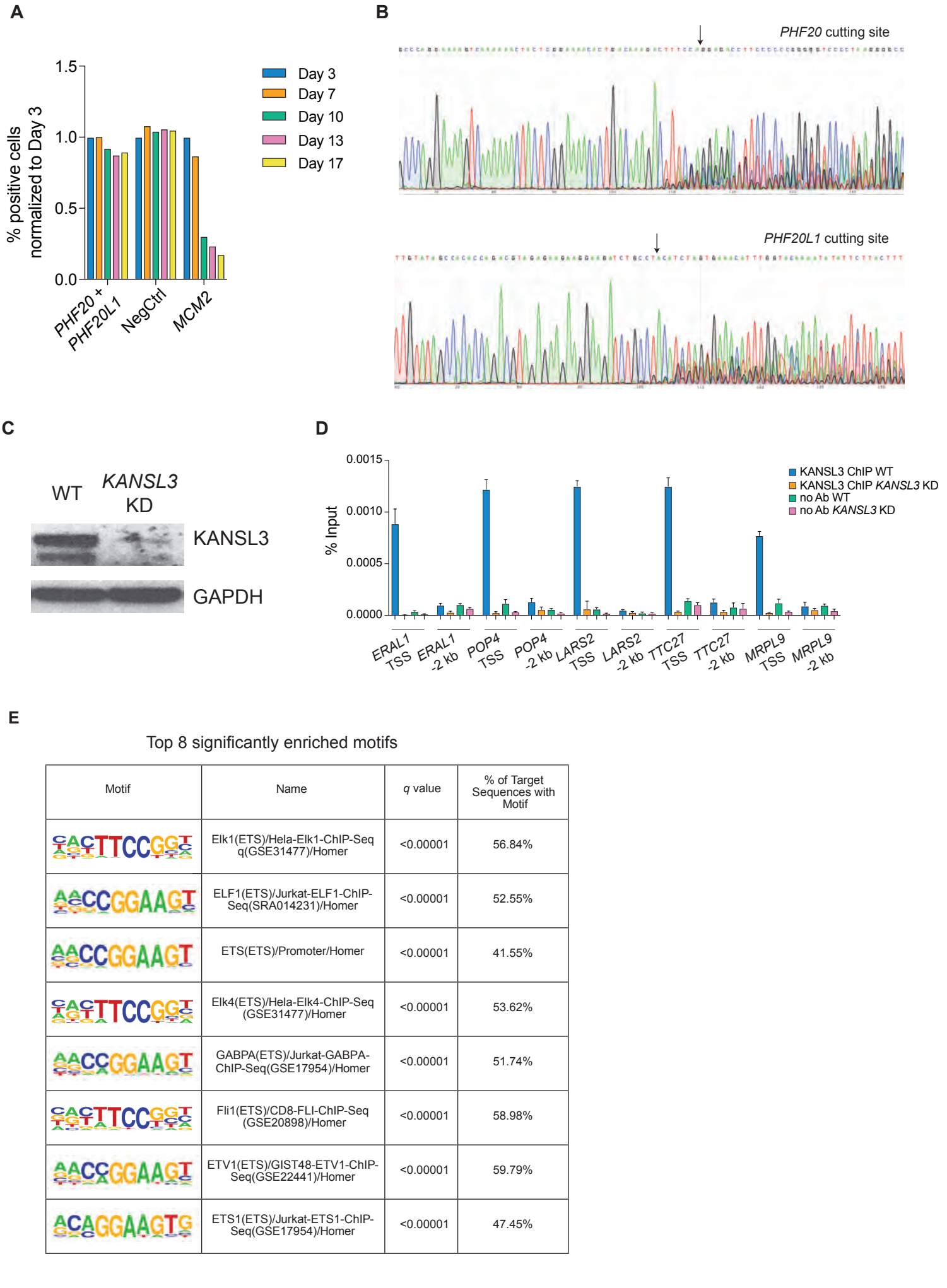
Figure S1



E

Top 8 significantly enriched motifs

Motif	Name	*q* value	% of Target Sequences with Motif
	Elk1(ETS)/Hela-Elk1-ChIP-Seq q(GSE31477)/Homer	<0.00001	56.84%
	ELF1(ETS)/Jurkat-ELF1-ChIP-Seq(SRA014231)/Homer	<0.00001	52.55%
	ETS(ETS)/Promoter/Homer	<0.00001	41.55%
	Elk4(ETS)/Hela-Elk4-ChIP-Seq (GSE31477)/Homer	<0.00001	53.62%
	GABPA(ETS)/Jurkat-GABPA-ChIP-Seq(GSE17954)/Homer	<0.00001	51.74%
	Fli1(ETS)/CD8-FLI-ChIP-Seq (GSE20898)/Homer	<0.00001	58.98%
	ETV1(ETS)/GIST48-ETV1-ChIP-Seq(GSE22441)/Homer	<0.00001	59.79%
	ETS1(ETS)/Jurkat-ETS1-ChIP-Seq(GSE17954)/Homer	<0.00001	47.45%

Figure S1. The NSL complex is essential for cell survival and predominantly localizes to transcription start sites. Related to Figure 1. (A) CRISPR knockout competition assays to test essentiality of *PHF20/PHF20L1* double KO. Double KO cells were gated for using BFP (sgPHF20) and GFP (sgPHF20L1) double positivity. **(B)** Sanger sequencing traces demonstrating efficient sgRNA targeting of the *PHF20* and *PHF20L1* loci (sgRNA cutting sites are arrowed). **(C)** Western blotting for KANSL3 and GAPDH in wild type (WT) and *KANSL3* KD cells used for ChIP experiments. The cells were collected 5 days after sgRNA transduction. **(D)** qRT-PCR validation of KANSL3 ChIP at selected TSSs of NSL complex target genes. Regions 2 kb upstream of the target gene TSSs were probed as a control, "no Ab" indicates no antibody ChIP control. Values are shown as mean ±SD. **(E)** Top significantly enriched motifs in the vicinity of KANSL3 ChIP-seq peaks.

Figure S2

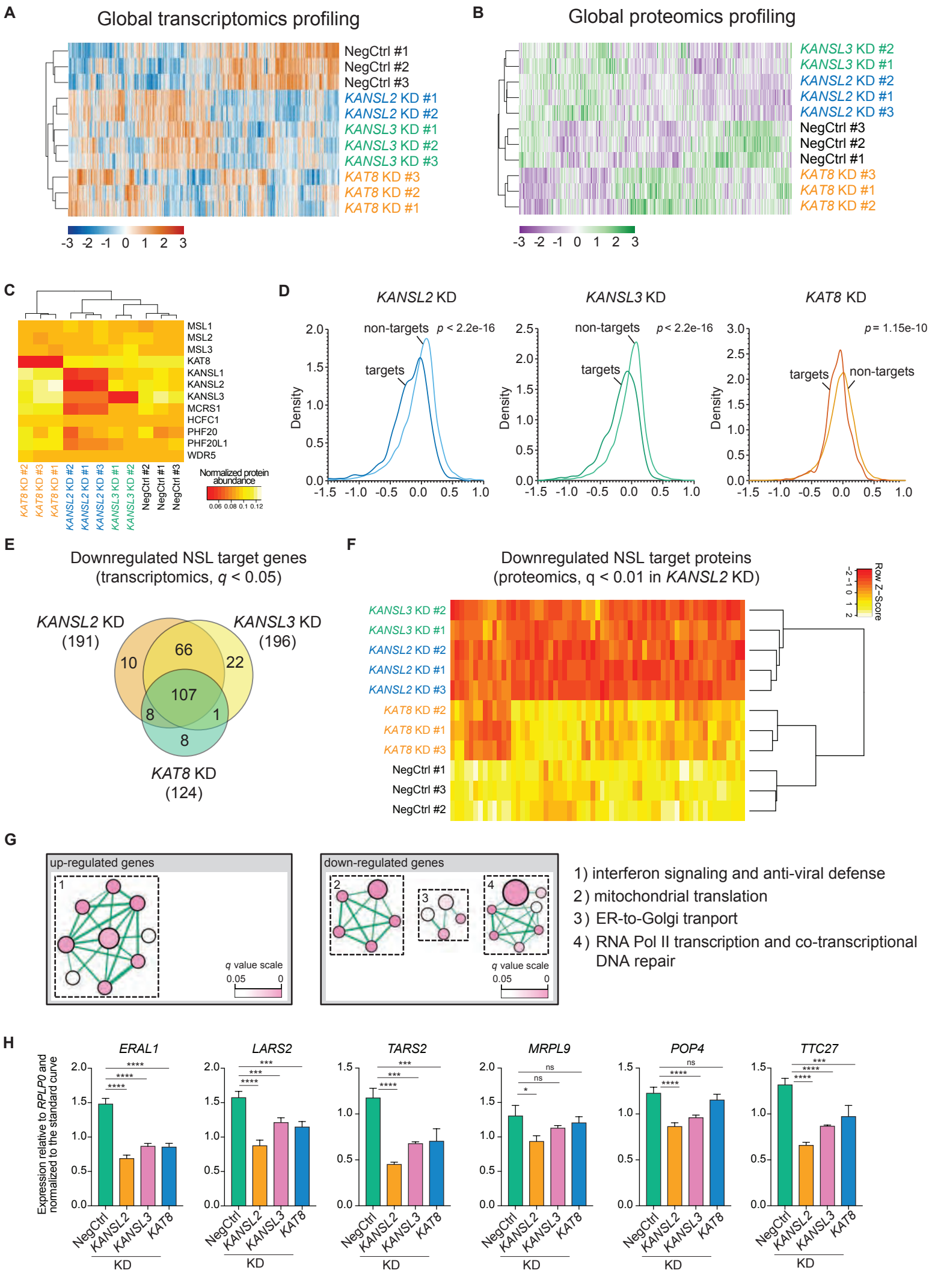

A
Global transcriptomics profiling
NegCtrl #1
NegCtrl #2
NegCtrl #3
KANSL2 KD #1
KANSL2 KD #2
KANSL3 KD #1
KANSL3 KD #2
KANSL3 KD #3
KAT8 KD #3
KAT8 KD #2
KAT8 KD #1
-3 -2 -1 0 1 2 3
B
Global proteomics profiling
KANSL3 KD #2
KANSL3 KD #1
KANSL2 KD #2
KANSL2 KD #1
KANSL2 KD #3
NegCtrl #3
NegCtrl #2
NegCtrl #1
KAT8 KD #3
KAT8 KD #1
KAT8 KD #2
-3 -2 -1 0 1 2 3
C
MSL1
MSL2
MSL3
KAT8
KANSL1
KANSL2
KANSL3
MCRS1
HCFC1
PHF20
PHF20L1
WDR5
KAT8 KD #2
KAT8 KD #3
KAT8 KD #1
KANSL2 KD #2
KANSL2 KD #1
KANSL2 KD #3
KANSL3 KD #1
KANSL3 KD #2
NegCtrl #2
NegCtrl #1
NegCtrl #3
Normalized protein abundance
0.06 0.08 0.1 0.12
D
KANSL2 KD
KANSL3 KD
KAT8 KD
non-targets
targets
p < 2.2e-16
p = 1.15e-10
Density
E
Downregulated NSL target genes
(transcriptomics, q < 0.05)
KANSL2 KD
(191)
KANSL3 KD
(196)
KAT8 KD
(124)
10
66
22
107
8
1
8
F
Downregulated NSL target proteins
(proteomics, q < 0.01 in KANSL2 KD)
Row Z-Score
KANSL3 KD #2
KANSL3 KD #1
KANSL2 KD #2
KANSL2 KD #1
KANSL2 KD #3
KAT8 KD #2
KAT8 KD #1
KAT8 KD #3
NegCtrl #1
NegCtrl #3
NegCtrl #2
G
up-regulated genes
down-regulated genes
q value scale
0.05 0
1) interferon signaling and anti-viral defense
2) mitochondrial translation
3) ER-to-Golgi tranport
4) RNA Pol II transcription and co-transcriptional DNA repair
H
ERAL1
LARS2
TARS2
MRPL9
POP4
TTC27
Expression relative to RPLP0 and normalized to the standard curve
NegCtrl
KANSL2
KANSL3
KAT8
KD

Figure S2. The NSL complex is required for expression of a subset of essential genes in human cells. Related to Figure 2. (A) Hierarchical clustering analysis of RNA-seq gene expression data from NSL complex KD series (# indicates replicate number). **(B)** Hierarchical clustering analysis of global proteomics data from NSL complex KD series (# indicates replicate number). **(C)** Heatmap of normalized protein abundances for MSL and NSL complex members in NSL complex KD series. (**D**) Distributions of gene expression changes of NSL complex target genes ("targets") and all other genes ("non-targets") after KDs of *KANSL2*, *KANSL3* and *KAT8*. **(E)** Venn diagram indicating overlap of genes in the vicinity of a KANSL3 peak that get downregulated after KDs of *KANSL2*, *KANSL3* or *KAT8* (q value < 0.05). **(F)** Hierarchical clustering analysis of changes in NSL complex target protein expression after KDs of *KANSL2*, *KANSL3* and *KAT8* (# indicates replicate number). Only proteins significantly downregulated in *KANSL2* KD are plotted (q value < 0.01). **(G)** Clusters of the Reactome gene sets enriched among differentially expressed proteins after *KANSL2* KD. **(H)** qRT-PCR analysis of selected essential NSL complex target genes (*ERAL1*, *LARS2*, *TARS2*, *MRPL9*, *POP4*, *TTC27*) after KD of key NSL complex subunits. Values are normalized to *RPLP0* and shown as mean ±SD.

Figure S3

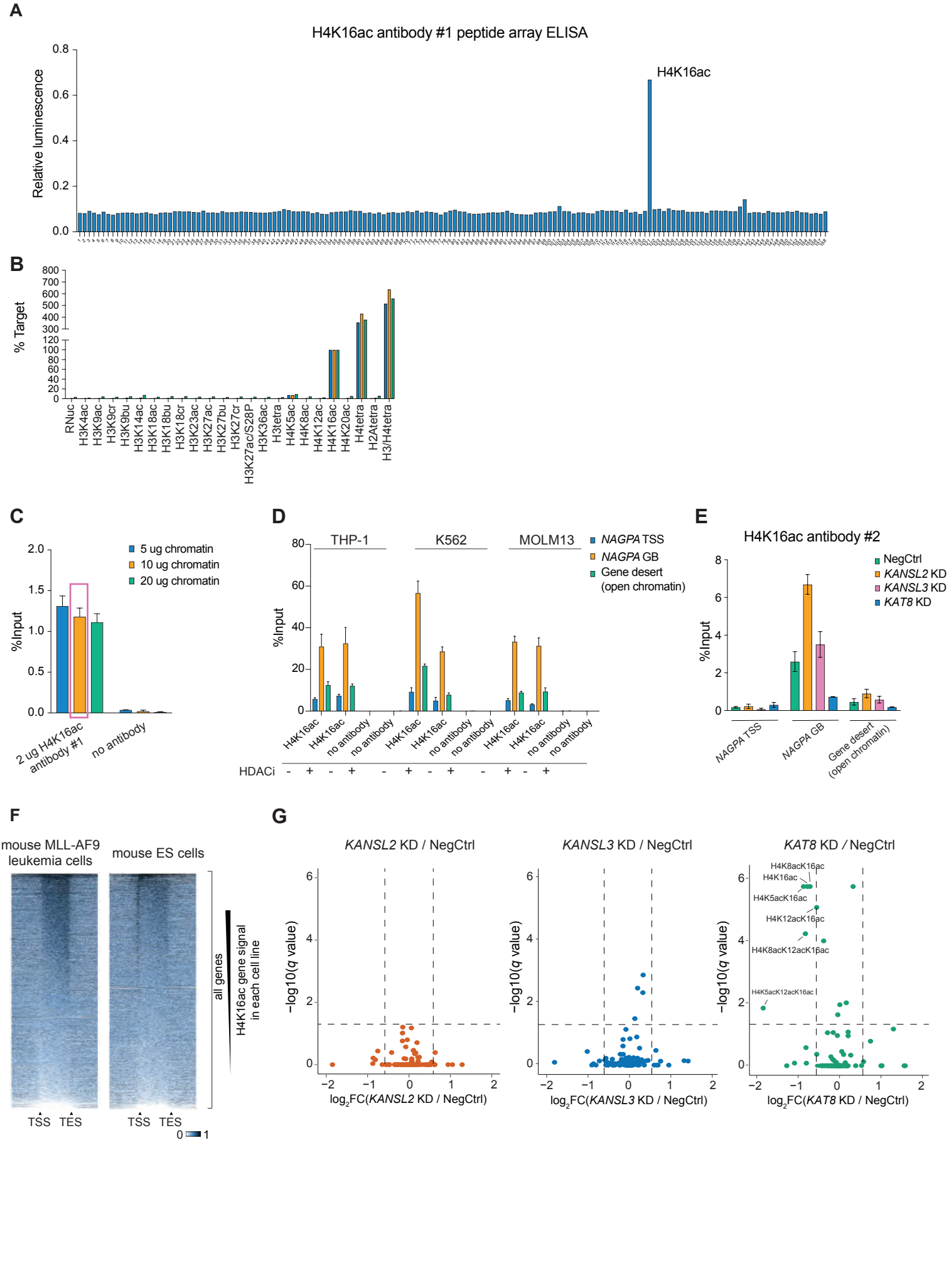

A
H4K16ac antibody #1 peptide array ELISA
Relative luminescence
H4K16ac
B
% Target
RNuc
H3K4ac
H3K9ac
H3K9cr
H3K9bu
H3K14ac
H3K18ac
H3K18bu
H3K18cr
H3K23ac
H3K27ac
H3K27bu
H3K27cr
H3K27ac/S28P
H3K36ac
H3tetra
H4K5ac
H4K8ac
H4K12ac
H4K16ac
H4K20ac
H4tetra
H2Atetra
H3/H4tetra
C
5 ug chromatin
10 ug chromatin
20 ug chromatin
%Input
2 ug H4K16ac antibody #1
no antibody
D
THP-1
K562
MOLM13
NAGPA TSS
NAGPA GB
Gene desert (open chromatin)
%Input
HDACi
E
H4K16ac antibody #2
NegCtrl
KANSL2 KD
KANSL3 KD
KAT8 KD
%Input
NAGPA TSS
NAGPA GB
Gene desert (open chromatin)
F
mouse MLL-AF9 leukemia cells
mouse ES cells
all genes
H4K16ac gene signal in each cell line
TSS
TES
G
KANSL2 KD / NegCtrl
KANSL3 KD / NegCtrl
KAT8 KD / NegCtrl
−log10(q value)
log2FC(KANSL2 KD / NegCtrl)
log2FC(KANSL3 KD / NegCtrl)
log2FC(KAT8 KD / NegCtrl)
H4K8acK16ac
H4K16ac
H4K5acK16ac
H4K12acK16ac
H4K8acK12acK16ac
H4K5acK12acK16ac

Figure S3. Depletion of KAT8, but not KANSL2 and KANSL3, leads to global loss of H4K16ac. Related to Figure 3. (A) Barplot of relative luminescence signals obtained in the ELISA assay with anti-H4K16ac #1 (Abcam, ab109463) and a home-made library of 158 histone peptides (see STAR Methods and Table S3). The H4K16ac-containing peptide is indicated. **(B)** Barplot of the substrates bound in a Luminex-based assay with anti-H4K16ac #1 (Abcam, ab109463) and a panel of 24 designer nucleosomes (EpiCypher). Data at each antibody dilution (250x, 1000x and 4000x) is normalized to the singly acetylated target residue (H4K16ac; set to 100). Key: ac, acetyl; bu, butryl; cr, crotonyl; H2Atetra, H2AK5ac9ac13ac15ac; H3tetra, H3K4ac9ac14ac18ac; H4tetra, H4K5ac8ac12ac16ac. **(C)** qRT-PCR quantitation of H4K16ac ChIP signal obtained at different starting chromatin amounts at the *NAGPA* TSS region (NSL complex target gene). 10 ug/2 ug chromatin/antibody ratio was chosen for subsequent experiments. Values are shown as mean ±SD. **(D)** qRT-PCR quantitation of H4K16ac ChIP signal obtained in three human cell lines (THP-1, K562 and MOLM13) with or without addition of two HDAC inhibitors during ChIP protocol (sodium butyrate and trichostatin A). Key: TSS, transcription start site; GB, gene body; no Ab, no antibody control. Values are shown as mean ±SD. **(E)** qRT-PCR quantitation of H4K16ac ChIP signal obtained using anti-H4K16ac #2 (Millipore, 07-329) after KDs of *KANSL2*, *KANSL3* or *KAT8* at the NSL complex target gene *NAGPA*. Values are shown as mean ±SD. **(F)** Heatmaps of H4K16ac ChIP-seq signal at all genes (gene length is normalized) in two mouse cell lines: E14 mouse ES cells and MLL-AF9 acute myeloid leukemia cell line. ChIP-seq signal is normalized to the total number of reads. **(G)** Volcano plots demonstrating changes in the abundances of 125 histone proteoforms in cells with KDs of *KANSL3*, *KANSL2* and *KAT8* relative to cells transduced with a negative control sgRNA. Significantly changed peptides (>= 1.5-fold change, *q* value <= 0.05) are labeled.

Figure S4

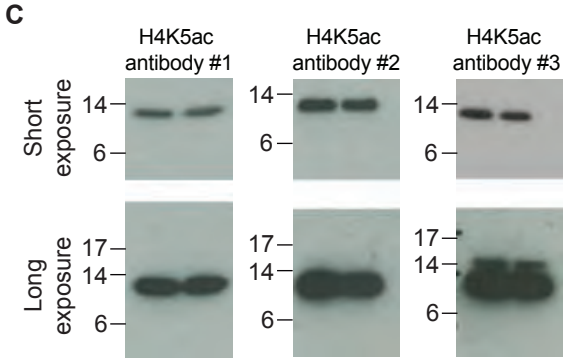
A

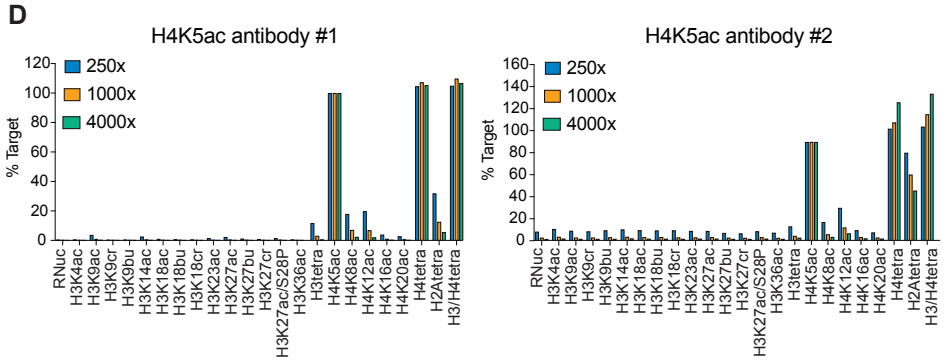
MSGRGKGGKG Histone H4 (1-10)
MSGRGKQGGK Histone H2A (1-10)

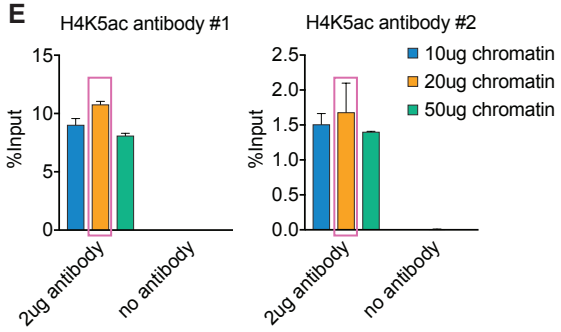
B

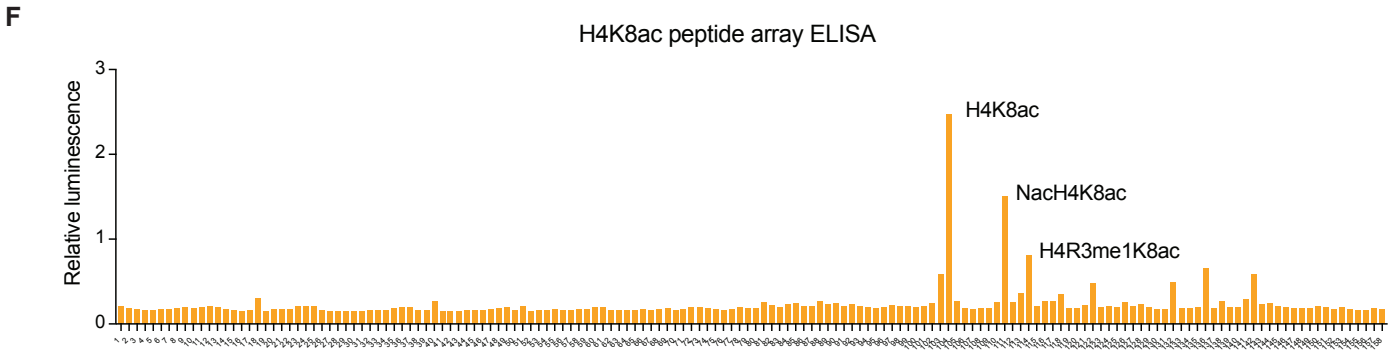
Summary of the ELISA assay results for H4K5ac antibodies

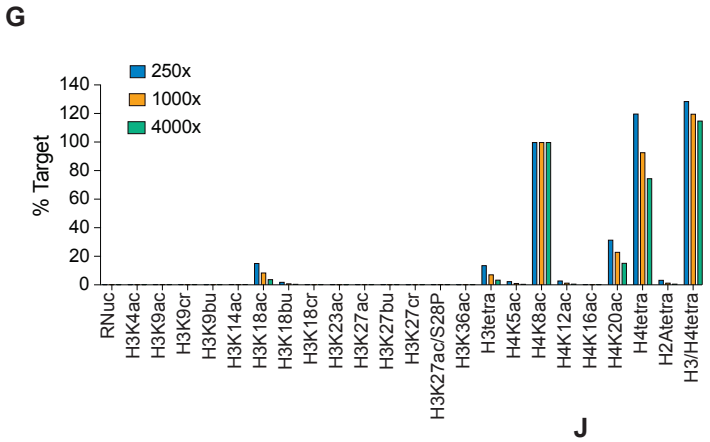
Antibody number	Supplier info	Recognition of the target epitope	Unspecific peptide recognition
H4K5ac #1	Abcam ab51997	Yes (assay performed by Abcam)	small cross-reactivity with H4K8ac (assay performed by Abcam)
H4K5ac #2	Abcam ab114146	Yes	N-acetylated H2AK5ac
H4K5ac #3	Diagenode C15410025	Yes	N-acetylated H2AK5ac, H2A5ac
H4K5ac #4	Millipore 17-10045	No	N-acetylated H2AK5ac
H4K5ac #5	Millipore 07-327	Yes	N-acetylated H2AK5ac, H2A5ac, H4K12ac, H4K15ac

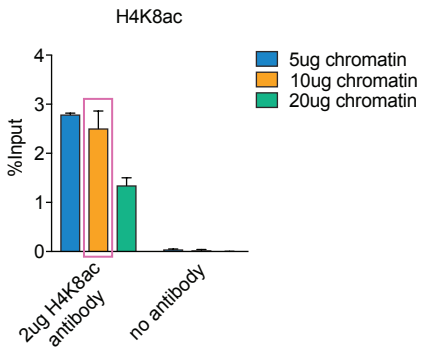


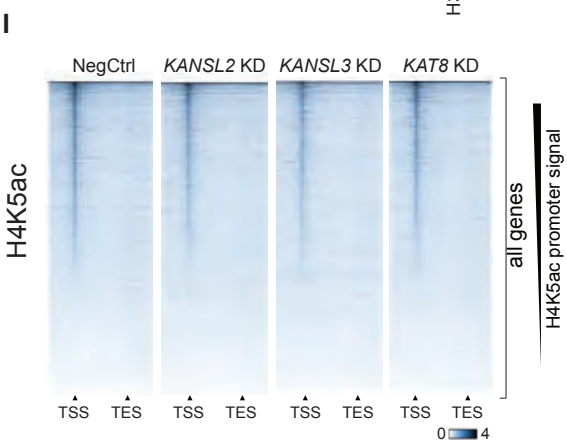


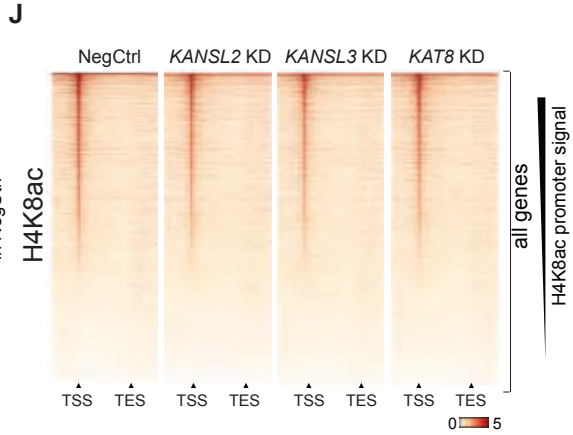












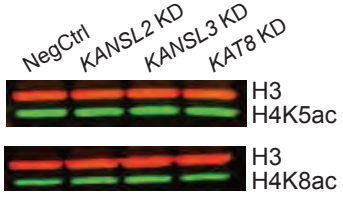


Figure S4. The NSL complex catalyzes H4K5ac and H4K8ac. Related to Figure 4. (A) Amino acid sequence alignment of histone H4 (1-10) and histone H2A (1-10). Lysine 5 is indicated in blue; distinct amino acids are indicated red. **(B)** Table summarizing ELISA assay results for commercial H4K5ac antibodies using a home-made library of 158 histone peptides. Antibodies #1-3 were chosen for western blotting validation. **(C)** Western blotting for H4K5ac using three commercial antibodies. Antibody #3 recognized an additional band of a larger molecular weight, likely acetylated H2A, and was not used for subsequent experiments. Two independent THP-1 cell lysates were used. **(D)** Barplot of the substrates bound in a Luminex-based assay with antibodies to H4K5ac and a panel of 24 designer nucleosomes (EpiCypher). Data at each antibody dilution (250x, 1000x and 4000x) is normalized to the singly acetylated target residue (H4K5ac; set to 100). Key: ac, acetyl; bu, butryl; cr, crotonyl. **(E)** qRT-PCR quantitation of H4K5ac ChIP signal obtained at different starting chromatin amounts at the *NAGPA* TSS region (NSL complex target gene). 20 ug/2 ug chromatin/antibody ratio was chosen for subsequent experiments. Values are shown as mean ±SD. **(F)** Barplot of relative luminescence signals obtained in the ELISA assay with H4K8ac antibody and a home-made library of 158 histone peptides (see STAR Methods and Table S3). The H4K8ac-containing peptides are indicated. **(G)** Barplot of the substrates bound in a Luminex-based assay with anti-H4K8ac and a panel of 24 designer nucleosomes (EpiCypher). Data at each antibody dilution (250x, 1000x and 4000x) is normalized to the singly acetylated target residue (H4K8ac; set to 100). **(H)** qRT-PCR quantitation of H4K8ac ChIP signal obtained at different starting chromatin amounts at the *NAGPA* TSS region (NSL complex target gene). 10 ug/2 ug chromatin/antibody ratio was chosen for subsequent experiments. Values are shown as mean ±SD. **(I-J)** Heatmaps of H4K5ac antibody#1 (I) and H4K8ac (J) ChIP-seq signal at all genes after KDs of *KANSL2*, *KANSL3* or *KAT8*. ChIP-seq signal is normalized to the total number of reads. **(K)** Western blotting analysis of H4K5ac, H4K8ac and H3 levels after KDs of *KANSL2*, *KANSL3* or *KAT8*.

Figure S5

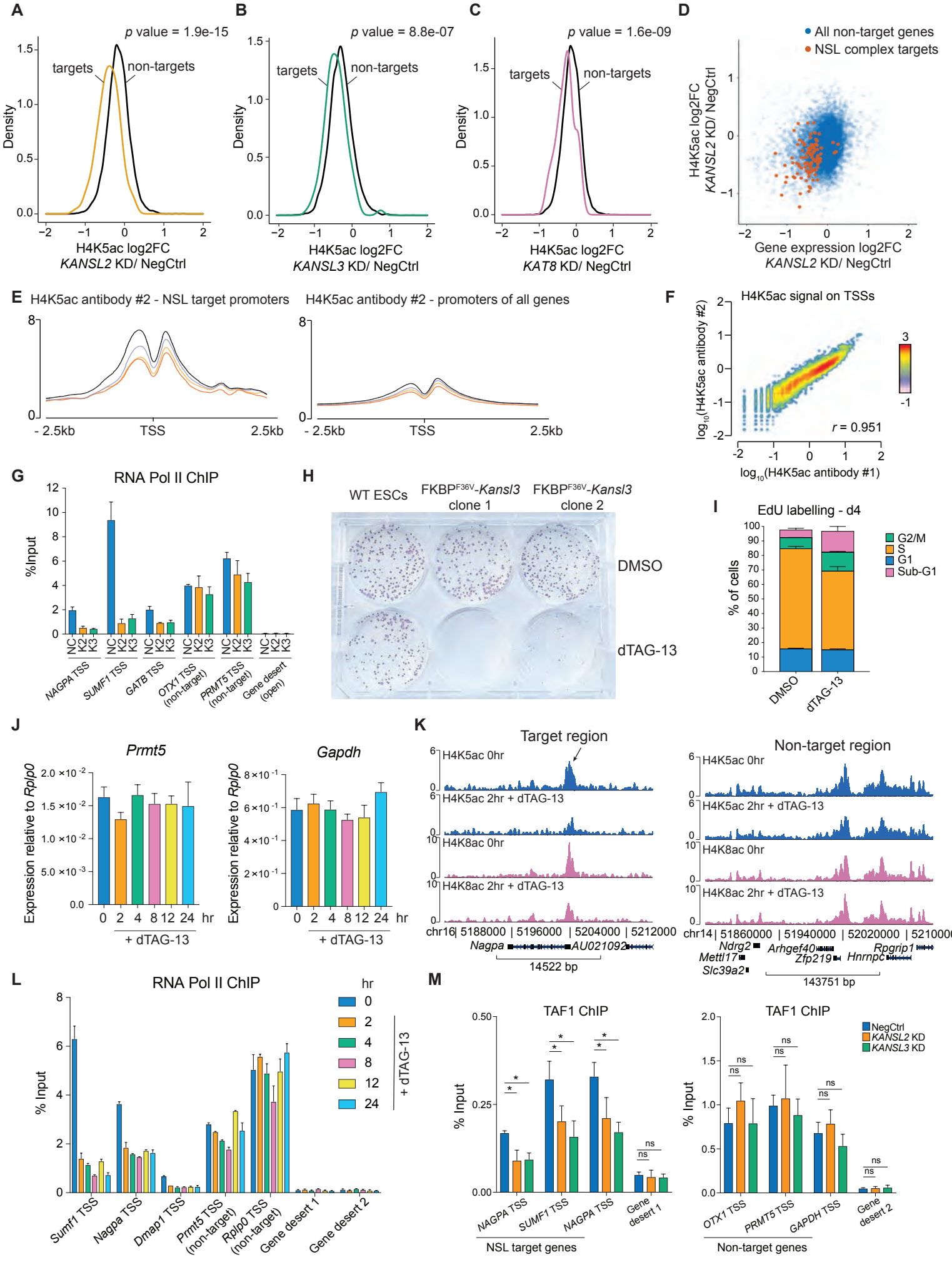

A
p value = 1.9e-15
targets
non-targets
Density
H4K5ac log2FC
KANSL2 KD/ NegCtrl
B
p value = 8.8e-07
targets
non-targets
Density
H4K5ac log2FC
KANSL3 KD/ NegCtrl
C
p value = 1.6e-09
targets
non-targets
Density
H4K5ac log2FC
KAT8 KD/ NegCtrl
D
All non-target genes
NSL complex targets
H4K5ac log2FC
KANSL2 KD/ NegCtrl
Gene expression log2FC
KANSL2 KD/ NegCtrl
E
H4K5ac antibody #2 - NSL target promoters
H4K5ac antibody #2 - promoters of all genes
- 2.5kb
TSS
2.5kb
F
H4K5ac signal on TSSs
log10(H4K5ac antibody #2)
log10(H4K5ac antibody #1)
r = 0.951
G
RNA Pol II ChIP
%Input
NC
K2
K3
NAGPA TSS
SUMF1 TSS
GATB TSS
OTX1 TSS (non-target)
PRMT5 TSS (non-target)
Gene desert (open)
H
WT ESCs
FKBP F36V-Kansl3 clone 1
FKBP F36V-Kansl3 clone 2
DMSO
dTAG-13
I
EdU labelling - d4
% of cells
G2/M
S
G1
Sub-G1
DMSO
dTAG-13
J
Prmt5
Gapdh
Expression relative to Rplp0
0 2 4 8 12 24 hr
+ dTAG-13
K
Target region
Non-target region
H4K5ac 0hr
H4K5ac 2hr + dTAG-13
H4K8ac 0hr
H4K8ac 2hr + dTAG-13
chr16| 5188000 | 5196000 | 5204000 | 5212000
Nagpa
AU021092
14522 bp
chr14 | 51860000 | 51940000 | 52020000 | 52100000
Ndrg2
Arhgef40
Rpgrip1
Mettl17
Zfp219
Hnrnpc
Slc39a2
143751 bp
L
RNA Pol II ChIP
% Input
hr
0
2
4
8
12
24
+ dTAG-13
Sumf1 TSS
Nagpa TSS
Dmap1 TSS
Prmt5 TSS (non-target)
Rplp0 TSS (non-target)
Gene desert 1
Gene desert 2
M
TAF1 ChIP
% Input
NAGPA TSS
SUMF1 TSS
NAGPA TSS
Gene desert 1
NSL target genes
TAF1 ChIP
NegCtrl
KANSL2 KD
KANSL3 KD
OTX1 TSS
PRMT5 TSS
GAPDH TSS
Gene desert 2
Non-target genes

Figure S5. The NSL complex catalyzes H4K5 and H4K8 acetylation and is required for transcriptional initiation. Related to Figure 4. (A-C) Distributions of the TSS H4K5ac changes after *KANSL2* (A), *KANLS3* (B) and *KAT8* (C) KDs at NSL complex target genes (defined as having a KANSL3 peak in the vicinity of promoter region) and all other genes. **(D)** Scatterplot describing relationship between gene expression changes and changes in promoter H4K5ac at NSL complex target genes and all other genes. **(E)** Average normalized H4K5ac antibody#2 (Abcam, ab114146) ChIP-seq profiles across NSL target genes or all genes. **(F)** Scatter plot showing correlation between H4K5ac ChIP-seq signals obtained using two independent antibodies (#1 and #2) at all TSSs. **(G)** qRT-PCR quantitation of Pol II ChIP signal at selected loci after KDs of *KANSL2* and *KANSL3*. *NAGPA*, *SUMF1* and *GATB* are NSL complex target genes, *OTX1* is a non-target gene control. Key: NC, NegCtrl; K2, *KANSL2* KD; K3, *KANSL3* KD. Values are shown as mean ±SD of two independent replicates. **(H)** Colony formation assay for wild type mouse ES cells and ES cells with the biallelic $FKBP^{F36V}$ degron knock-in at the *Kansl3* locus. Cells were plated, treated with either DMSO or dTAG-13, and stained for alkaline phosphatase after 6 days. **(I)** Cell cycle analysis using EdU labeling of *Kansl3* degron knock-in mouse ES cells after 4 days of dTAG-13 treatment. **(J)** qRT-PCR gene expression analysis of *Prmt5* and *Gapdh* (non-NSL target genes) in *Kansl3* degron knock-in mouse ES cell line at different times of dTAG-13 treatment. Values are normalized to *Rplp0* and shown as mean ±SD. **(K)** H4K5ac and H4K8ac ChIP-seq tracks at the *Nagpa* gene locus (NSL target gene) in the *Kansl3* degron knock-in mouse ES cell line at 0 and 2 hours after dTAG-13 treatment. Arrow indicates *Nagpa* TSS. **(L)** qRT-PCR quantitation of Pol II ChIP signal at selected loci in the *Kansl3* degron knock-in mouse ES cell line at different times after dTAG-13 treatment. *Nagpa*, *Sumf1* and *Dmap1* are NSL complex target genes, *Prmt5* and *Rplp0* are non-target gene controls. Values are shown as mean ±SD. **(M)** qRT-PCR quantitation of TAF1 ChIP signal at selected loci after KDs of *KANSL2* and *KANSL3* in THP-1 cells. *NAGPA*, *SUMF1* and *GATB* are NSL complex target genes, *OTX1, PRMT5* and *GAPDH* are non-target gene controls. Values are shown as mean ±SD of three independent replicates.

Figure S6

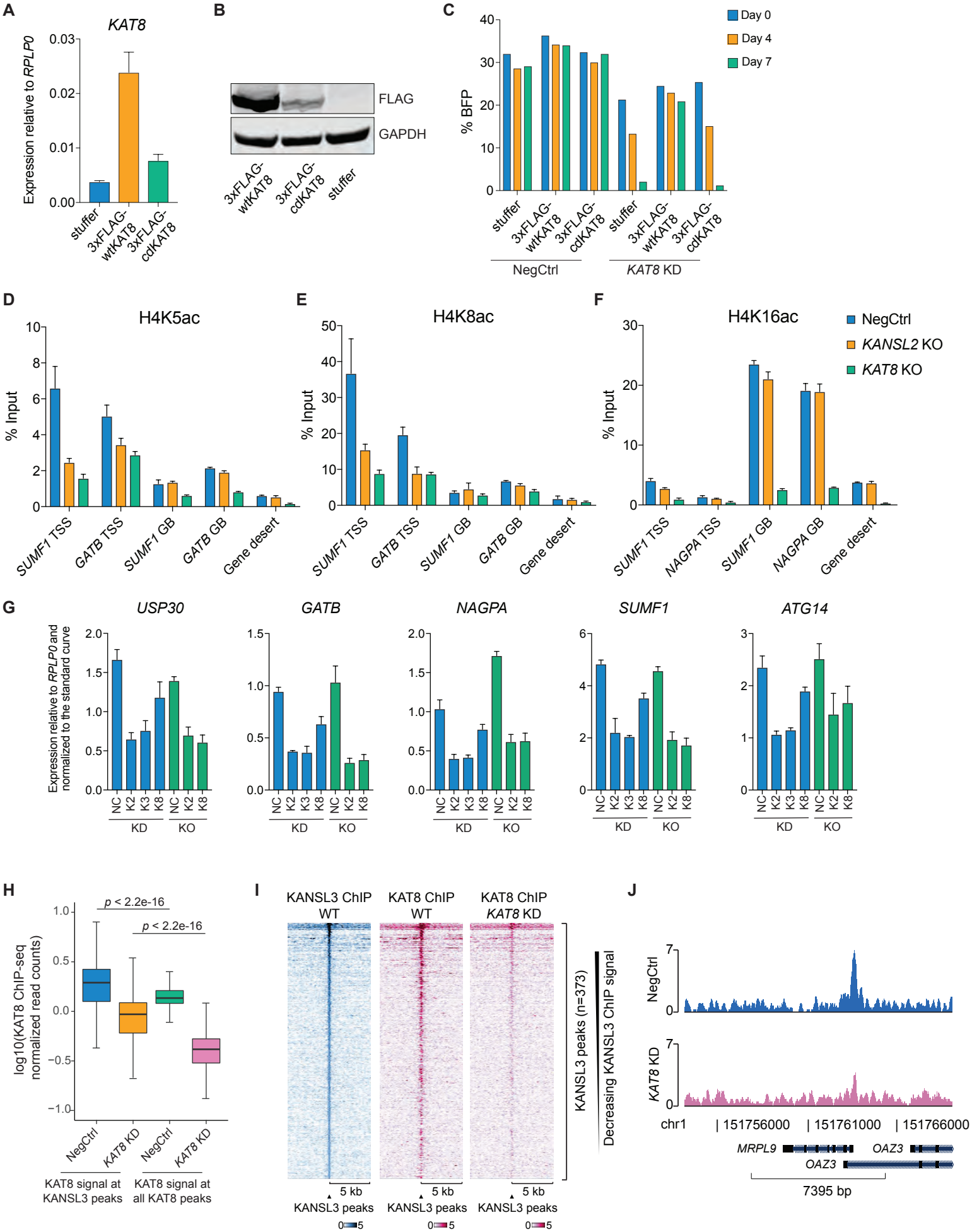

A
KAT8
Expression relative to RPLP0
stuffer
3xFLAG-wtKAT8
3xFLAG-cdKAT8
B
FLAG
GAPDH
3xFLAG-wtKAT8
3xFLAG-cdKAT8
stuffer
C
% BFP
Day 0
Day 4
Day 7
NegCtrl
KAT8 KD
D
H4K5ac
% Input
SUMF1 TSS
GATB TSS
SUMF1 GB
GATB GB
Gene desert
E
H4K8ac
F
H4K16ac
NegCtrl
KANSL2 KO
KAT8 KO
NAGPA TSS
NAGPA GB
G
USP30
GATB
NAGPA
SUMF1
ATG14
Expression relative to RPLP0 and normalized to the standard curve
NC
K2
K3
K8
KD
KO
H
p < 2.2e-16
log10(KAT8 ChIP-seq normalized read counts)
KAT8 signal at KANSL3 peaks
KAT8 signal at all KAT8 peaks
I
KANSL3 ChIP WT
KAT8 ChIP WT
KAT8 ChIP KAT8 KD
KANSL3 peaks (n=373)
Decreasing KANSL3 ChIP signal
5 kb
KANSL3 peaks
J
NegCtrl
KAT8 KD
chr1
151756000
151761000
151766000
MRPL9
OAZ3
7395 bp

Figure S6. The catalytic activity of KAT8 is required for cell survival. Related to Figure 4. (A) qRT-PCR analysis of KAT8 expression in THP-1 cells transduced with empty vector (stuffer), 3xFLAG-wtKAT8 or 3xFLAG-cdKAT8 open reading frames. Values are normalized to *RPLP0* and shown as mean ±SD. **(B)** Western blot of FLAG and GAPDH expression in THP-1 cells transduced with empty vector (stuffer), 3xFLAG-wtKAT8 or 3xFLAG-cdKAT8 open reading frames. **(C)** Competition assay to assess rescue of the *KAT8* KD-induced proliferation defect by wild type (wt) or catalytic dead (cd) (E350Q) KAT8 expression. Cell lines stably expressing wild type or mutant version of KAT8 were transduced with non-targeting sgRNA (NegCtrl) or an sgRNA against *KAT8* and percentage of BFP positive (sgRNA-expressing) cells was monitored over 7 days. **(D-F)** qRT-PCR quantitation of H4K5ac (D), H4K8ac (E) and H4K16ac (F) ChIP signal at gene bodies (GB) and TSSs of *SUMF1* and *GATB* (NSL complex target genes). Values are shown as mean ±SD. **(G)** qRT-PCR analysis of NSL complex target gene expression in cells with either CRISPR KO or CRISPRi KD of *KANSL2* (K2), *KANSL3* (K3) and *KAT8* (K8). Values are normalized to *RPLP0* and shown as mean ±SD. **(H)** Boxplot comparing KAT8 ChIP-seq signal intensity in wild type or *KAT8* KD cells either at all KAT8 peaks or at KANSL3 peaks. **(I)** Heatmaps of normalized KANSL3 and KAT8 ChIP-seq signals at 10 kb regions surrounding all significant KANSL3 peaks. **(J)** KAT8 ChIP-seq example tracks.

Figure S7

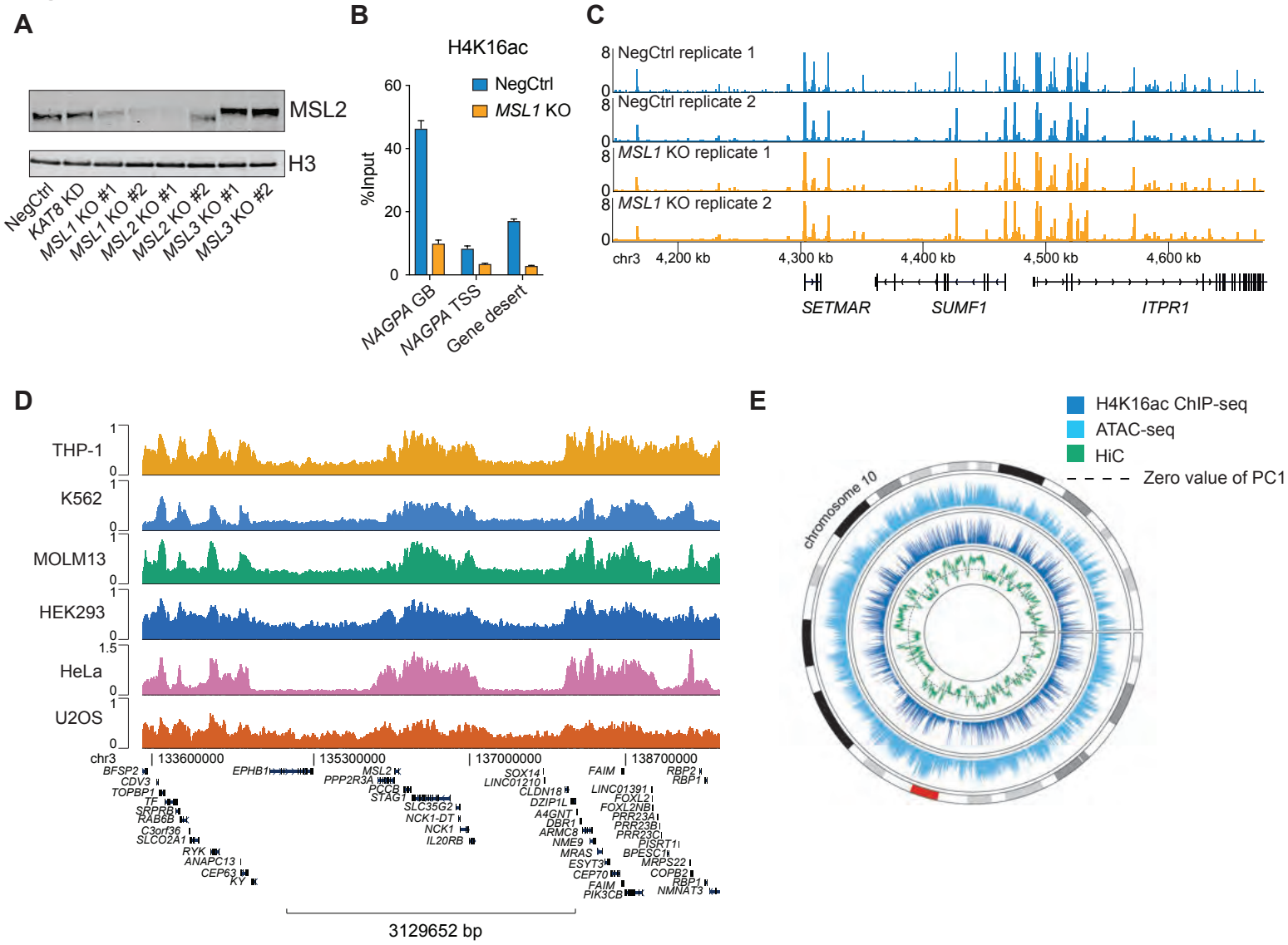

A
MSL2
H3
NegCtrl
KAT8 KD
MSL1 KO #1
MSL1 KO #2
MSL2 KO #1
MSL2 KO #2
MSL3 KO #1
MSL3 KO #2
B
H4K16ac
NegCtrl
MSL1 KO
%Input
60
40
20
0
NAGPA GB
NAGPA TSS
Gene desert
C
NegCtrl replicate 1
NegCtrl replicate 2
MSL1 KO replicate 1
MSL1 KO replicate 2
chr3
4,200 kb
4,300 kb
4,400 kb
4,500 kb
4,600 kb
SETMAR
SUMF1
ITPR1
D
THP-1
K562
MOLM13
HEK293
HeLa
U2OS
chr3
133600000
135300000
137000000
138700000
3129652 bp
E
H4K16ac ChIP-seq
ATAC-seq
HiC
Zero value of PC1
chromosome 10

Figure S7. H4K16ac broadly correlates with open chromatin and does not regulate chromatin accessibility *in vivo*. Related to Figure 5. (A) Western blot analysis of MSL2 expression in THP-1 cells transduced with sgRNAs targeting MSL complex members. **(B)** qRT-PCR analysis of H4K16ac ChIP signal in wild type (NegCtrl) or *MSL1* KO cells at representative TSS, gene body (GB) and intergenic regions. Values are shown as mean ±SD. **(C)** Representative ATAC-seq tracks of wild type and *MSL1* KO cells. **(D)** H4K16ac ChIP-seq tracks at representative megabase-scale locus of the human genome in six cell lines (THP-1, K562, MOLM13, HEK293, HeLa, U2OS). **(E)** Chromosome-scale distribution of H4K16ac ChIP-seq signal, ATAC-seq peaks and HiC-defined chromatin A and B compartments across chromosome 10 in wild type THP-1 cells.