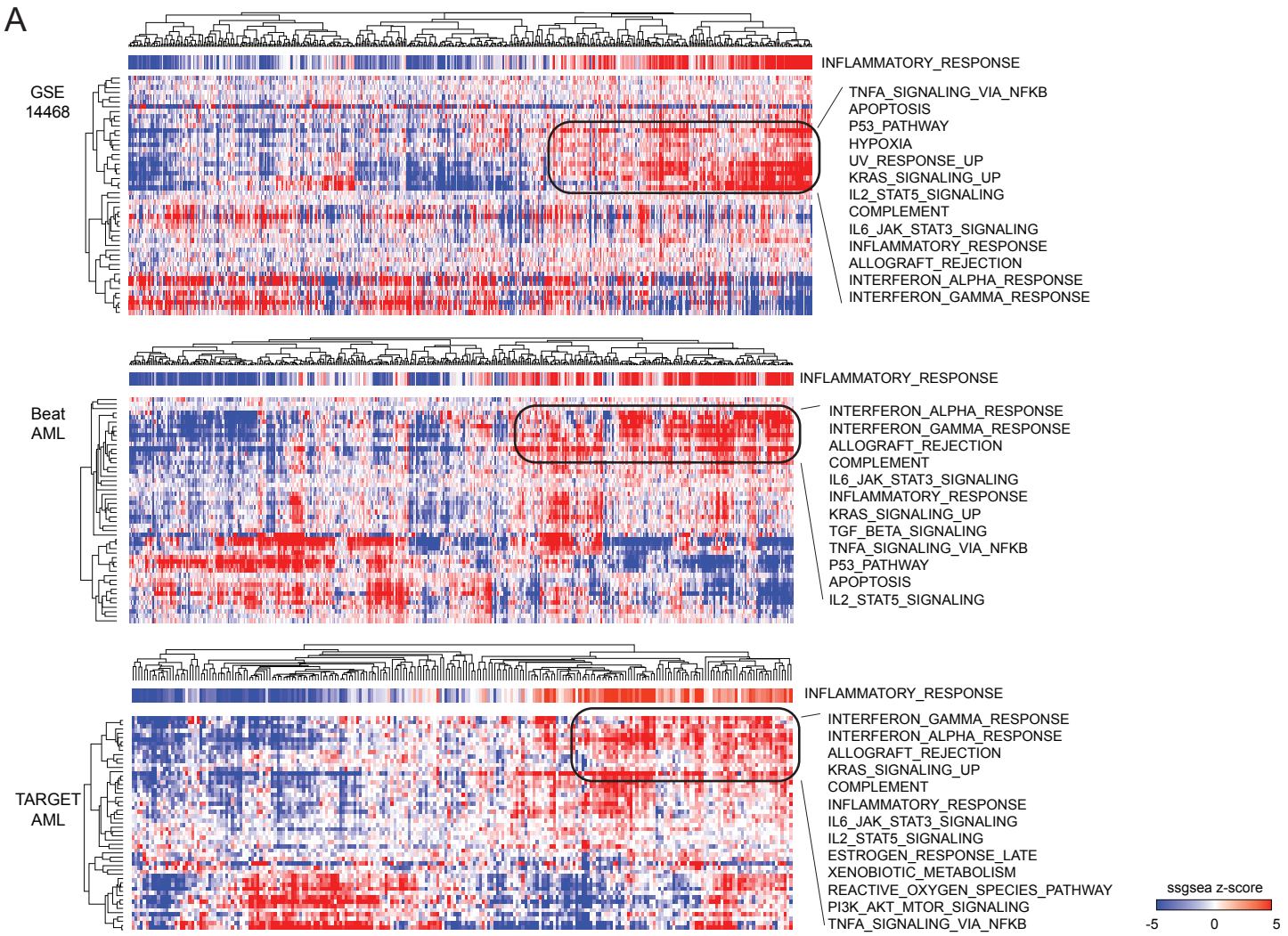
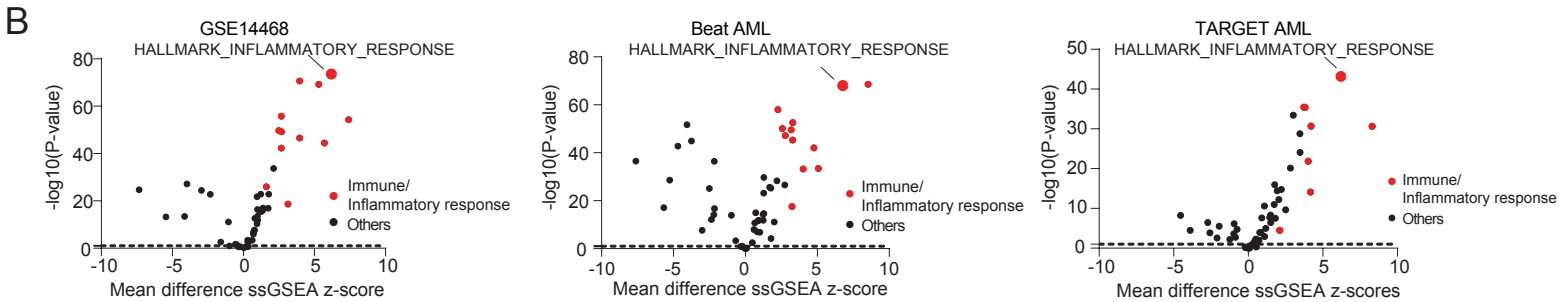
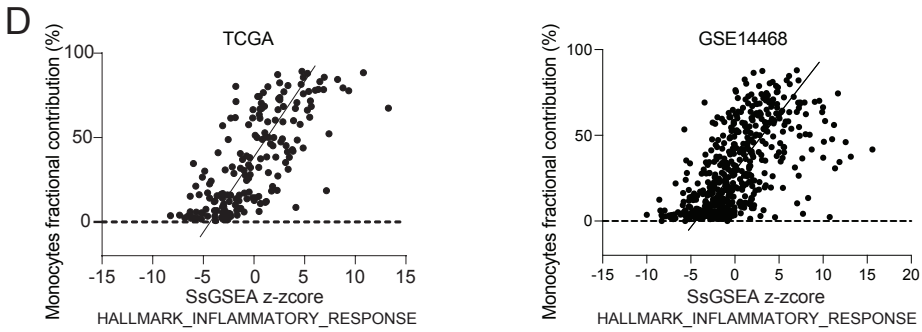
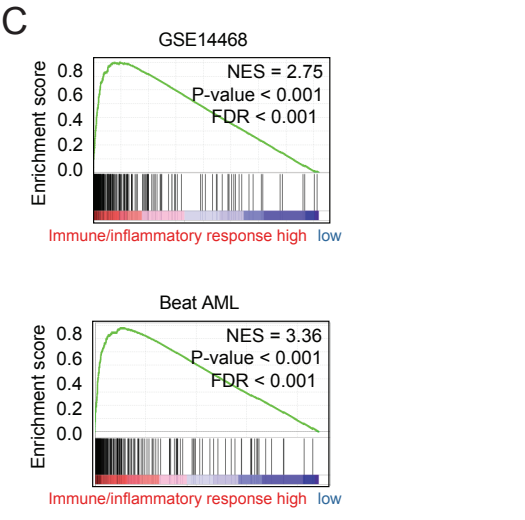
Figure S1


A
GSE 14468
INFLAMMATORY_RESPONSE
TNFA_SIGNALING_VIA_NFKB
APOPTOSIS
P53_PATHWAY
HYPOXIA
UV_RESPONSE_UP
KRAS_SIGNALING_UP
IL2_STAT5_SIGNALING
COMPLEMENT
IL6_JAK_STAT3_SIGNALING
INFLAMMATORY_RESPONSE
ALLOGRAFT_REJECTION
INTERFERON_ALPHA_RESPONSE
INTERFERON_GAMMA_RESPONSE
Beat AML
INFLAMMATORY_RESPONSE
INTERFERON_ALPHA_RESPONSE
INTERFERON_GAMMA_RESPONSE
ALLOGRAFT_REJECTION
COMPLEMENT
IL6_JAK_STAT3_SIGNALING
INFLAMMATORY_RESPONSE
KRAS_SIGNALING_UP
TGF_BETA_SIGNALING
TNFA_SIGNALING_VIA_NFKB
P53_PATHWAY
APOPTOSIS
IL2_STAT5_SIGNALING
TARGET AML
INFLAMMATORY_RESPONSE
INTERFERON_GAMMA_RESPONSE
INTERFERON_ALPHA_RESPONSE
ALLOGRAFT_REJECTION
KRAS_SIGNALING_UP
COMPLEMENT
INFLAMMATORY_RESPONSE
IL6_JAK_STAT3_SIGNALING
IL2_STAT5_SIGNALING
ESTROGEN_RESPONSE_LATE
XENOBIOTIC_METABOLISM
REACTIVE_OXYGEN_SPECIES_PATHWAY
PI3K_AKT_MTOR_SIGNALING
TNFA_SIGNALING_VIA_NFKB
ssgsea z-score
-5
0
5


B
GSE14468
HALLMARK_INFLAMMATORY_RESPONSE
-log10(P-value)
Immune/ Inflammatory response
Others
Mean difference ssGSEA z-score
Beat AML
HALLMARK_INFLAMMATORY_RESPONSE
-log10(P-value)
Immune/ Inflammatory response
Others
Mean difference ssGSEA z-score
TARGET AML
HALLMARK_INFLAMMATORY_RESPONSE
-log10(P-value)
Immune/ Inflammatory response
Others
Mean difference ssGSEA z-scores


D
TCGA
Monocytes fractional contribution (%)
SsGSEA z-zcore
HALLMARK_INFLAMMATORY_RESPONSE
GSE14468
Monocytes fractional contribution (%)
SsGSEA z-zcore
HALLMARK_INFLAMMATORY_RESPONSE


C
GSE14468
Enrichment score
NES = 2.75
P-value < 0.001
FDR < 0.001
Immune/inflammatory response high low
Beat AML
Enrichment score
NES = 3.36
P-value < 0.001
FDR < 0.001
Immune/inflammatory response high low

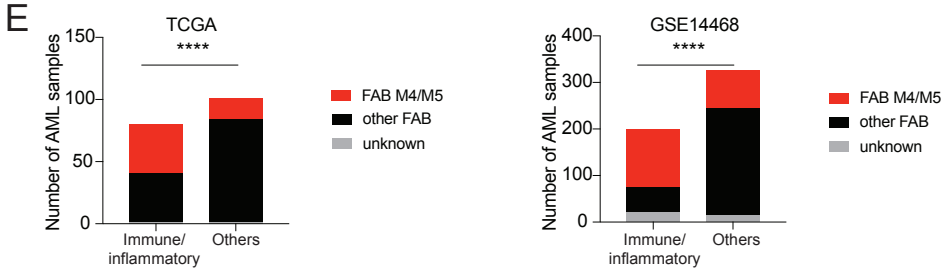

E
TCGA
****
Number of AML samples
FAB M4/M5
other FAB
unknown
Immune/ inflammatory
Others
GSE14468
****
Number of AML samples
FAB M4/M5
other FAB
unknown
Immune/ inflammatory
Others

Figure S1

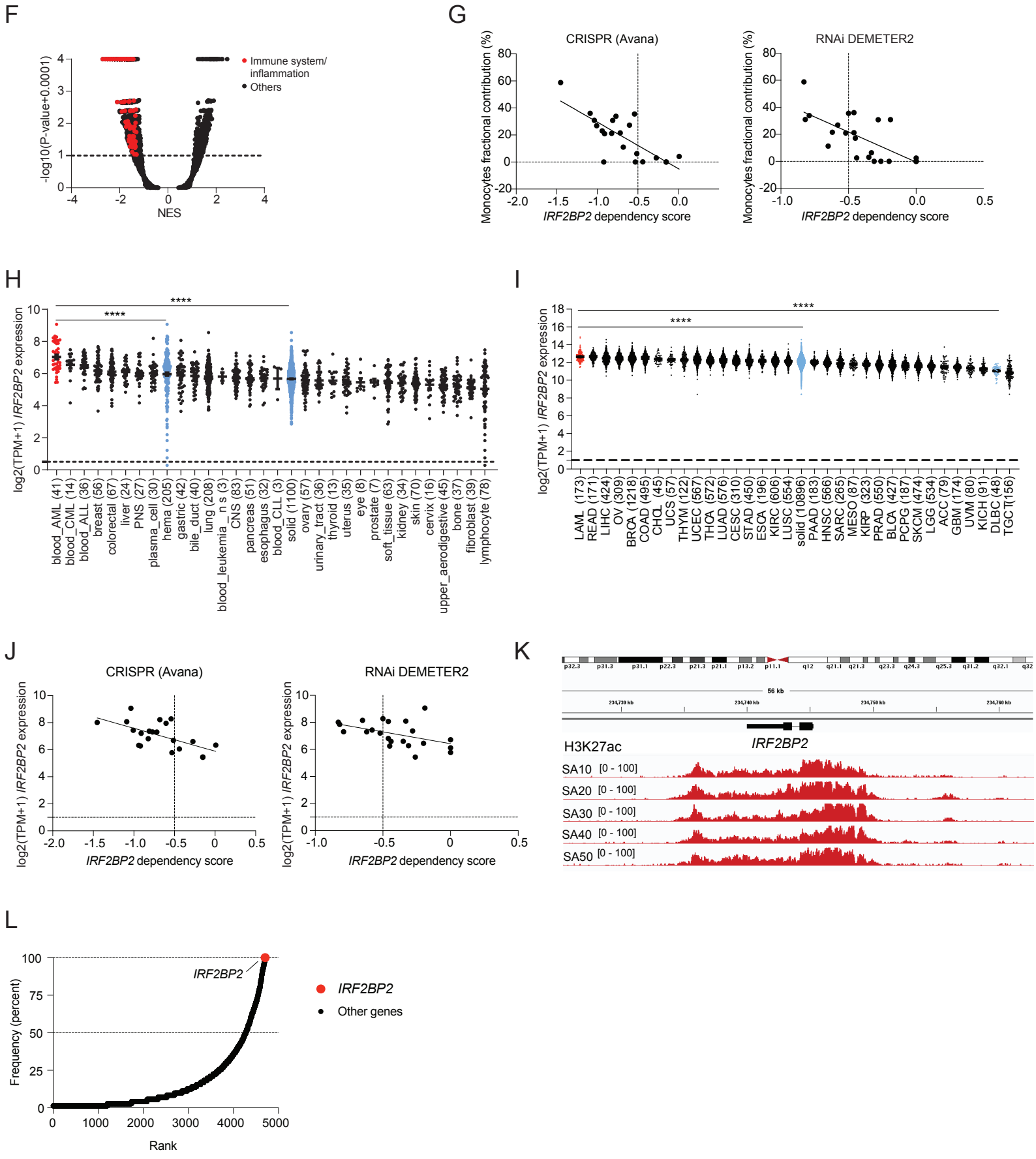

F
-log10(P-value+0.0001)
NES
Immune system/ inflammation
Others
G
CRISPR (Avana)
RNAi DEMETER2
Monocytes fractional contribution (%)
IRF2BP2 dependency score
H
log2(TPM+1) IRF2BP2 expression
****
blood_AML (41)
blood_CML (14)
blood_ALL (36)
breast (56)
colorectal (67)
liver (24)
PNS (27)
plasma_cell (30)
hema (205)
gastric (42)
bile_duct (40)
lung (208)
blood_leukemia_ n s (3)
CNS (83)
pancreas (51)
esophagus (32)
blood_CLL (3)
solid (1100)
ovary (57)
urinary_tract (36)
thyroid (13)
uterus (35)
eye (8)
prostate (7)
soft_tissue (63)
kidney (34)
skin (70)
cervix (16)
upper_aerodigestive (45)
bone (37)
fibroblast (39)
lymphocyte (78)
I
log2(TPM+1) IRF2BP2 expression
LAML (173)
READ (171)
LIHC (424)
OV (309)
BRCA (1218)
COAD (495)
CHOL (45)
UCS (57)
THYM (122)
UCEC (567)
THCA (572)
LUAD (576)
CESC (310)
STAD (450)
ESCA (196)
KIRC (606)
LUSC (554)
solid (10896)
PAAD (183)
HNSC (566)
SARC (265)
MESO (87)
KIRP (323)
PRAD (550)
BLCA (427)
PCPG (187)
SKCM (474)
LGG (534)
ACC (79)
GBM (174)
UVM (80)
KICH (91)
DLBC (48)
TGCT (156)
J
CRISPR (Avana)
RNAi DEMETER2
log2(TPM+1) IRF2BP2 expression
IRF2BP2 dependency score
K
56 kb
H3K27ac
IRF2BP2
SA10 [0 - 100]
SA20 [0 - 100]
SA30 [0 - 100]
SA40 [0 - 100]
SA50 [0 - 100]
L
Frequency (percent)
Rank
IRF2BP2
Other genes

Figure S1. GSEA identifies an AML group enriched for inflammatory pathways across independent patient sample collections

A, Heatmaps depicting the single sample GSEA (ssGSEA) projection of the GSE14468 expression data for 526 AML samples (top panel), the Beat AML data for 451 AML samples (middle panel) and the TARGET AML expression data for 232 AML samples (bottom panel) on the collection of 50 hallmark gene sets (MSigDB v7.0), defining a cluster of AML samples enriched for immune/ inflammatory pathways. AML samples are annotated with the ssGSEA scores for HALLMARK_INFLAMMATORY_RESPONSE. Data are clustered according to the hierarchical clustering for Spearman rank correlation. Top scoring hallmark gene sets within the cluster with strong enrichment for immune/ inflammatory response are listed next to the heatmap.

B, Volcano plot depicting the differential enrichment of the ssGSEA projection on hallmark gene sets for AML samples enriched for immune/ inflammatory pathways (defined as shown in panel A) compared to all other AML samples within the GSE14468 collection (left panel), the Beat AML study (middle panel), and the TARGET AML data set (right panel). Highlighted in red are the immune/ inflammatory hallmark gene sets. Limma eBayes, $|\text{effect size}| \geq 0.5$, p-value ≤ 0.10.

C, GSEA plot for HALLMARK_INFLAMMATORY_RESPONSE enrichment in the genome-wide list of genes ranked by differential expression in AML samples enriched for immune/ inflammatory pathways (defined as shown in panel A) compared to all other AML samples within the GSE14468 (upper panel) and the Beat AML collection (lower panel). NES ≥ 1.3, p-value ≤ 0.05, FDR ≤ 0.25.

D, Scatter plot depicting the correlation between the HALLMARK_INFLAMMATORY_RESPONSE ssGSEA z-scores and the monocytic lineage enrichment (%) per AML sample within the TCGA LAML data (Pearson R = 0.72; $p < 0.0001$; left panel) and within the GSE14468 sample collection (Pearson R = 0.62; $p < 0.0001$; right panel).

E, Barplots depicting the number of AML samples divided into FAB M4/M5, other FAB or unknown with enrichment for immune/ inflammatory pathways compared to all other AML samples within the TCGA LAML collection (left panel) and GSE14468 (right panel).

F, Volcano plot depicting the GSEA in the genome-wide differential CERES dependency scores for AML cell lines compared to all other cancer cell lines within the CRISPR (Avana) 20Q3 data (789 cell lines) within the GOBP (Gene Ontology Biological Process) gene set collection (MSigDB v7.0). Gene sets classified as immune system/ inflammation related are highlighted red.

G, Scatter plots depicting the negative linear associations between IRF2BP2 dependency scores and the monocytic lineage enrichment (%) across AML cell lines within the CRISPR (Avana) dependency data (Pearson R = - 0.76; $p < 0.0001$; left panel) and the RNAi DEMETER2 dependency data (Pearson R = - 0.70; $p < 0.0001$; right panel).

H, Boxplots depicting the *IRF2BP2* log2(TPM+1) expression scores across the 32 lineages in the CCLE 20Q3 data on 1,305 cell lines. The number of cell lines per lineage is indicated in parentheses. Lineages are ranked by average *IRF2BP2* expression scores. AML cell lines are highlighted in red. Differential expression was tested using one-way ANOVA, Tukey's multi-comparisons test, **** $p < 0.0001$.

I, Boxplots depicting the *IRF2BP2* log2(TPM+1) expression scores across the 34 lineages in the panTCGA data on 11,069 tumors (sample identifiers:

http://gdac.broadinstitute.org). The number of samples per lineage is indicated in parentheses. Lineages are ranked by the average *IRF2BP2* expression scores. AML samples (LAML) are highlighted in red. Differential expression was assessed using one-way ANOVA, Tukey's multi-comparisons test, **** $p < 0.0001$.

J, Scatter plots depicting the linear associations between IRF2BP2 dependency scores and *IRF2BP2* expression across AML cell lines within the CRISPR (Avana) dependency data (Pearson R = -0.51; not significant; left panel) and the RNAi DEMETER2 dependency data (Pearson R = -0.47; not significant; right panel).

K, Integrative Genomics Viewer (IGV) tracks depicting the H3K27ac super-enhancer mark in the *IRF2BP2* genomic region in five exemplary samples from a study collection of 71 samples from patients with AML (32).

L, Hockey stick plot depicting the frequency of the 4704 super-enhancer nearest gene targets across the collection of H3K27ac super-enhancers identified for the 71 AML patient samples from the McKenow et al. collection. IRF2BP2 has a 100% frequency as it is a super-enhancer target gene for each of the AML samples in the data set.