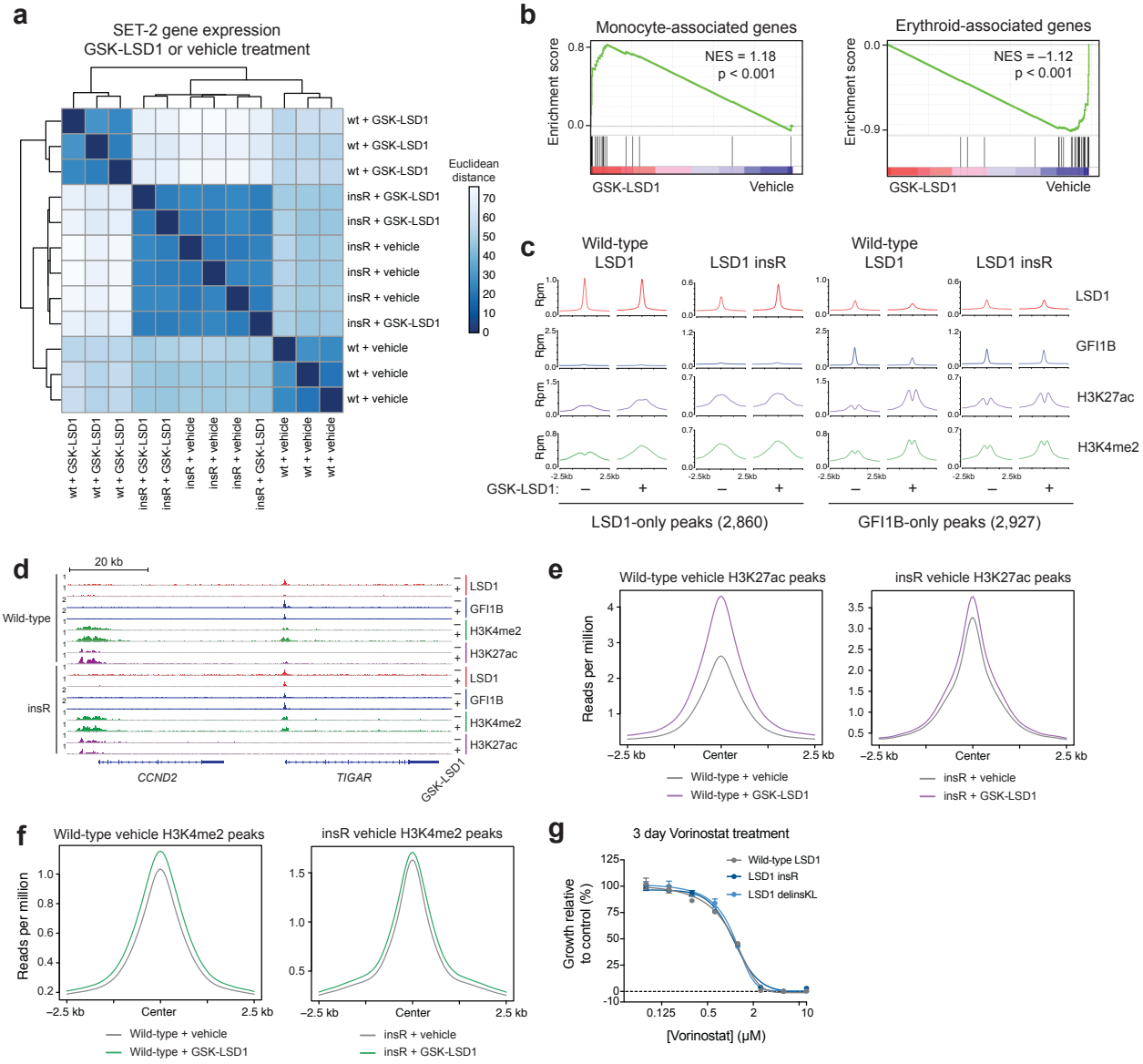
Supplementary Figure 4 | Characterization of wild-type and mutant LSD1 variants

a. Illustration depicting the workflow for sgRNA validation through a competitive growth assay monitored by GFP fluorescence signal.
b. Schematic indicating the genotypes of two single-cell derived, clonally-expanded SET-2 cell lines containing drug-resistant LSD1 mutations.
c. Schematic depicting the coding mutations identified in the SET-2 drug-resistant cell lines.
d. Immunoblots showing that LSD1 mutants are expressed in drug-resistant cell lines. See **Supplementary Fig. 7a** for uncropped blot.
e. SDS-PAGE confirming successful protein expression at the expected molecular weight of wild-type LSD1, LSD1 (K661A), and identified mutant LSD1 variants. One of two independent replicates is shown.
f. UV-vis spectra for wild-type LSD1, LSD1 (K661A), LSD1 delinsKL, and LSD1 insR are shown. Data represents a single measurement. One of two independent replicates is shown.
g. Circular dichroism spectra for wild-type LSD1, LSD1 (K661A), LSD1 delinsKL, and LSD1 insR are shown. Data represents mean value across five individual scans. Error bars are omitted for clarity. Experiment performed once.
h. Thermal stability curves showing relative fluorescence of FAD (y axis) in wild-type LSD1, LSD1 (K661A), LSD1 delinsKL, and LSD1 insR as a function of temperature (x axis). Data represent mean values across five replicates. Error bars are omitted for clarity. Calculated FAD release temperatures (T_m) are shown, where error represents s.e. across five replicates. One of two independent replicates is shown.



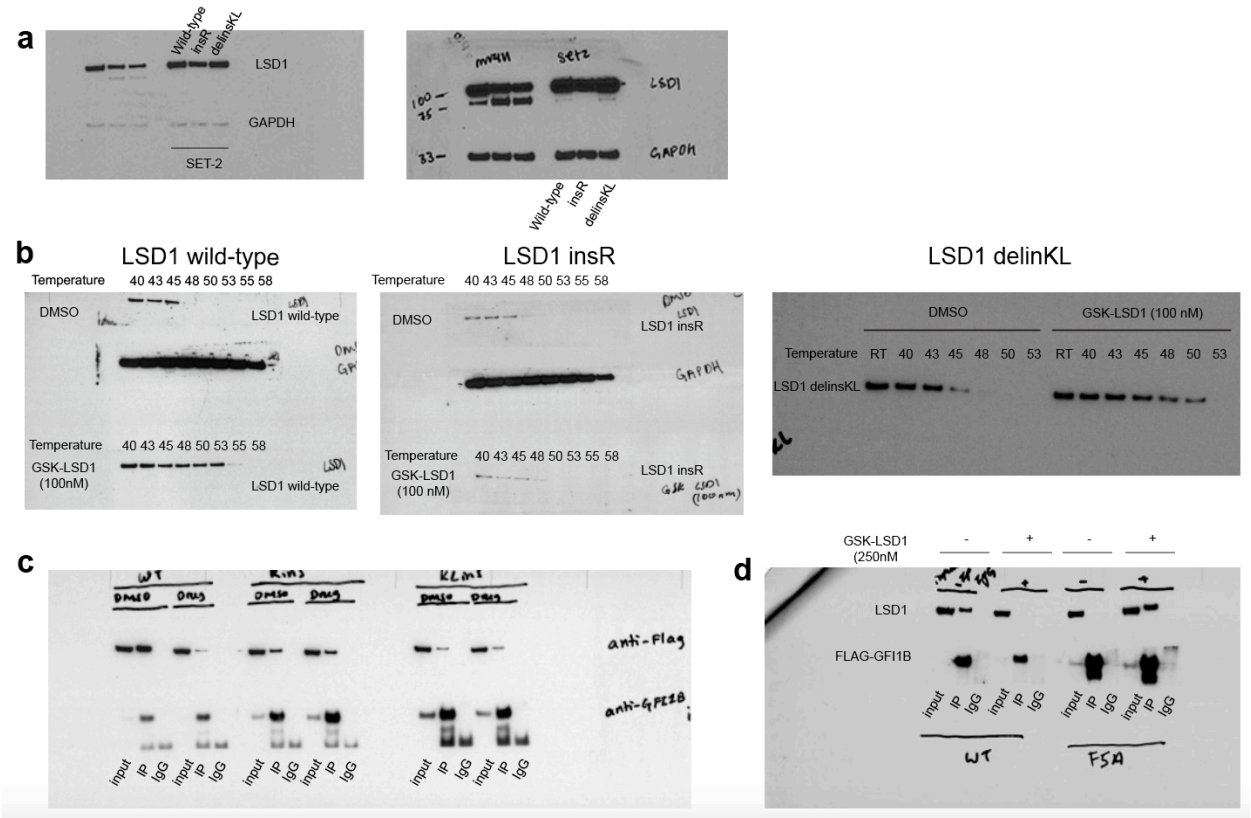
Supplementary Figure 5 | GFI1/GFI1B (F5A) is a GSK-LSD1 orthogonal drug-complementary allele

a. Structural view of the LSD1 catalytic site labeled by the GSK-LSD1 adduct modeled with a SNAG (F5A) peptide.
b. Dose-response curves for MV4;11 cells expressing wild-type GFI1B-FLAG or GFI1B-FLAG (F5A) treated with GSK-LSD1 are shown. Data represent mean values ± s.e. across three technical replicates. One of two independent replicates is shown.
c. Dose-response curves for SET-2 cells expressing wild-type GFI1B-FLAG or GFI1B-FLAG (F5A) treated with AW2 are shown. Data represent mean values ± s.e. across three technical replicates. One of two independent replicates is shown.
d. Dose-response curves for SET-2 cells expressing wild-type GFI1-FLAG or GFI1-FLAG (F5A) treated with GSK-LSD1 are shown. Data represent mean values ± s.e. across three technical replicates. One of two independent replicates is shown.



Supplementary Figure 6 | Gene expression analysis and chromatin profiling of wild-type SET-2 and SET-2 LSD1 insR in the presence of GSK-LSD1 or vehicle treatment

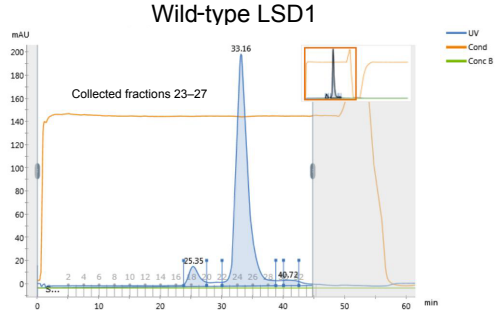
a. Heat map showing Euclidean distance between log-transformed gene expression vectors for wild-type (wt) SET-2 and SET-2 LSD1 insR cells treated with vehicle or GSK-LSD1 (250 nM, 48 h). Samples are grouped by hierarchical clustering with all three cell culture replicates depicted separately.
b. Gene set enrichment analysis (GSEA) showing enrichment of genes expressed in monocytes in genes up-regulated by GSK-LSD1, and enrichment of genes expressed in erythroid cells in genes down-regulated by GSK-LSD1. Normalized Enrichment Scores (NES) and nominal p-value are depicted and calculated as previously described.
c. ChIP-seq composite plots showing average signal (*y* axis, rpm) for LSD1, GFI1B, H3K27ac, and H3K4me2 ChIP-seq in wild-type SET-2 and SET-2 LSD1 insR cells treated with vehicle or GSK-LSD1 (250 nM, 48 h) centered around LSD1-only and GFI1B-only peaks identified in wild-type SET-2. The *x* axis shows flanking regions of ±2.5 kb around the peak center. Experiment performed once.
d. ChIP-seq composite plots showing average peak signal (*y* axis, rpm) for LSD1, GFI1B, H3K4me2, and H3K27ac in wild-type SET-2 and SET-2 LSD1 insR cells under vehicle or GSK-LSD1 treatment (250 nM, 48 h). Experiment performed once.
e-f. ChIP-seq profile plots showing average peak signal (*y* axis, rpm) for H3K27ac (**e**) and H3K4me2 (**f**) in wild-type SET-2 and SET-2 LSD1 insR under vehicle or GSK-LSD1 treatment (250 nM, 48 h). The *x* axis depicts flanking regions of ±2.5 kb around the peak center.
g. Dose-response curves for cell lines treated with the pan-HDAC inhibitor, vorinostat, are shown. Data represent mean values ± s.e. across three replicates. One of two independent replicates is shown.



Supplementary Figure 7 | Uncropped immunoblot images

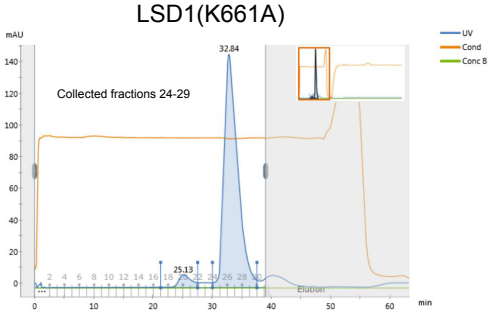
a. Uncropped immunoblot detecting LSD1 expression levels in SET-2 wild-type, insR, and delinKL cells with short exposure (left) and long exposure (right). One of two independent replicates is shown.
b. Uncropped CETSA immunoblot displaying LSD1 levels at the indicated temperatures in SET-2 wild-type LSD1 (left), SET-2 LSD1 insR (middle), and SET-2 LSD1 delinsKL (right) in vehicle or GSK-LSD1 (100 nM, 1 h). One of two independent replicates is shown.
c. Uncropped Co-IP immunoblot of wild-type FLAG-LSD1 and FLAG-LSD1 mutant variants with GFI1B after anti-GFI1B IP in transiently transfected HEK 293T cells after vehicle or GSK-LSD1 treatment (250 nM, 48 h). One of two independent replicates is shown.
d. Uncropped Co-IP immunoblot of wild-type GFI1B-FLAG, GFI1B-FLAG (F5A), and LSD1 after anti-FLAG IP in SET-2 overexpressing wild-type GFI1B-FLAG or GFI1B-FLAG (F5A) under vehicle or GSK-LSD1 treatment (250 nM, 48 h). One of two independent replicates is shown.

Wild-type LSD1



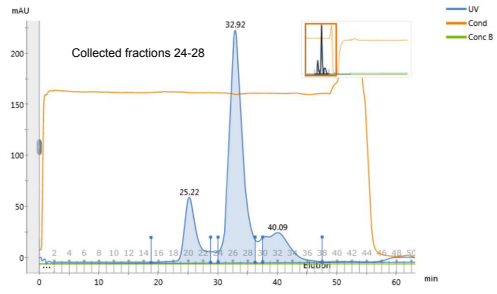
Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	25.347	32.51	6.84	14.67	17 - 19	3.753
Peak B	33.160	431.6	90.83	14.67	22 - 28	8.753
Peak C	40.723	11.04	2.32	14.67	30 - 31	2.500

LSD1(K661A)



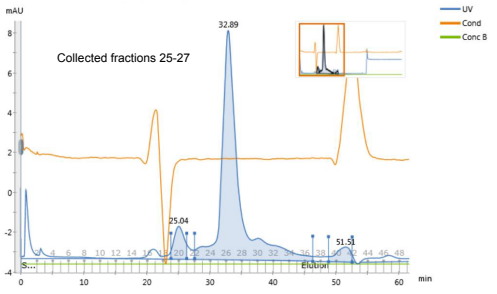
Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	25.133	20.47	5.09	14.16	17 - 21	6.247
Peak B	32.840	381.9	94.91	14.16	24 - 29	7.507

LSD1 N660delinsKL



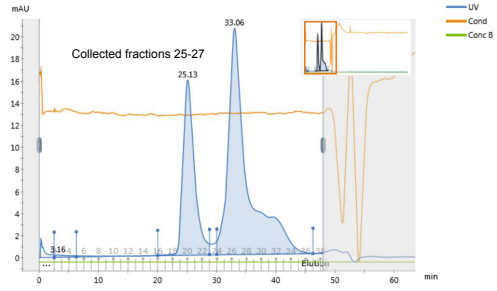
Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	25.222	154.4	17.89	14.81	15 - 22	10.007
Peak B	32.922	571.0	66.16	14.78	24 - 28	6.253
Peak C	40.088	137.7	15.95	14.78	30 - 37	10.007

LSD1 L659_N660insK



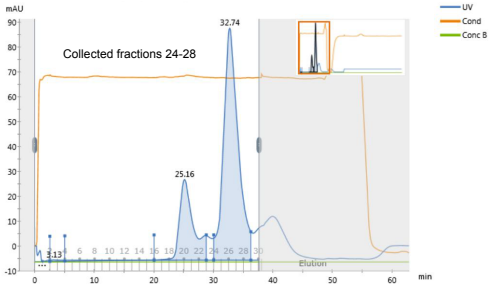
Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	25.040	2.904	7.30	14.12	19 - 20	2.503
Peak B	32.887	34.88	87.67	14.17	22 - 36	18.757
Peak C	51.510	2.002	5.03	14.48	39 - 41	3.750

LSD1 N660delinsRD



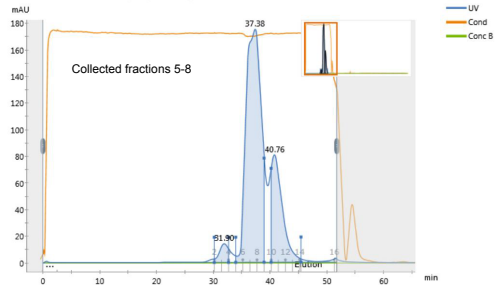
Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	3.158	0.5693	0.51	16.50	2 - 4	3.760
Peak B	25.125	34.78	31.20	16.48	16 - 22	8.753
Peak C	33.058	76.14	68.29	16.49	24 - 36	16.253

LSD1 L659_N660insR



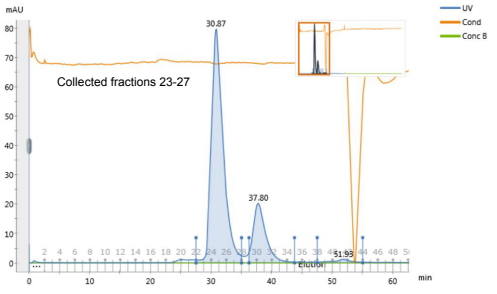
Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	3.127	1.215	0.35	14.65	2 - 3	2.507
Peak B	25.160	86.64	24.86	14.65	16 - 22	8.753
Peak C	32.740	260.7	74.79	14.62	24 - 28	6.253

CoREST1



Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	31.900	23.48	3.22	16.67	2 - 3	2.513
Peak B	37.380	527.7	72.42	16.65	5 - 8	5.000
Peak C	40.760	177.4	24.35	16.67	10 - 14	5.253

LSD1-CoREST1



Peak	Retention (min)	Area (min*mAU)	Area %	Conductivity (mS/cm)	Fraction(s)	Width (min)
Peak A	30.873	174.8	76.75	14.11	22 - 27	7.507
Peak B	37.800	48.78	21.42	14.11	29 - 34	7.507
Peak C	51.933	4.155	1.82	13.92	38 - 43	7.507

Supplementary Figure 8 | FPLC traces of LSD1 variants, CoREST1, and LSD1-CoREST1 purified by size-exclusion chromatography.

Supplementary Dataset Legends

Supplementary Dataset 1. sgRNA sequences used for LSD1 CRISPR-suppressor scanning

Supplementary Dataset 2. Log_2+1 transformed sgRNA read-count normalized reads for SET-2 treated with GSK-LSD1

Supplementary Dataset 3. Log_2+1 transformed sgRNA read-count normalized reads for MV4;11 treated with GSK-LSD1

Supplementary Dataset 4. Log_2+1 transformed sgRNA read-count normalized reads for SET-2 CRISPR-suppressor scanning screen

Supplementary Dataset 5. PCR primers employed for genomic DNA amplification

Supplementary Dataset 6. Cluster classification for differentially expressed genes in Figure 6a

Supplementary Dataset 7. Gene signatures used in Gene Set Enrichment Analysis