Cancer Cell, Volume *33*

Supplemental Information

lncRNA Epigenetic Landscape Analysis Identifies *EPIC1* as an Oncogenic lncRNA that Interacts with MYC and Promotes Cell-Cycle Progression in Cancer

Zehua Wang, Bo Yang, Min Zhang, Weiwei Guo, Zhiyuan Wu, Yue Wang, Lin Jia, Song Li, The Cancer Genome Atlas Research Network, Wen Xie, and Da Yang

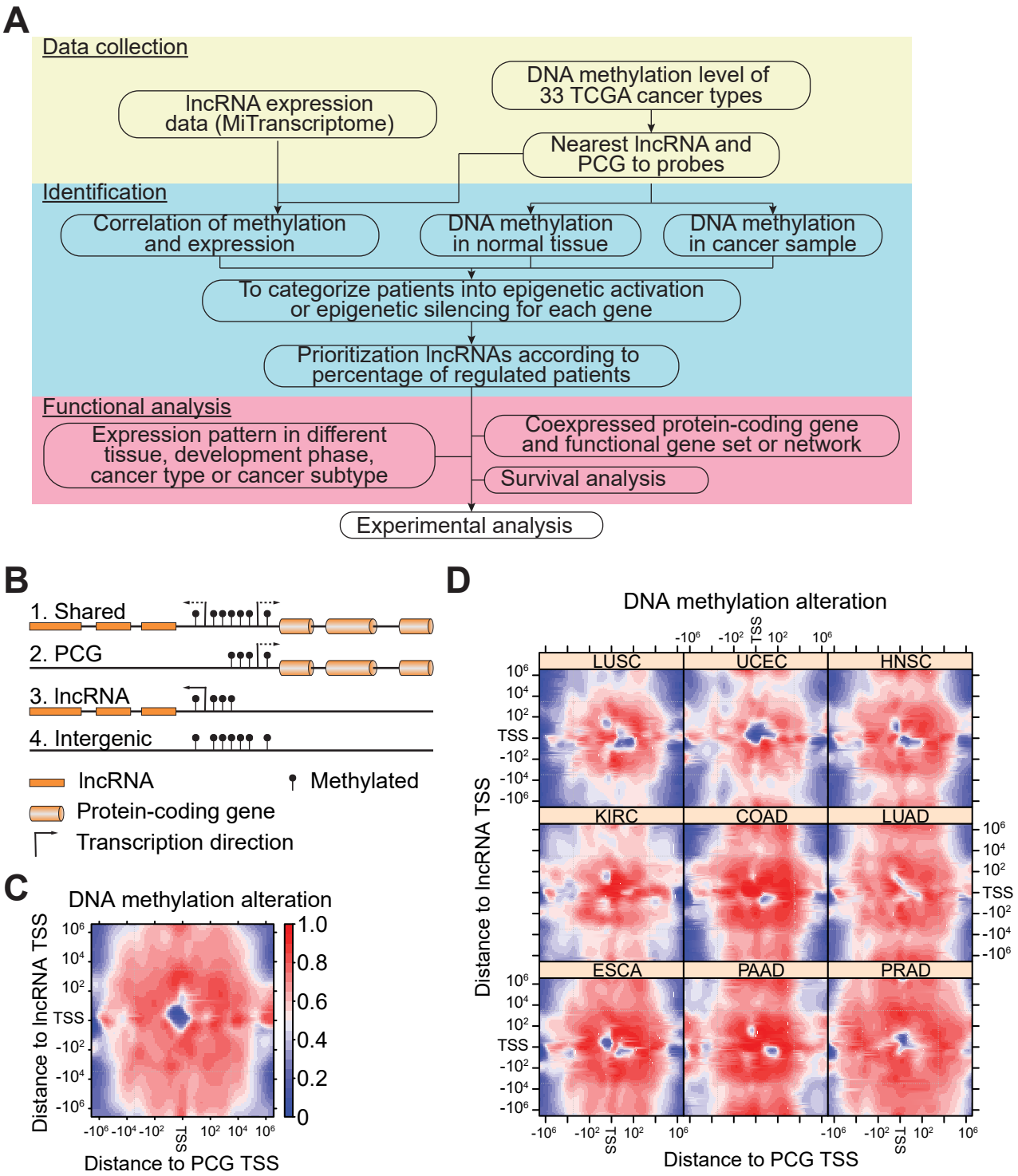
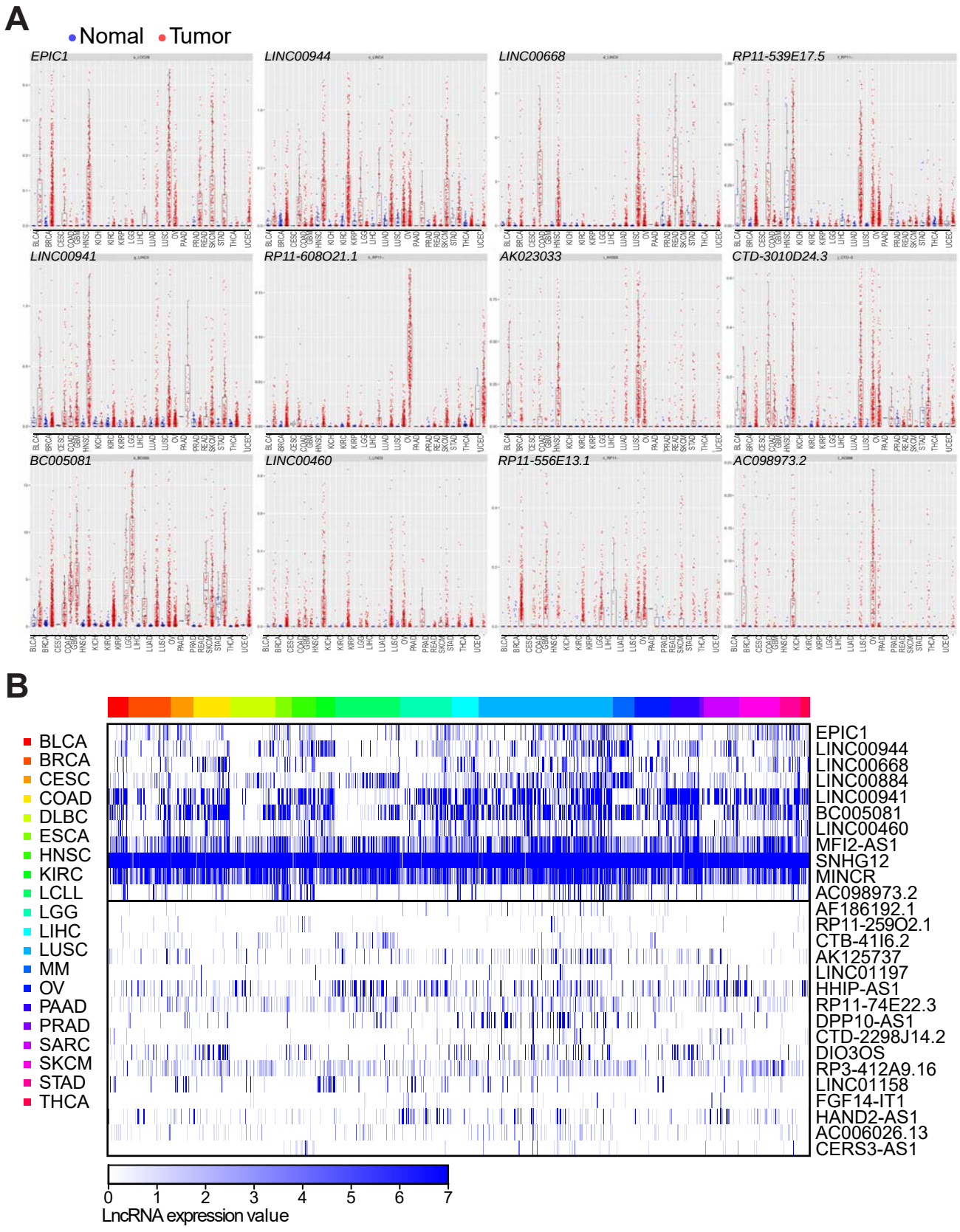
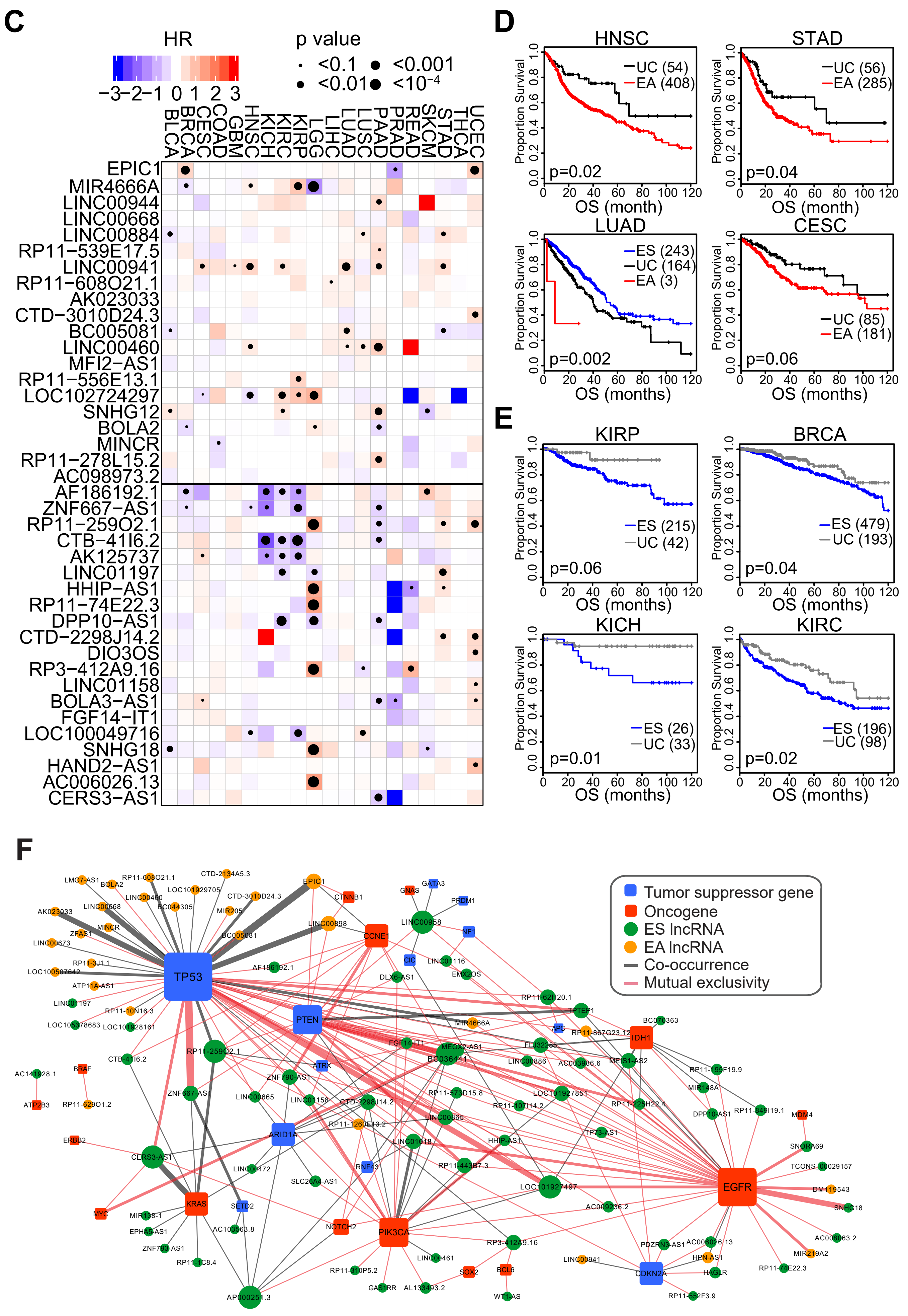


Figure S1. LncRNAs are both epigenetically activated and silenced by DNA methylation alteration in the promoter region, Related to Figure 1. (**A**) Flow chart of identification and functional analysis of EA and ES lncRNAs in cancer. (**B**) Schematic of the annotation of DNA methylation probes to protein-coding genes and lncRNA genes. (**C**) Differential DNA methylation between breast cancer and normal tissues. Density plot of average differential DNA methylation (indicated by beta values) within 100 windows in ± 1000 kb from TSS sites are shown. The windows are arranged based on their distances to protein coding gene TSS (x-axis) and lncRNA gene TSS (y-axis). The hypermethylation region in tumor is shown as red, whereas the hypomethylation region is shown as blue. (**D**) Differential DNA methylation between cancer and matched normal tissues in nine cancer types.


A
Nomal
Tumor
EPIC1
LINC00944
LINC00668
RP11-539E17.5
LINC00941
RP11-608O21.1
AK023033
CTD-3010D24.3
BC005081
LINC00460
RP11-556E13.1
AC098973.2
B
BLCA
BRCA
CESC
COAD
DLBC
ESCA
HNSC
KIRC
LCLL
LGG
LIHC
LUSC
MM
OV
PAAD
PRAD
SARC
SKCM
STAD
THCA
EPIC1
LINC00944
LINC00668
LINC00884
LINC00941
BC005081
LINC00460
MFI2-AS1
SNHG12
MINCR
AC098973.2
AF186192.1
RP11-259O2.1
CTB-41I6.2
AK125737
LINC01197
HHIP-AS1
RP11-74E22.3
DPP10-AS1
CTD-2298J14.2
DIO3OS
RP3-412A9.16
LINC01158
FGF14-IT1
HAND2-AS1
AC006026.13
CERS3-AS1
0
1
2
3
4
5
6
7
LncRNA expression value



C
HR
−3 −2 −1 0 1 2 3
p value
<0.1
<0.01
<0.001
<10^-4
BLCA BRCA CESC COAD GBM HNSC KICH KIRC KIRP LGG LIHC LUAD LUSC PAAD PRAD READ SKCM STAD THCA UCEC
EPIC1
MIR4666A
LINC00944
LINC00668
LINC00884
RP11−539E17.5
LINC00941
RP11−608O21.1
AK023033
CTD−3010D24.3
BC005081
LINC00460
MFI2−AS1
RP11−556E13.1
LOC102724297
SNHG12
BOLA2
MINCR
RP11−278L15.2
AC098973.2
AF186192.1
ZNF667−AS1
RP11−259O2.1
CTB−41I6.2
AK125737
LINC01197
HHIP−AS1
RP11−74E22.3
DPP10−AS1
CTD−2298J14.2
DIO3OS
RP3−412A9.16
LINC01158
BOLA3−AS1
FGF14−IT1
LOC100049716
SNHG18
HAND2−AS1
AC006026.13
CERS3−AS1
D
HNSC
UC (54)
EA (408)
p=0.02
STAD
UC (56)
EA (285)
p=0.04
LUAD
ES (243)
UC (164)
EA (3)
p=0.002
CESC
UC (85)
EA (181)
p=0.06
Proportion Survival
OS (month)
OS (months)
E
KIRP
ES (215)
UC (42)
p=0.06
BRCA
ES (479)
UC (193)
p=0.04
KICH
ES (26)
UC (33)
p=0.01
KIRC
ES (196)
UC (98)
p=0.02
Proportion Survival
OS (months)
F
Tumor suppressor gene
Oncogene
ES lncRNA
EA lncRNA
Co-occurrence
Mutual exclusivity
TP53
PTEN
ARID1A
KRAS
PIK3CA
EGFR
IDH1
CCNE1
CDKN2A

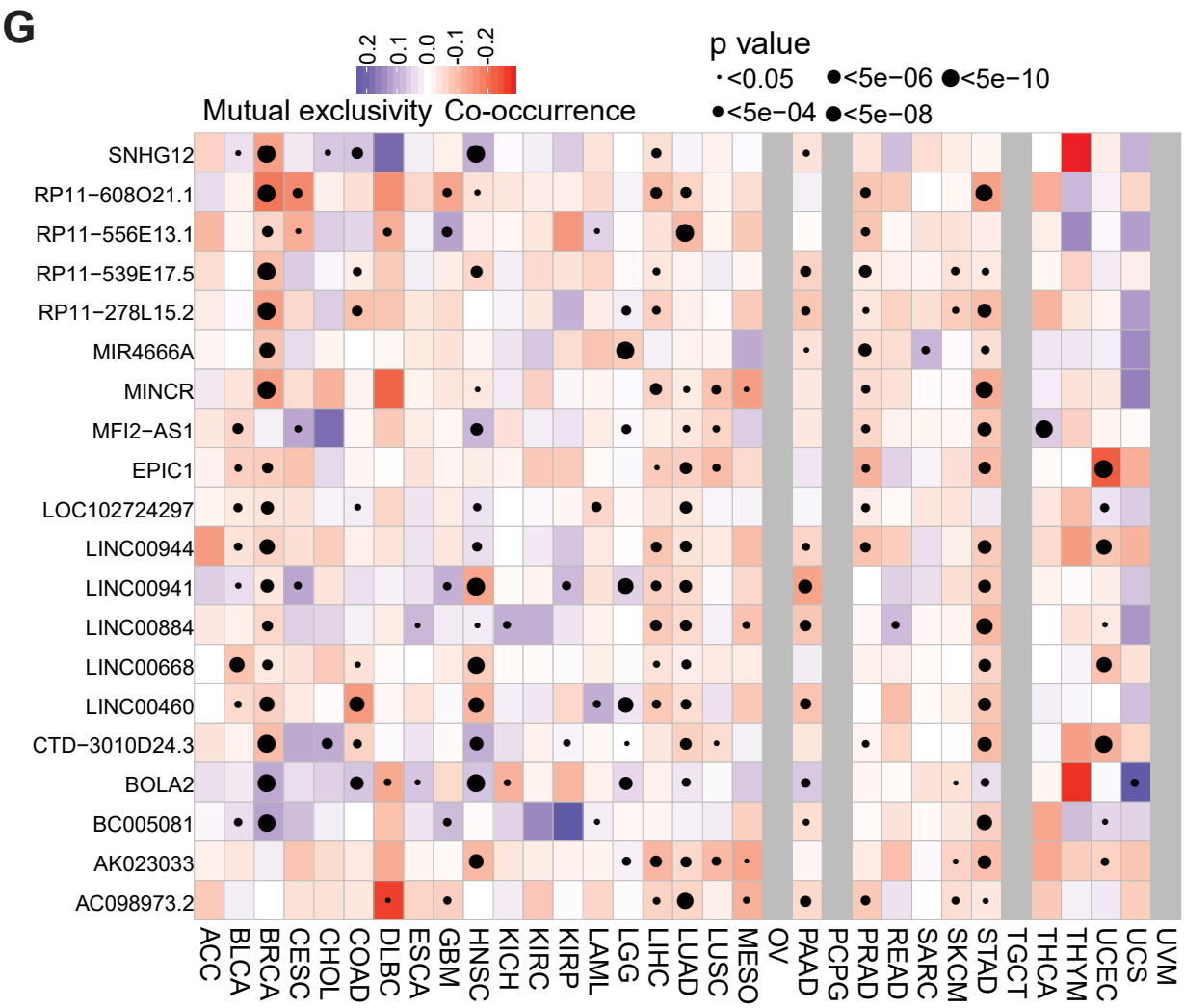
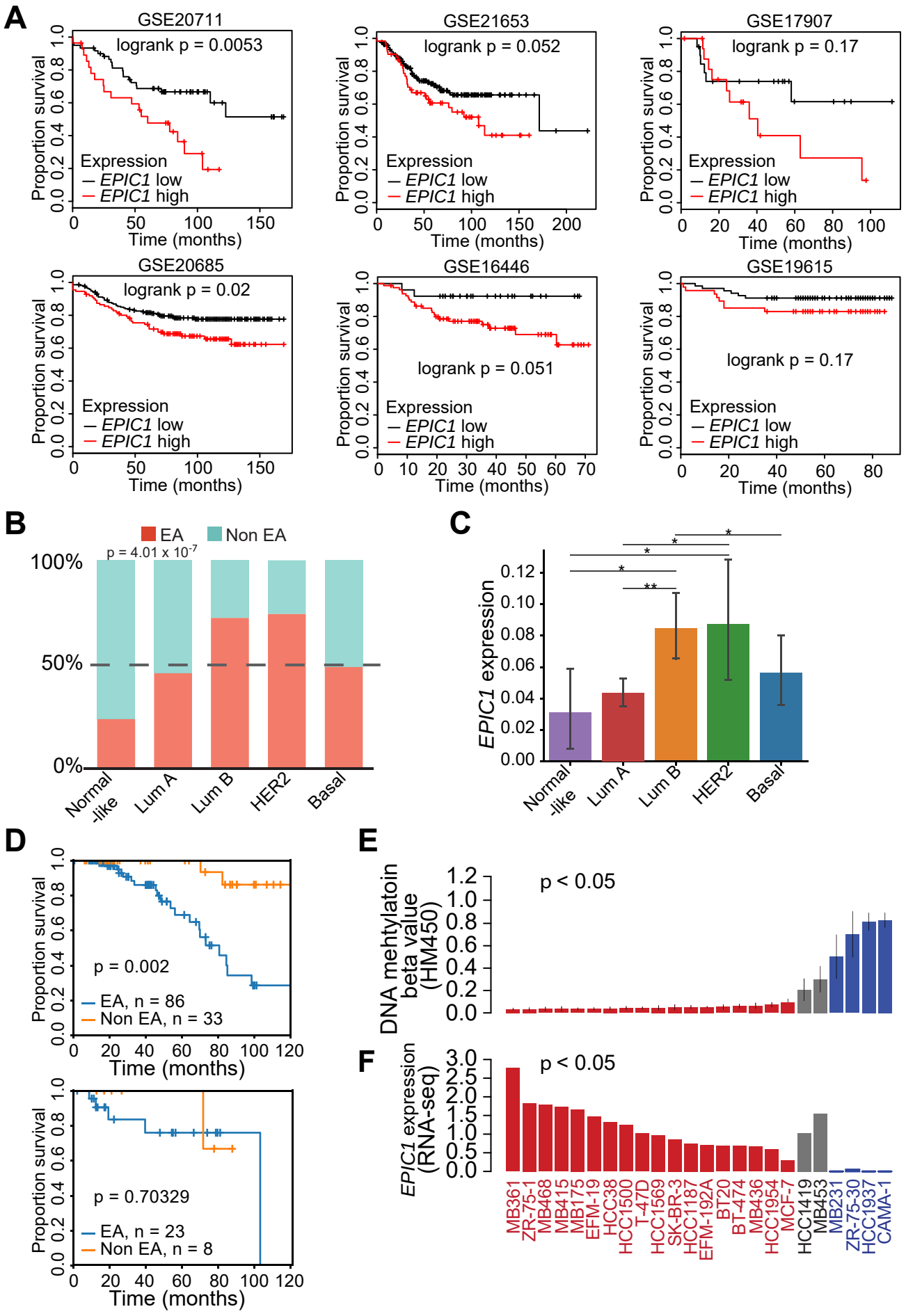


Figure S2. EA lncRNAs exhibit an "on or off" pattern with completely no expression in normal tissues, and are associated with tumor survival and tumor gene alterations, Related to Figure 2. (**A**) Representative expression pattern of EA lncRNAs in multiple cancer types, compared to normal tissues. Blue dot denotes normal tissue, and red dot denotes tumor. (**B**) Expression pattern of EA and ES lncRNAs in cancer cell lines. White denotes low expression, and dark blue denotes high expression. (**C**) Correlation of EA lncRNAs (top panel) and ES lncRNAs (bottom panel) with survival. The size of each cycle indicates the p value calculated by Cox regression survival analysis. The heatmap indicates the hazard ratio (HR). (**D, E**) Overall survival (OS) of representative EA lncRNA *LINC00941* (**D**) and ES lncRNA *AF186192.1* (**E**) in multiple cancer types. UC, unchanged; EA, epigenetic activation; ES, epigenetic silencing. (**F**) Mutual exclusivity and co-occurrence network for EA and ES lncRNAs with tumor suppressor genes and oncogenes. (**G**) Mutual exclusivity and co-occurrence between 20 EA lncRNAs and *TP53* mutation in 33 cancer types. The size of each circle indicates p value calculated by Fisher's exact test. The color indicates mutual exclusivity (blue) or co-occurrence (red).


A
GSE20711
logrank p = 0.0053
GSE21653
logrank p = 0.052
GSE17907
logrank p = 0.17
GSE20685
logrank p = 0.02
GSE16446
logrank p = 0.051
GSE19615
logrank p = 0.17
Proportion survival
Time (months)
Expression
EPIC1 low
EPIC1 high
B
EA
Non EA
p = 4.01 x 10-7
100%
50%
0%
Normal -like
Lum A
Lum B
HER2
Basal
C
EPIC1 expression
D
p = 0.002
EA, n = 86
Non EA, n = 33
p = 0.70329
EA, n = 23
Non EA, n = 8
E
DNA mehtylatoin beta value (HM450)
p < 0.05
F
EPIC1 expression (RNA-seq)
p < 0.05
MB361
ZR-75-1
MB468
MB415
MB175
EFM-19
HCC38
HCC1500
T-47D
HCC1569
SK-BR-3
HCC1187
EFM-192A
BT20
BT-474
MB436
HCC1954
MCF-7
HCC1419
MB453
MB231
ZR-75-30
HCC1937
CAMA-1

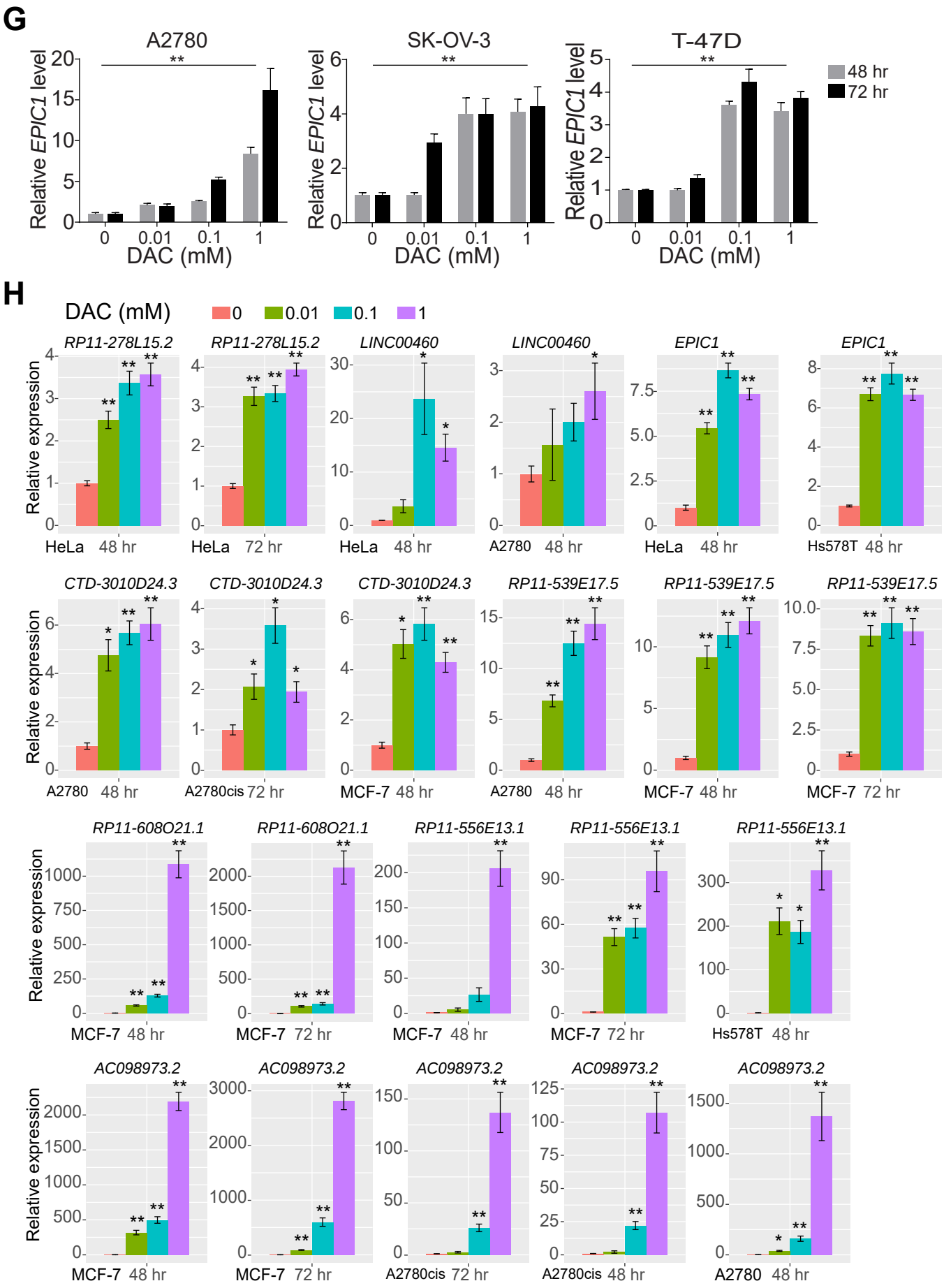
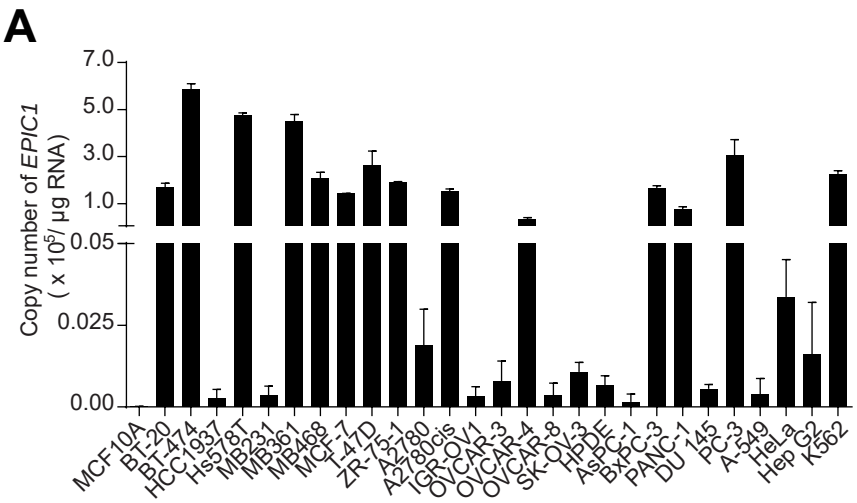
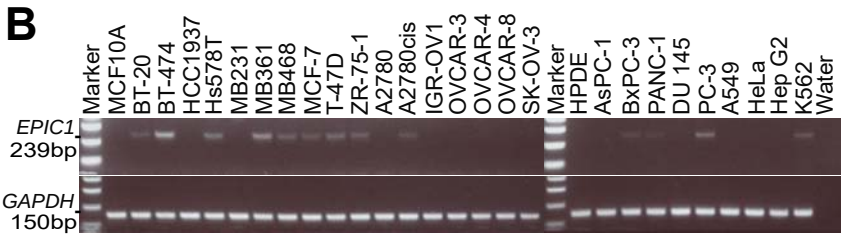

G
A2780
SK-OV-3
T-47D
Relative EPIC1 level
DAC (mM)
0
0.01
0.1
1
48 hr
72 hr
H
DAC (mM)
0
0.01
0.1
1
RP11-278L15.2
LINC00460
EPIC1
CTD-3010D24.3
RP11-539E17.5
RP11-608O21.1
RP11-556E13.1
AC098973.2
Relative expression
HeLa 48 hr
HeLa 72 hr
A2780 48 hr
Hs578T 48 hr
A2780cis 72 hr
MCF-7 48 hr
MCF-7 72 hr
A2780cis 48 hr

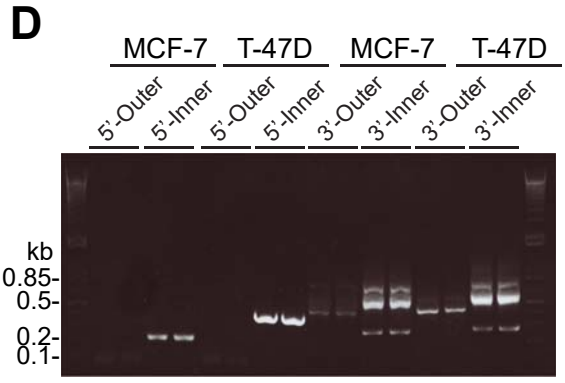
Figure S3. *EPIC1* is robustly correlated with poor survival in 1892 breast cancer samples and can be epigenetically activated by decitabine, Related to Figure 3. (**A**) The association between *EPIC1* expression and breast cancer survival in six independent breast cancer cohorts. (**B**) The association between *EPIC1*'s EA status and five breast cancer subtypes. The overall p value is calculated by chi-square test. (**C**) Comparison of *EPIC1*'s expression in five subtypes (**B**). The p value is calculated by pair-wise Wilcoxon Rank Sum test. *$p < 0.05$, **$p < 0.01$. (**D**) *EPIC1*'s correlation with survival in luminal B (top) and HER2 (bottom) subtypes. (**E, F**) Association of DNA methylation status (**E**) and *EPIC1* expression (**F**) in 24 breast cancer cell lines based on lncRNA expression (CCLE) and genome-wide DNA methylation data (GSE44837). (**G**) qRT-PCR analysis of *EPIC1* expression in A2780, SK-OV-3, and T-47D cells treated with decitabine (DAC) as indicated time and dosage. (**H**) qRT-PCR analysis of eight EA lncRNA expression in five cell lines treated with DAC. Error bars indicate mean ± SD, n = 3 for technical replicates. *$p < 0.05$, **$p < 0.01$.

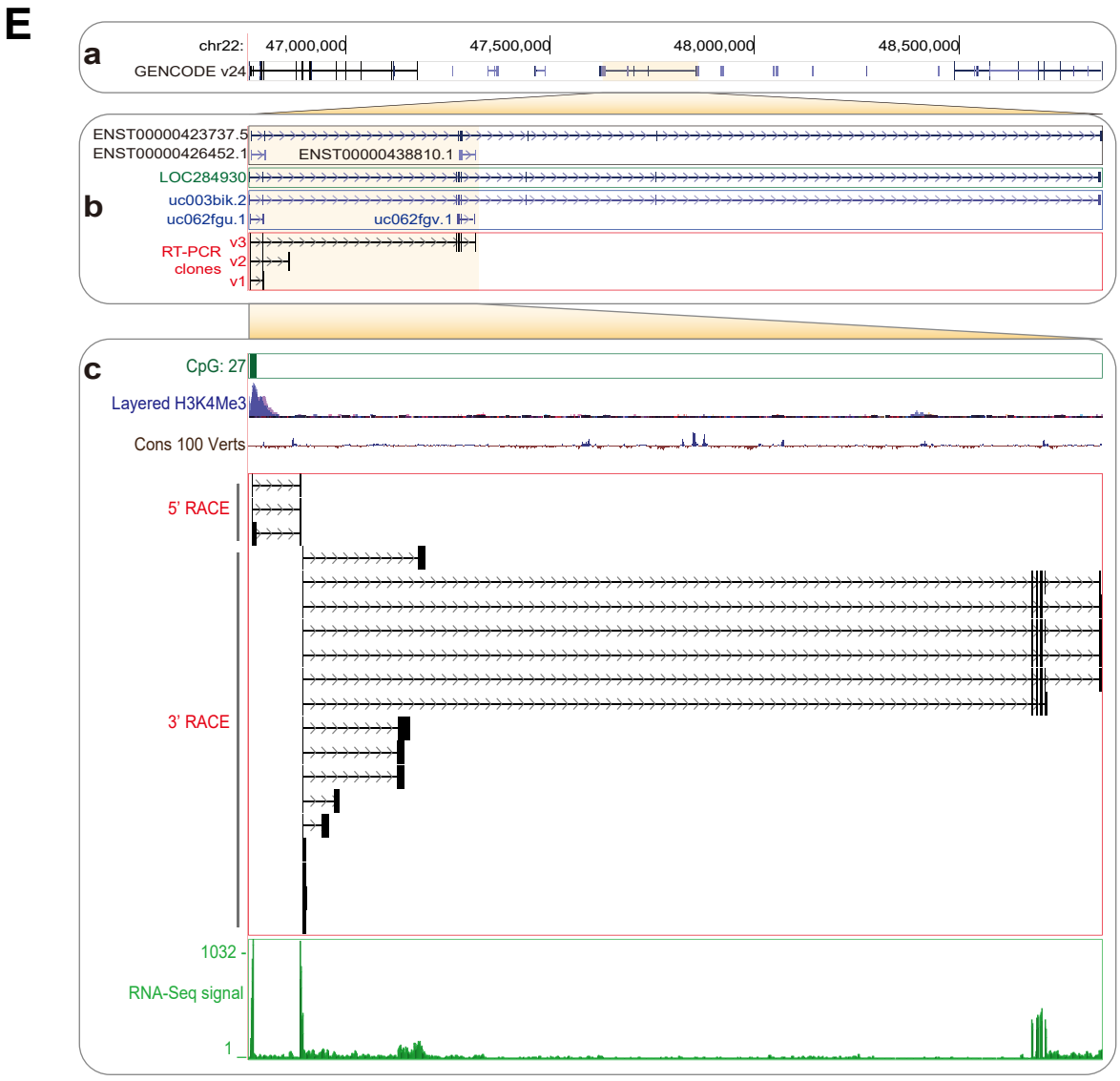


C

Cell line	RNA copy # per cell
MCF10A	0.50 ± 0.26
MCF-7	82.51 ± 0.63
T-47D	53.95 ± 3.10
ZR-75-1	70.26 ± 4.34
MB231	5.36 ± 0.64
A2780	2.56 ± 0.31
A2780cis	19.60 ± 0.83
OVCAR-4	20.87 ± 4.35







F

>EPIC1 v1

AGTCCGCCATTGCAAACACGAAGCTCTTCCAGAAACGCCCTCACAGACACCCCGGAAGTCACGTACCCACTCTGTAGGTGCCCCGGGGCACAGGCAAGCG
GACGAGCCAGTTATCCCTCAGAGCTCCTGCTGCCTCGCCCGCTTTCTCTCGGAAACGTGAAGTGTGGCCTCAGCTGAAAGTGAGGTGGGCCTCATTCAAT
CAGTTGAATTCTTCAAGAGAGAAAAACTGAAGTCCCTTAGAAGGAAAGAGTTCTGCCTTCAGACTGTCTTTGAACTTAAGACTGTAGCGTCGACTCCTGC
CGGAATTTCCAGCCTGCTGGCCAGCTCTGCAGATTCACACTTGCCAGCCTCCACAATCGTGTGAGCCAATTCCTTAACTTCTCTTTCTCCGTGTATCCCT
TTGGTGCTGCCTCTCTGGGGAGCCCTGACTAATATGCATGCAGATGATACGGTGCCTGGCATTCTGAATACATGCACTAAATCCACCACTTTTCCCCATT
TATAGATTTGGATTAACACACTAACTTACTCATATCTGCAAGTATAAATAAAAAAAATTGCTGGTGC

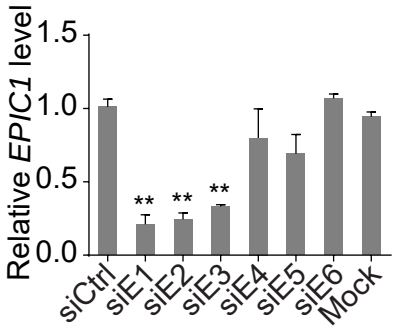
>EPIC1 v2

AGTCCGCCATTGCAAACACGAAGCTCTTCCAGAAACGCCCTCACAGACACCCCGGAAGTCACGTACCCACTCTGTAGGTGCCCCGGGGCACAGGCAAGCG
GACGAGCCAGTTATCCCTCAGAGCTCCTGCTGCCTCGCCCGCTTTCTCTCGGAAACGTGAAGTGTGGCCTCAGCTGAAAGTGAGGTGGGCCTCATTCAAT
CAGTTGAATTCTTCAAGAGAGAAAAACTGAAGTCCCTTAGAAGGAAAGAGTTCTGCCTTCAGACTGTCTTTGAACTTAAGACTGTAGCGTCGACTCCTGC
CGGAATTTCCAGCCTGCTGGCCAGCTCTGCAGATTCACACTTGCCAGCCTCCACAATCTTCCTGGATTTGAAACTGAAGAAGCAAGCAATCTGGAAATGT
CAGTGGATGCACACAAAGAAACAACCGCAAAAGCCTGCTCGCTCTAGCCAAGGGACAAGAATAGGGGCAGTCCATCAAGACAGAATCCTTTTAAAAAATA
ACCACTCCACTCCAGCAATACCACAGAAGAATCTGGCTGTACCCCAGGTACATCAGCAAAGATAACCTTTACCTAGCAGTAAAGAGGTCCCCCTTACACT
GGGAGCCCTAGTGAAGAGCAGGGACTTTCACCCCCACTTAGCAGTGATGGGGCCCCACCCACCACAGTGCCAGCAGAGACCATGTGGGAGCCAGAATCCT
CATCCCTACCCAGCAGTAACAAGGAGCCCTCCTCACTGCGGGCATCAAGGGTGAGTGAGTGCAAAACCTGGGTGTCACTCGGAAGGGAAGAATGGTGTCT
CCTTCCTTCCCATCCCCTGCCAGAGTGATATCACTAGGAAAAAG

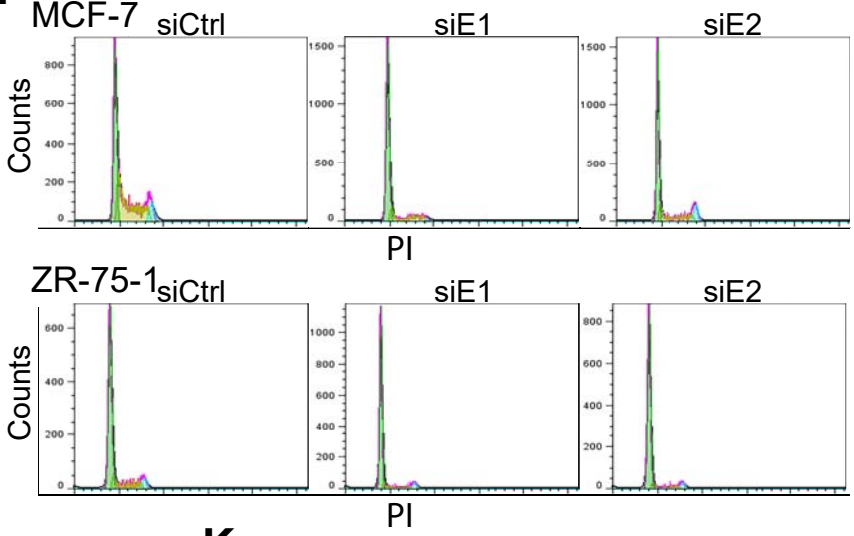
>EPIC1 v3

AGTCCGCCATTGCAAACACGAAGCTCTTCCAGAAACGCCCTCACAGACACCCCGGAAGTCACGTACCCACTCTGTAGGTGCCCCGGGGCACAGGCAAGCG
GACGAGCCAGTTATCCCTCAGAGCTCCTGCTGCCTCGCCCGCTTTCTCTCGGAAACGTGAAGTGTGGCCTCAGCTGAAAGTGAGGTGGGCCTCATTCAAT
CAGTTGAATTCTTCAAGAGAGAAAAACTGAAGTCCCTTAGAAGGAAAGAGTTCTGCCTTCAGACTGTCTTTGAACTTAAGACTGTAGCGTCGACTCCTGC
CGGAATTTCCAGCCTGCTGGCCAGCTCTGCAGATTCACACTTGCCAGCCTCCACAATCGCAGCTGAGGCGGAGGAACCCTAAGGGCTCATTGAGATCATG
GATTTGCCCTTCTATGCATTGATGGAGCACCTGCTGCCCACAGCGTCTGTATTTGGTGCTGGGATGCTGAGCCTCCTTCTTTATGAATTTTTAAAAGGAC
ACTGAGATCTTCAAACAGAGGCTGCCACTCTAAGCAAACAGATCCCGAGTCCTGGACTCTGAAGCTTGGGCCCAGTTCTCCTTTTCTCCGGGTTTCAGAT
CCCACTGTAAAGTGAGGGGGCCCTTCTGATTCAGGACCCGGGGAAGCCAGGGGCATGAGCATCGGTGCCTCTTCTCTATTTCAAGGACCCTTCTGGGTGT
AAAGTTCTCTGAGATGCCTTACATGGATTCCCACCACTGCAAGATAACCATCGTATGTAAAGTGTTATGACCAGCAGAGTGTAATTGAAGTGCATTCCAG
AGGGAAAGACAGCGGCTCAGATTCTATTGAAAGAAACATGACATAATGATACCACAGCAAAAGCCAATCTTGCTCCTTTTTA

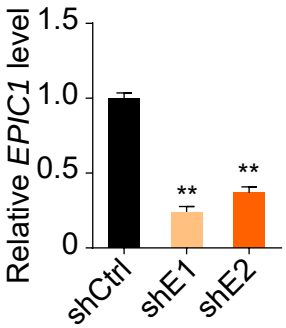
G



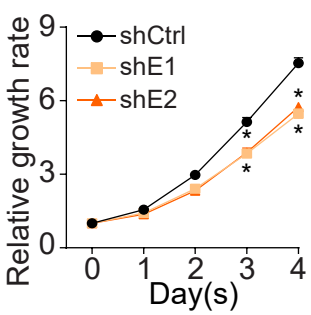
H



I



J



K

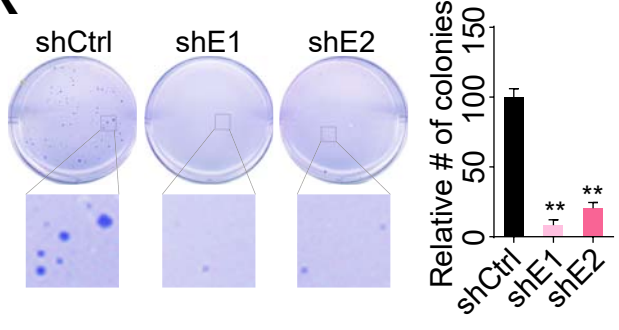
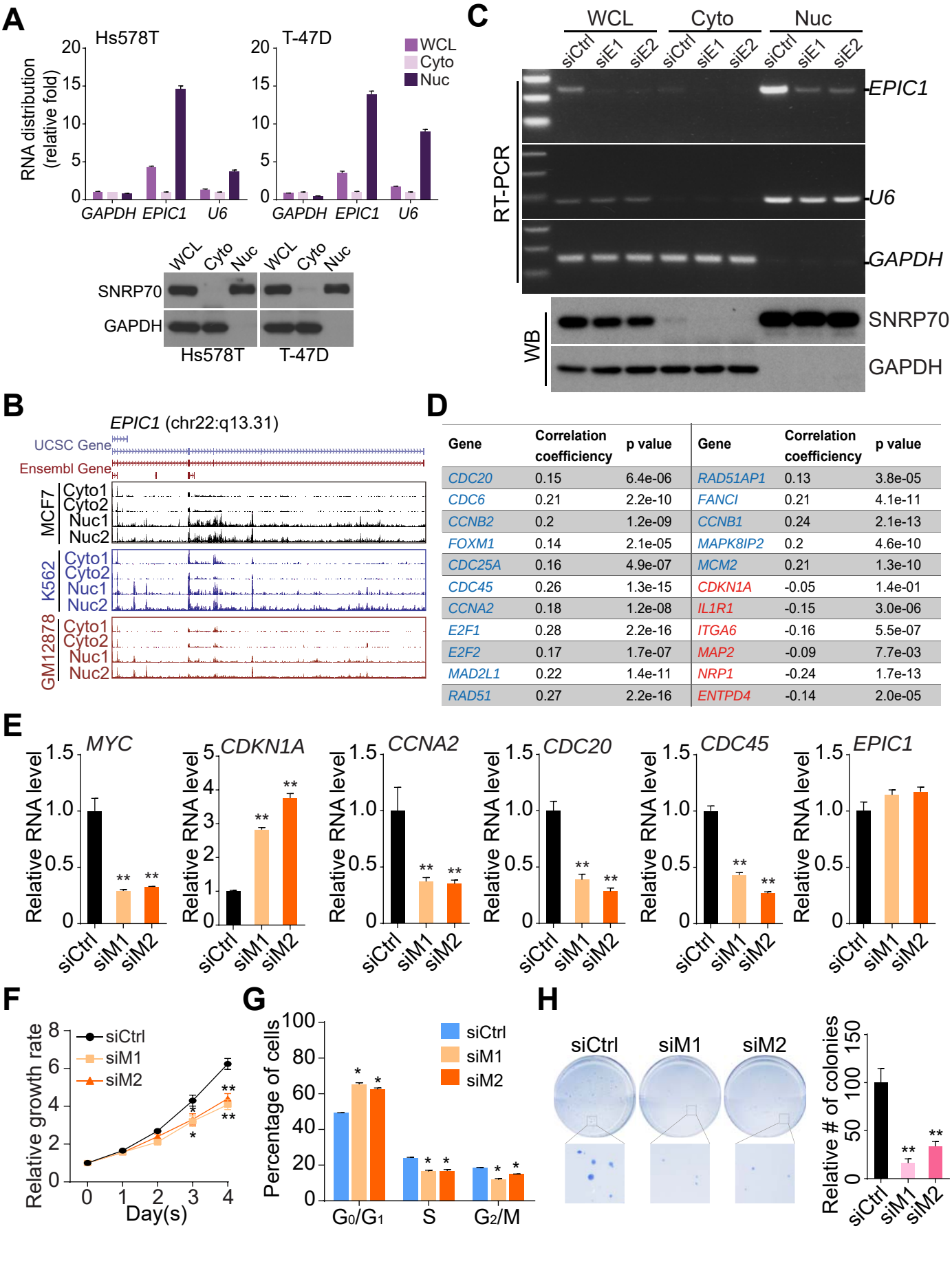


Figure S4. *EPIC1* functions as an oncogenic lncRNA regulating cell cycle progression, Related to Figure 4. (**A, B**) *EPIC1* copy numbers by qRT-PCR (A) and RT-PCR analysis of the *EPIC1* products (B) in multiple cancer-type cell lines. Copy numbers were calculated according to a standard curve of a serial dilutions of cDNA from *in vitro*-transcribed *EPIC1* ranging from 10^2 to 10^9 molecules. *GAPDH* was used as the internal control. (**C**) Quantification of *EPIC1* RNA copy number/cell according to a standard curve of *in vitro*-transcribed *EPIC1*. (**D**) 5'-RACE and 3'-RACE cloning of *EPIC1* in MCF-7 and T-47D cells. (**E**) Alignment of *EPIC1* with UCSC browser. a, Genomic location of *EPIC1* in GENCODE is highlighted in background color of light yellow. The nearest protein coding genes (upstream TBC1D22A and downstream FAM19A5) are also shown in the two ends. b, *EPIC1*'s gene structure, isoforms from GENCODE (black), RefSeq (green) and UCSC (blue) annotation are enlarged. *EPIC1* isoforms, i.e. v1, v2, and v3, are also listed in the red window. c, The CpG island, H3K4Me3 signal from ENCODE project and conservation tracks are presented at top. Sequences derived from 5'RACE and 3'RACE are listed in red window. RNA-Seq signal from CCLE breast cancer cell lines are shown at bottom. (**F**) Sequences of *EPIC1* isoforms cloned are shown. (**G**) qRT-PCR analysis of knockdown efficiency of *EPIC1* siRNAs in MCF-7 cells. (**H**) Cell cycle profiles of MCF-7 cells and ZR-75-1 cells treated with *EPIC1* siRNAs. (**I-K**) qRT-PCR analysis of *EPIC1* expression (**I**), MTT assay (**J**), and anchorage-independent colony formation assays and representative images (**K**) of MCF-7 cells stably expressing shCtrl and sh*EPIC1* RNA, respectively. Error bars indicate mean ± SD, n = 3 for technical replicates. *$p < 0.05$, **$p < 0.01$.



Gene	Correlation coefficiency	p value	Gene	Correlation coefficiency	p value
CDC20	0.15	6.4e-06	*RAD51AP1*	0.13	3.8e-05
CDC6	0.21	2.2e-10	*FANCI*	0.21	4.1e-11
CCNB2	0.2	1.2e-09	*CCNB1*	0.24	2.1e-13
FOXM1	0.14	2.1e-05	*MAPK8IP2*	0.2	4.6e-10
CDC25A	0.16	4.9e-07	*MCM2*	0.21	1.3e-10
CDC45	0.26	1.3e-15	*CDKN1A*	-0.05	1.4e-01
CCNA2	0.18	1.2e-08	*IL1R1*	-0.15	3.0e-06
E2F1	0.28	2.2e-16	*ITGA6*	-0.16	5.5e-07
E2F2	0.17	1.7e-07	*MAP2*	-0.09	7.7e-03
MAD2L1	0.22	1.4e-11	*NRP1*	-0.24	1.7e-13
RAD51	0.27	2.2e-16	*ENTPD4*	-0.14	2.0e-05

Figure S5. *EPIC1* is a nuclear lncRNA regulating MYC targets expression, Related to Figure 5. (**A**) qRT-PCR analysis of *EPIC1* expression and Western blot of subcellular fractionation in Hs578T (left panel) and T-47D (right panel) cells. (**B**) RNA-seq analysis of *EPIC1* subcellular localization in MCF7, K562, and GM12878 cells. Subcellular localization RNA-seq data are downloaded from ENCODE. (**C**) RT-PCR analysis of *EPIC1* expression levels in different subcellular fractionation of MCF-7 cells treated with *EPIC1* siRNAs. (**D**) Correlation between *EPIC1*-regulated genes (22 qRT-PCR validated genes in Figure 5F) and *EPIC1* expression in TCGA tumors. (**E**) qRT-PCR analysis of MYC targets' expression in MCF-7 cells treated with *MYC* siRNAs. (**F-H**) MTT assay (**F**), cell cycle analysis (**G**), and anchorage-independent colony formation assays (**H**) of MCF-7 cells treated with *MYC* siRNAs (i.e., siM1 and siM2). Error bars indicate mean ± SD, n = 3 for technical replicates. *$p < 0.05$, **$p < 0.01$.

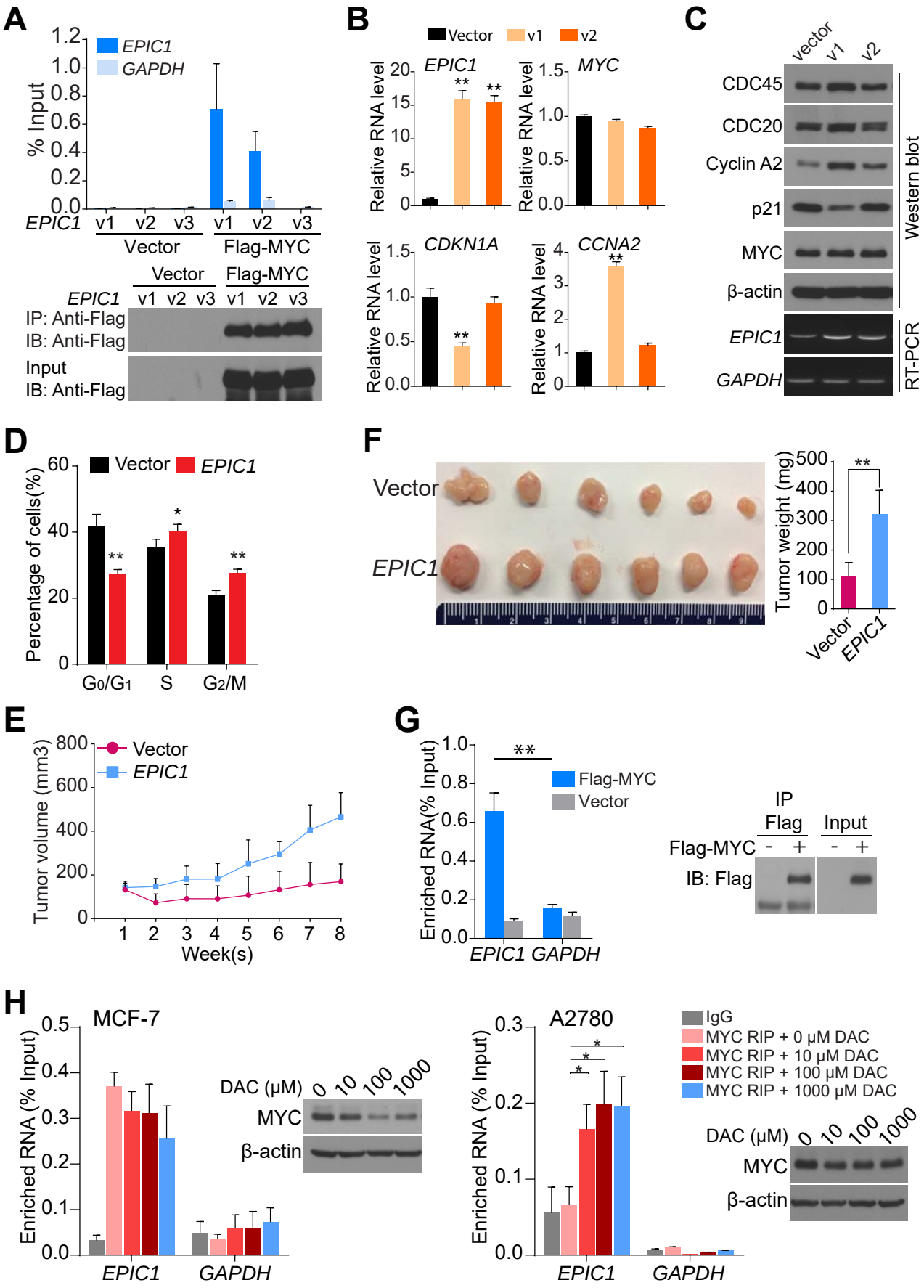
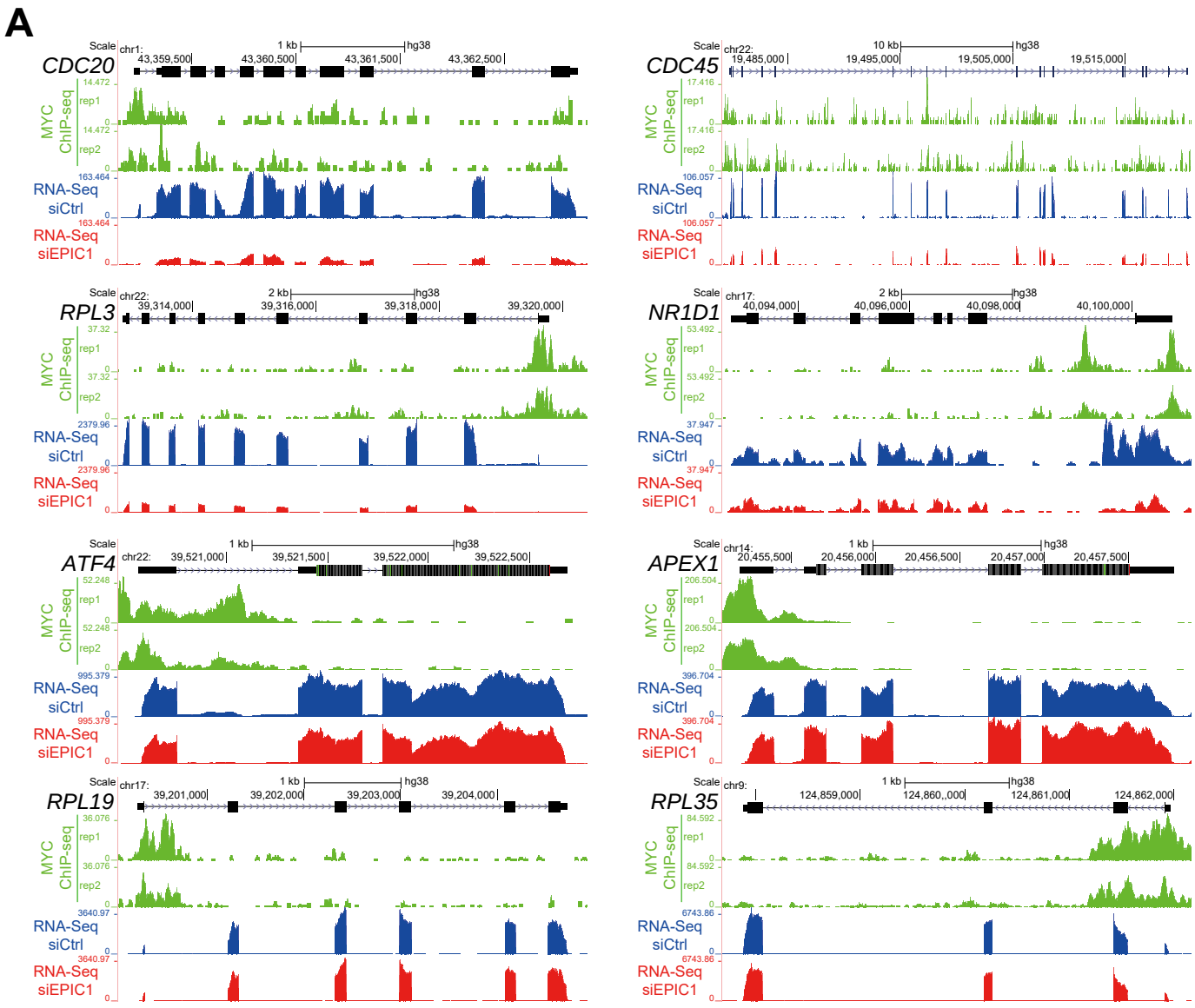
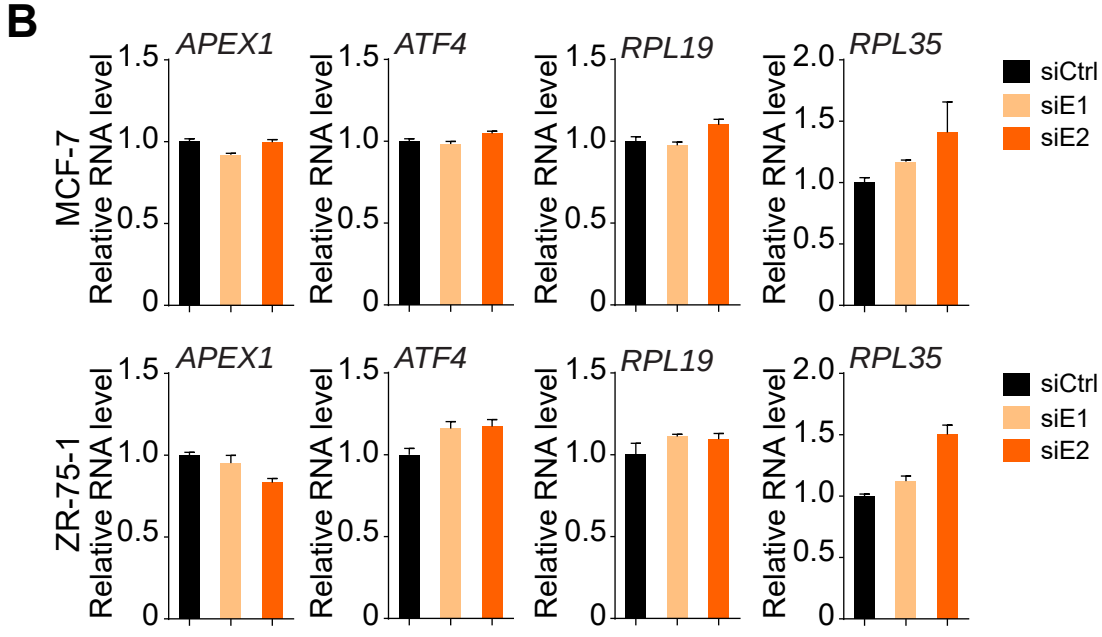

A
EPIC1
GAPDH
% Input
EPIC1 v1 v2 v3 v1 v2 v3
Vector
Flag-MYC
IP: Anti-Flag
IB: Anti-Flag
Input
IB: Anti-Flag
B
Vector v1 v2
EPIC1
MYC
CDKN1A
CCNA2
Relative RNA level
C
Vector v1 v2
CDC45
CDC20
Cyclin A2
p21
MYC
β-actin
EPIC1
GAPDH
Western blot
RT-PCR
D
Vector EPIC1
Percentage of cells(%)
G0/G1 S G2/M
E
Vector
EPIC1
Tumor volume (mm3)
Week(s)
F
Vector
EPIC1
Tumor weight (mg)
G
Enriched RNA(% Input)
Flag-MYC
Vector
EPIC1 GAPDH
IP
Flag Input
Flag-MYC - + - +
IB: Flag
H
MCF-7
Enriched RNA (% Input)
DAC (μM) 0 10 100 1000
MYC
β-actin
EPIC1 GAPDH
A2780
IgG
MYC RIP + 0 μM DAC
MYC RIP + 10 μM DAC
MYC RIP + 100 μM DAC
MYC RIP + 1000 μM DAC

Figure S6. *EPIC1* directly binds with MYC, Related to Figure 6. (**A**) Binding of *EPIC1* isoforms with MYC proteins in 293T cells. *EPIC1* isoforms retrieved by exogenous Flag-tagged MYC protein with Flag RIP were detected by qRT-PCR, and Western blot of Flag-MYC is shown (right). (**B, C**) qRT-PCR analysis of MYC targets (**B**), RT-PCR of *EPIC1* products, and Western blot of MYC targets in MCF-7 stably over-expressing *EPIC1* isoforms (**C**). (**D**) Cell cycle analysis of MCF-7 cells with stable overexpression of *EPIC1* and an empty vector. (**E, F**) Tumor growth (**E**), tumor size and weight (**F**) in MCF-7 cells with stable overexpression of *EPIC1* and an empty vector. Error bars indicate mean ± SD (n = 6). *p < 0.05, **p < 0.01. (**G**) Flag-tagged MYC retrieves *EPIC1* RNA (left). Western blot of Flag-MYC is shown (right). (**H**) MYC IP retrieves *EPIC1* in MCF-7 and A2780 cells after decitabine (DAC) treatment as indicated. Western blot of MYC is also shown. Error bars indicate mean ± SD, n = 3 for technical replicates. *p < 0.05, **p < 0.01.


A
CDC20
CDC45
RPL3
NR1D1
ATF4
APEX1
RPL19
RPL35
MYC ChIP-seq
rep1
rep2
RNA-Seq siCtrl
RNA-Seq siEPIC1


B
MCF-7
ZR-75-1
Relative RNA level
APEX1
ATF4
RPL19
RPL35
siCtrl
siE1
siE2

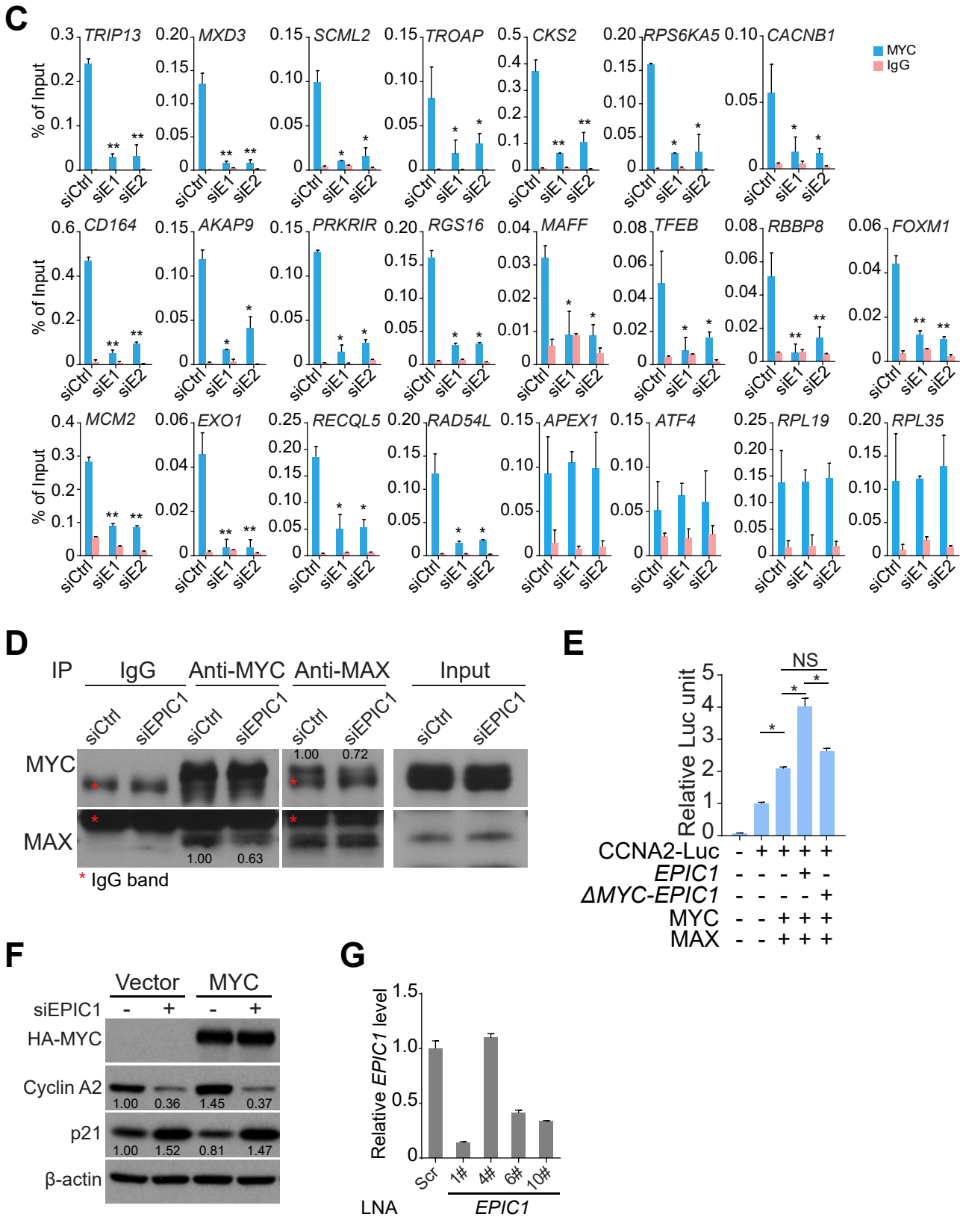

C
TRIP13
MXD3
SCML2
TROAP
CKS2
RPS6KA5
CACNB1
CD164
AKAP9
PRKRIR
RGS16
MAFF
TFEB
RBBP8
FOXM1
MCM2
EXO1
RECQL5
RAD54L
APEX1
ATF4
RPL19
RPL35
% of Input
MYC
IgG
siCtrl
siE1
siE2
D
IP
IgG
Anti-MYC
Anti-MAX
Input
siCtrl
siEPIC1
MYC
MAX
1.00
0.72
1.00
0.63
* IgG band
E
Relative Luc unit
NS
CCNA2-Luc
EPIC1
ΔMYC-EPIC1
MYC
MAX
F
Vector
MYC
siEPIC1
HA-MYC
Cyclin A2
p21
β-actin
1.00
0.36
1.45
0.37
1.00
1.52
0.81
1.47
G
Relative EPIC1 level
Scr
1#
4#
6#
10#
LNA
EPIC1

Figure S7. MYC is required for the regulatory role of *EPIC1* in cancer, Related to Figure 7. (**A**) Alignment of two biological replicates of MYC ChIP-seq in MCF-7 cells (green) and RNA-seq from siCtrl (blue) and siEPIC1 (red) RNA treated MCF-7 cells. *EPIC1*-regulated MYC targets (*CDC20*, *CDC45*, *RPL3*, and *NR1D1*) and non *EPIC1*-regualted MYC targets (*ATF4*, *APEX1*, *RPL19*, and *RPL35*) are shown. (**B**) qRT-PCR validation of non *EPIC1*-regualted MYC targets expression in MCF-7 and ZR-75-1 cells with *EPIC1* siRNAs treatment. (**C**) ChIP-qPCR analysis of MYC occupancy on the promoters of target genes in MCF-7 cells treated with *EPIC1* siRNAs. Non *EPIC1*-regulated MYC targets, i.e., *APEX1*, *ATF4*, *RPL19*, and *RPL35* are chosen as negative controls. (**D**) Effect of *EPIC1* on MYC and MAX complexes. Same quantity of lysates from MCF-7 cells with *EPIC1* siRNAs treatment was used for co-IP with anti-MYC or anti-MAX antibodies. (**E**) Reporter assay of WT-*EPIC1* and *ΔMYC-EPIC1* truncated mutant on CCNA2-Luc. (**F**) Western blot of p21 and Cyclin A2 in MCF-7 cells transfected with *EPIC1* siRNAs followed by overexpression of indicated MYC vectors. (**G**) qRT-PCR analysis of knockdown efficiency of *EPIC1* LNA in MCF-7 cells. Error bars indicate mean ± SD, n = 3 for technical replicates. *$p < 0.05$, **$p < 0.01$. NS, not significant.