

Supplementary Information

Histone variant H2AZ1 drives lung cancer progression through the RELA-HIF1A-EGFR signaling pathway

Huijie Zhao^{2, 1}, Xing Wu¹, Yinghan Wang¹, Xiuling Li², Yuhui Du¹, Zhiqing Zhou¹, Yu Li¹, Yue Liu¹, Xiaofei Zeng^{1,3}, Guoan Chen^{1,4}

¹Department of Human Cell Biology and Genetics, Joint Laboratory of Guangdong-Hong Kong Universities for Vascular Homeostasis and Diseases, School of Medicine, Southern University of Science and Technology, Shenzhen, China;

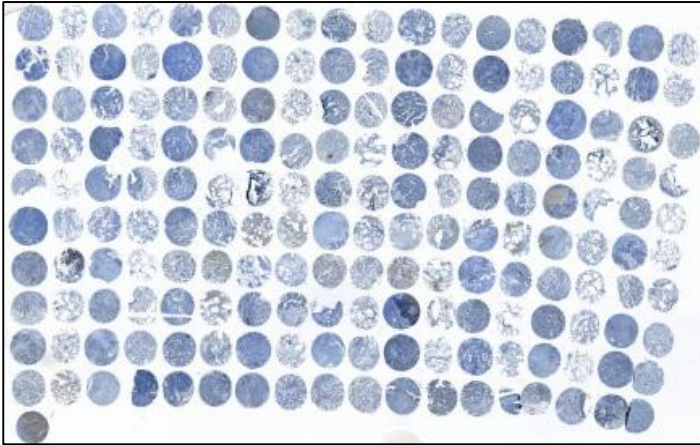
²Department of Oncology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China;

³National Key Laboratory for Tropical Crop Breeding, Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, Guangdong 518120, China

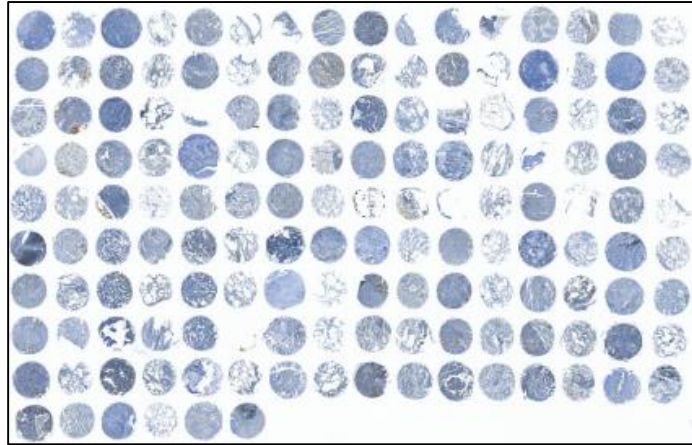
⁴The First Affiliated Hospital of Southern University of Science and Technology, Shenzhen, China

Figure S1

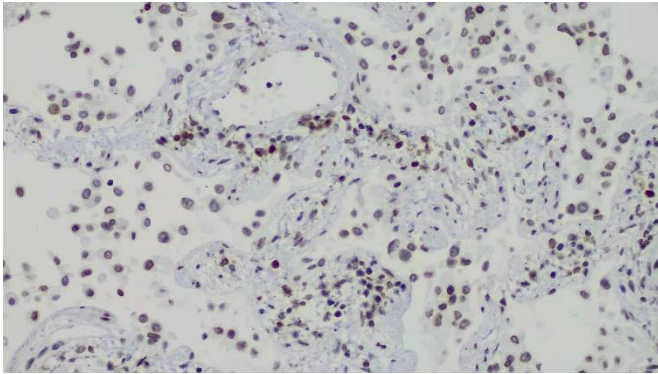
A H2AZ1 IHC of lung tissues on TMA 180



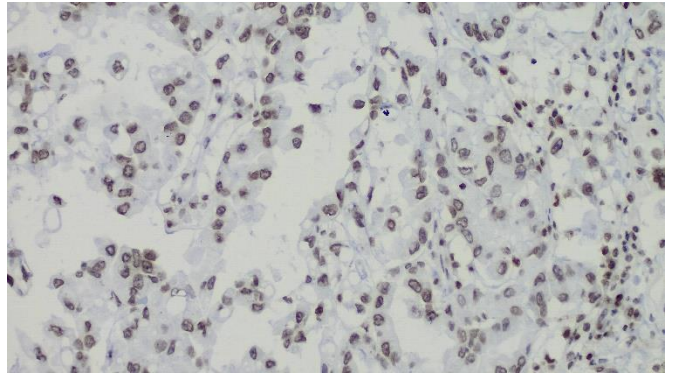
B H2AZ1 IHC of lung tissues on TMA 150



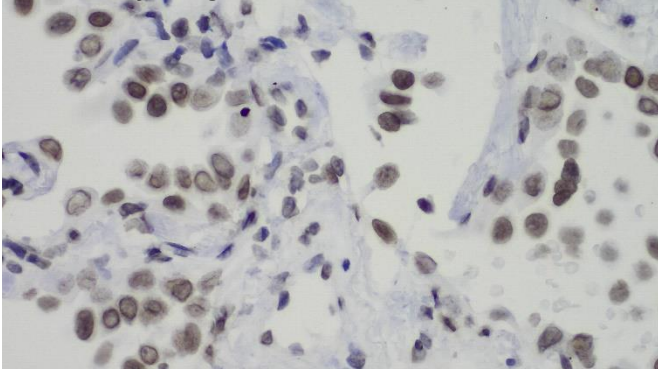
C Primary lung carcinoma



D Brain metastatic tumor from lung carcinoma



E Primary lung carcinoma



F Brain metastatic tumor from lung carcinoma

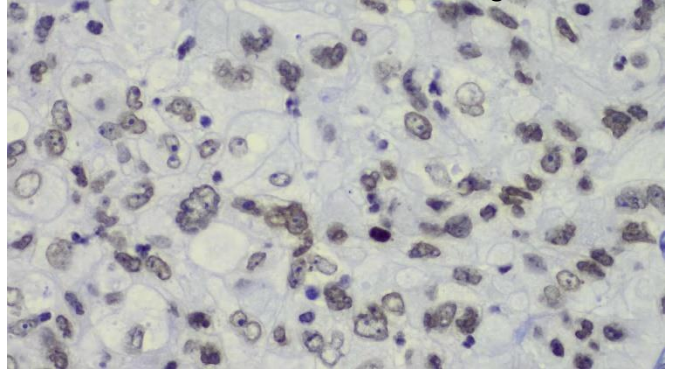
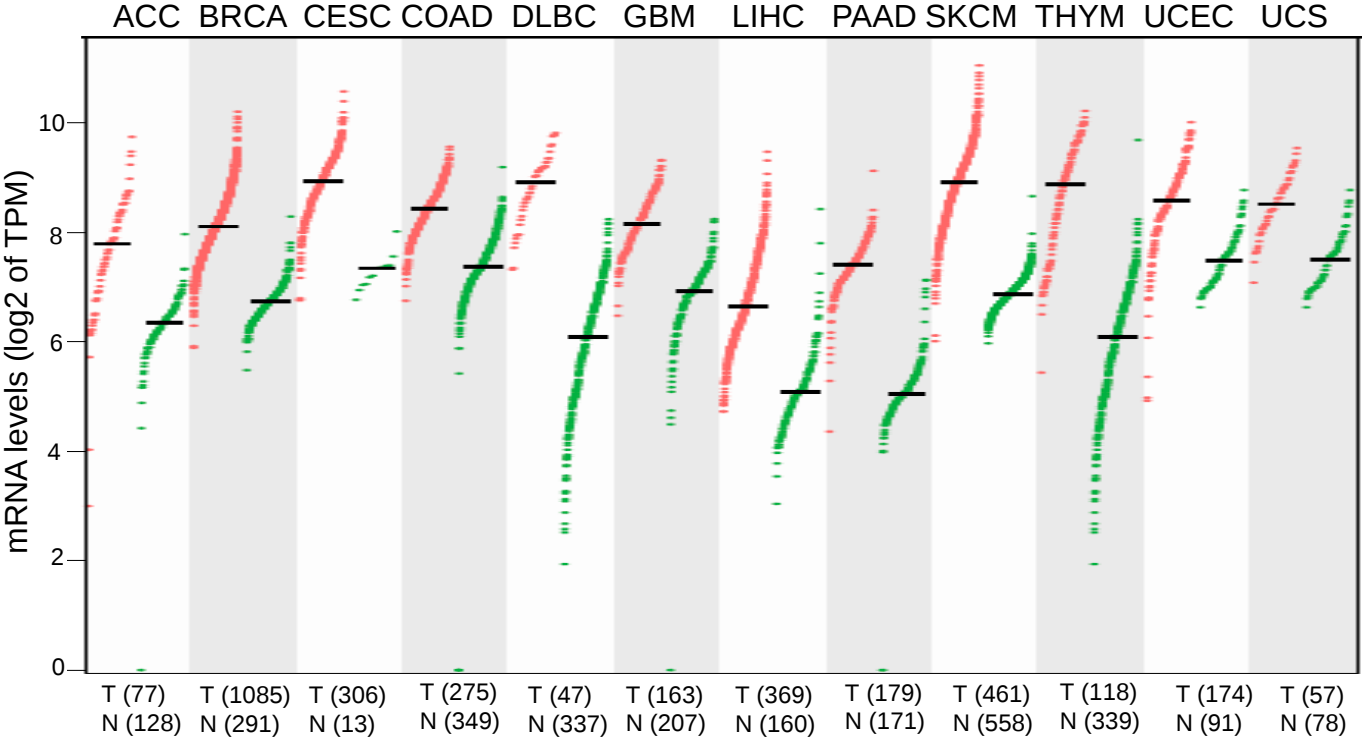
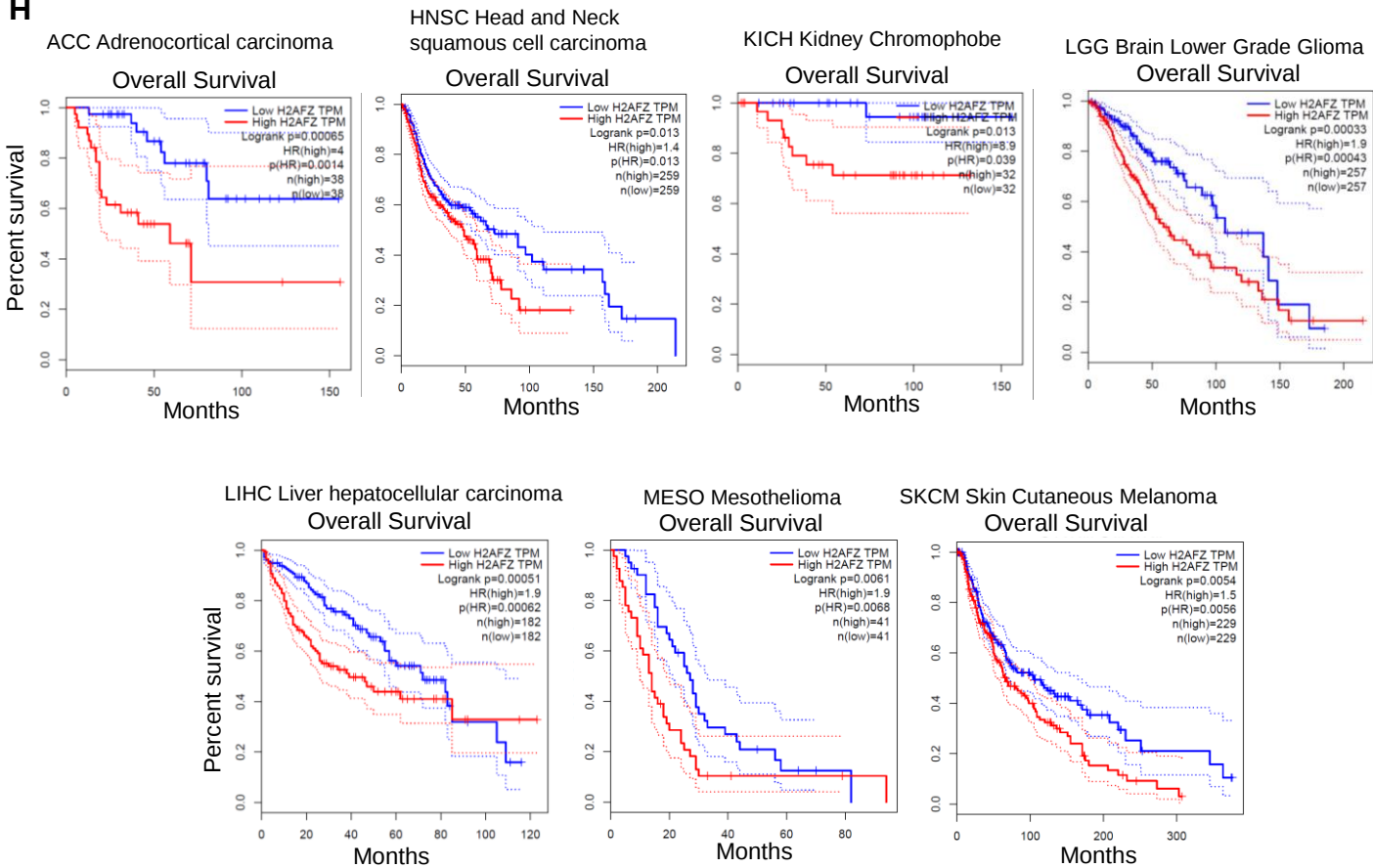


Figure S1

G



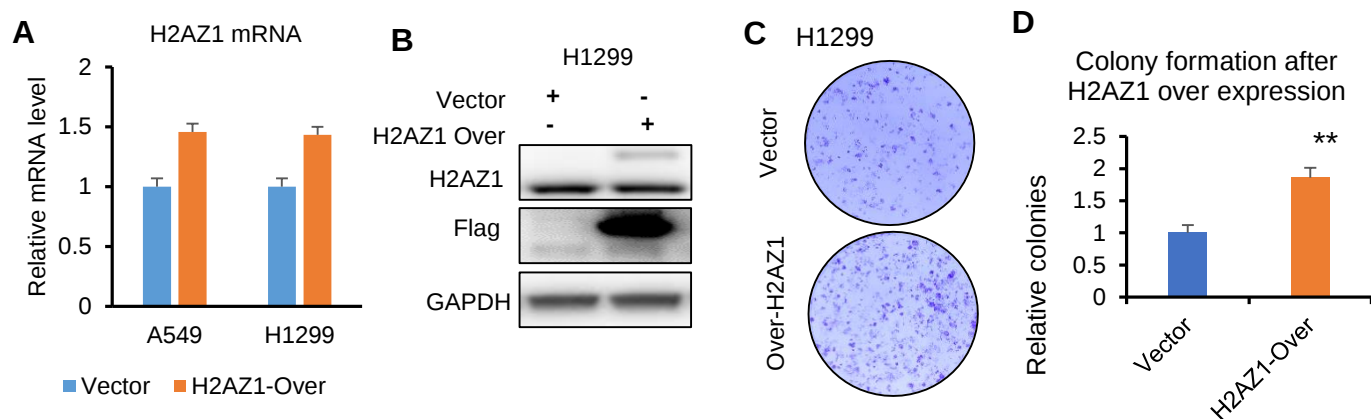
H



Supplementary Figure S1. H2AZ1 mRNA expression in different types of cancers. (A, B)

Full microphotographs of H2AZ1 IHC on two sets of lung TMA (180 and 150). (C-F) Representative image of H2AZ1 IHC on primary lung carcinoma (C, E) and brain metastatic cancer from lung carcinoma (D, F). (G) H2AZ1 mRNA expressions were significantly higher in different types of cancers as compared to their normal tissues, respectively ($p < 0.001$, data from the GEPIA website). ACC, Adrenocortical carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma, and endocervical adenocarcinoma; COAD, Colon adenocarcinoma; DLBC, Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; GBM, Glioblastoma multiforme; LIHC, Liver hepatocellular carcinoma; PAAD, Pancreatic adenocarcinoma; SKCM, Skin Cutaneous Melanoma; THYM, Thymoma; UCEC, Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma. (H) Higher expression of H2AZ1 mRNAs was correlated with poor patient survival in different types of cancers ($p < 0.05$, data from the GEPIA website).

Figure S2



Supplementary Figure S2. The ability of cell colony formation after the overexpression of H2AZ1. (A, B) Overexpression of H2AZ1 was measured by qRT-PCR or Western blot. (C, D) The ability of cell colony formation was increased after the stable overexpression of H2AZ1 (D was the quantitative value of C), ** $p < 0.01$.

Figure S3

A CUT&Tag: H2AZ1 binding signals around TSS and gene body (TES)

B

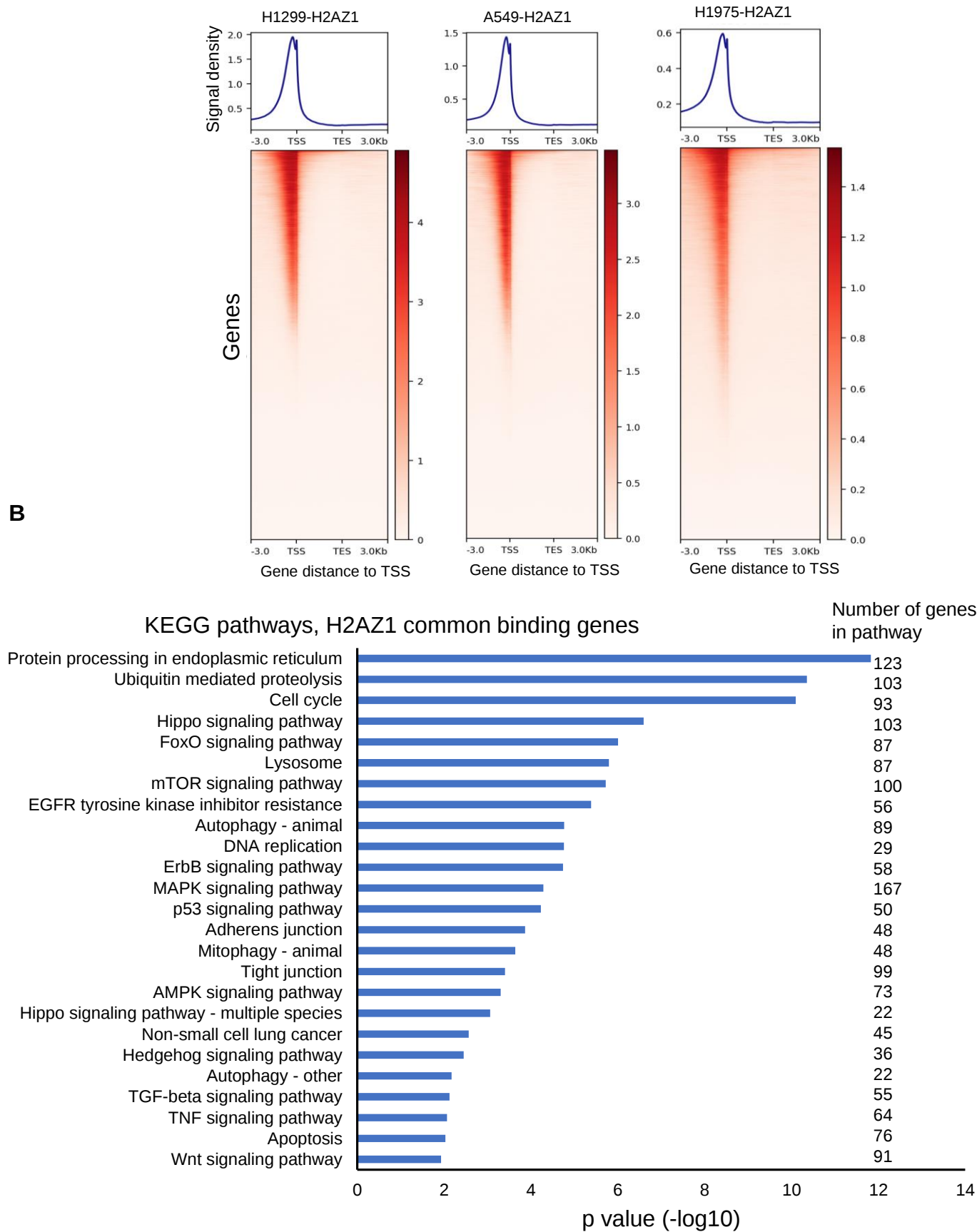
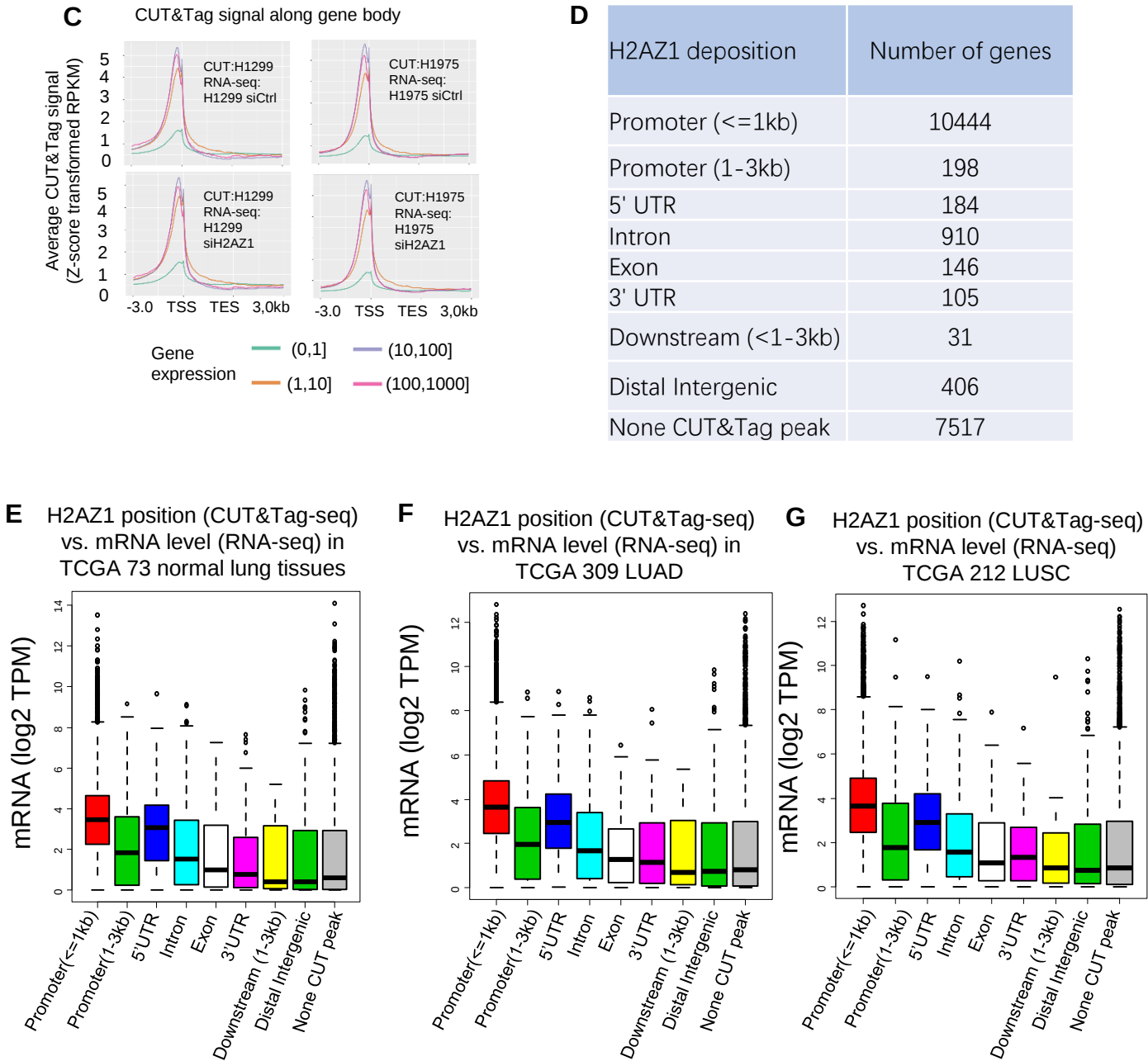


Figure S3



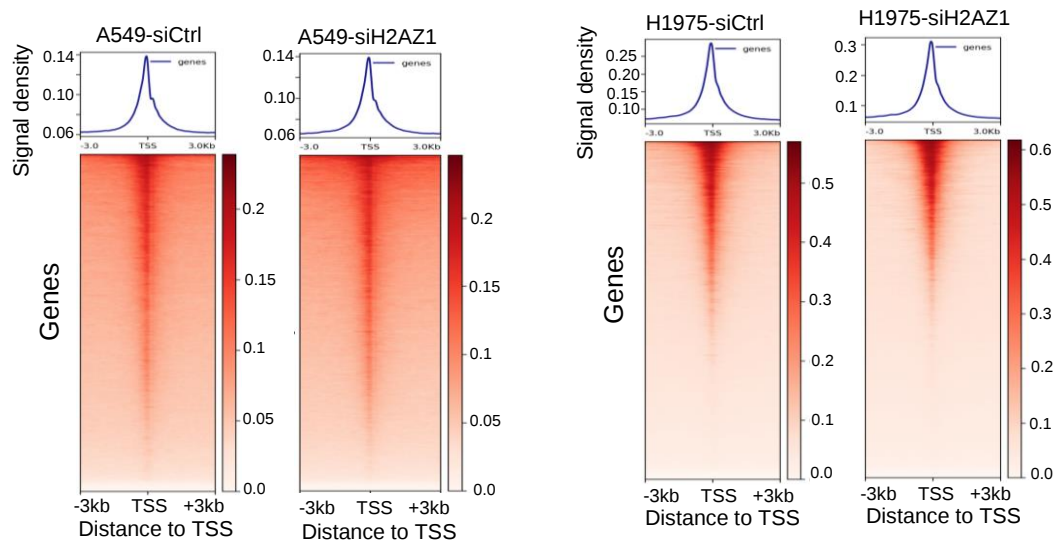
H CUT&Tag vs RNA-seq

Compare Name	H2AZ1 bond -vs - UP DEG	H2AZ1 bond -vs - Down DEG
H1299 siCtrl vs siH2AZ1	826	642
H1975siCtrl vs siH2AZ1	785	640
A549siCtrl vs siH2AZ1	813	866

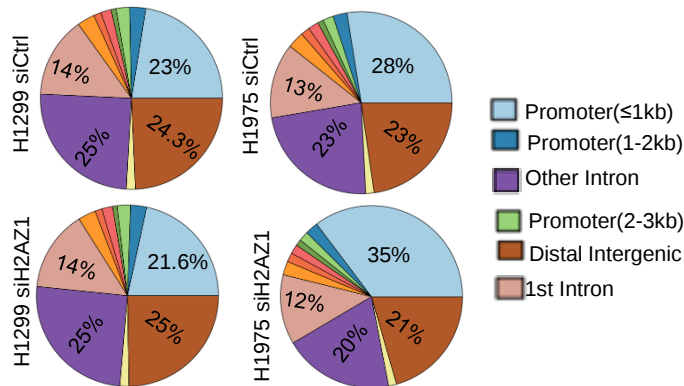
Supplementary Figure S3. H2AZ1 binding to TSS causes higher gene expression and affects multiple cancer-related signaling pathways revealed by CUT&Tag-seq and RNA-seq. (A) CUT&Tag: The H2AZ1 binding peaks in gene body in 3 lung cancer cell lines. (B) CUT&Tag-seq: KEGG enrichment demonstrated the genes corresponding to the common peaks of 3 cell lines were significantly enriched in important tumor-related signaling pathways. (C) The relationship between H2AZ1 binding density by CUT&Tag-seq and gene expression of RNA-seq data in H1299 and H1975 cell lines. (D) The gene number distribution from H2AZ1 CUT&Tag-seq in the H1975 cell line. (E-G) The relationship of H2AZ1 chromatin deposition and gene expression in lung tissues from TCGA RNA-seq data. (H) CUT&Tag vs. RNA-seq showing the changing of H2AZ1 bound genes upon H2AZ1 knockdown.

Figure S4

A ATAC-seq signals enriched around TSS



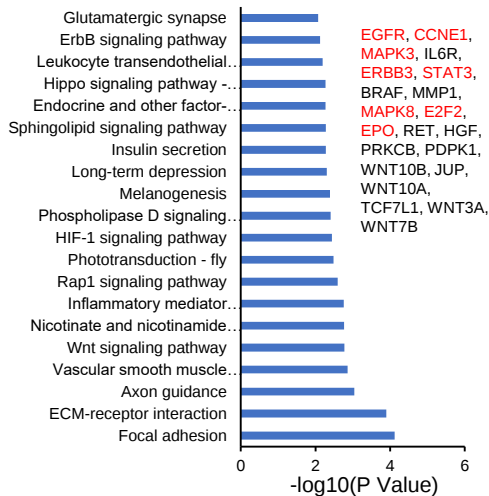
B ATAC-seq: Distribution of peaks on chromatin accessibility



C ATAC-seq: Number of DAR associated promoter genes after H2AZ1 knockdown

Group Name	Promoter Gene (Gain DAR)	Promoter Gene (Loss DAR)
H1299 siCtrl vs H1299 siH2AZ1	3,560	3,023
H1975 siCtrl vs H1975 siH2AZ1	1,379	4,377
A549 siCtrl vs A549 siH2AZ1	2,740	2,912

D KEGG, H1299 (ATAC-seq)



E ATAC-seq signal vs. gene expression

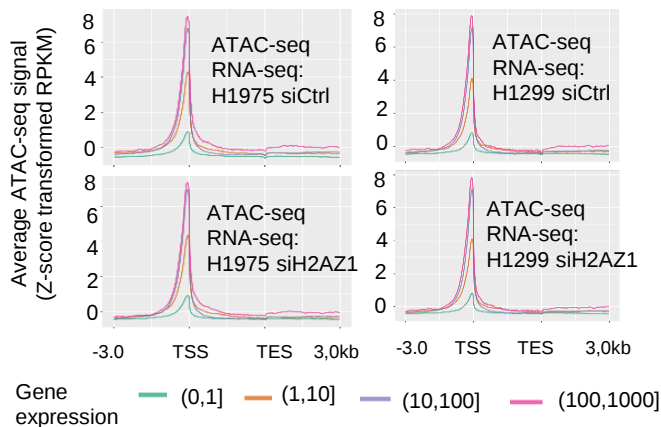


Figure S4

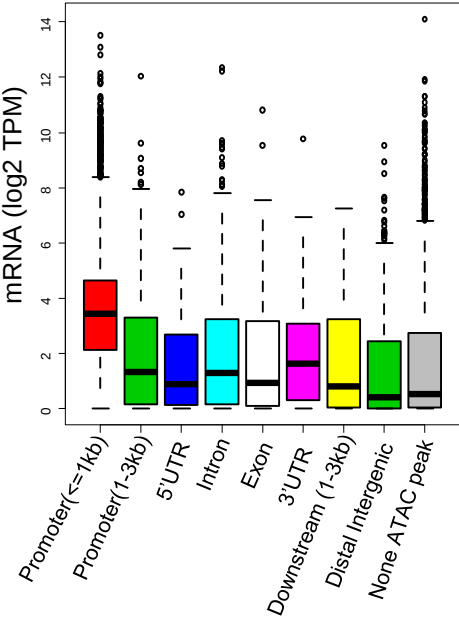
F Gene number distribution of ATAC-seq

ATAC peak position	Number of genes
Promoter(<=1kb)	11260
Promoter(1-3kb)	422
5UTR	72
Intron	1516
Exon	215
3UTR	135
Downstream (1-3kb)	78
Distal Intergenic	577
None ATAC peak	5666

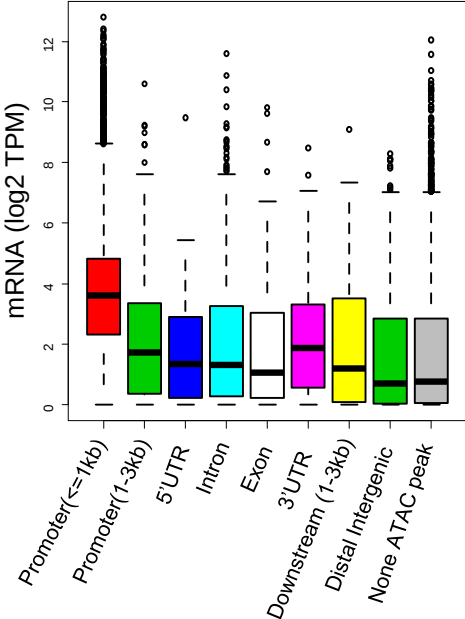
G Number of genes of CUT&Tag and ATAC-seq in promoter (<1kb) region

Status	Number of genes
CUT&Tag+ATAC	9195
CUT&Tag only	1249
ATAC only	2065
None CUT&Tag+ATAC	7432

H ATAC-seq peak/gene position vs. mRNA level in TCGA 73 normal lung tissues



I ATAC-seq peak/gene position vs. mRNA level TCGA 309 LUAD



J ATAC-seq peak/gene position vs. mRNA level TCGA 212 LUSC

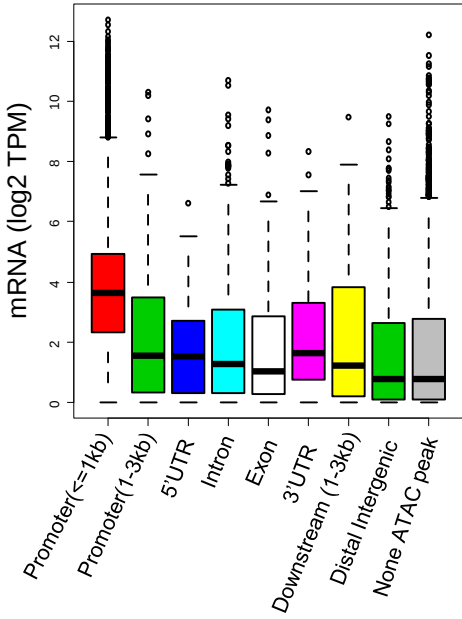
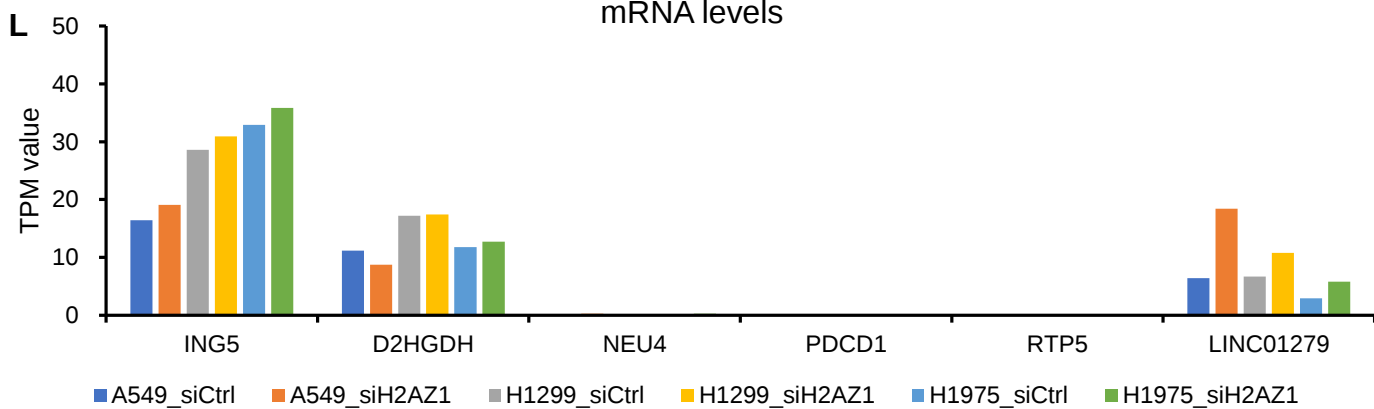
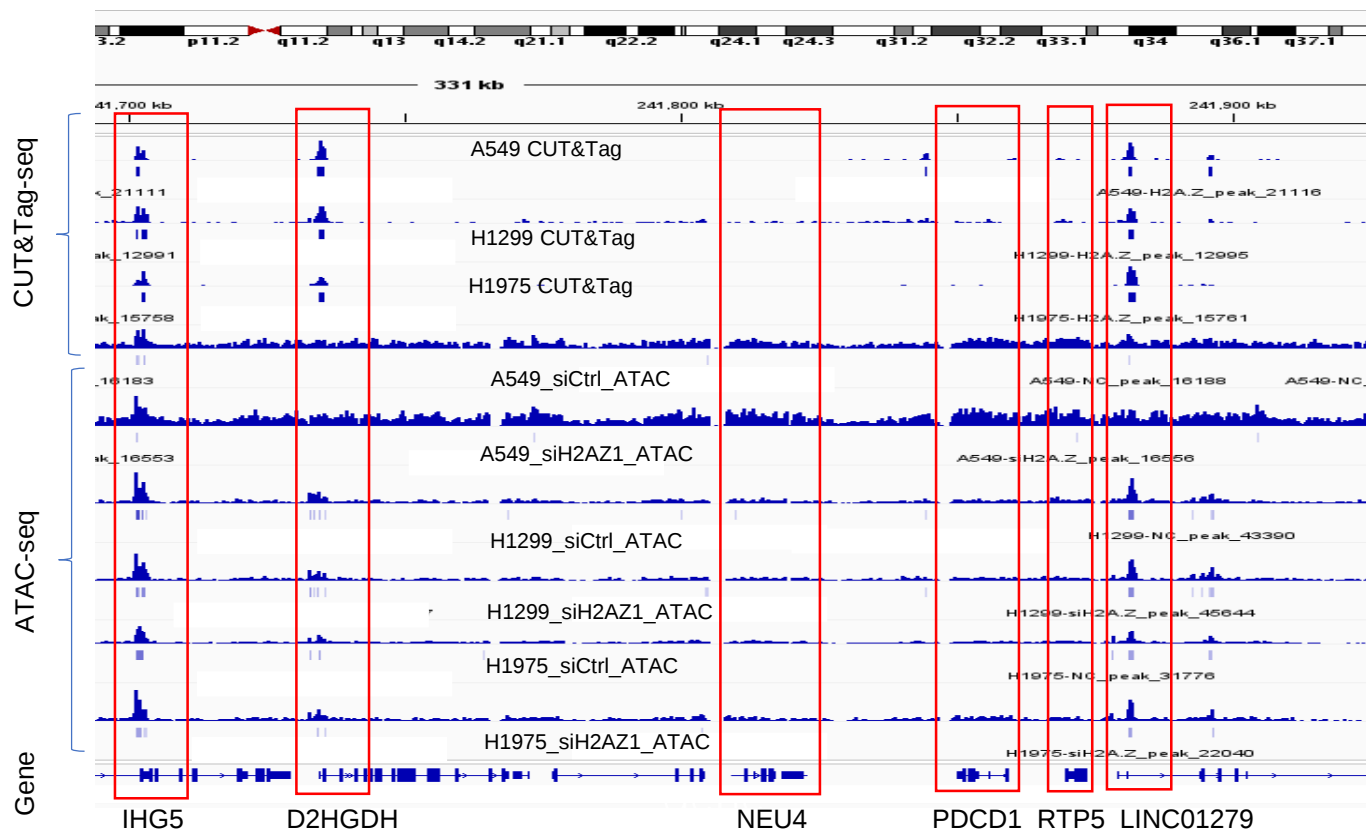
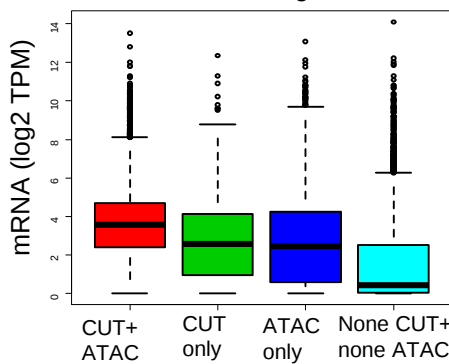
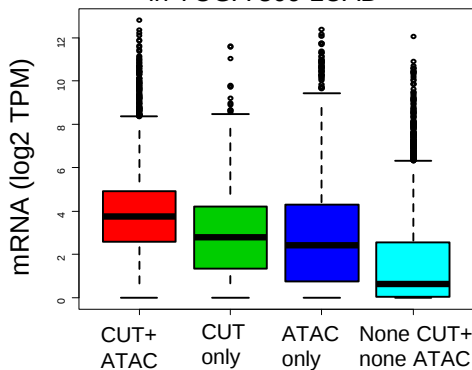
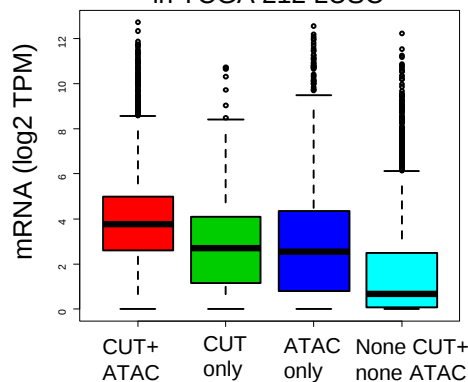


Figure S4**K** Correlation of CUT&Tag, ATAC-seq and mRNA levels**M** CUT&Tag and/or ATAC vs. mRNA level in promoter (<1kb) in TCGA 73 normal lung tissues**N** CUT&Tag and/or ATAC vs. mRNA level in promoter (<1kb) in TCGA 309 LUAD**O** CUT&Tag and/or ATAC vs. mRNA level in promoter (<1kb) in TCGA 212 LUSC

Supplementary Figure S4. Open chromatin accessibility at TSS leads to higher gene expression revealed by ATAC-seq and RNA-seq. (A) ATAC-seq signaling enriched around TSS with H2AZ1 silencing or not in 3 lung cancer cell lines. (B) Distribution of peaks on chromatin accessibility in H1299 and H1975 cell lines. (C) ATAC-seq showing the number of gain or loss DAR-associated promoter genes upon H2AZ1 silencing in 3 lung cancer cell lines. (D) ATAC-seq: After H2AZ1 knockdown, KEGG enrichments of loss DAR promoter genes in the H1299 cell line (since H1299 had the highest ATAC-seq signaling density/peaks). (E) The correlation between chromatin accessibility (ATAC-seq signal) and gene expression (RNA-seq) in H1975 and H1299 cell lines. (F) The gene number distribution from ATAC-seq in the H1975 siCtrl. (G) The number of genes from CUT&Tag-seq and ATAC-seq in the promoter (≤ 1 kb) region in the H1975 siCtrl. (H-J) The correlation of the chromatin accessibility and gene expression levels in lung tissues from TCGA RNA-seq data. (K, L) Example of the correlation among CUT&Tag-seq, ATAC-seq, and mRNA expression of RNA-seq. ING5, D2HGDH, and LINC01279 having H2AZ1 binding signaling (CUT&Tag) and open chromatin around TSS (ATAC-seq) have mRNA expression (RNA-seq) in all 3 lung cancer cell lines, while NEU4, PDCD1 (PD1) and RTP5 having not H2AZ1 binding signaling and open chromatin around TSS have not mRNA expression in all 3 lung cancer cell lines. (M-O) The correlation of location of H2AZ1 bound (CUT&Tag-seq), chromatin accessibility (ATAC-seq) in the promoter (≤ 1 kb), and gene expression (RNA-seq) in lung tissues from TCGA RNA-seq data.

Figure S5

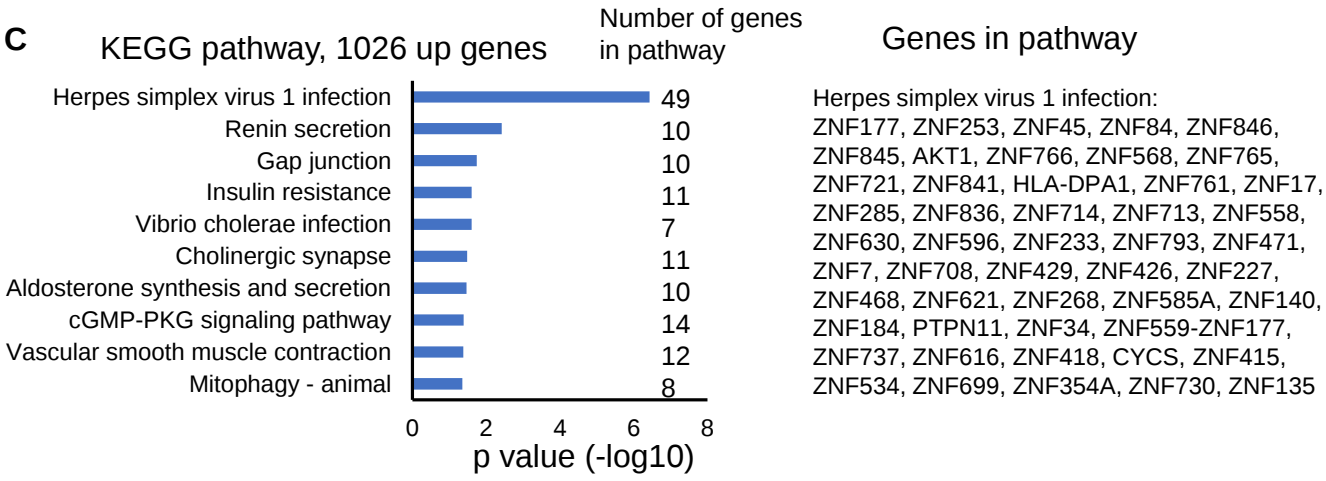
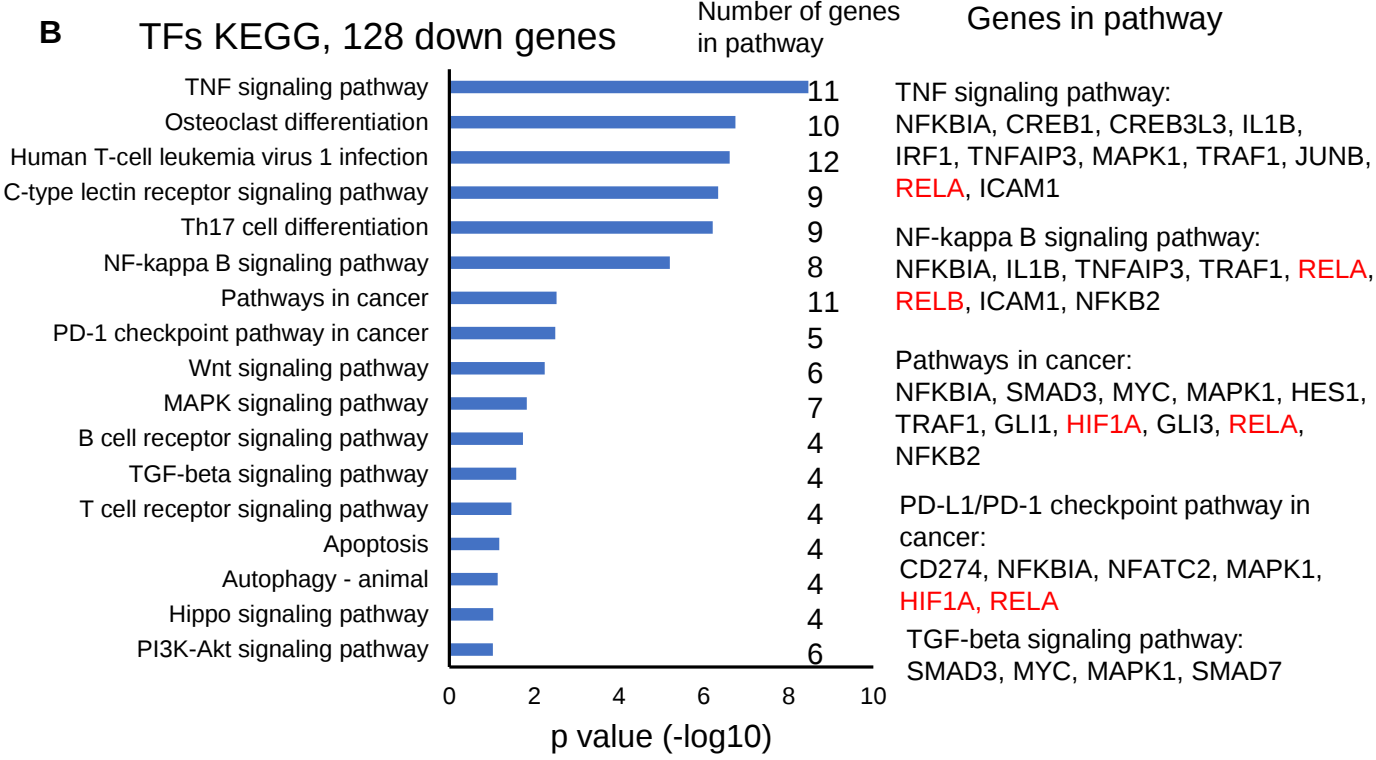
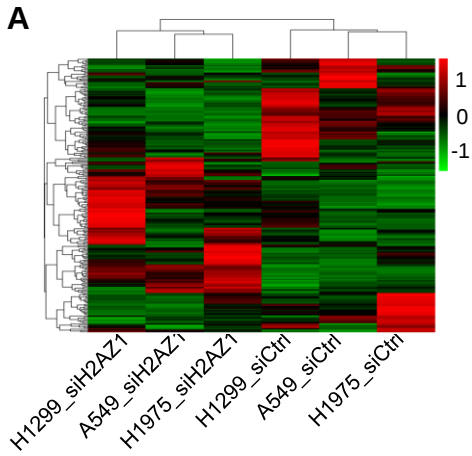
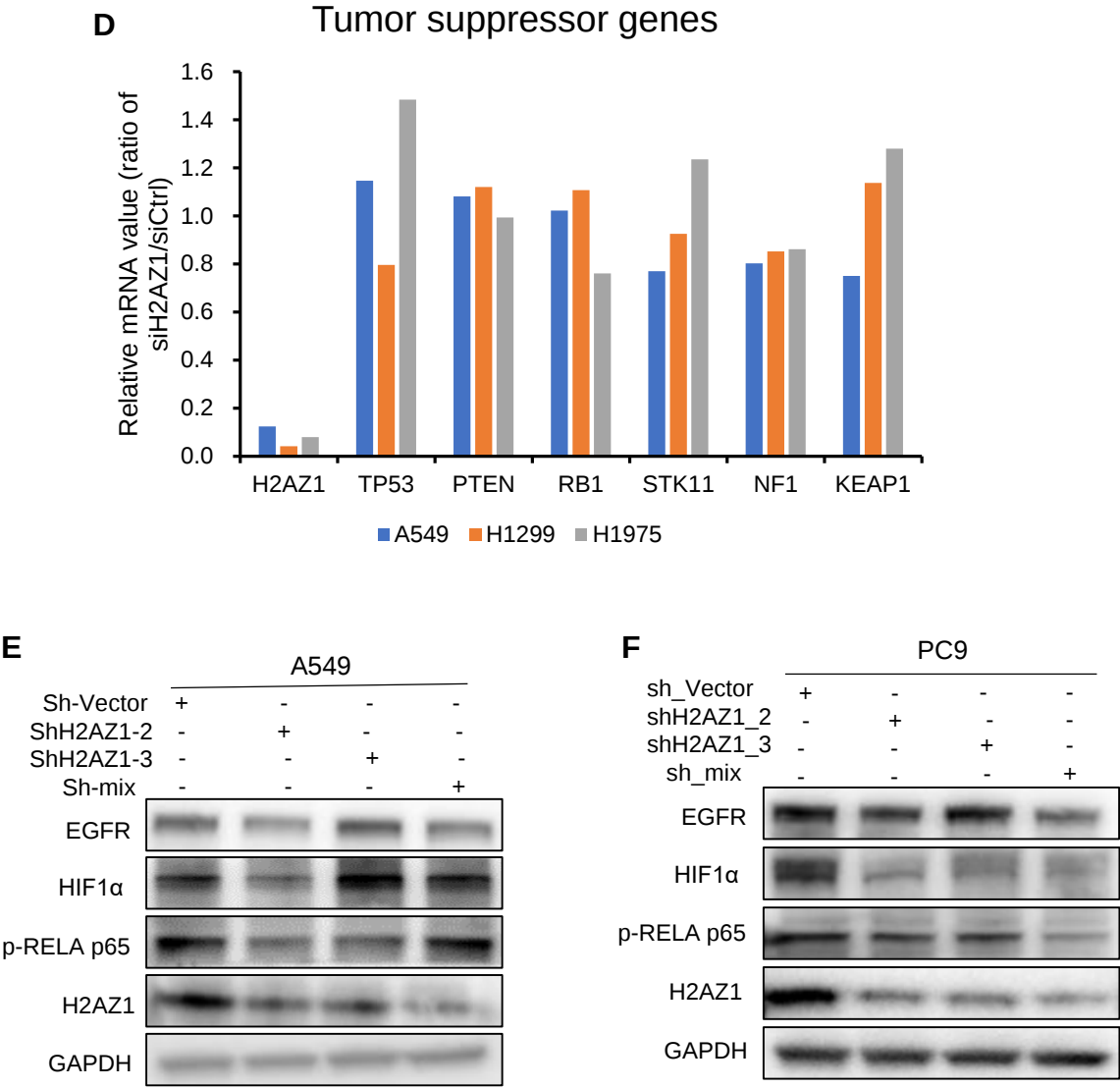


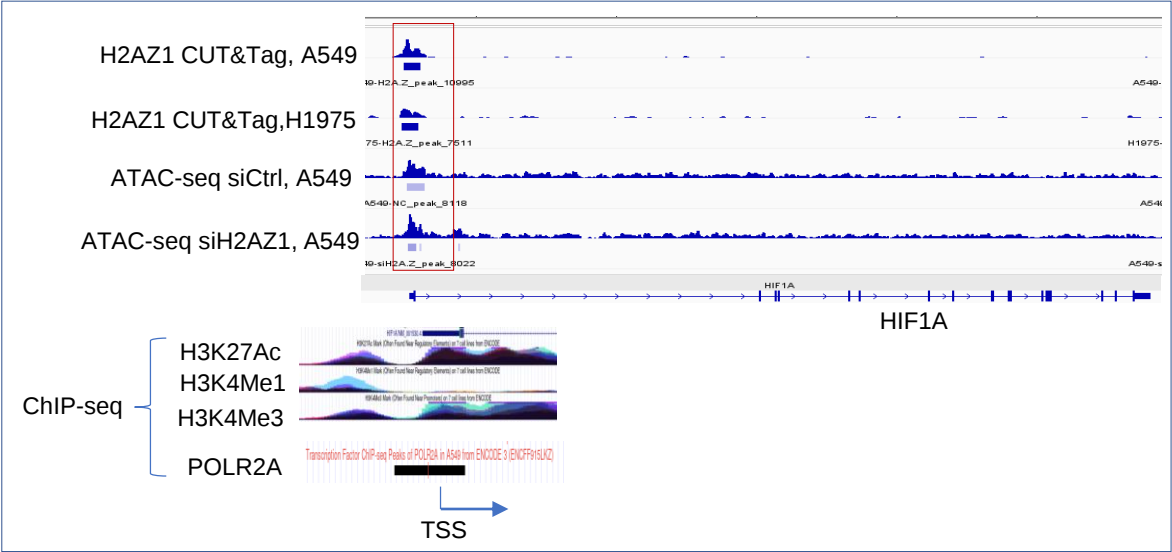
Figure S5



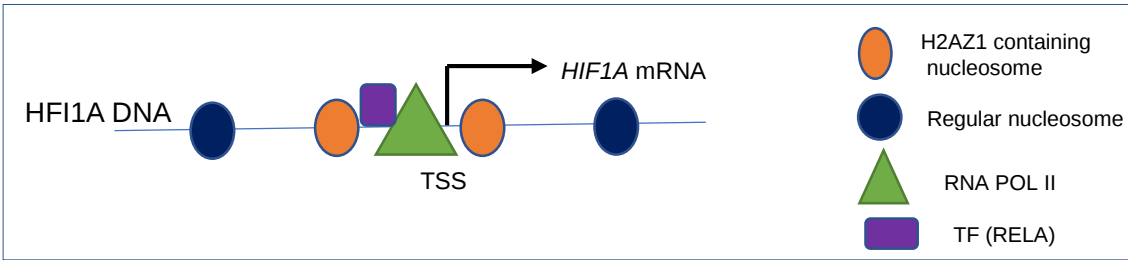
Supplementary Figure S5. H2AZ1 knockdown affects cancer-related pathways. (A) RNA-seq: heat map of differentially expressed genes after H2AZ1 knockdown in A549, H1975, and H1299 cell lines. (B) KEGG pathway enrichment using 128 down-regulated TFs after H2AZ1 knockdown. (C) KEGG pathway enrichment using 1026 up-regulated genes after H2AZ1 knockdown. (D) Upon H2AZ1 knockdown, the key tumor suppressor genes were not changed. (E, F) Stably H2AZ1 knockdown by shRNAs was constructed using lentivirus. The expression of EGFR, HIF1- α , and p-RELA p65 were reduced in A549 and PC9 cell lines.

Figure S6

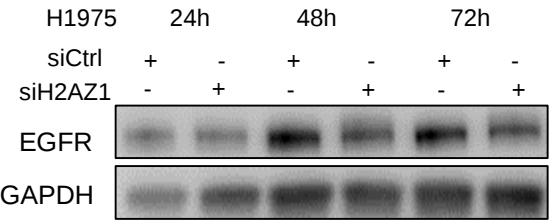
A H2AZ1 position on HIF1A promoter



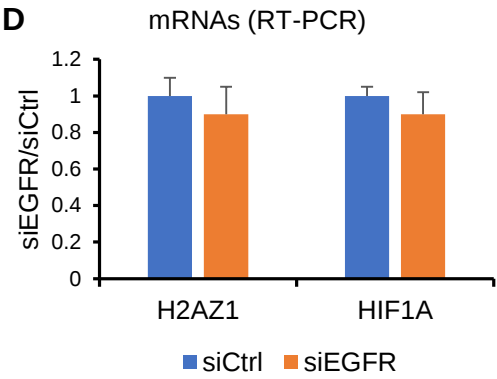
B Model of H2AZ1 deposits on HIF1A promoter



C

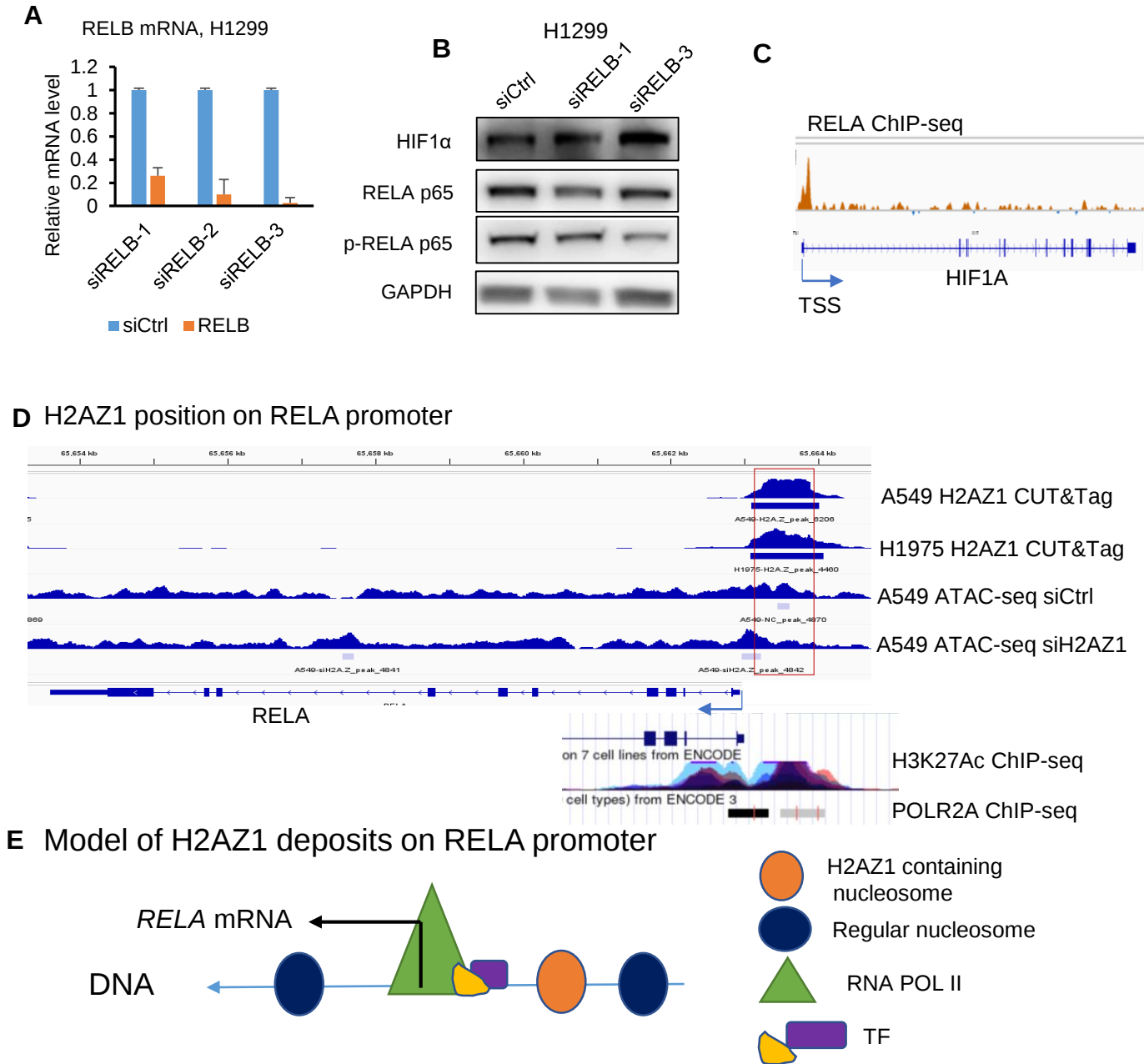


D



Supplementary Figure S6. H2AZ1 regulates EGFR through HIF1A. (A) CUT&Tag of H2AZ1 indicating H2AZ1 was bonded to HIF1A promoter with open chromatin for RNA polymerase binding. This region was also highly acetylated. The image of ChIP-seq of H3K27Ac and POLR2A was downloaded from the UCSC website. (B) Working model of H2AZ1 deposits on HIF1A promotor to regulate its expression. (C) EGFR protein was decreased at 24, 48, and 72 hours after H2AZ1 knockdown with siRNA in the H1975 cell line. (D) RNA-seq results show that mRNAs of H2AZ1 and HIF1A were not changed after EGFR knockdown with siRNA.

Figure S7



Supplementary Figure S7. H2AZ1 regulates HIF1A via RELA. (A) qRT-PCR assay validated the knockdown efficiency of RELB siRNAs in H1299 cells. (B) HIF1 α was not affected by RELB knockdown with siRNAs in H1299 cells. (C) RELA protein was bound to the HIF1A promoter by ChIP-seq in lymphoblastoid B cells. (data source: GSE117250 and GSE55105). (D) CUT&Tag of H2AZ1 indicating H2AZ1 was bonded to RELA promoter with open chromatin for RNA polymerase binding. This region was also highly acetylated. An image of ChIP-seq of H3K27Ac and POLR2A was downloaded from the UCSC website. (E) Working model of H2AZ1 deposits on RELA promoter.

Table S1 siRNA sequence, PCR primer and reagents

Reagent or Resource(5'-3')	Source	Identifier
siRNA/shRNA target sequence		
GCAGAAGUUUAUAGUACAATT	GenePharma	H2AZ1 siRNA target sequence-2-F
UUGUUACUUAACUUCUGCTT	GenePharma	H2AZ1 siRNA target sequence-2-R
CUGGUGGAUAAGUUCAAUATT	GenePharma	H2AZ1 siRNA target sequence-3-F
UAUUGAACUUAUCCACCAGTT	GenePharma	H2AZ1 siRNA target sequence-3-R
GCAGAAGTTATAGTAACAA	GenePharma	H2AZ1 shRNA target sequence-2
CTGGTGGATAAGTTCAATA	GenePharma	H2AZ1 shRNA target sequence-3
GCUGGAGACACAUAUAUUTT	GenePharma	HIF1A siRNA target sequence-1-F
AUAUGAUUGUGUCUCCAGCTT	GenePharma	HIF1A siRNA target sequence-1-R
GCCGCUCAUUUAUGAAUATT	GenePharma	HIF1A siRNA target sequence-2-F
UAUUCAUAAAUUGAGCGGCTT	GenePharma	HIF1A siRNA target sequence-2-R
CCACCACUGAUGAAUAAATT	GenePharma	HIF1A siRNA target sequence-3-F
UUUAAUUCAUACAGUGGUGGTT	GenePharma	HIF1A siRNA target sequence-3-R
GCAACAUGUCGAUGGACUUTT	GenePharma	EGFR siRNA target sequence-2-F
AAGUCCAUCGACAUGUUGCTT	GenePharma	EGFR siRNA target sequence-2-R
GGAGAUAAUGUGAGGAGAUTT	GenePharma	EGFR siRNA target sequence-3-F
AUCUCCAUCACUUAUCUCCTT	GenePharma	EGFR siRNA target sequence-3-R
GCACUACGGAUUCUGGUGTT	GenePharma	RELA siRNA target sequence-1-F
CACCAGAAUCCGUAAGUGCTT	GenePharma	RELA siRNA target sequence-1-R
UCAAAGCACUUGGACUCUUTT	GenePharma	RELA siRNA target sequence-3-F
AAGAGUCCAAGUCUUUGATT	GenePharma	RELA siRNA target sequence-3-R
CCAGGAGCACAGAUGAAUUTT	GenePharma	RELB siRNA target sequence-1-F
AAUUAUCUGUGCUCCUGGTT	GenePharma	RELB siRNA target sequence-1-R
GGAAGAUUCAACUGGGCAUTT	GenePharma	RELB siRNA target sequence-2-F
AUGCCCAGUUGAAUCUUCCTT	GenePharma	RELB siRNA target sequence-2-R
CCGUGACAGUCAACGUCUUTT	GenePharma	RELB siRNA target sequence-3-F
AAGACGUUGACUGUCACGGTT	GenePharma	RELB siRNA target sequence-3-R
GCUUGUACCUGCAGGAUCUTT	GenePharma	MYC siRNA target sequence-1-F
AGAUCUGCAGGUACAAGCTT	GenePharma	MYC siRNA target sequence-1-R
GGAAGAAUUCGAUGUUGUUTT	GenePharma	MYC siRNA target sequence-2-F
AACAACAUCGAUUUCUUCCTT	GenePharma	MYC siRNA target sequence-2-R
GUGCCACUAACUACAUUUATT	GenePharma	MET siRNA target sequence-1-F
UAAAUGUAGUUAGUGGCACTT	GenePharma	MET siRNA target sequence-1-R
GUCCCCGAGAAUGGUCAUAATT	GenePharma	MET siRNA target sequence-3-F
UUAUGACCAUUCUCGGGACTT	GenePharma	MET siRNA target sequence-3-R

Stable overexpressed sequences

ATGGCTGGCGGTAAGGCTGGAAAGGACTCCGGAAA	this study	Stable overexpressed sequence-H2AZ1
GGCCAAGACAAAGGCGTTTCCCGCTCGCAGAGAG		
CCGGCTTGCACTTCCAGTGGGCCGTATTCATCGAC		
ACCTAAATCTAGGACGACCAGTCATGGACGTGTGG		
GCGCGACTGCCGCTGTGTACAGCGCAGCCATCCTGG		
AGTACCTACCGCAGAGGTACTGAACTGGCAGGA		
AATGCATCAAAGACTTAAAGGTAAAGCGTATTACC		
CCTCGTCACTTGCAACTTGCTATTCGTGGAGATGAA		
GAATTGGATTCTCTCATCAAGGCTACAATTGCTGGT		
GGTGGTGTATTCCACACATCCACAAATCTCTGATTG		
GGAAGAAAGGACAACAGAAGACTGTC		

RT-PCR primers

CCTCACCGCAGAGGTTACTTG	BGI Genomics	H2AZ1-F
GTTGCAAGTGACGAGGGGTA	BGI Genomics	H2AZ1-R
CCACTAGGCGCTCACTGTT	BGI Genomics	GAPDH-F
CCAATACGACCAATCCGTTGAC	BGI Genomics	GAPDH-R
ATGTGGAGATCATTGAGCAGC	BGI Genomics	RELA-F
CCTGGTCCTGTGTAGCCATT	BGI Genomics	RELA-R
GAACGTCGAAAAGAAAAGTCTCG	BGI Genomics	HIF1A-F
CCTTATCAAGATGCGAACTCACA	BGI Genomics	HIF1A-R
AGGCACGAGTAACAAGCTCAC	BGI Genomics	EGFR-F
ATGAGGACATAACCAGCCACC	BGI Genomics	EGFR-R

Regents	Source	Identifier
Antibodies		
H2A.Z	Cell Signaling Technology (CST)	2718
H2A.Z (CUT&Tag)	Abcam	ab4174
GAPDH	Cell Signaling Technology (CST)	5174s
EGF Receptor (D38B1) XP® Rabbit mAb	Cell Signaling Technology (CST)	4267s
HIF-1α (D1S7W) XP® Rabbit mAb	Cell Signaling Technology (CST)	36169s
NF-κB p65 (D14E12) XP® Rabbit mAb	Cell Signaling Technology (CST)	8242s
Phospho-NF-κB p65(Ser536)(93H1) Rabbit mAb	Cell Signaling Technology (CST)	3033s
Anti-Flag	Sigma	SLCJ4124
Anti-rabbit IgG	Cell Signaling Technology (CST)	7074s
Anti-mouse IgG	Cell Signaling Technology (CST)	7076

Chemical reagents

Pierce™ BCA Protein Assay kit	Thermo Fisher Scientific	23225
RPMI 1640	Gibco	C11875500BT
Fetal bovine serum	Gibco	10270-106
Penicillin-streptomycin	Gibco	15140-122
Phosphate Buffered Saline	Gibco	C10010500BT
Puromycin	Genechem	REVG1001
Matrigel Basement Membrane Matrix	Corning	356234
RIPA buffer	Cell Signaling Technology	9806S
PMSF	Beyotime	ST506-2
Protease & phosphatase inhibitor	Thermo Fisher Scientific	1861280
RNA Lipofectamine RNAi max	Invitrogen	13778-150
Opti-MEM medium	Gibco	31985-062
Trypsin-EDTA (1X)	Gibco	25200-056
Methyl alcohol	Guangdong Guanghua	CAS67-56-1
Trizol reagent	Ambion	15596026
Tris-MOPS-SDS Running Buffer Powder	GenScript	M00138
SDS-PAGE gel	GenScript	M00654

TBS-T buffer (10x)	Boster Biological Technology	AR0195-10
Prestained Protein Ladder	Gene Tech	R1001-002
BSA	Sigma-Aldrich	9048468
UltraSignal ECL Reagent	4A Biotech	4AW011-500
Cell Counting Kit	Yeasen	40203ES92
FastPure Cell/Tissue Total RNA Isolation kit V2	Vazyme	RC112-01
PrimeScript™ RT reagent Kit with gDNA Eraser	Takara	RR047A
TB Green Premix EX Taq II	Takara	RR820L
PVDF Western Blotting Membranes	Roche	3010040001
Tris-glycine-SDS transfer buffer (10x)	Boster Biological Technology	AR1152-10
Cell lines		
H838	Cell Bank of the Chinese Academy of Sciences	CC0224
A549	Cell Bank of the Chinese Academy of Sciences	CC0202
H1975	Cell Bank of the Chinese Academy of Sciences	CC0206
PC9	Cell Bank of the Chinese Academy of Sciences	CC0204
H1299	Cell Bank of the Chinese Academy of Sciences	CC0203