Super-enhancer-driven long non-coding RNA LINC01503, regulated by TP63, is over-expressed and oncogenic in squamous cell carcinoma

Xie, Jian-Jun; Jiang, Yan-Yi; Jiang, Yuan; Li, Chun-Quan; Lim, Mei-Chee; An, Omer; Mayakonda, Anand; Ding, Ling-Wen; Long, Lin; Sun, Chun; Lin, Le-Hang; Chen, Li; Wu, Jian-Yi; Wu, Zhi-Yong; Cao, Qi; Fang, Wang-Kai; Yang, Wei; Soukiasian, Harmik; Meltzer, Stephen J.; ...Koeffler, H. Phillip

2018

Xie, J.-J., Jiang, Y.-Y., Jiang, Y., Li, C.-Q., Lim, M.-C., An, O., . . . Koeffler, H. P. (2018). Super-enhancer-driven long non-coding RNA LINC01503, regulated by TP63, is over-expressed and oncogenic in squamous cell carcinoma. Gastroenterology, 154(8), 2137-2151.e1. doi:10.1053/j.gastro.2018.02.018

https://hdl.handle.net/10356/105429

https://doi.org/10.1053/j.gastro.2018.02.018

Increased Expression of the Long Non-coding RNA LINC01503, Regulated by TP63, in Squamous Cell Carcinoma and Effects on Oncogenic Activities of Cancer Cell Lines

Jian-Jun Xie[1,2,*,†], Yan-Yi Jiang[3,*], Yuan Jiang[3,*], Chun-Quan Li[4], Lim-Mei Chee[3], Omer An[3], Anand Mayakonda[3], Ling-Wen Ding[3], Lin Long[1], Chun Sun[1], Le-Hang Lin[2], Li Chen[2], Jian-Yi Wu[1], Zhi-Yong Wu[5], Qi Cao[2], Wang-Kai Fang[1], Wei Yang[6], Stephen J. Meltzer[7], Henry Yang[3], Melissa Fullwood[3,8], Li-Yan Xu[9,†], En-Min Li[1,†], De-Chen Lin[2,†], and H. Phillip Koeffler[2,3,10]

[1]Department of Biochemistry and Molecular Biology, Medical College of Shantou University, Shantou, China;

[2]Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, USA;

[3]Cancer Science Institute of Singapore, National University of Singapore, Singapore;

[4]School of Medical Informatics, Daqing Campus, Harbin Medical University, Daqing, China;

[5]Department of Oncologic Surgery, Shantou Central Hospital, Affiliated Shantou Hospital of Sun Yat-Sen University, Shantou, China;

[6]Departments of Surgery and Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, USA;

[7]Departments of Medicine and Oncology, the Johns Hopkins University School of Medicine, Sidney Kimmel Comprehensive Cancer Center, Baltimore, USA;

[8]School of Biological Sciences, Nanyang Technological University, Singapore;

[9]Institute of Oncologic Pathology, Medical College of Shantou University, Shantou, China;

[10]National University Cancer Institute, National University Hospital Singapore, Singapore;

[*]These authors contributed equally to this work.

[†]Correspondence should be addressed to D-C.L. (De-Chen.Lin@cshs.org), J-J.X. (g_jjxie@stu.edu.cn), L-Y. X. (lyxu@stu.edu.cn) and E-M.L. (nmli@stu.edu.cn).

Short titles

Super-enhancer-associated lncRNA in ESCC

Conflict of interest statement

The authors declare no conflict of interest.

Author Contributions

D-C.L. conceived and devised the study. D-C.L., J-J.X. designed experiments and analysis. J-J.X., Y-Y.J., Y.J., L-M.C., O.A., L-W.D., L.L., C.S., L-H.L., C.L., J-Y.W., Q.C., W-K.F., S.M., H.Y., W.Y. and M.F. performed the experiments. J-J.X., A.M., O.A, C-Q.L. and H.Y. performed bioinformatics and statistical analysis. L-Y.X., E-M.L and Z-Y.W. contributed reagents and materials. D-C.L., J-J.X., Y-Y.J. and H.P.K. analyzed the data. D-C.L. and H.P.K. supervised the research and, together with J-J.X and S.J.M., wrote the manuscript.

Funding

This research is supported by the National Research Foundation Singapore under its Singapore Translational Research Investigator Award (NMRC/STaR/0021/2014) and administered by the Singapore Ministry of Health's National Medical Research Council (NMRC), the NMRC Centre Grant awarded to National University Cancer Institute, the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiatives to H.P.K. D-C.L received support from National Center for Advancing Translational Sciences UCLA

CTSI Grant UL1TR000124, CURE:DDRC DK P30 41301 Pilot and Feasibility Award, and the Tower Cancer Research Foundation. S.J.M. was supported by NIH grants CA190040, CA211457, and a Johns Hopkins University Discovery Award; he is the American Cancer Society Clinical Research Professor and the Harry and Betty Myerberg Professor of Gastroenterology. J-J.X was supported by the National Natural Science Foundation of China (No.81472342) and the Fok Ying-Tong Education Foundation (No.141034). This study was funded by the RNA Biology Center at the Cancer Science Institute of Singapore, NUS, as part of funding under the Singapore Ministry of Education's Tier 3 grants (MOE2014-T3-1-006). This work is also partially supported by National Natural Science Foundation of China (No.81672786).

Abstract

Background & Aims: Long non-coding RNAs (lncRNAs) are expressed in tissue-specific pattern, but it is not clear how these are regulated. We aimed to identify squamous cell carcinoma (SCC)-specific lncRNAs and investigate mechanisms that control their expression and function.

Methods: We studied expression patterns and functions of 4 SCC-specific lncRNAs. We obtained 113 esophageal SCC (ESCC) and matched non-tumor esophageal tissues from a hospital in Shantou City, China, and performed quantitative reverse transcription PCR assays to measure expression levels of LINC01503. We collected clinical data from patients and compared expression levels with survival times. LINC01503 was knocked down using small interfering RNAs and oligonucleotides in TE7, TE5 and KYSE510 cell lines and overexpressed in KYSE30 cells. Cells were analyzed by chromatin immunoprecipitation sequencing, luciferase reporter assays, colony formation, migration and invasion, and mass spectrometry analyses. Cells were injected into nude mice and growth of xenograft tumors was measured. LINC01503 interaction with proteins was studied using fluorescence in situ hybridization, RNA pulldown, and RNA immunoprecipitation analyses.

Results: We identified a lncRNA, *LINC01503*, which is regulated by a super-enhancer and is expressed at significantly higher levels in esophageal and head and neck SCCs than in non-tumor tissues. High levels in SCCs correlated with shorter survival times of patients. The transcription factor TP63 bound to the super-enhancer at the LINC01503 locus and activated its transcription. Expression of LINC01503 in

ESCC cell lines increased their proliferation, colony formation, migration and invasion. Knockdown of LINC01503 in SCC cells reduced their proliferation, colony formation, migration and invasion, and growth of xenograft tumors in nude mice. Expression of LINC01503 in ESCC cell lines reduced ERK2 dephosphorylation by DUSP6, leading to activation of ERK signaling via MAPK. LINC01503 disrupted the interaction between EBP1 and the p85 subunit of PI3K, increasing AKT signaling.

Conclusions: We identified a lncRNA, LINC01503, that is increased in SCC cells compared with non-tumor cells. Increased expression of LINC01503 promotes ESCC cell proliferation, migration, invasion, and growth of xenograft tumors. It might be developed as a biomarker of aggressive SCCs in patients.

KEY WORDS: oncongenesis, gene regulation, signal transduction, epigenetic regulation

Introduction

Recent reports have suggested that lineage-specific genes, which play important roles in establishing cell identity or tissue homeostasis, are also important for promoting or maintaining the disease state.[1] In cancer biology, tissue-specific oncogenes or tumor suppressor genes are crucial for mediating the neoplastic properties of various types of tumors, such as MITF in melanoma, TP63 in certain SCCs, and CEBPA in myeloid leukemia.[2,3,4] Interestingly, long non-coding RNAs (lncRNAs) have been found to exhibit higher tissue- and cell-specific expression than do protein-coding genes.[5,6] This tissue-specific pattern of lncRNA expression has also been observed in cancer cells. For example, one melanoma-specific lncRNA, *SAMMSON*, was found to be a master regulator of mitochondrial homeostasis and metabolism, and it conferred strong oncogenic functions only on melanoma cells.[7] However, the mechanisms underlying such specific transcriptional regulation of lncRNAs, as well as their biologic significance, are unknown for most mammalian cell types.

SCC is a histologic category of epithelial malignancies involving many anatomical sites; this group of cancers is highly capable of metastatic spread. Non-melanoma squamous cell cancer of the skin, HNSC, lung SCC (LUSC) and ESCC comprise the majority of SCC cases.[8] As we and others have demonstrated previously, SCCs originating from different organs share a number of phenotypic and molecular characteristics, and several oncogenes and tumor suppressor genes have been identified as tissue-specific to SCCs.[3,9,10] For example, TP63 and SOX2 are amplified oncogenes in ESCC and HNSC, while ZNF750 is a tumor suppressor gene specific to SCCs.[3, 11,12]

Locally clustered enhancers form Super-Enhancers (SEs), which recruit a large number of transcription factors and cofactors, confer particularly strong

transcriptional regulation.[13,14] Recently, using ESCC as a model, we identified many SE-associated, SCC-specific master regulators and novel oncogenic transcripts.[15] Notably, we also discovered SE-associated lncRNAs whose functions remain unknown. In the current study, we characterized these SE-associated lncRNAs and identified a novel SCC-specific lncRNA (LINC01503) that promotes the oncogenic phenotype of ESCC cells.

Materials and Methods

Human Cell Lines

Sources of ESCC cell lines have been described previously.[15,16] OE33 and FLO-1 cell lines were obtained from ATCC and the European Type Culture Collection. Esophageal cancer cells and HEK293T cells were maintained in Dulbecco's modified Eagle medium (DMEM, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (OMEGA), penicillin (100 U/mL) and streptomycin (100 mg/mL).

Chip-Seq analysis

Chip-seq was performed as described previously.[15,17] Cells were initially crosslinked with 1% formaldehyde, and followed by nuclear extraction. Chromatin/DNA complex was sheared in a Bioruptor Sonicator (Diagenode). Sonicated lysates were cleared and incubated overnight at 4℃ with magnetic beads coupled with either H3K27ac antibody (Abcam, ab4729) or TP63 antibody (Santa Cruz, sc-8431). Precipitated immune-complex was washed, and DNA was eluted and sequenced in an Illumina Hiseq4000 at Beijing Genomic Institute (BGI). Sequencing reads were aligned to the reference genome (GRCh37/hg19) with Bowtie Aligner (1.2.2), and Chip peaks were identified using MACS (Model-based Analysis of ChIP-seq, 2.1.1.20160309). Wiggle files were generated and normalized at the unit of reads per million reads (rpm). These files were then converted into bigwig files using wigToBigWig tool (http://hgdownload.cse.ucsc.edu/admin/exe/) and visualized in Integrative Genomics Viewer (http://www.broadinstitute.org/igv/home). Chip-seq data has been deposited to Gene Expression Omnibus (GSE106432).

Luciferase reporter assays

SE constituent enhancers were cloned into Firefly luciferase reporter vector pGL3-Promoter (Promega). The indicated cells were transfected using Bio-T

transfection reagent (Bioland Scientific) and A *Renilla* luciferase plasmid was co-transfected as a normalization control. Luciferase reporter activity was measured using the Dual-Luciferase Reporter Assay System (Promega).

Plasmids, siRNAs and transfection

LINC01503 was cloned into the pLVX-DsRed-Monomer-N1 Vector (lentiviral expression vector, Clontech). The pLKO.1-LINC01503-shRNA was generated by inserting double-stranded oligonucleotides into pLKO.1-puro lentiviral vector (Addgene #8453), and was confirmed by DNA sequencing. Recombinant lentiviral vectors and packaging vectors, psPAX2 (Addgene #12260) and pMD2.G (Addgene #12259), were co-transfected into 293T cells using PEI transfection reagent. Supernatants containing viral particles were harvested 48h after transfection; indicated cells were infected with the lentiviruses in the presence of 8 mg/ml Polybrene (Sigma-Aldrich).

Expression plasmids including pCMV-HA-Ebp1 (Plasmid #67792) and DUSP6 (Plasmid #27975) were obtained from Addgene. siRNA Smartpools against LINC01503, FAM83H-AS, MIR205HG, PVT1, DUSP-6 and EBP-1 were obtained from GE Dharmacon. DNA vectors and siRNAs were transfected into cells using Bio-T transfection reagent (Bioland Scientific) or Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's instructions. The cloning primers, shRNA sequence and siRNA sequence are listed in Supplementary Table 1.

MTT Assay and colony formation assay

In MTT assays, 5,000 cells were seeded onto 96-well plates, and cultured for 5 days. Cell viability was assessed daily using MTT (3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyl tetrazolium bromide) staining method. In colony formation assay, 1,000

cells were seeded onto 6-well plates, and cultured for 2 weeks. Cells were fixed with methanol and stained with crystal violet.[18]

Transwell migration assay and Matrigel invasion assay

Cell migration assay was performed in Boyden Chamber. Briefly, $1x10^5$ cells in FBS-free medium were seeded onto the membrane with 8-μm pores of the top chamber (Thermo Fisher Scientific, Inc), with the bottom chamber containing regular medium with 10% FBS. After 24 hours, the membranes were washed, fixed and stained with crystal violet, and migrated cells were quantified.[19] In Matrigel invasion assays, the top chamber was coated with Matrigel.

FISH

A pool of 12 FISH probes against LINC01503 was designed using Stellaris probe designer program (Biosearch Technologies, Supplementary Table 2). Cells were grown on slides and fixed in 3.7% formaldehyde and permeabilized in 70% ethanol. Hybridization was carried out overnight at 37 °C in 2×SSC, 10% formamide and 10% dextran. Cells were counterstained with DAPI and visualized using a Leica fluorescence microscope.

RNA immunoprecipitation (RIP) assay

RIP assay was performed with the Magna RIP RNA-Binding Protein Immunoprecipitation Kit (Millipore). Briefly, cells were lysed in RIP lysis buffer, and immunoprecipitated with either indicated antibody or IgG with protein magnetic beads. After proteinase K digestion, precipitated nucleotides were purified by phenol chloroform extraction. RNAs were then re-suspended in RNase-free water and cDNA

was synthesized and subjected to qPCR to measure LINC01503 or β-actin (internal control) transcripts.

RNA-pull down assays

Biotin-labeled RNAs were synthesized by Scientific TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific, Inc). Briefly, LINC01503 was cloned into pGEM-T vector with T7 promoter (Promega), and then used to amplify DNA templates for RNA synthesis. RNA was transcribed in vitro with T7 RNA polymerase and biotin labeling mix, treated with RNase-free DNase I, and purified with the RNeasy Mini Kit (Qiagen). Cell lysates were incubated with biotin-labeled RNAs overnight. Proteins associated with biotin-labeled RNAs were immunoprecipitated with streptavidin magnetic beads (Thermo Fisher Scientific, Inc), and subjected to Mass spectrometric analysis or Western blotting analysis.

Mass spectrometric analysis

Mass spectrometric analysis was performed as reported.[20] Briefly, eluted proteins were resolved by 12% SDS-PAGE, in-gel reduced, alkylated, and digested by trypsin. Tryptic peptides were extracted, concentrated, reconstituted in 0.2% formic acid, separated on a 15 cm EASY-Spray C_{18} column, and analyzed by an LTQ Orbitrap Elite mass spectrometer (Thermo Scientific). After each survey scan, up to 20 collision-induced dissociation (CID) spectra were acquired in the rapid CID scan mode. For protein identification and quantitation, raw mass spectrometric data were searched against the Uniprot_Human database (released on 01/22/16, including 20985 sequences) with MaxQuant (v1.5.2.8) and Andromeda. Searching parameters were as

follows: trypsin, up to two missed cleavage; carbamidomethylation of cysteines as fixed modification; acetylation of protein N-term, oxidation of methionines and deamidation of asparagines and glutamines as variable modifications. A stringent 1% false discovery rate was set to filter peptide and protein identifications.

Chromatin isolation by RNA purification (ChIRP)

ChIRP for LINC01503 was performed in TE7 cells using the published protocol.[21] Briefly, biotin TEG antisense oligos were generated using Biosearch Technologies' Stellaris Probe Designer (Biosearch Technologies) for LINC01503 and A330023F24Rik (a lncRNA specifically expressed in mouse, Negative control).[22] Cells cross-linked in 1% glutaraldehyde were lysed and sonicated. The biotinylated probes were hybridized overnight. RNA was extracted from a small aliquot of post-ChIRP beads, and qPCR was performed on the RNA samples to confirm lncRNA retrieval. For protein elution, beads were collected on a magnetic stand, resuspended in biotin elution buffer (12.5 mM biotin, 7.5 mM HEPES, 75 mM NaCl, 1.5 mM EDTA, 0.15% SDS, 0.075% sarkosyl, and 0.02% Na-Deoxycholate). Final protein samples were assayed by Western blotting. Probe sequences are described in Supplementary Table 2.

Sample collection and preparation

113 ESCC samples were collected from the Department of Oncological Surgery of the Central Hospital of Shantou City, China. Informed consent was obtained. Tumor and paired non-tumor esophageal epithelium were collected from each patient who

underwent surgical resection. Tissues were snap frozen in liquid nitrogen and stored at −80 °C.

Xenograft assay in nude mice

Xenograft assays were performed as described previously.[23] In brief, ten 6-week-old nude mice (Taconic Biosciences) were subcutaneously injected with $2x10^6$ ESCC cells on their dorsal flanks, and each mouse carried four explants. Tumor size was measured every three days. After 20 days, mice were sacrificed and their tumors were dissected, weighed and analyzed. For pathways analysis, tumor samples were lysed in RIPA buffer and the lysates were subjected to Western blotting. This animal study was approved by the Institutional Animal Care and Use Committee (IACUC) at Medical College of Shantou University.

Data availability

LncRNA expression from TCGA samples and different types of cancer cell lines were examined by MiTranscriptome RNA-seq compendium (URL: https:// http://mitranscriptome.org, ref. 24) and Cancer Cell Line Encyclopedia (CCLE, URL: http://www.broadinstitute.org/ccle; ref. 25), respectively. GSE53622 and GSE53624; were obtained from Gene Expression Omnibus.[26]

Results

Identification of SE-associated oncogenic lncRNAs in ESCC

Our group previously characterized SE-associated oncogenes (protein-coding genes) in ESCC cells using an integrative analysis of both their expression levels and sensitivity to transcriptional inhibition.[15] Using the same method, four candidate oncogenic lncRNAs were identified (FAM83H-AS, MIR205HG, PVT1 and LINC01503) (Supplementary Figures 1-3). Importantly, compared to the corresponding non-malignant tissues, all 4 lncRNAs were particularly up-regulated in TCGA HNSC and LUSC samples (Figure 1C, Supplementary Figure 4), suggesting an SCC-specific feature.[24] As an initial step to test their functional significance, each lncRNA was silenced individually by siRNAs and cell proliferation was monitored. This approach revealed that LINC01503 was the only candidate required for ESCC cell proliferation (Supplementary Figure 5, Figures 4A-B).

We next analyzed the protein-coding potential of LINC01503 by using two independent mathematical methods based on various sequence features, including Coding Potential Calculator (http://cpc.cbi.pku.edu.cn/) and Coding Potential Assessment Tool (http://lilab.research.bcm.edu/cpat/index.php).[27,28] Both methods predicted that the protein-coding potential of LINC01503 was very weak. We also analyzed LINC01503 sequence in Conserved Domain Database (CDD, https://www.ncbi.nlm.nih.gov/cdd/) and Pfam (http://pfam.xfam.org/) and found that no conserved domains were indentified.[29,30] Moreover, RNA pull-down experiment showed no interaction between LINC01503 and ribosomal protein LP0, which is a component of the 60S subunit of ribosome[31] (Supplementary Figure 6). These data strongly suggest that LINC01503 does not encode a protein.

Characterization of the expression of LINC01503

Little is known regarding the regulation or function of LINC01503 in human cells. Both our RNA-Seq and qRT-PCR analysis confirmed the expression of this lncRNA in ESCC cells (Supplementary Figure 7). Supporting the RNA-seq results, cDNA microarray data from the Cancer Cell Line Encyclopedia (CCLE) project showed that LINC01503 was expressed at the highest level in SCC cells, including HNSC and ESCC (Figure 1A). Bladder carcinoma (BLCA), a significant proportion of which exhibits SCC genomic and epigenomic features,[32,33] also expressed a high level of LINC01503, indicating its SCC-specific regulation. Our qRT-PCR data validated the CCLE results (Figure 1B). Notably, LINC01503 was almost undetectable in esophageal adenocarcinoma (EAC) cells, corroborating this SCC-specific regulation. Two additional public datasets validated the up-regulation of LINC01503 in ESCC samples *vs*. matched nonmalignant esophageal mucosa (Figure 1D and E). Moreover, we performed qRT-PCR analysis on 113 tumor/normal matched ESCC cases, and confirmed that the expression of LINC01503 was significantly higher in cancerous tissues (Supplementary table 3, Figure 1F). Notably, high LINC01503 expression significantly correlated with shorter overall survival (OS, $P = 0.023$, Figure 1G) and disease-free survival (DFS, $P = 0.035$, Figure 1H), indicating the biological significance of LINC01503. Multivariate regression analysis based on the Cox proportional hazard model showed that LINC01503 was an independent prognostic factor for poor OS ($P = 0.02$, Supplementary table 4) but not for poor DFS ($P = 0.063$, Supplementary table 5).

LINC01503 is a SCC-specific SE-driven gene

To characterize the transcriptional regulation of LINC01503, we analyzed H3k27ac profiles in 5 ESCC and 3 EAC cell lines (Manuscript in preparation). An SE upstream of *LINC01503* was found in 3 (TE7, TE5 and TT) of 5 ESCC cell lines but was absent

in all 3 EAC cell lines (Figure 2A, lower panel). Interestingly, a number of SEs and typical-enhancers (TEs) in the genomic neighborhood also displayed an ESCC-specific pattern (Figure 2A, lower panel). We employed circularized chromosome conformation capture assays (4C) to explore the interaction landscape of the LINC01503-SE in TE7 cells. Using 2 different constituent enhancers as viewpoints (E1 and E2), we observed the associations between this SE with *LINC01503* TSS, as well as with its gene body (Figure 2A, upper panel, Figure 2B and C). The interactions were confirmed by independent replicate experiments (Supplementary Figure 8). Moreover, this approach revealed extensive interactions between LINC01503-SE with the majority of neighboring SEs and TEs, a characteristic of active SEs that we recently observed in gastric cells.[34] This suggests strong enhancer activity and transcriptional co-regulation within this genomic region (Figure 2A, upper panel, Supplementary Figure 8). Given the SCC-specific feature of LINC01503-SE, we hypothesized that SCC-specific transcription factors could be responsible for the establishment and maintenance of its activity. We therefore performed Chip-seq to map the genome-wide occupancy of TP63 and SOX2, two of the most prominent master transcriptional regulators described in SCC cells.[3,10] Notably, TP63 Chip-seq profile showed several prominent peaks in this genomic segment, with the strongest one overlapping with LINC01503-SE (Figure 2A and B), suggesting the involvement of TP63 in the transcription regulation of LINC01503. We next cloned individual constituent enhancers (E1, E2 and E3) of the LINC01503-SE into enhancer-reporter vectors and measured their activities by luciferase reporter assay (Figure 2B). E2 and E3 were active in all 4 ESCC cell lines tested (E2 and E3 was TEs in KYSE510 and KYSE140 cells, Figure 2A), while E1 was active only in TE7 and TE5 cells, consistent with the H3K27ac profiles (Figures 2D and E;

Supplementary Figure 9A and B). In contrast, all constituents were inactive in EAC cell lines (Supplementary Figure 9C and D), again underscoring the SCC-specific nature of this SE.

TP63 activates the transcription of LINC01503

To investigate the role of TP63 in the transcriptional regulation of LINC01503, enhancer E2 was further divided into two constituents (E2A and E2B) for luciferase assay. Enhancer E2 was selected because the TP63 binding site was located in E2A, and a TP63 motif was observed in this constituent (Figure 2B). Importantly, the enhancer activity of E2A, but not E2B, was significantly decreased upon TP63 knockdown in ESCC cells (Figure 3A, Supplementary Figure 10A). As expected, TP63 silencing had no effect on the activity of either E2A or E2B in EAC cells, which are not squamous cells (Supplementary Figure 10B). Using Chip-qPCR, we validated that TP63 bound to E2A specifically in squamous cells (Figure 3B). Moreover, the expression of LINC01503 was consistently decreased upon silencing of TP63 (but not SOX2), validating the Chip-seq data (Figure 3C, Supplementary Figure 10C, Supplementary Figure 11). In addition, a strongly positive correlation between the expression levels of TP63 and LINC01503 was found in ESCC cell lines (Figure 3D, left) and HNSC cell lines (Figure 3D, right). This significant correlation was confirmed in ESCC primary tissues in a public dataset (Figure 3E, left) and in an in-house cohort (Figure 3E, right). Importantly, we transfected the expression vector of ΔNp63 (SCC-specific isoform) into the TP63-depleted ESCC cells,[3,35] and found that restoration of ΔNp63 rescued the expression of LINC01503 (Figure 3F, Supplementary Figure 10D and E). These findings suggested that TP63 activates the transcription of LINC01503 by binding to its SCC-specific SE.

We next measured expression levels of all 13 transcripts located within 1Mb of the SE.

These experiments revealed that the expression levels of 12 transcripts were not altered by TP63 knockdown (Supplementary Figure 12A and B). The only exception was CCBL1, whose level was decreased in TP63-depleted cells (Supplementary Figure 12B); moreover, restoration of ΔNp63 rescued its expression (Supplementary Figure 12C), suggesting that CCBL1 may also be regulated by TP63. We next asked if CCBL1 also has functional significance in ESCC cells. MTT assay showed that CCBL1 knockdown did not affect cellular growth (Supplementary Figure 13).

LINC01503 promotes malignant phenotypes in ESCC cells

To investigate the biological significance of LINC01503 in ESCC cells, both loss-of-function and gain-of-function approaches were employed. Specifically, pooled siRNAs were applied to ESCC cell lines exhibiting high LINC01503 expression levels (TE7, TE5 and KYSE510), while an overexpression vector was used in KYSE30 cells, which showed low levels of this lncRNA (Supplementary Figure 14, Figure 4A). Importantly, knockdown of LINC01503 greatly reduced proliferation, clonogenicity, migration and invasion, while ectopic expression of LINC01503 produced the opposite effects (Figures 4B-E, Supplementary Figure 15). To validate these results, we silenced the expression of LINC01503 using locked nucleic acid (LNA)-modified antisense oligonucleotides (which trigger RNase-H-mediated degradation of their targets,[36] independent of the RNAi mechanism), and similar cellular effects were observed (Supplementary Figure 16). Moreover, restoration of LINC01503 in LINC01503 stably-depleted cells reversed decreased cell proliferation and migration (Supplementary Figure 17). These data identified LINC01503 as a functionally oncogenic lncRNA in ESCC.

LINC01503 interacts with both EBP-1 and ERK2

RNA Fluorescence in situ hybridization (RNA-FISH) revealed that LINC01503 was

principally localized to the cytoplasm, with only weak expression in the nucleus (Figure 5A). qRT-PCR analysis of fractionated nuclear and cytoplasmic RNA confirmed its predominant presence in the cytoplasm (Figure 5B).

Since the majority of LINC01503 was located in the cytoplasm, we next focused on its potential cytoplasmic functions in ESCC cells. Cytoplasmic lncRNAs have the ability to bind to large biomolecules, such as RNAs or proteins.[37] To investigate this phenomenon, we performed *in vitro* RNA pull-down of LINC01503, then subjected the precipitants to mass spectrometry (MS) analysis (Figure 5C and D). Among the top 10 proteins identified, several have been reported to have RNA binding capacity, including 14-3-3σ, 14-3-3ζ, and proteins in the heat shock protein family, supporting the validity of our approach.[38] We were particularly interested in two candidates: EBP-1 (ErbB3 binding protein 1) and ERK2 (Mitogen-activated protein kinase 1), since EBP-1 is a crucial regulator of the PI3K/Akt pathway and ERK2 is an important component of ERK/MAPK signaling.[39,40] We first selected 4 candidates (EBP-1, ERK2, 14-3-3σ and 14-3-3ζ) and performed RNA pull-down followed by Western blotting; results of these experiments confirmed these interactions (Figure 5E and Supplementary Figure 18). Importantly, the binding of EBP-1 and ERK2 to LINC01503 was further validated by an independent approach, chromatin isolation by RNA purification (ChIRP) (Figure 5F). Finally, qPCR analysis following RNA immunoprecipitation assays confirmed an enrichment of LINC01503 in the complex with either EBP-1 or ERK2, compared with IgG control (Figure 5G).

LINC01503 activates both PI3K/Akt and ERK/MAPK signaling pathways

We next explored the potential role of LINC01503 in the activation of the PI3K/Akt and ERK/MAPK signaling pathways. Notably, LINC01503 depletion significantly attenuated both basal levels and EGF-induced phosphorylation of ERK1/2, Akt,

p70S6K and mTOR (Figure 6A, Supplementary Figure 19). Similar effects were observed in cells treated with IGF (Supplementary Figure 20). In contrast, phosphorylation levels of EGFR and MEK1/2 were unchanged upon silencing of LINC01503 (Figure 6A).

ERK1/2 are phosphorylated by MEK1/2 and dephosphorylated by phosphatases, including Dual Specificity Phosphatase 6 (DUSP-6).[41] We postulated that the interaction between LINC01503 and ERK2 could weaken the dephosphorylation of ERK1/2, since activity of MEK1/2 was unaltered following LINC01503 knockdown. To address this hypothesis, we performed co-immunoprecipitation (IP), which revealed that the interaction between DUSP-6 and ERK2 was increased in LINC01503 knockdown cells with or without EGF treatment, suggesting that binding of LINC01503 to ERK2 suppresses the interaction between DUSP-6 and ERK2 and thus inhibits ERK2 dephosphorylation (Figure 6B, Supplementary Figure 21A). We confirmed that DUSP-6 was responsible for ERK2 dephosphorylation in ESCC cell lines. Specifically, RNAi-mediated knockdown of DUSP-6 prominently enhanced the phosphorylation of ERK1/2, while DUSP-6 overexpression strongly suppressed ERK1/2 phosphorylation (Figures 6C and D).

EBP-1 (p42) directly interacts with the p85 subunit of PI3K, leading to its degradation through ubiquitination-proteasome system.[42] Importantly, following LINC01503 silencing, we found that binding of EBP-1 to p85 was increased (Figure 6E, Supplementary Figure 21B). We further treated cells with MG132 which, as expected, elevated steady-state levels of p85. In the absence of LINC01503, binding between EBP-1 to p85 was consistently enhanced. Moreover, depletion of LINC01503 reduced the protein level of p85, which was rescued by MG132 treatment (Figure 6E, input, comparing lane 4 versus 5). We validated the finding that EBP-1

silencing in ESCC cells led to increased levels of PI3K p85 and p-Akt; and that overexpression of EBP-1 inhibited this pathway (Figures 6F and G). These data suggest that LINC01503 prevents EBP-1 from binding to p85, suppresses the deubiquitination of PI3K, and subsequently activates PI3K/Akt signaling.

Taken together, the above results suggest that both DUSP-6 and EBP-1 are important mediators of the SCC-specific functions of LINC01503. To confirm this hypothesis further at the cellular phenotypic level, we silenced either DUSP-6 or EBP-1, which promoted both cell proliferation and migration (Supplementary Figure 22), phenocopying the effects of LINC01503 upregulation in these cells.

Therapeutic potential of targeting LINC01503

We next established ESCC cells with stable depletion of LINC01503, and further examined this gene's oncogenic effects by soft-agar colony formation and xenograft assays. In agreement with earlier MTT results, silencing of LINC01503 led to decreased anchorage-independent growth (Figure 7A and B). Moreover, both xenograft volume and weight were reduced upon LINC01503 depletion (Figures 7C and D). Analyses of the tumor samples by Western blotting confirmed downregulation of both the PI3K/Akt and the ERK/MAPK signaling pathways in LINC01503-deficient tumors (Figures 7E and F). The effect of LINC01503 in tumorigenesis was further confirmed in KYSE510 cells (Supplementary Figure 23).

To test the therapeutic potential of targeting LINC01503, we used MTT assay to evaluate the impact of LINC01503 knockdown on the efficacy of inhibitors against Akt (MK2206) or MEK (U0126). IC50 value of MK2206 was about 6.068 μM in TE7 cells and 7.027μM in KYSE510 cells. Strikingly, these values reduced more than 10 fold in LINC01503 knockdown cells (Figure 7G, left panels). Such substantial decrease of IC50 value was also noted in the MEK inhibitor, U0126 (Figure 7G, right

panels). We further confirmed these results by long-term colony formation assay (Supplementary Figure 24). These data indicated that LINC01503-deficient cells were more sensitive than untreated control cells to small-molecule inhibitors of AKT and MEK.

Discussion

LncRNAs unique to SCCs have not been extensively studied. Based on our previous results,[15] we identified novel SE-lncRNAs with specific expression patterns in SCC tissues. Among them, LINC01503 exhibited strong oncogenic potential specifically in SCC cells. Moreover, LINC01503 overexpression was associated with poor clinical outcome in ESCC patients (Supplementary Figure 25).

ΔNp63 is the predominant isoform of TP63 expressed in SCC, and it exerts oncogenic functions.[3,35] Tp63 binds to promoters and enhancers to determine gene expression programs in a cell-type-specific manner.[43] However, lncRNAs targeted by TP63 have not been discovered thus far, and the regulation of TP63 on SE has not been studied. We found that ΔNp63 bound to the LINC015030-SE and contributed to its SCC-specific transcription activation, suggesting that LINC01503 is an important down-stream effector of TP63.

LncRNAs operate through multiple different mechanisms.[5] Most of the well-characterized lncRNAs, such as HOTAIR, MALAT1 and XIST, are chromatin-associated lncRNAs.[44,45,46] They regulate chromatin modification, transcription activation and processing of RNA. In comparison, the functions of cytoplasmic lncRNAs, particularly those involved in the regulation of signalling pathways, remain less well understood. Here, we identified and validated several interacting protein partners of LINC01503 including 14-3-3, ERK2 and EBP-1. Notably, 14-3-3 family proteins have recently been observed to have RNA-binding activity.[38] Since 14-3-3 also binds to a multitude of functionally-diverse signaling proteins, including kinases, phosphatases and transmembrane receptors,[47] the association between LINC01503 and 14-3-3 may link this lncRNA to multiple additional cellular processes. In addition, our results showed that a small proportion of

LINC01503 was located in nucleus, and that certain nuclear proteins such as YBOX3 were identified as potential interaction partners of LINC01503. These data imply that LINC01503 may have potential additional nuclear roles.

Both PI3K/Akt and ERK/MAPK pathways are key signaling pathways driving cancer proliferation and survival. However, whether and how these pathways are regulated by RNA molecules is incompletely understood. Here, we found that LINC01503 regulated both PI3K/Akt and ERK/MAPK pathways by interacting with ERK2 and EBP-1. Our findings support recent studies suggesting that lncRNAs can function as important regulators of Akt activity. [48,49]

Initially identified as an ErbB3 receptor-binding protein, Ebp1 has two alternatively spliced isoforms, p48 and p42.[50,51] Ebp1 p42 inhibits cancer cell growth in a number of tumors such as retinoblastoma, prostate cancer and glioma partially through regulating cancer-related factors such as Rb, HDAC2 and Sin3A.[39] Recently, Ebp1 p42 was observed to interact with the p85 subunit of PI3K, subsequently recruiting p85 to HSP70/CHIP-mediated proteasomal degradation, substantially reducing the phosphorylation of Akt.[42] In support of this finding, we showed that Ebp1 p42 inhibited the PI3K/Akt pathway in ESCC cells. Importantly, Ebp1 p42 associated directly with LINC01503, and this interaction released p85 from the Ebp1/PI3K complex, leading to increased Akt activity.

DUSP-6 is a MAPK phosphatase that plays a critical role in the negative regulation of ERK phosphorylation.[41] In ESCC, DUSP-6 was found to be down-regulated and to inhibit tumor progression.[52] We confirmed that DUSP-6 functions as a tumor suppressor in ESCC by inactivating the ERK/MAPK signaling pathway. Importantly, LINC01503 bound directly to ERK2 and prevented the dephosphorylation of ERK2 by DUSP-6.

To summarize, we identified a TP63-dependent, SE-mediated mechanism for the upregulation of a novel oncogenic LncRNA, LINC01503, in SCCs. Considering this gene's SCC-specific nature, its association with poor patient survival, and its oncogenic functions, LINC01503 represents a potential biomarker and/or therapeutic target in this group of deadly cancers.

Acknowledgements

We thank Dr. Bo Zhou (Cedars-Sinai Medical Center) for help doing MS analysis.

References

1. **Hnisz D, Abraham BJ, Lee TI**, et al. Super-enhancers in the control of cell identity and disease. Cell 2013;155(4):934-947.
2. **Yokoyama S, Woods SL, Boyle GM**, et al. A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. Nature 2011;480(7375):99-103.
3. Cancer Genome Atlas Research Network. Integrated genomic characterization of oesophageal carcinoma. Nature 2017;541(7636):169-175.
4. Pulikkan JA, Tenen DG, Behre G. C/EBPα deregulation as a paradigm for leukemogenesis. Leukemia 2017;31(11):2279-2285.
5. Huarte M. The emerging role of lncRNAs in cancer. Nat Med 2015;21(11):1253-61.
6. **Schmitt AM, Chang HY.** Long Noncoding RNAs in Cancer Pathways. Cell 2016;29(4):452-463.
7. Leucci E, Vendramin R, Spinazzi M, et al. Melanoma addiction to the long non-coding RNA SAMMSON. Nature 2016;531(7595):518-522.
8. Yan W, Wistuba II, Emmert-Buck MR, et al. Squamous Cell Carcinoma - Similarities and Differences among Anatomical Sites. Am J Cancer Res 2011;1(3):275-300.
9. **Lin DC, Hao JJ, Nagata Y**, et al. Genomic and molecular characterization of esophageal squamous cell carcinoma. Nat Genet 2014;46(5):467-473.
10. **Bass AJ, Watanabe H**, Mermel CH, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet 2009;41(11):1238-1242.
11. **Hazawa M, Lin DC**, Handral H, et al. ZNF750 is a lineage-specific tumour suppressor in squamous cell carcinoma. Oncogene 2017;36(16):2243-2254.
12. Lin DC, Wang MR, Koeffler HP. Genomic and Epigenomic Aberrations in Esophageal Squamous Cell Carcinoma and Implications for Patients. Gastroenterology. 2017 Jul 27. pii: S0016-5085(17)35953-X. doi: 10.1053/j.gastro.2017.06.066.
13. Amaral PP, Bannister AJ. Re-place your BETs: the dynamics of super enhancers. Mol Cell 2014;56(2):187-189.
14. Chipumuro E, Marco E, Christensen CL, et al. CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. Cell

2014;159(5):1126-1139.

15. Jiang YY, Lin DC, Mayakonda A, et al. Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. Gut 2017;66(8):1358-1368.
16. **Lin DC, Q. Dinh H, Xie JJ**, et al. Identification of distinct mutational patterns and new driver genes in esophageal squamous cell carcinomas and adenocarcinomas. Gut 2017 August 31. doi: 10.1136/gutjnl-2017-314607.
17. **Yuan J, Jiang YY**, Mayakonda A, et al. Super-Enhancers Promote Transcriptional Dysregulation in Nasopharyngeal Carcinoma. Cancer Res 2017;77(23):6614-6626.
18. **Sun H, Lin DC, Cao Q**, et al. CRM1 Inhibition Promotes Cytotoxicity in Ewing Sarcoma Cells by Repressing EWS-FLI1-Dependent IGF-1 Signaling. Cancer Res 2016;76(9):2687-2697.
19. Xie J, Lin D, Lee DH, et al. Copy number analysis identifies tumor suppressive lncRNAs in human osteosarcoma. Int J Oncol 2017;50(3):863-872.
20. Morley S, You S, Pollan S, et al. Regulation of microtubule dynamics by DIAPH3 influences amoeboid tumor cell mechanics and sensitivity to taxanes. Sci Rep 2015;5:12136.
21. Chu C, Zhang QC, da Rocha ST, et al. Systematic discovery of Xist RNA binding proteins. Cell 2015;161(2):404-416.
22. **Yue F, Cheng Y, Breschi A**, et al. A comparative encyclopedia of DNA elements in the mouse genome. Nature 2014;515(7527):355-364.
23. **Xu L, Chen Y, Dutra-Clarke M**, et al. BCL6 promotes glioma and serves as a therapeutic target. Proc Natl Acad Sci U S A 2017;114(15):3981-3986.
24. **Iyer MK, Niknafs YS**, Malik R, et al. The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 2015;47(3):199-208.
25. **Barretina J, Caponigro G, Stransky N**, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483(7391):603-607.
26. **Li J, Chen Z, Tian L**, et al. LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut 2014;63(11):1700-1710.
27. Kong L, Zhang Y, Ye ZQ, et al. CPC: assess the protein-coding potential of transcripts using

sequence features and support vector machine. Nucleic Acids Res 2007;35(Web Server issue):W345-349.

28. **Wang L, Park HJ**, Dasari S, et al. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res 2013;41(6):e74.
29. Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. Nucleic Acids Res 2004;32(Web Server issue):W327-331.
30. Finn RD, Coggill P, Eberhardt RY, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 2016;44(D1):D279-285.
31. Francisco-Velilla R, Fernandez-Chamorro J, Ramajo J, et al. The RNA-binding protein Gemin5 binds directly to the ribosome and regulates global translation. Nucleic Acids Res 2016;44(17):8335-8351.
32. **Hoadley KA, Yau C, Wolf DM**, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell 2014;158(4):929-944.
33. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature 2014;507(7492):315-322.
34. **Ooi WF, Xing M, Xu C**, et al. Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. Nat Commun 2016;7:12983.
35. Hu H, Xia SH, Li AD, et al. Elevated expression of p63 protein in human esophageal squamous cell carcinomas. Int J Cancer 2002;102(6):580-583.
36. Kauppinen S, Vester B, Wengel J. Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. Drug Discov Today Technol 2005;2(3):287-290.
37. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell 2013;154(1):26-46.
38. Castello A, Fischer B, Eichelbaum K, et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 2012;149(6):1393-1406.
39. Ko HR, Chang YS, Park WS, et al. Opposing roles of the two isoforms of ErbB3 binding protein 1 in human cancer cells. Int J Cancer 2016;139(6):1202-1208.
40. McKay MM, Morrison DK. Integrating signals from RTKs to ERK/MAPK. Oncogene 2007;26(22):3113-3121.

41. Caunt CJ, Keyse SM. Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling. FEBS J 2013;280(2):489-504.
42. Ko HR, Kim CK, Lee SB, et al. P42 Ebp1 regulates the proteasomal degradation of the p85 regulatory subunit of PI3K by recruiting a chaperone-E3 ligase complex HSP70/CHIP. Cell Death Dis 2014;5:e1131.
43. **Heintzman ND, Hon GC, Hawkins RD**, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 2009;459(7243):108-112.
44. Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010;464(7291):1071-1076.
45. **West JA, Davis CP**, Sunwoo H, et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol Cell 2014;55(5):791-802.
46. Engreitz JM, Pandya-Jones A, McDonel P, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 2013;341(6147):1237973.
47. Morrison DK. The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. Trends Cell Biol 2009;19(1):16-23.
48. **Lin A, Hu Q, Li C**, et al. The LINK-A lncRNA interacts with PtdIns(3,4,5)P3 to hyperactivate AKT and confer resistance to AKT inhibitors. Nat Cell Biol 2017;19(3):238-251.
49. Koirala P, Huang J, Ho TT, et al. LncRNA AK023948 is a positive regulator of AKT. Nat Commun 2017;8:14422.
50. Yoo JY, Wang XW, Rishi AK, et al. Interaction of the PA2G4 (EBP1) protein with ErbB-3 and regulation of this binding by heregulin. Br J Cancer 2000;82(3):683-690.
51. Liu Z, Ahn JY, Liu X, et al. Ebp1 isoforms distinctively regulate cell survival and differentiation. Proc Natl Acad Sci U S A 2006;103(29):10917-10922.
52. Ma J, Yu X, Guo L, et al. DUSP6, a tumor suppressor, is involved in differentiation and apoptosis in esophageal squamous cell carcinoma. Oncol Lett 2013;6(6):1624-1630.

Author names in bold designate shared co-first authorship

Figure Legends

Figure 1. LINC01503 is uniquely up-regulated in SCC tumors. (A) Expression levels of LINC01503 in cancer cell lines were derived from the CCLE project. (B) mRNA levels of LINC01503 in different cancer cell lines were measured by qRT-PCR. (C) Expression levels of LINC01503 in TCGA normal and tumor samples. Each point represents one tissue sample. (D, E) mRNA levels of LINC01503 were derived from the indicated GEO datasets measuring matched ESCC and adjacent non-malignant esophageal epithelium. (F) qRT-PCR analysis of LINC01503 expression levels in 113 pairs of ESCC and adjacent non-malignant esophageal epithelium. (G, H) Kaplan–Meier plots showing the association between LINC01503 expression and either overall survival (G) or disease-free survival (H) in ESCC patients.

Figure 2. *LINC01503* is a SCC-specific SE-driven gene. (A) Identification of *LINC01503* as an SE-associated gene. Upper panel, 4C-seq showing the long-range interactions anchored on the SE upstream of *LINC01503* in TE7 cells. Deeper red color indicates higher interaction frequency. Lower panel, Chip-seq profiles for H3K27ac, TP63 and SOX2 at *LINC01503* locus in ESCC and EAC cell lines. Predicted super-enhancers are depicted as blue bars. The Y axis indicated reads per million (Rpm). (B) H3K27ac Chip-seq signals at the *LINC01503* locus. Three constituent enhancers (E1, E2, and E3) within the SE were separately cloned into luciferase reporter vector pGL3-promoter. (C) Connecting lines showing the interactions detected by 4C. (D and E) Enhancer activity measured by luciferase reporter assays in ESCC cells. Mean ± s.d. are shown, n = 3.

Figure 3. TP63 activates the transcription of *LINC01503* by binding to its

super-enhancer. (A) E2 segment was divided into two constituents (E2A and E2B) and the luciferase activities of E2, E2A and E2B were measured in ESCC cells upon TP63 silencing. (B) Chip-qPCR analysis of the interaction between TP63 and E2A constituent. IgG antibody was used as a negative control. (C) Expression of LINC01503 and TP63 was determined by qRT-PCR in TP63 knockdown ESCC cells using two independent shRNAs (shTP63-1 and sh-TP63-3). (D, E) Expression correlation between LINC01503 and TP63 in HNSCC and ESCC cell lines (D) and ESCC tissues (E). (F) ΔNTP63 expression plasmid was transfected into TP63 knockdown ESCC cells and the levels of LINC01503 were recovered upon ΔNTP63 expression. Mean ± s.d. are shown, n = 3. *, $P \leq 0.05$. n.s., not significant.

Figure 4. LINC01503 is required for the growth, migration and invasion of ESCC cells. (A) qRT-PCR analysis validating the silencing (TE7 and KYSE510 cells) or overexpression (KYSE30 cells) of LINC01503. EV, empty vector. (B) MTT and (C) colony formation assays in ESCC cells upon LINC01503 knockdown or overexpression. (D) Transwell migration assays and (E) Transwell Matrigel invasion assays were performed to determine the effect of LINC01503 on cell migration and invasiveness. Mean ± s.d. are shown, n = 3. n.s., not significant. *, $P \leq 0.05$.

Figure 5. LINC01503 interacts with EBP-1 and ERK2. (A) Subcellular distribution of LINC01503 (red) in ESCC cells was determined by RNA-FISH. PBS without probes was used as the negative control. DAPI was used to stain nucleus. (B) Nuclear and cytoplasmic fractions of TE7 and KYSE510 cells were subjected to qRT-PCR assay (mean ± s.d. , n = 3 biologically independent experiments). GAPDH and U1

were served as positive controls for cytoplasmic and nuclear expression, respectively. (C) Silver staining of biotinylated LINC01503-associated proteins after RNA pull-down. Poly(A)25 RNA was used as a negative RNA control. (D) LINC01503-bound proteins with top matching scores identified by mass spectrometry. (E) Western blotting of EBP-1 and ERK2 following pull-down of LINC01503 or negative RNA control. (F) ChIRP assay confirming the interactions between LINC01503 and EBP-1 or ERK2. Probes against LncRNA A330023F24Rik were served as the negative control. (G) Cell lysates were immunoprecipitated with EBP-1, ERK2 or IgG antibody. Expression levels of LINC01503 in either immunoprecipitates or cell lysates (input) of TE7 and KYSE510 cells were measured by qRT-PCR (mean ± s.d. , n = 3).

Figure 6. LINC01503 activates ERK/MAPK and PI3K/Akt pathways through binding to EBP-1 and ERK2. (A) TE7 or KYSE510 cells transfected with either control siRNA (Scramble) or LINC01503 siRNA (siLINC01503) were serum starved and stimulated with 10 ng/ml of EGF for indicated durations. Cell lysates were analyzed by Western blotting. (B) KYSE510 cells transfected with LINC01503 siRNA were cultured in medium with 10 ng/ml of EGF (3h). Cell lysates were subjected to immunoprecipitation assay following by Western blotting. Bands were quantified by Image J software and the ratio of DUSP-6 to ERK2 was shown as the average of three replicate experiments. (C, D) Effect of DUSP-6 knockdown (C) and overexpression (D) on the phosphorylation of ERK1/2. (E) KYSE510 cells transfected with LINC01503 siRNA were cultured in medium with 10 μM MG132

(6h). Cell lysates were subjected to immunoprecipitation assay following by Western blotting analysis. Bands were quantified by Image J software and the ratio of PI3K p85 to EBP-1 was shown as the average of three replicate experiments. (F, G) Effect of EBP-1 knockdown (F) and overexpression (G) on the expression of PI3K p85.

Figure 7. Therapeutic potential of targeting LINC01503. (A, B) TE7 and KYSE510 cells stably expressing either control shRNA (Scramble) or LINC01503 shRNAs (shLINC01503-1 and shLINC01503-2 are two independent shRNAs against LINC01503) were seeded in soft agar. Bars show the mean numbers of colonies (mean ± s.d., n = 3). (C, D) TE7 cells with stable depletion of LINC01503 were injected subcutaneously into nude mice (10 in each group). Tumor volumes (C) and weights (D) were determined. Data are means ± s.d. (E, F) Activation of the PI3K/Akt and ERK/MAPK pathways in paired ESCC xenografts was determined by Western blotting (E). Signal intensity for the levels of p-Akt and p-ERK1/2 was quantified by densitometry and normalized by the total amounts of each protein (F). (G) Cell viability assay (MTT) shows the effect of inhibitors against Akt (MK2206) or MEK (U0126) on LINC01503 knockdown cells. Cells were treated with the inhibitors for 72 hours. Mean ± s.d. are shown, n = 3. *, $P \leqslant 0.05$.

Figure 1

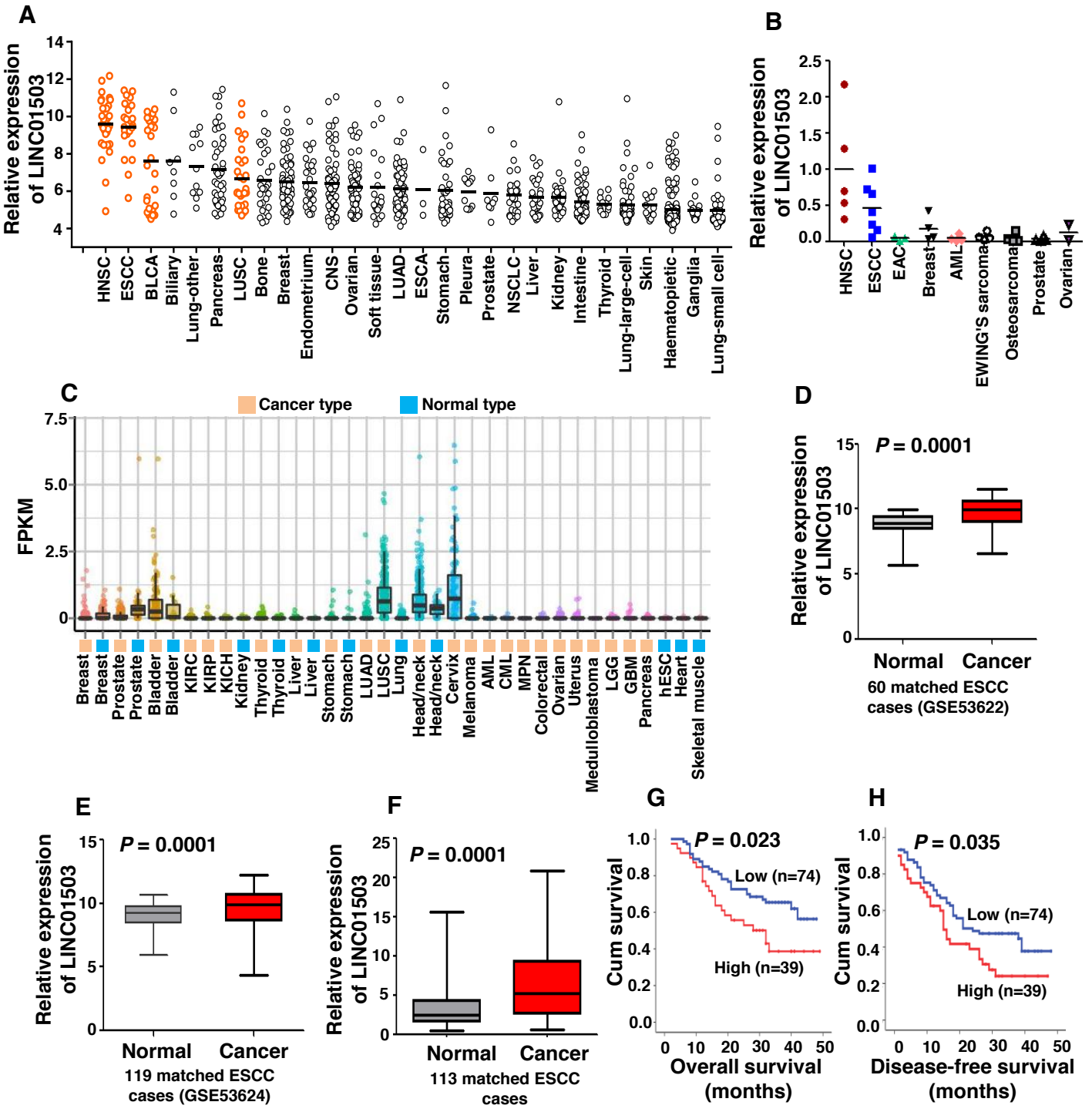

A
Relative expression of LINC01503
HNSC
ESCC
BLCA
Biliary
Lung-other
Pancreas
LUSC
Bone
Breast
Endometrium
CNS
Ovarian
Soft tissue
LUAD
ESCA
Stomach
Pleura
Prostate
NSCLC
Liver
Kidney
Intestine
Thyroid
Lung-large-cell
Skin
Haematopietic
Ganglia
Lung-small cell
B
Relative expression of LINC01503
HNSC
ESCC
EAC
Breast
AML
EWING'S sarcoma
Osteosarcoma
Prostate
Ovarian
C
Cancer type
Normal type
FPKM
Breast
Breast
Prostate
Prostate
Bladder
Bladder
KIRC
KIRP
KICH
Kidney
Thyroid
Thyroid
Liver
Liver
Stomach
Stomach
LUAD
LUSC
Lung
Head/neck
Head/neck
Cervix
Melanoma
AML
CML
MPN
Colorectal
Ovarian
Uterus
Medulloblastoma
LGG
GBM
Pancreas
hESC
Heart
Skeletal muscle
D
P = 0.0001
Relative expression of LINC01503
Normal
Cancer
60 matched ESCC cases (GSE53622)
E
P = 0.0001
Relative expression of LINC01503
Normal
Cancer
119 matched ESCC cases (GSE53624)
F
P = 0.0001
Relative expression of LINC01503
Normal
Cancer
113 matched ESCC cases
G
P = 0.023
Low (n=74)
High (n=39)
Cum survival
Overall survival (months)
H
P = 0.035
Low (n=74)
High (n=39)
Cum survival
Disease-free survival (months)

Figure 2

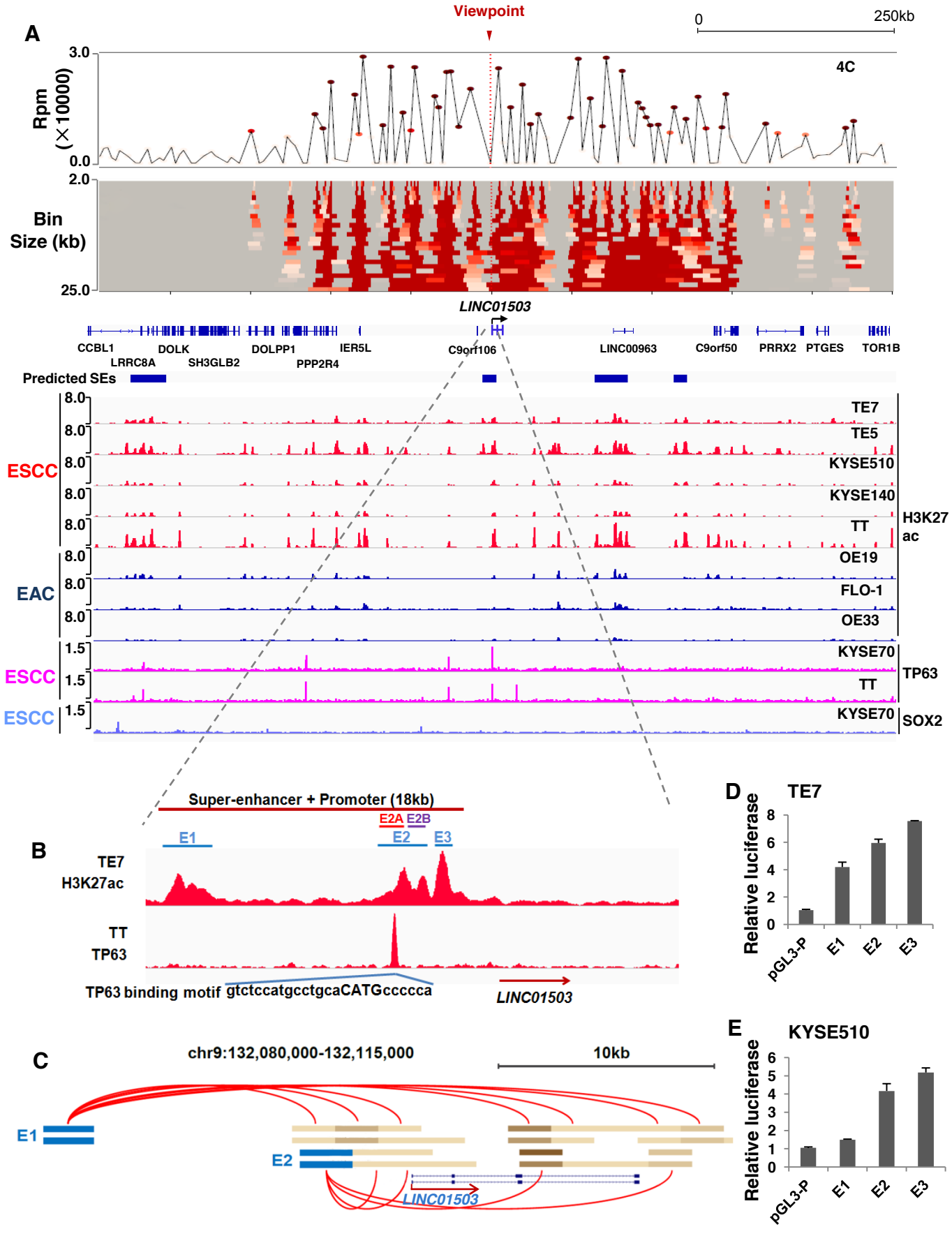

A
Viewpoint
0
250kb
4C
Rpm (×10000)
3.0
0.0
Bin Size (kb)
2.0
25.0
LINC01503
CCBL1
LRRC8A
DOLK
SH3GLB2
DOLPP1
PPP2R4
IER5L
C9orf106
LINC00963
C9orf50
PRRX2
PTGES
TOR1B
Predicted SEs
ESCC
EAC
ESCC
ESCC
TE7
TE5
KYSE510
KYSE140
TT
H3K27ac
OE19
FLO-1
OE33
KYSE70
TT
TP63
KYSE70
SOX2
B
Super-enhancer + Promoter (18kb)
E1
E2A
E2B
E2
E3
TE7
H3K27ac
TT
TP63
TP63 binding motif gtctccatgcctgcaCATGccccca
LINC01503
C
chr9:132,080,000-132,115,000
10kb
E1
E2
LINC01503
D
TE7
Relative luciferase
pGL3-P
E1
E2
E3
E
KYSE510
Relative luciferase
pGL3-P
E1
E2
E3

Figure 3

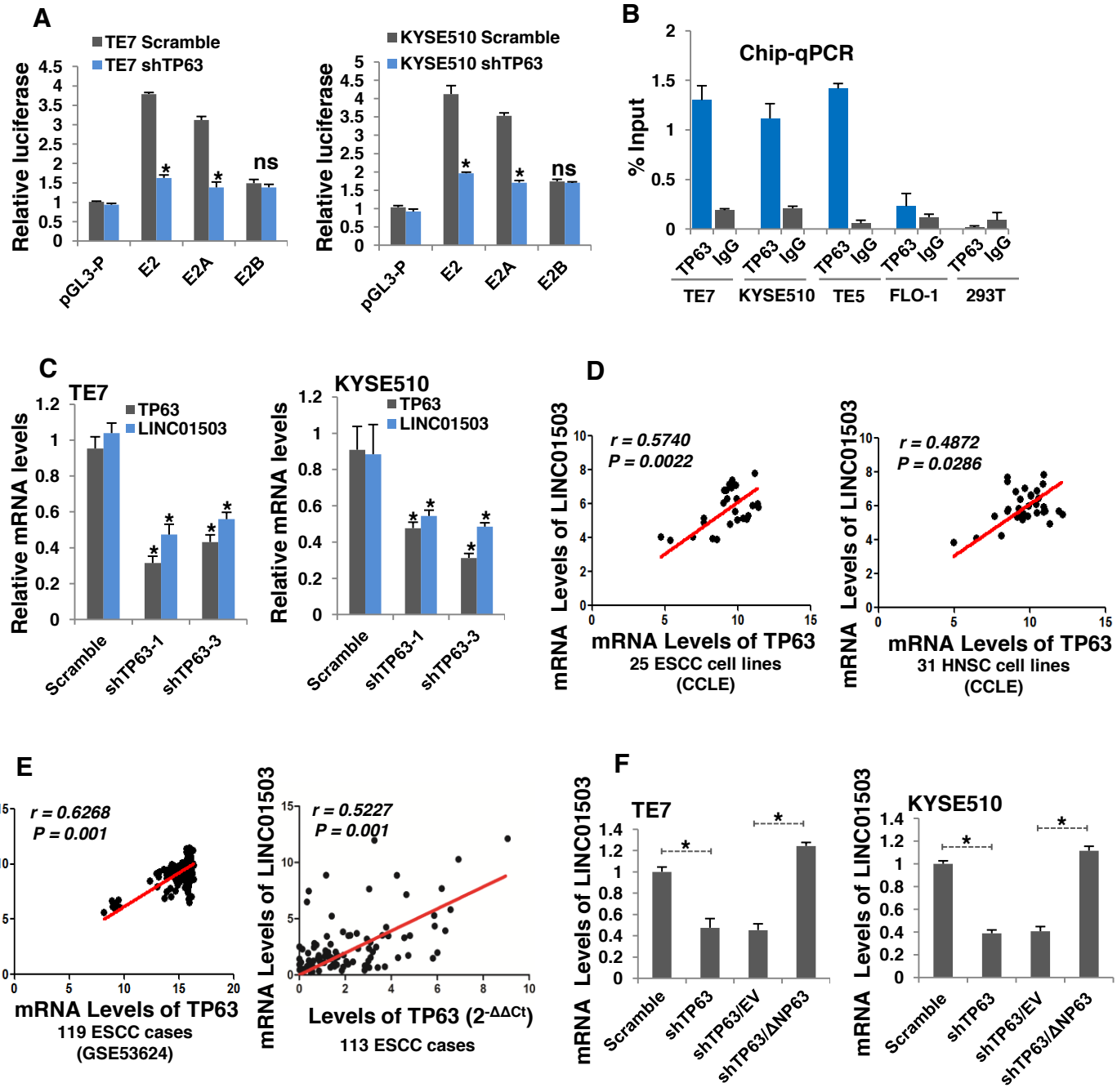

A
TE7 Scramble
TE7 shTP63
Relative luciferase
pGL3-P
E2
E2A
E2B
ns
KYSE510 Scramble
KYSE510 shTP63
B
Chip-qPCR
% Input
TP63
IgG
TE7
KYSE510
TE5
FLO-1
293T
C
TE7
KYSE510
TP63
LINC01503
Relative mRNA levels
Scramble
shTP63-1
shTP63-3
D
r = 0.5740
P = 0.0022
mRNA Levels of LINC01503
mRNA Levels of TP63
25 ESCC cell lines
(CCLE)
r = 0.4872
P = 0.0286
31 HNSC cell lines
E
r = 0.6268
P = 0.001
119 ESCC cases
(GSE53624)
r = 0.5227
P = 0.001
Levels of TP63 (2-ΔΔCt)
113 ESCC cases
F
TE7
KYSE510
shTP63
shTP63/EV
shTP63/ΔNP63

Figure 4

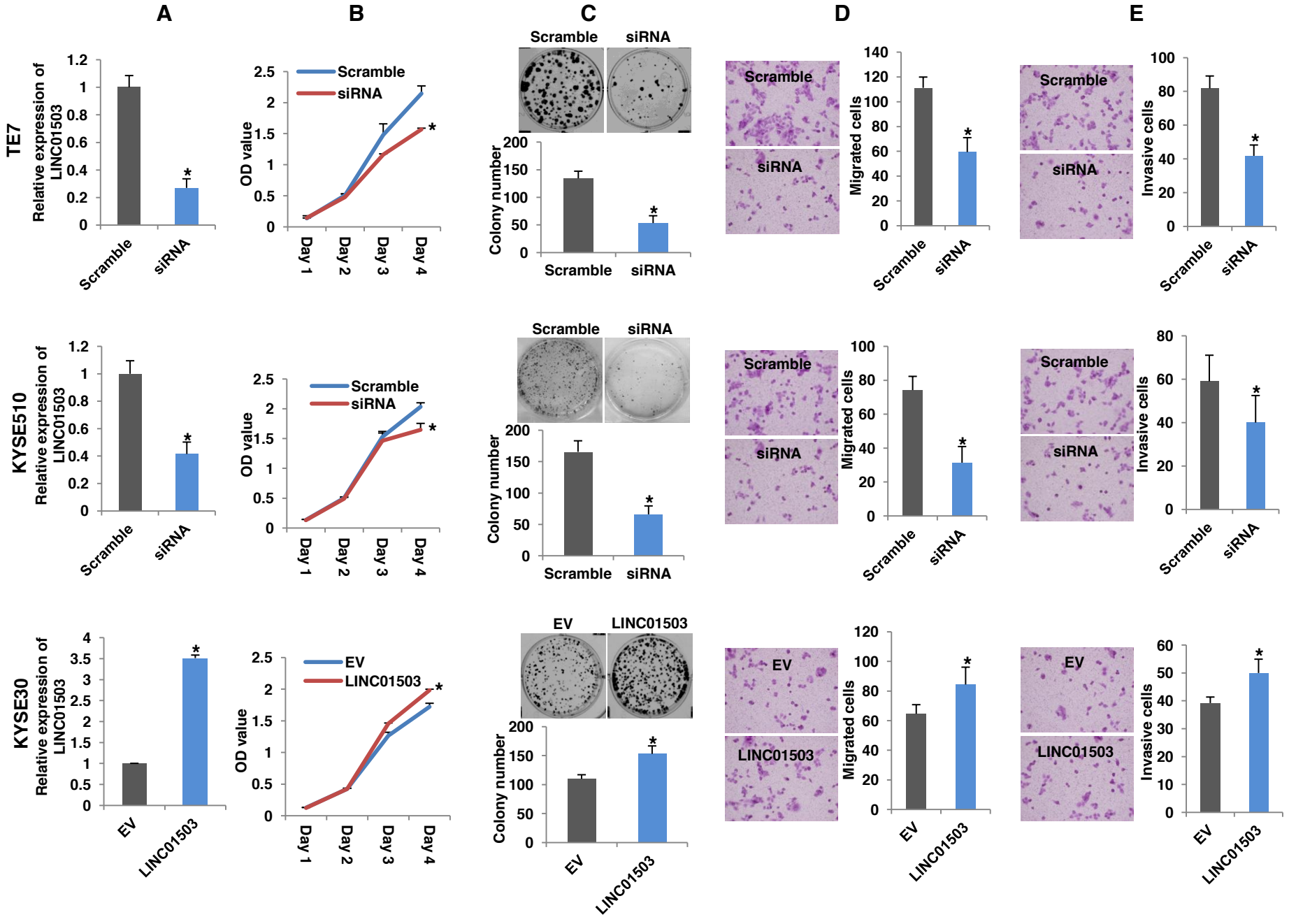
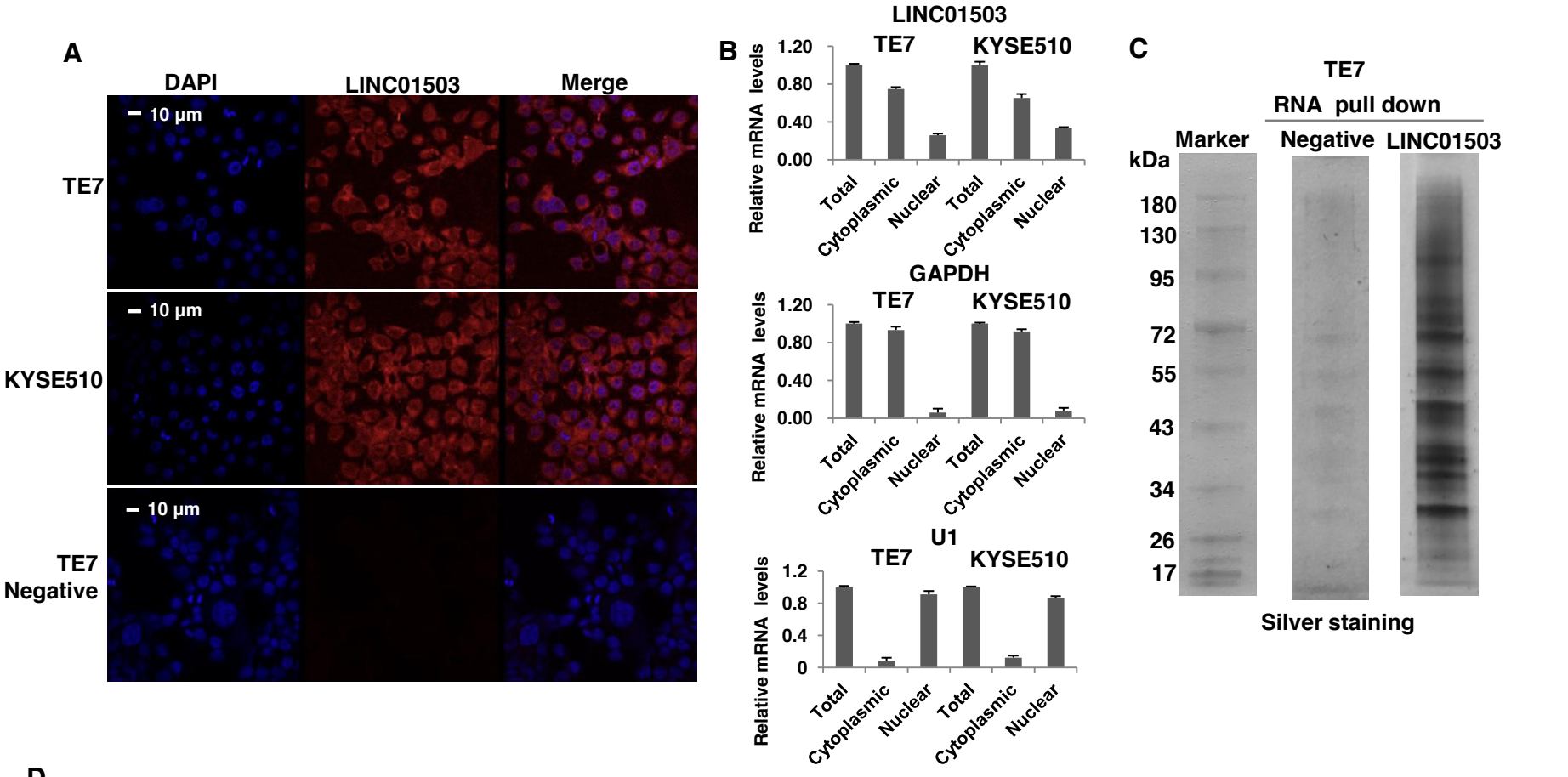

A
B
C
D
E
TE7
KYSE510
KYSE30
Relative expression of LINC01503
Scramble
siRNA
EV
LINC01503
OD value
Day 1
Day 2
Day 3
Day 4
Colony number
Migrated cells
Invasive cells
*

Figure 5



D

Protein name	Sequence coverage (%)	Counts (Negative)	Counts (01503)
YBOX3	13.7	0	37
COF1	32.5	2	10
UBQL1	5.8	0	10
PPIA	13.3	0	10
EBP-1	4.8	1	7
14-3-3σ*	9.7	0	8
HS71B*	8	1	7
TBA1C	12.5	0	8
ERK2	3.9	1	5
14-3-3ζ*	15.1	0	4

*, predicted or known RNA-binding proteins (Ref 38)

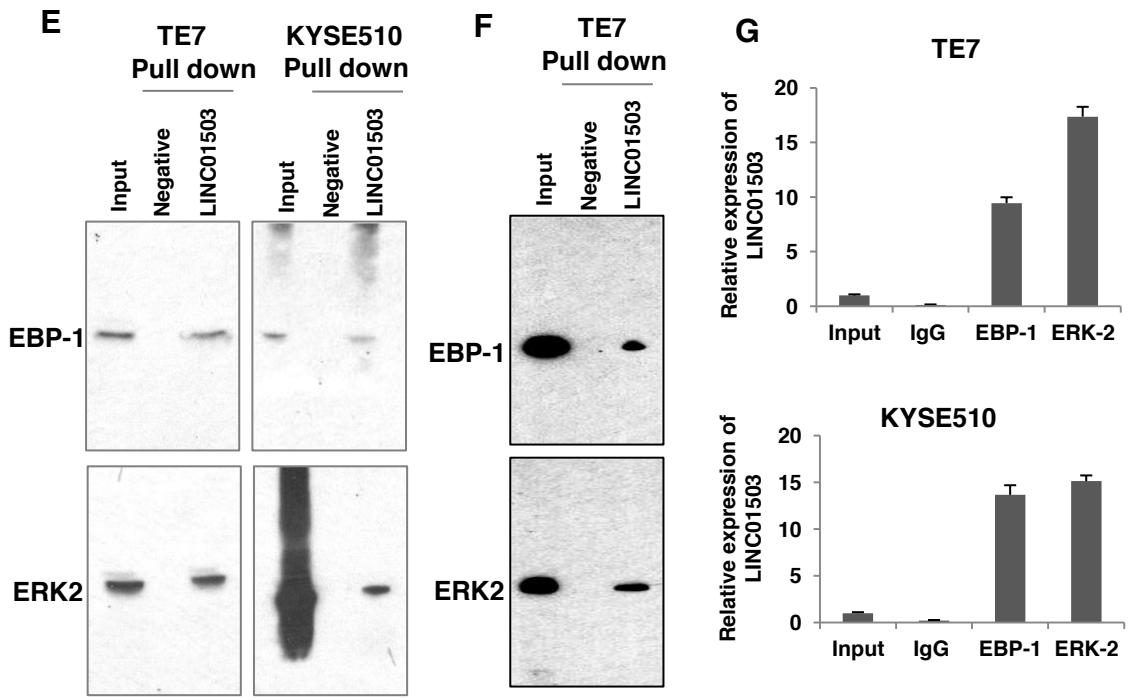


Figure 6

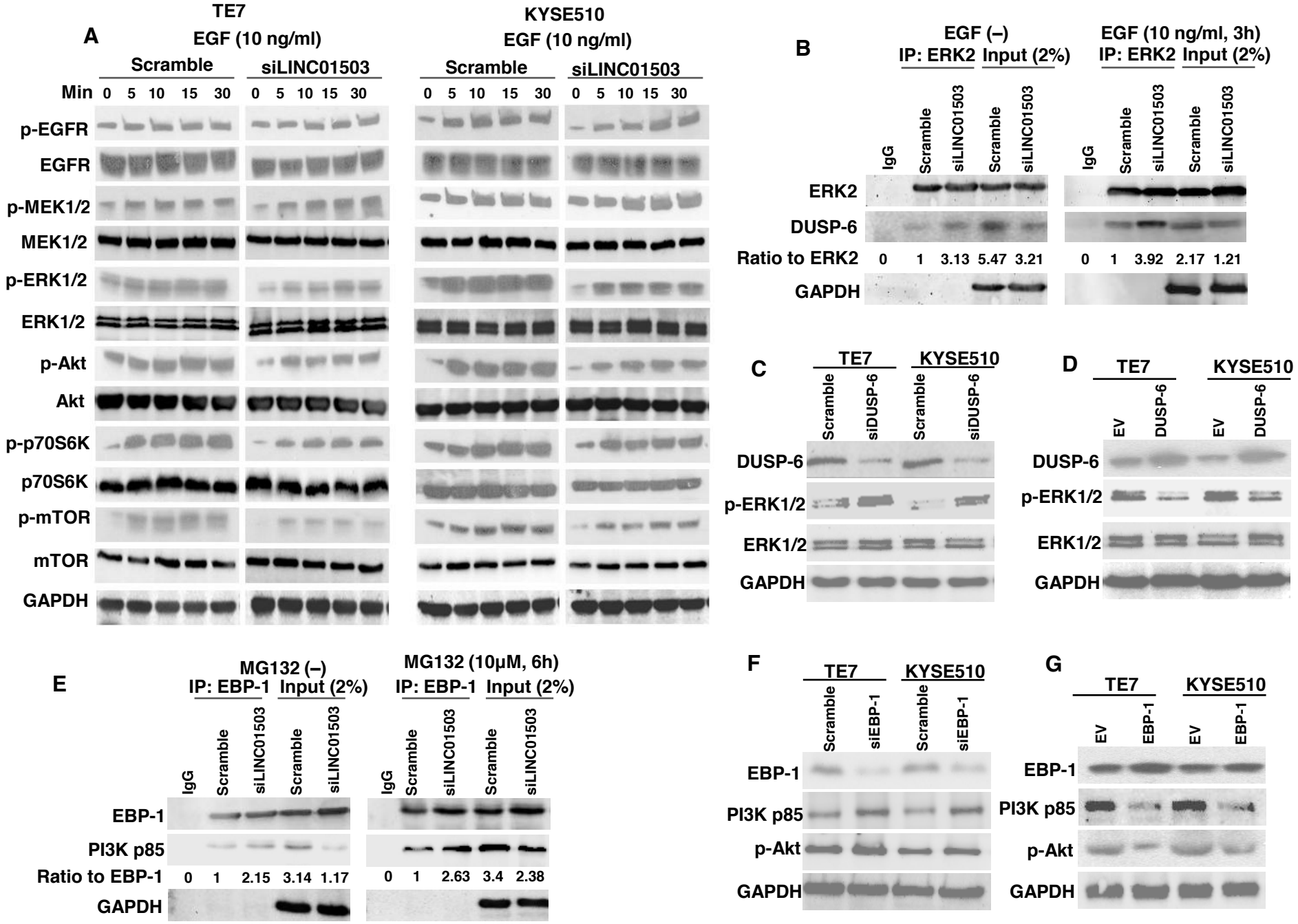
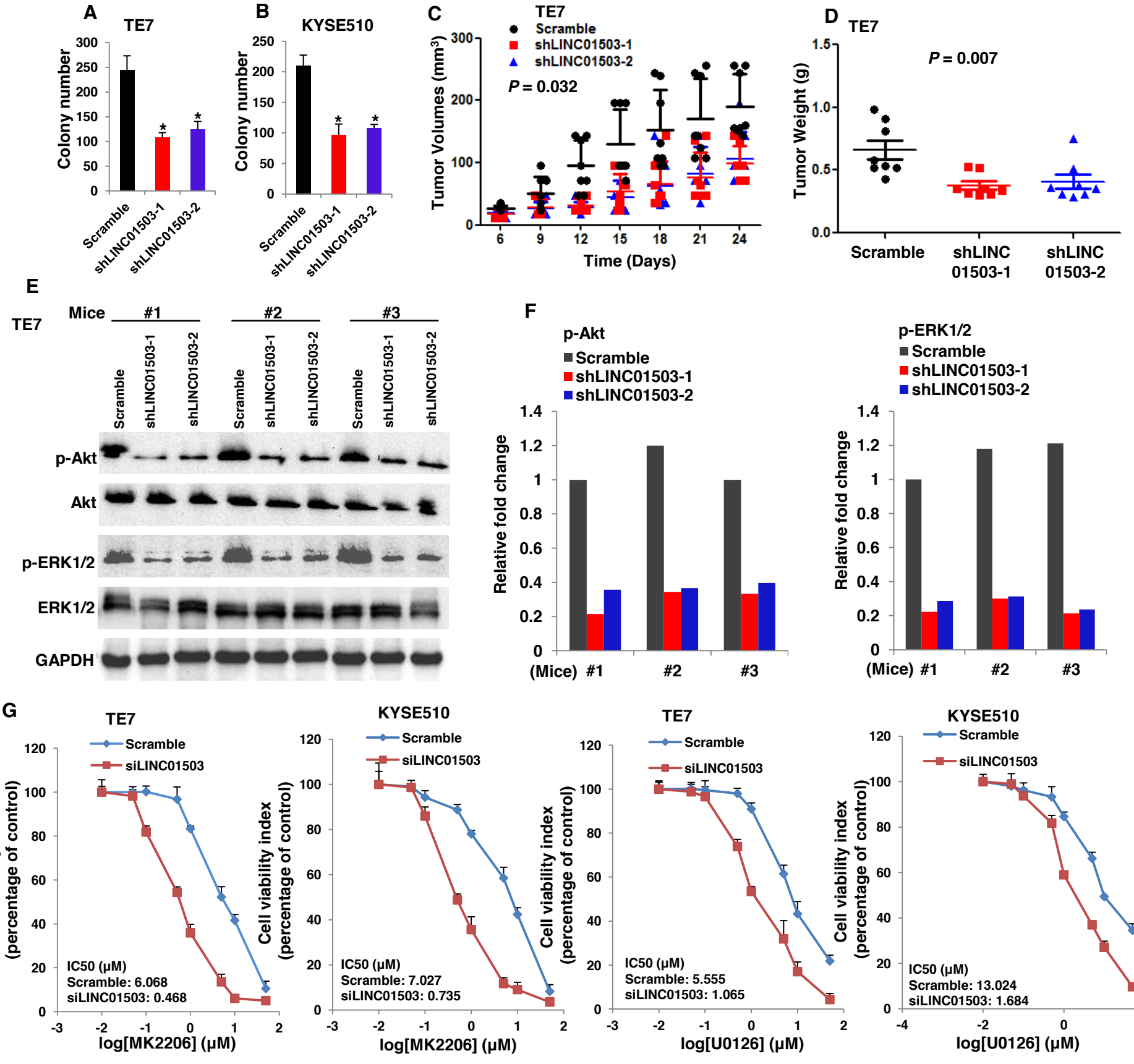
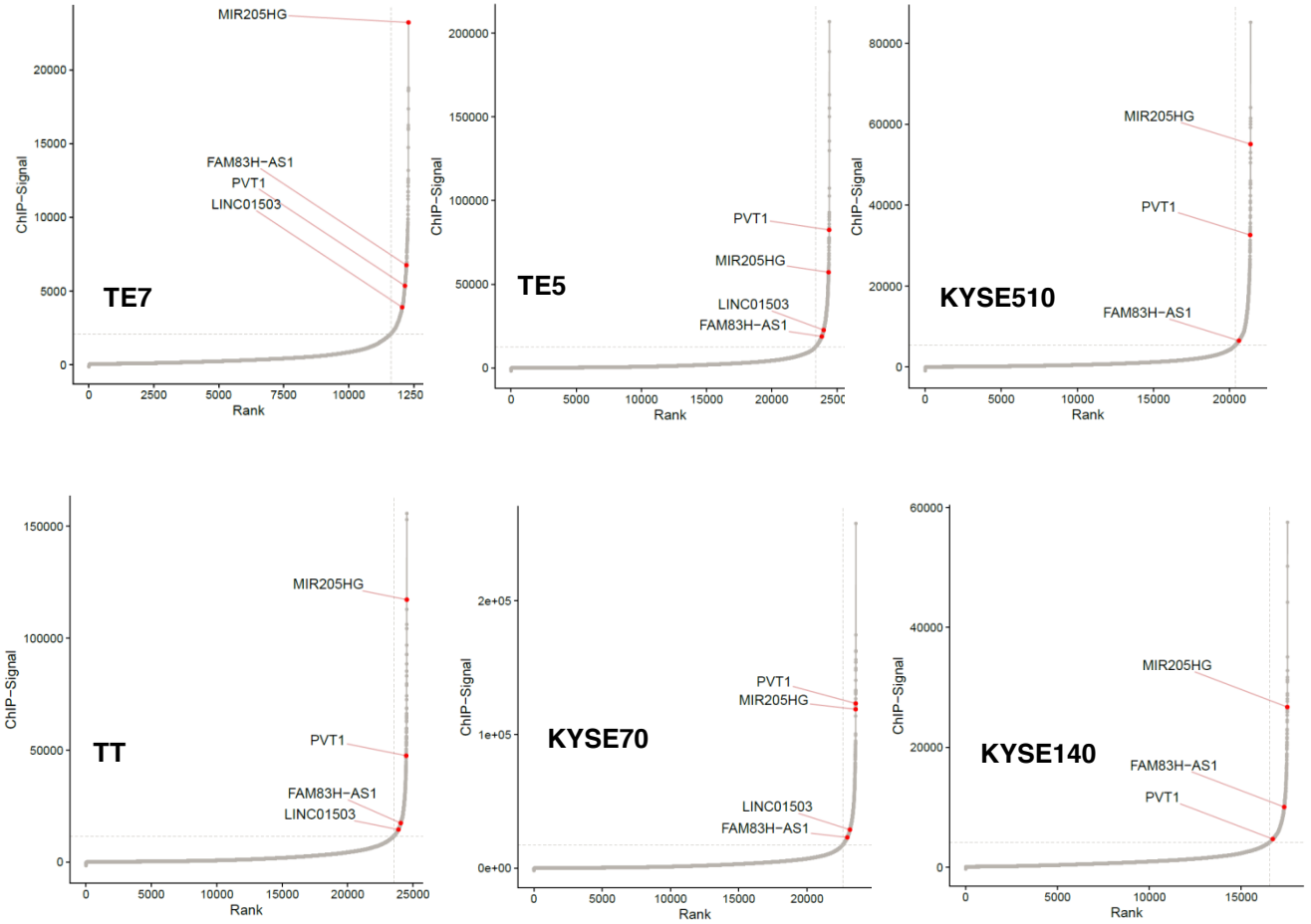

A
TE7
EGF (10 ng/ml)
Scramble
siLINC01503
KYSE510
EGF (10 ng/ml)
Scramble
siLINC01503
Min 0 5 10 15 30 0 5 10 15 30
Min 0 5 10 15 30 0 5 10 15 30
p-EGFR
EGFR
p-MEK1/2
MEK1/2
p-ERK1/2
ERK1/2
p-Akt
Akt
p-p70S6K
p70S6K
p-mTOR
mTOR
GAPDH
B
EGF (–)
IP: ERK2 Input (2%)
EGF (10 ng/ml, 3h)
IP: ERK2 Input (2%)
IgG
Scramble
siLINC01503
Scramble
siLINC01503
IgG
Scramble
siLINC01503
Scramble
siLINC01503
ERK2
DUSP-6
Ratio to ERK2 0 1 3.13 5.47 3.21 0 1 3.92 2.17 1.21
GAPDH
C
TE7
KYSE510
Scramble
siDUSP-6
Scramble
siDUSP-6
DUSP-6
p-ERK1/2
ERK1/2
GAPDH
D
TE7
KYSE510
EV
DUSP-6
EV
DUSP-6
DUSP-6
p-ERK1/2
ERK1/2
GAPDH
E
MG132 (–)
IP: EBP-1 Input (2%)
MG132 (10μM, 6h)
IP: EBP-1 Input (2%)
IgG
Scramble
siLINC01503
Scramble
siLINC01503
IgG
Scramble
siLINC01503
Scramble
siLINC01503
EBP-1
PI3K p85
Ratio to EBP-1 0 1 2.15 3.14 1.17 0 1 2.63 3.4 2.38
GAPDH
F
TE7
KYSE510
Scramble
siEBP-1
Scramble
siEBP-1
EBP-1
PI3K p85
p-Akt
GAPDH
G
TE7
KYSE510
EV
EBP-1
EV
EBP-1
EBP-1
PI3K p85
p-Akt
GAPDH

Figure 7

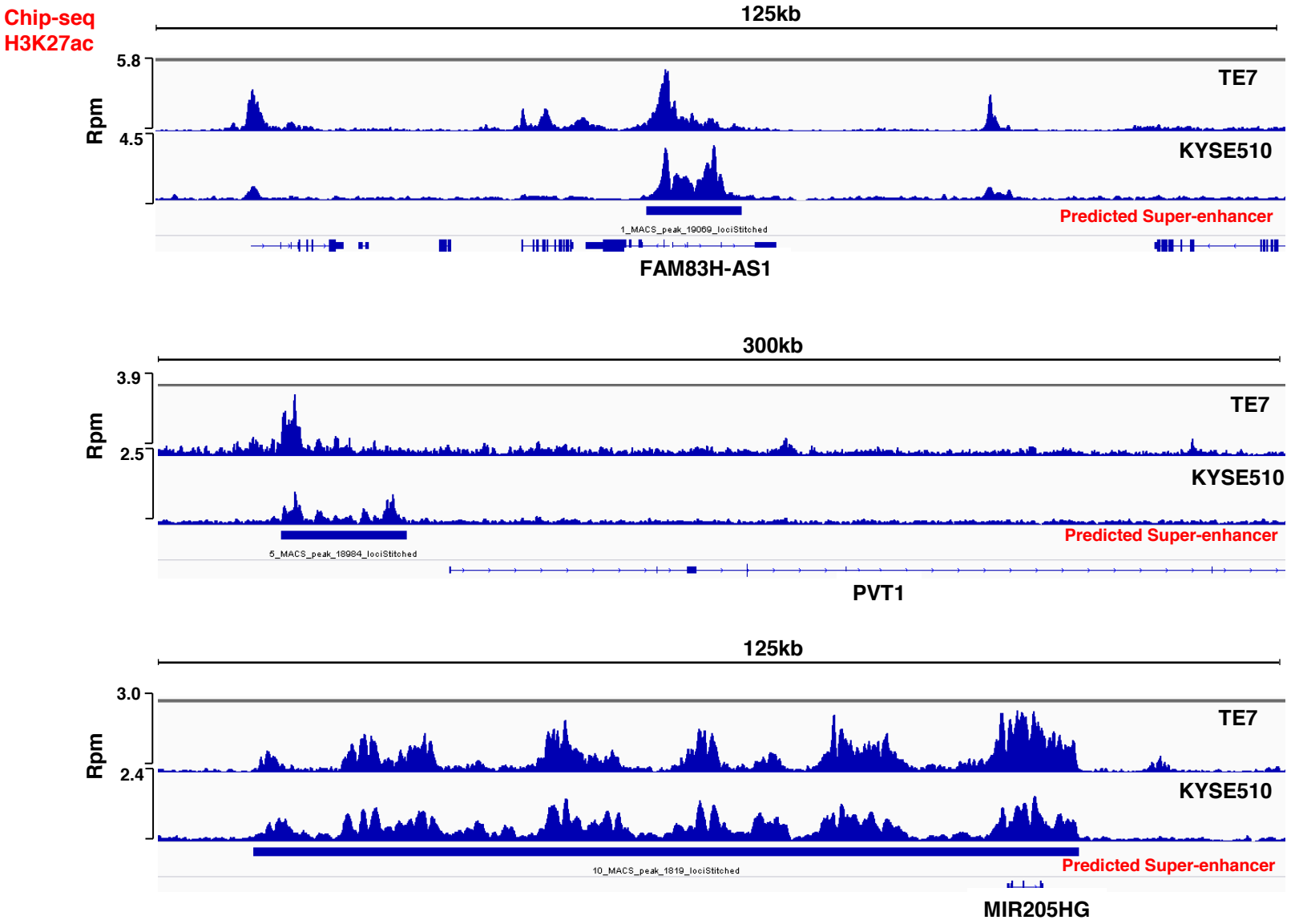

A
TE7
Colony number
Scramble
shLINC01503-1
shLINC01503-2
B
KYSE510
Colony number
C
TE7
Scramble
shLINC01503-1
shLINC01503-2
P = 0.032
Tumor Volumes (mm3)
Time (Days)
D
TE7
P = 0.007
Tumor Weight (g)
Scramble
shLINC 01503-1
shLINC 01503-2
E
TE7
Mice
#1
#2
#3
p-Akt
Akt
p-ERK1/2
ERK1/2
GAPDH
F
p-Akt
p-ERK1/2
Relative fold change
(Mice) #1
#2
#3
G
TE7
KYSE510
siLINC01503
Cell viability index (percentage of control)
IC50 (μM)
Scramble: 6.068
siLINC01503: 0.468
Scramble: 7.027
siLINC01503: 0.735
Scramble: 5.555
siLINC01503: 1.065
Scramble: 13.024
siLINC01503: 1.684
log[MK2206] (μM)
log[U0126] (μM)

Supplementary figures

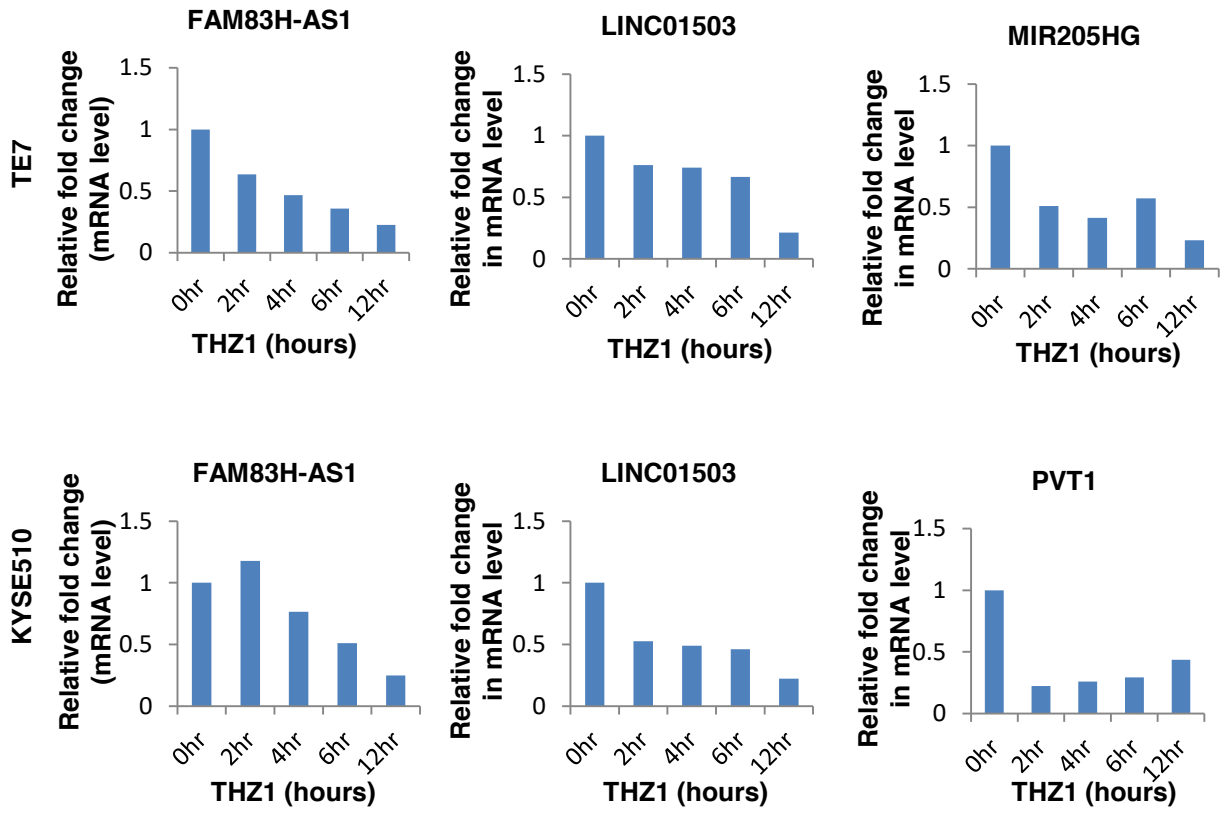
Click here to download Supporting Document supplementary figures-revised.pdf



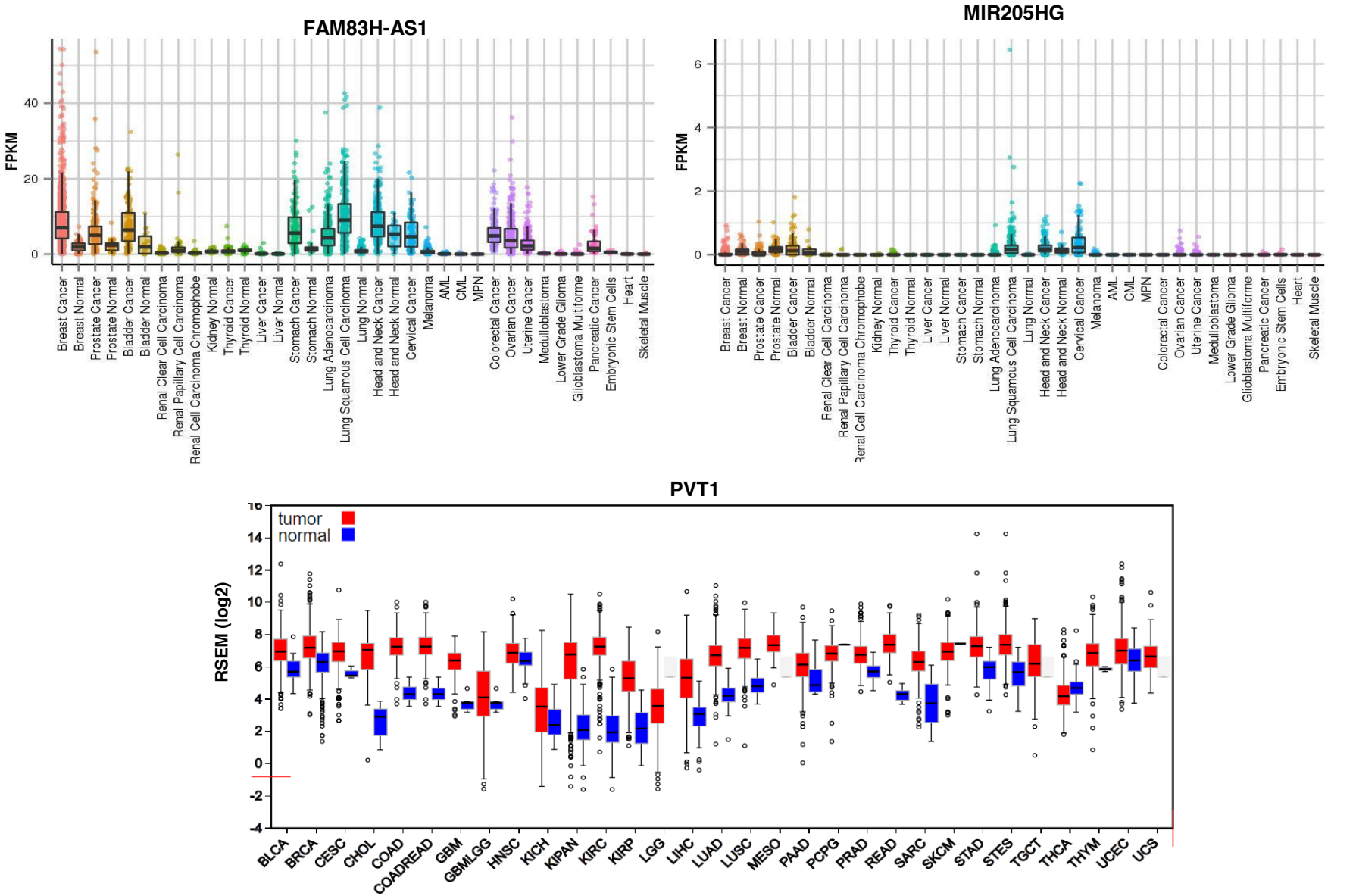
Supplementary Figure 1. Hockey stick plots showing input normalized, rank ordered H3K27ac signals for the candidate SE-associated lncRNAs in this study.



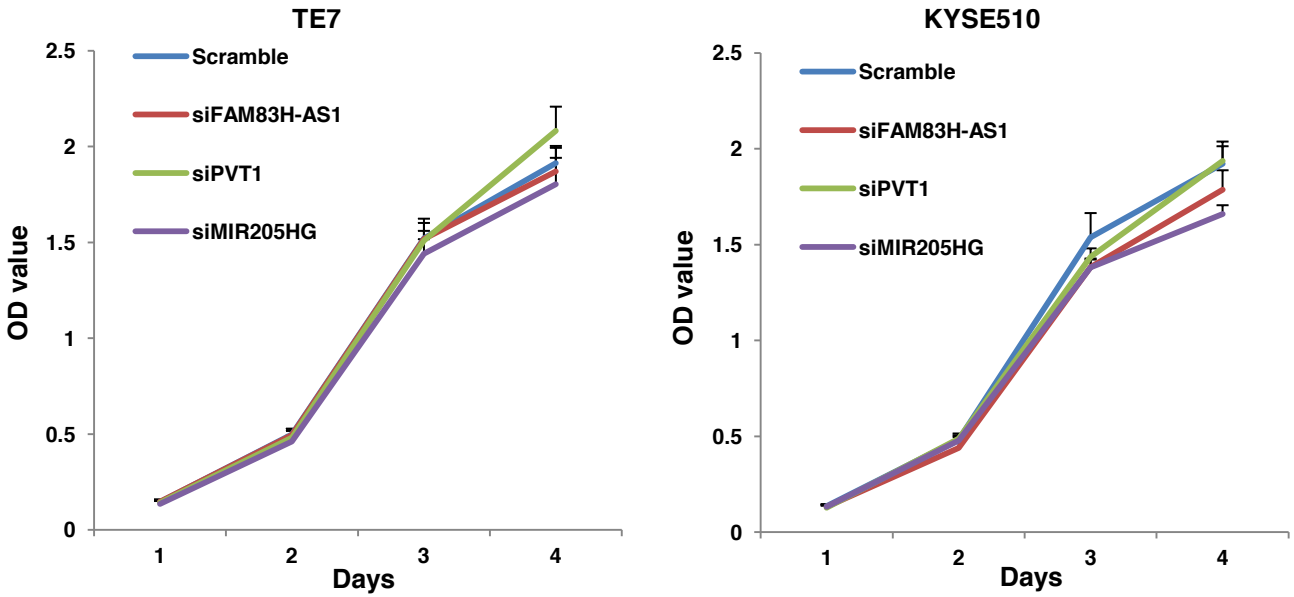
Supplementary Figure 2. H3K27ac ChIP-Seq binding profiles of representative SE-associated lncRNAs in ESCC cell lines.



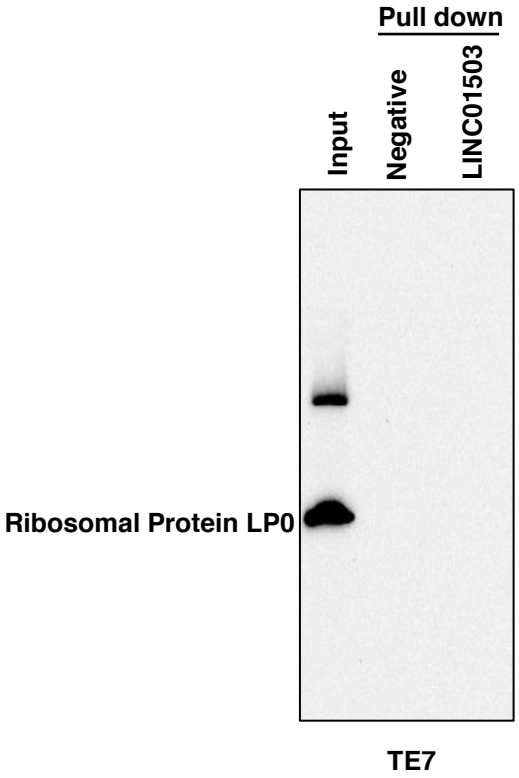
Supplementary Figure 3. Real-time PCR validation of RNA-Seq results of candidate SE-associated lncRNAs upon THZ1 (CDK7 inhibitor, 50nM) treatment in TE7 and KYSE510 cells.



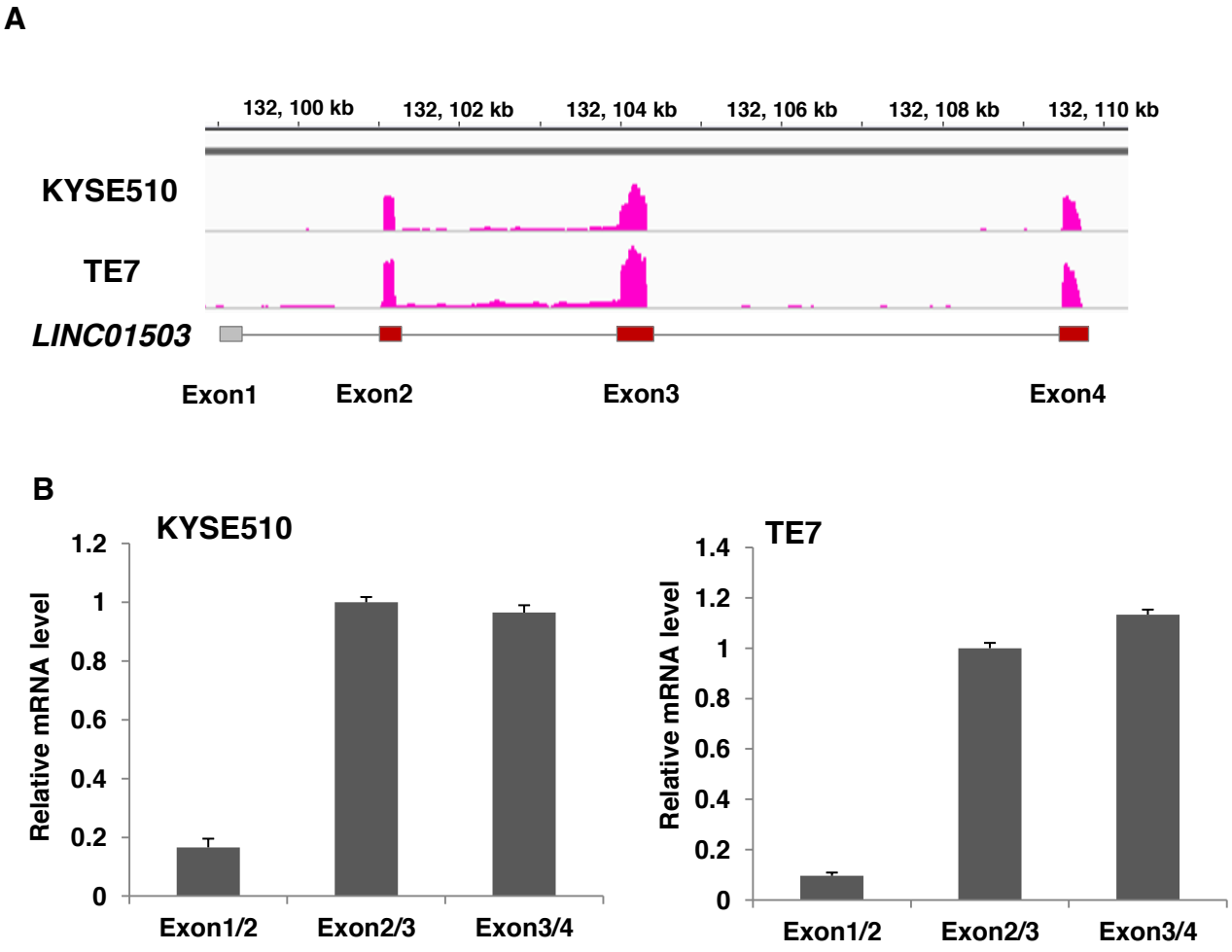
Supplementary Figure 4. Expression characteristics of candidate SE-associated LncRNAs identified in the study. Plots highlighting the expression of FAM83H-AS1, MIR205HG (http://mitranscriptome.org/) and PVT1 (http://firebrowse.org/) in different cancer/tissue types. Each point represents one RNA-seq tissue sample.



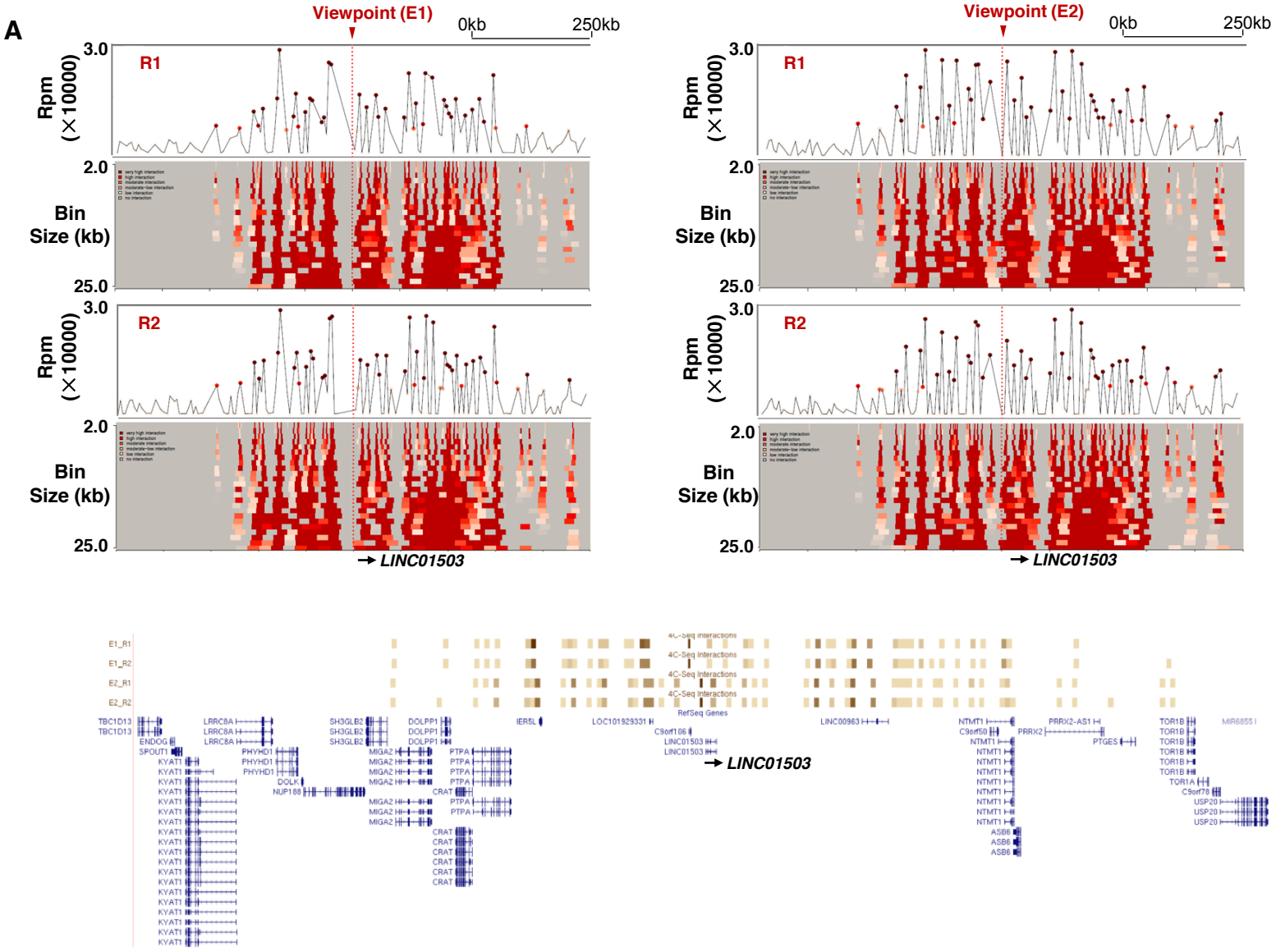
Supplementary Figure 5. Effect of lncRNAs on cell proliferation. siRNA-pools against FAM83H-AS1, PVT1 or MIR205HG were transfected into TE7 and KYSE510 cells, and MTT assays were performed to measure cell proliferation. Mean ± s.d. are shown, n = 3.



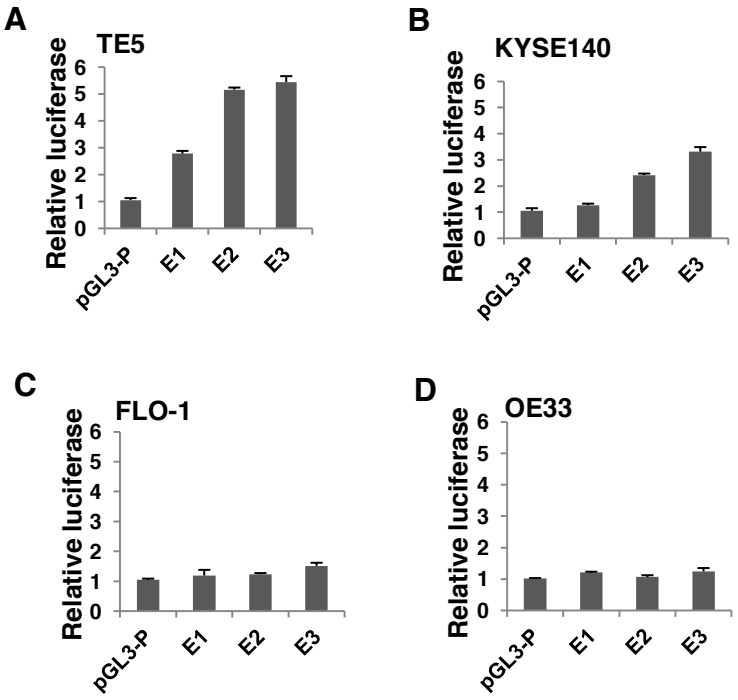
Supplementary Figure 6. Western blotting analysis of Ribosomal Protein LP0 (RPLP0) following pull-down of LINC01503 or negative RNA (LncRNA A330023F24Rik).



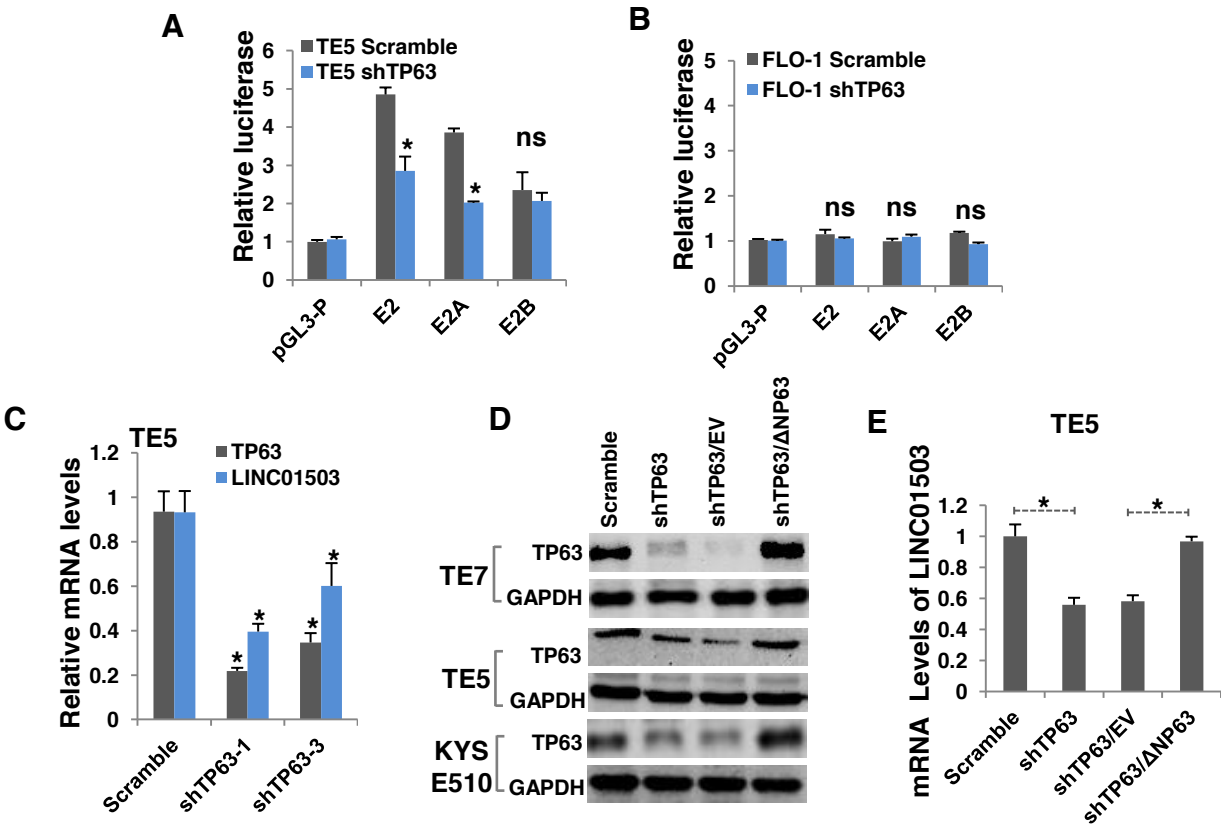
Supplementary Figure 7. A new splice isoform of LINC01503 identified in ESCC cell lines. A, RNA-Seq analysis of the new splice isoform of LINC01503 containing Exons 2-4. B, qRT-PCR validation.



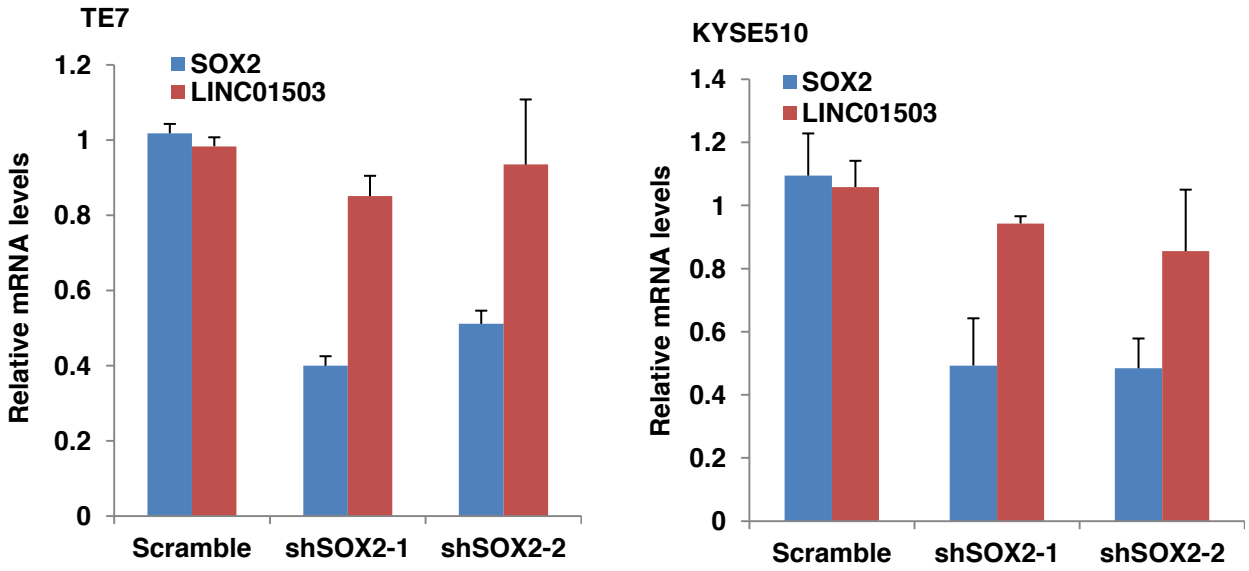
Supplementary Figure 8. (A) 4C-seq showing long-range interactions anchored on the SE upstream of LINC01503 in TE7 cells. Two repeats (R1 and R2) in each viewpoint (E1 and E2). (B) Significant interactions determined by 4C. Deeper red colors indicate higher interaction frequency.



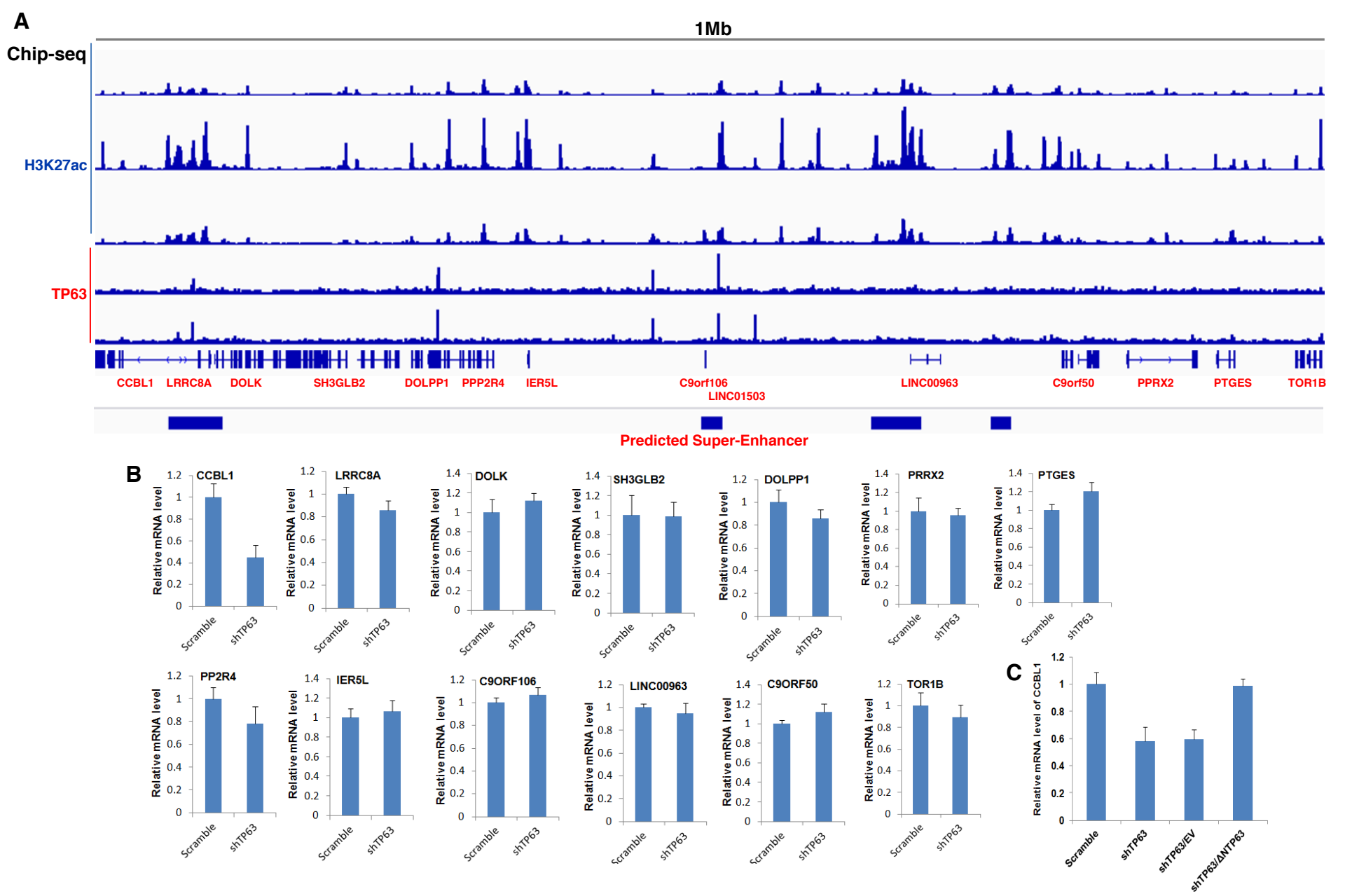
Supplementary Figure 9. Enhancer activity measured by luciferase reporter assays in ESCC cells (A and B) and EAC cells (C and D). Mean ± s.d. are shown, n = 3.



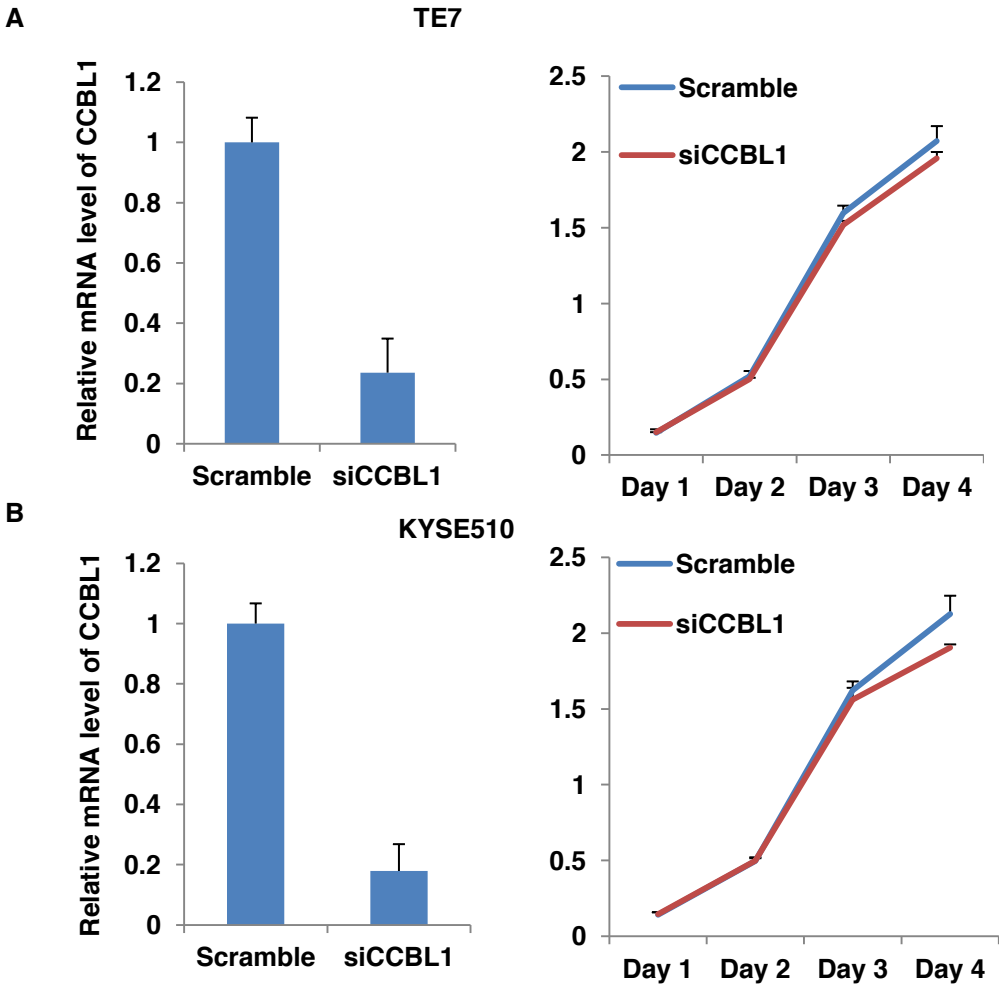
Supplementary Figure 10. TP63 activates the transcription of LINC01503 by mediating the activity of its super-enhancer. (A and B) E2 segment was divided into two constituents (E2A and E2B) and the luciferase activities of E2, E2A and E2B were measured in ESCC cells (A) and EAC cells (B) upon TP63 silencing. (C) Expression of LINC01503 and TP63 was determined by qRT-PCR in TP63 knockdown ESCC cells using two independent shRNAs (shTP63-1 and sh-TP63-3). (D, E) ΔNTP63 expression plasmid was transfected into TP63 knockdown ESCC cells; Western blotting was used to detect the expression of ΔNTP63 (D). The levels of LINC01503 were recovered upon ΔNTP63 expression (E). Mean ± s.d. are shown, n = 3. n.s., not significant. *, $P \leq 0.05$.



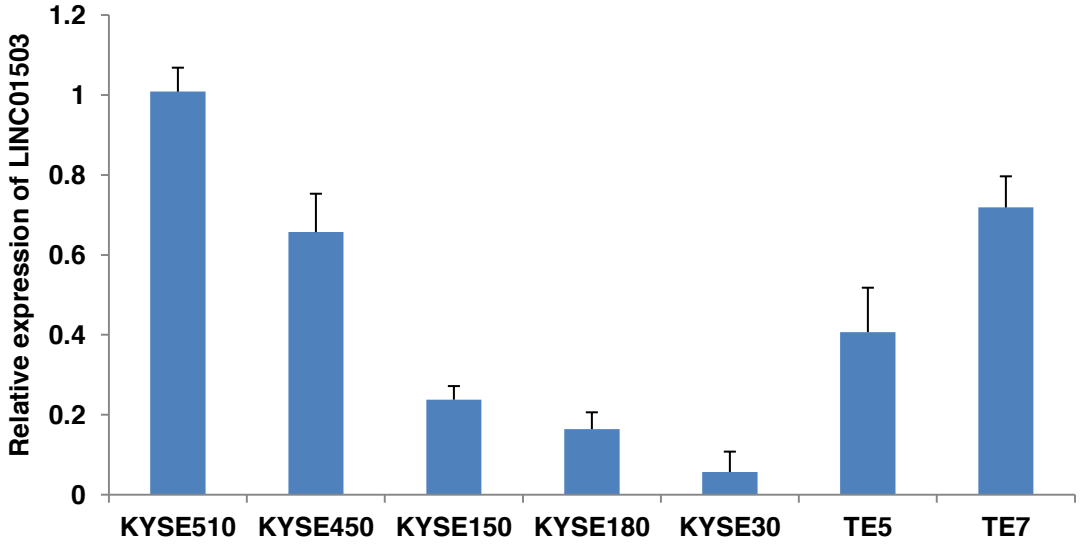
Supplementary Figure 11. Effect of SOX2 knockdown on the expression levels of LINC01503 was determined by qRT-PCR.



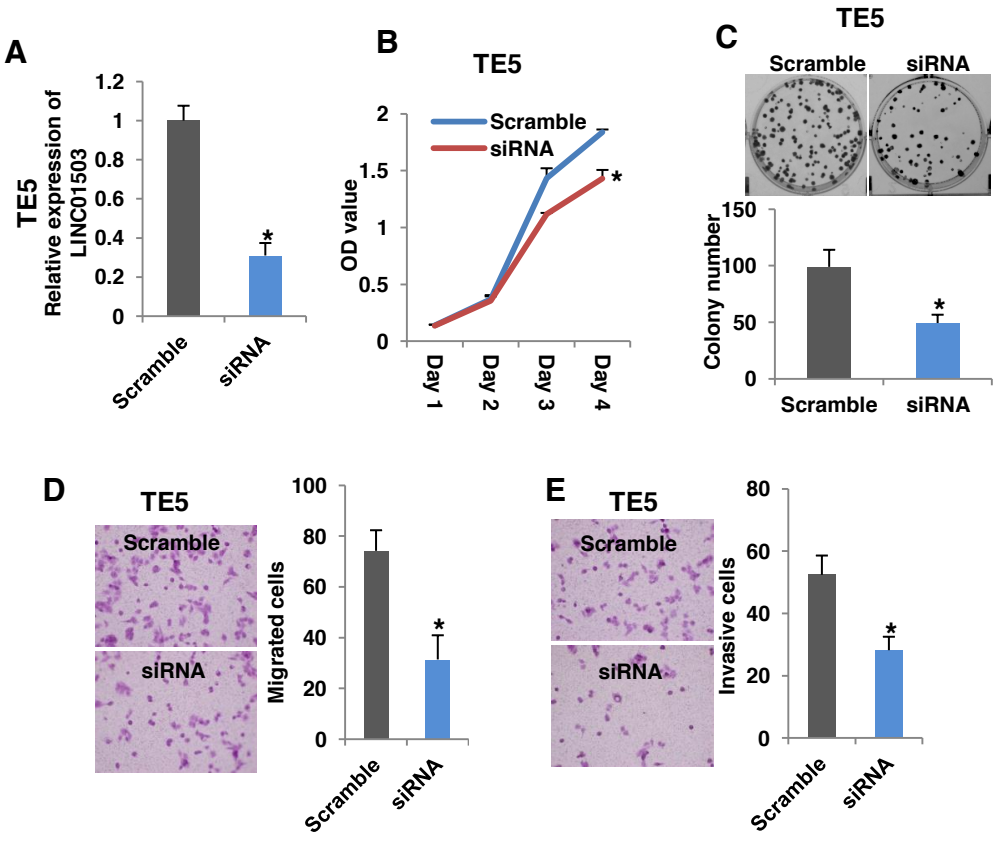
Supplementary Figure 12. Effect of TP63 silencing on the expressions of 13 transcripts within 1Mb of the LINC01503-SE. (A) Transcripts located in the region within 1Mb of the LINC01503-SE. (B) qPCR analysis for the expressions of these transcripts in TP63 silencing cells. (C) ΔNTP63 expression plasmid was transfected into TP63 knockdown ESCC cells and the level of CCBL1 was recovered upon ΔNTP63 expression.



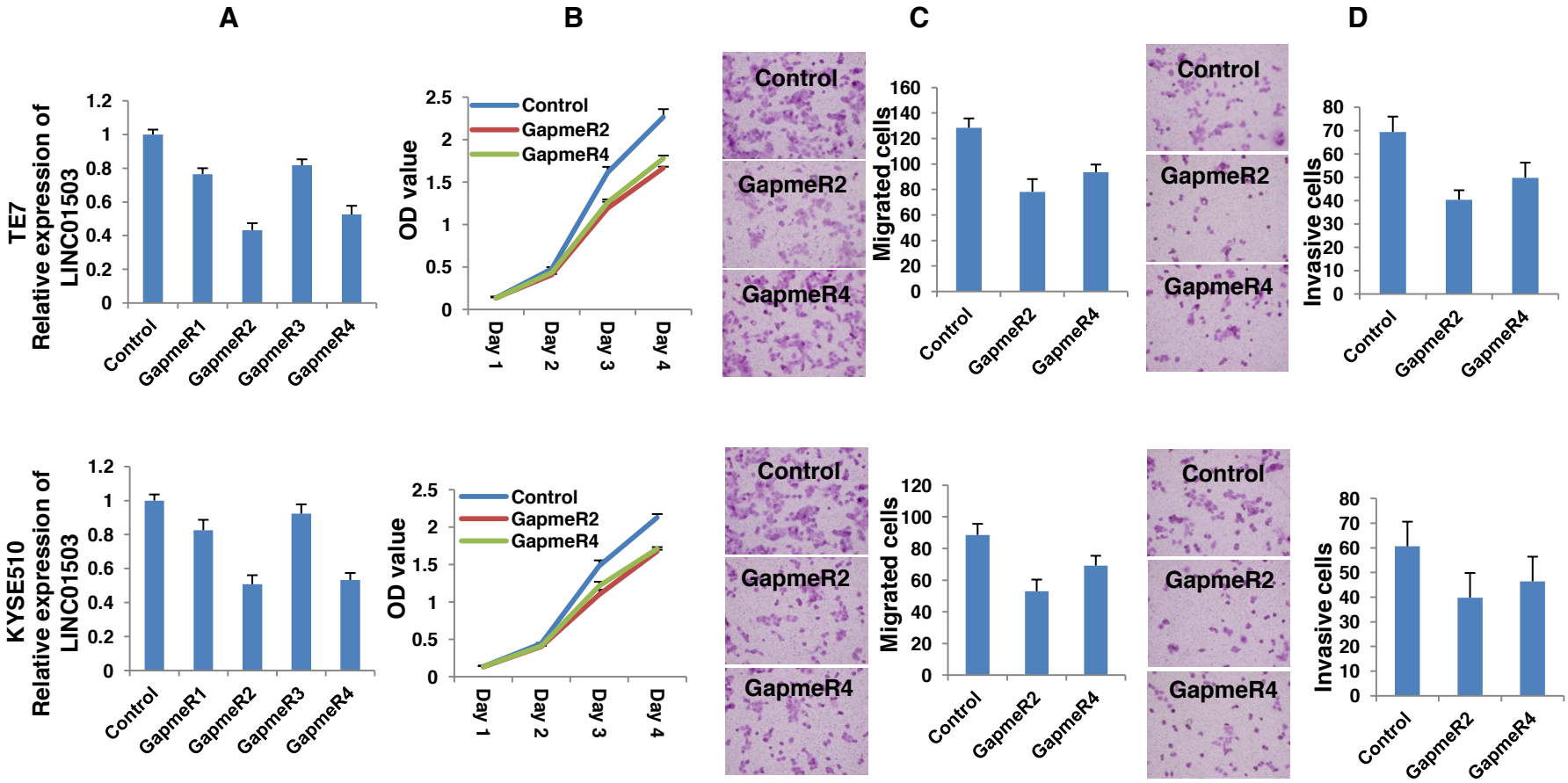
Supplementary Figure 13. Effect of CCBL1 knockdown on the proliferation of ESCC cells. (A) TE7 cells. (B) KYSE510 cells. Left, qRT-PCR analysis for the silencing of CCBL1. Right, MTT assays. Mean ± s.d. are shown, n = 3.



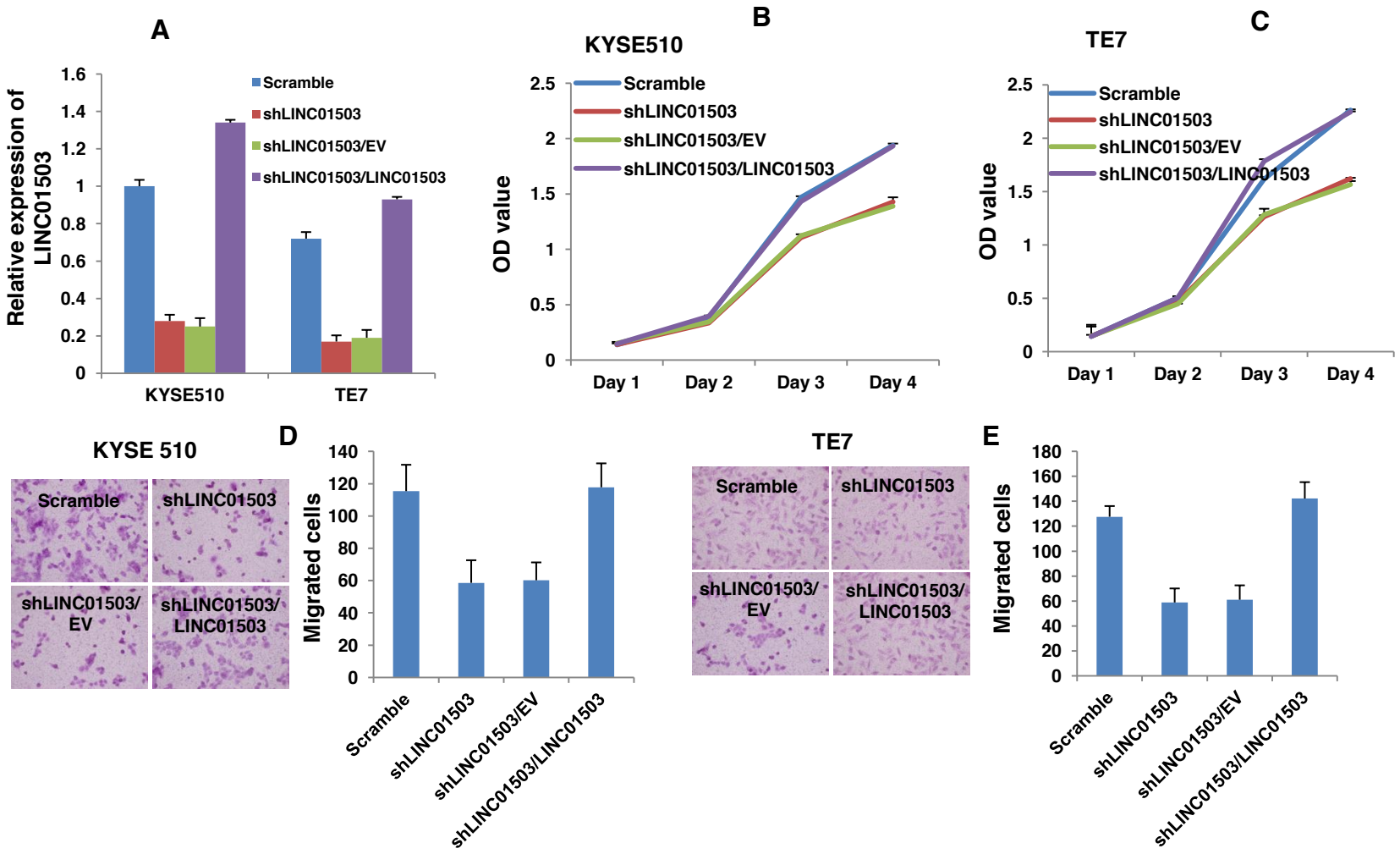
Supplementary Figure 14. qRT-PCR analysis for the expression levels of LINC01503 in ESCC cell lines. . Mean ± s.d. are shown, n = 3.



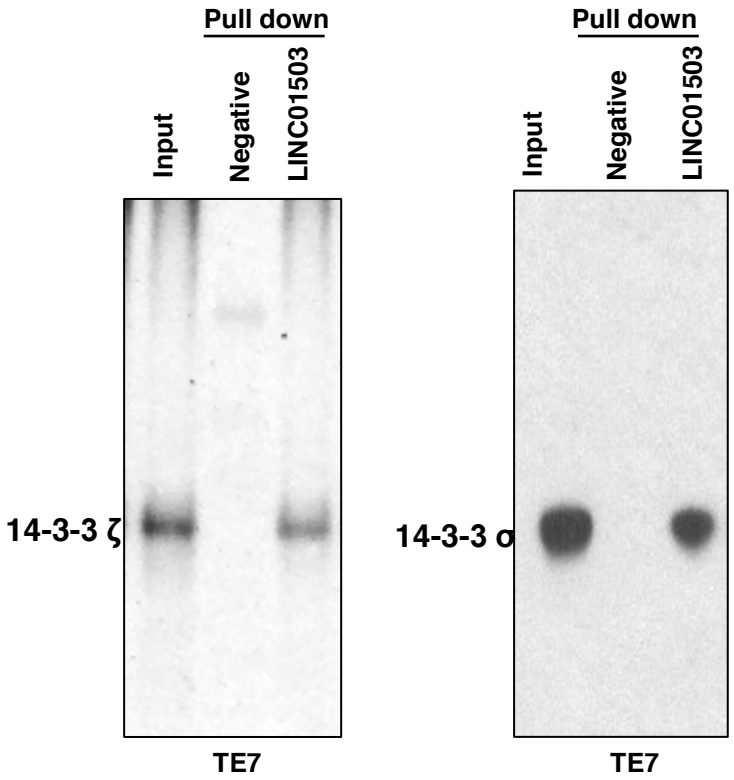
Supplementary Figure 15. LINC01503 is required for the growth, migration and invasion of ESCC cells. (A) qRT-PCR analysis validating the silencing of LINC01503 in TE5 cells. EV, empty vector. (B) MTT and (C) colony formation assays in TE5 cells upon LINC01503 knockdown. (D) Transwell migration assays and (E) Transwell Matrigel invasion assays were performed to determine the effect of LINC01503 on the migration and invasiveness of TE5 cells. Mean ± s.d. are shown, n = 3. n.s., not significant. *, $P \leq 0.05$.



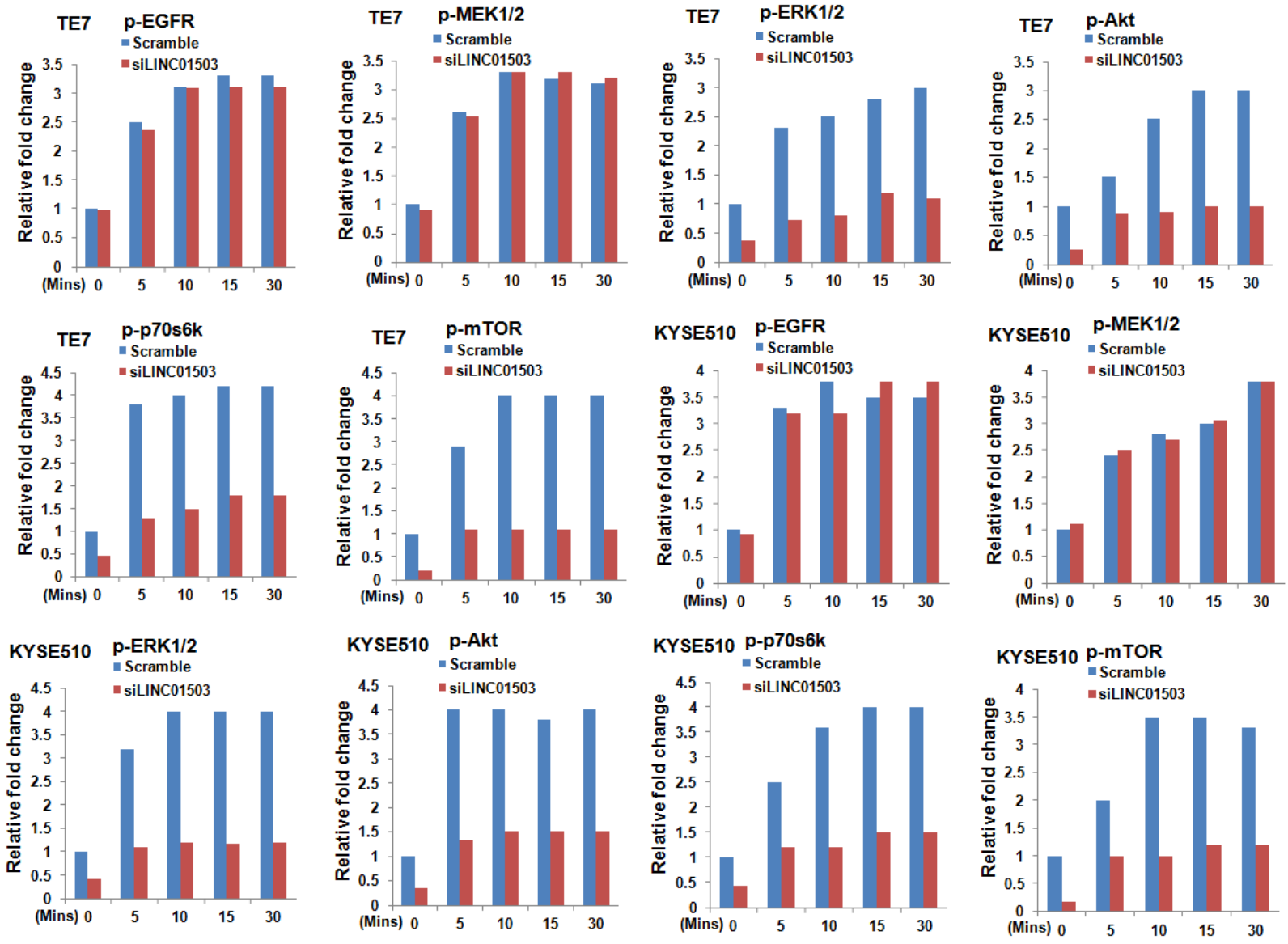
Supplementary Figure 16. LINC01503 is required for the growth, migration and invasiveness of ESCC cells. A, qRT-PCR analysis for the GapmeR-mediated silencing (TE7 and KYSE510 cells) of LINC01503. B, MTT assays in ESCC cells upon LINC01503 knockdown or overexpression. C and D, Transwell migration assays (C) and Transwell Matrigel invasion assays (D) were performed upon silencing of LINC01503. Mean ± s.d. are shown, n = 3.



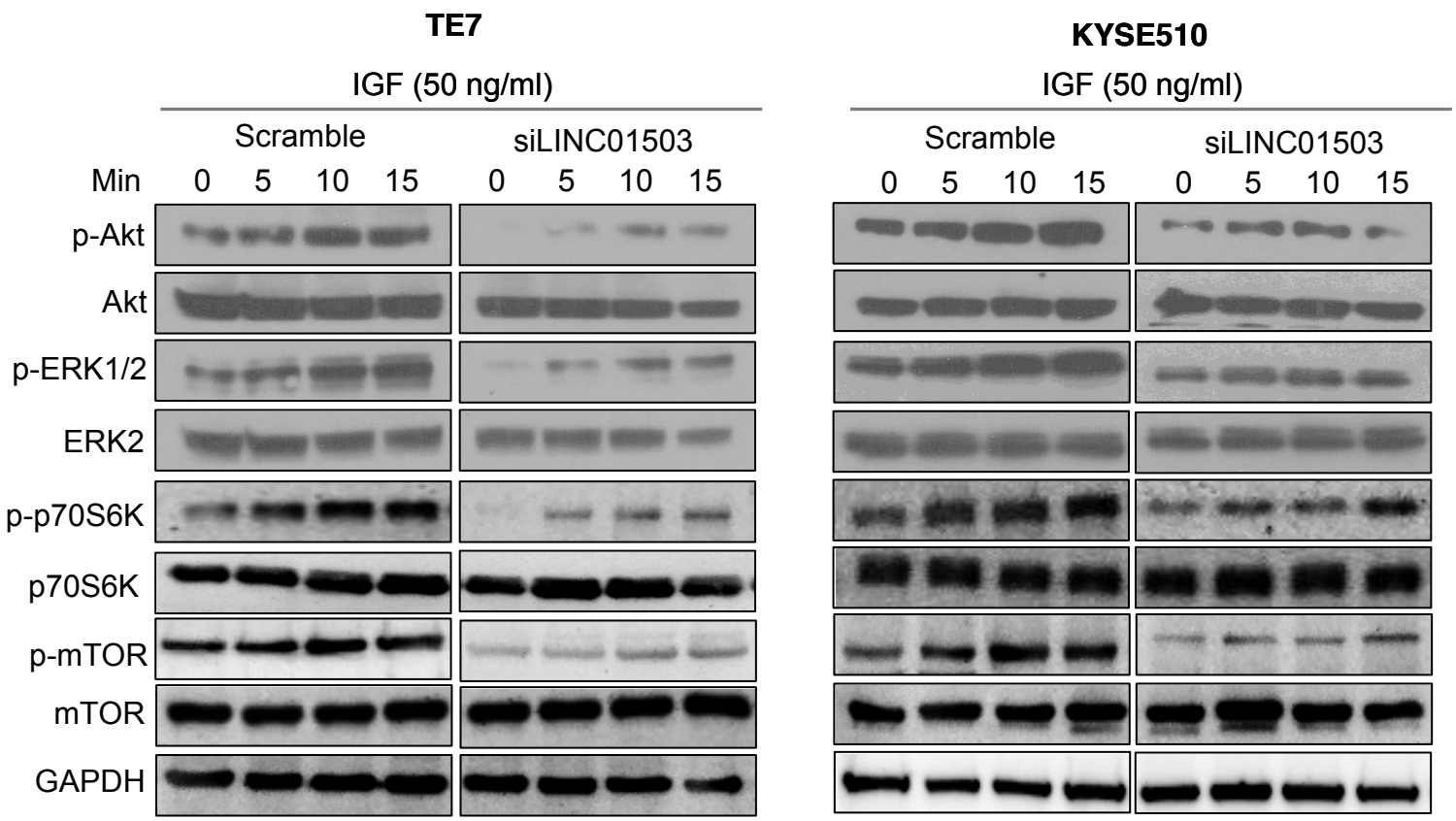
Supplementary Figure 17. Re-expression of LINC01503 rescued decreased cell growth and migration mediated by LINC01503 knockdown. KYSE510 and TE7 cells were infected by shRNA plasmid against LINC01503 (shLINC01503) and control shRNA (Scramble), and stable knockdown cells was selected. This was followed by transfection of over-expression plasmid of LINC01503 into the knockdown cells. A, qPCR analysis for the silencing and re-expression of LINC01503. B and C, MTT assays. D and E, Transwell migration assays.



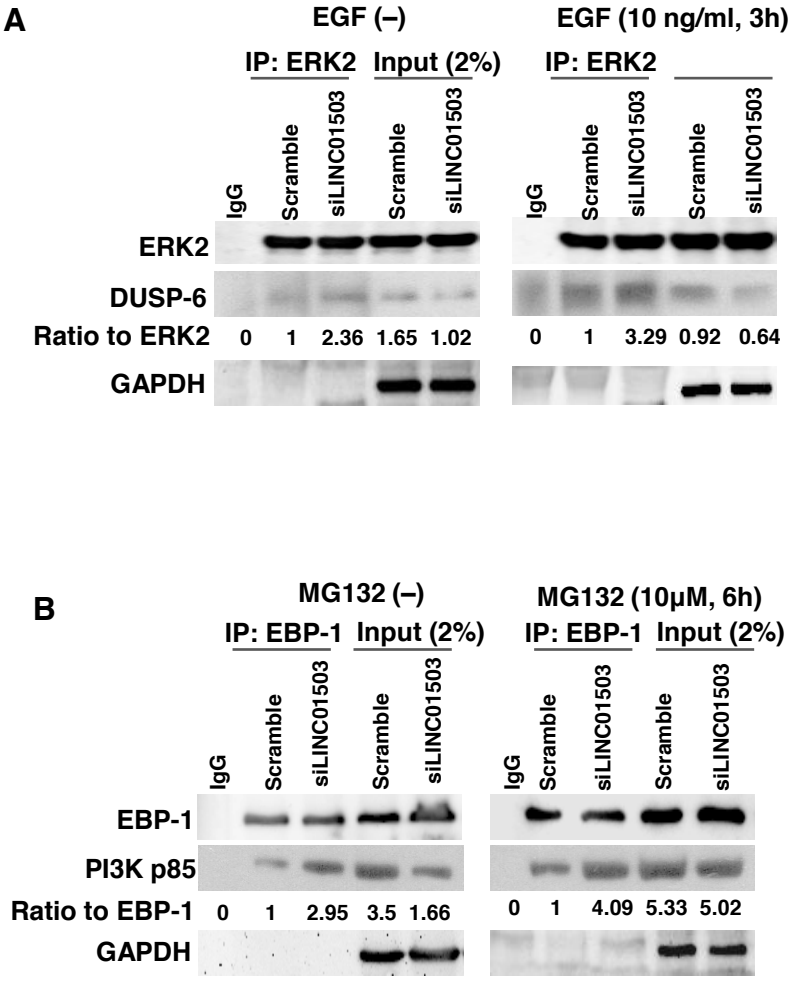
Supplementary Figure 18. Western blotting analysis of 14-3-3 proteins following pull-down of LINC01503 or negative RNA control.



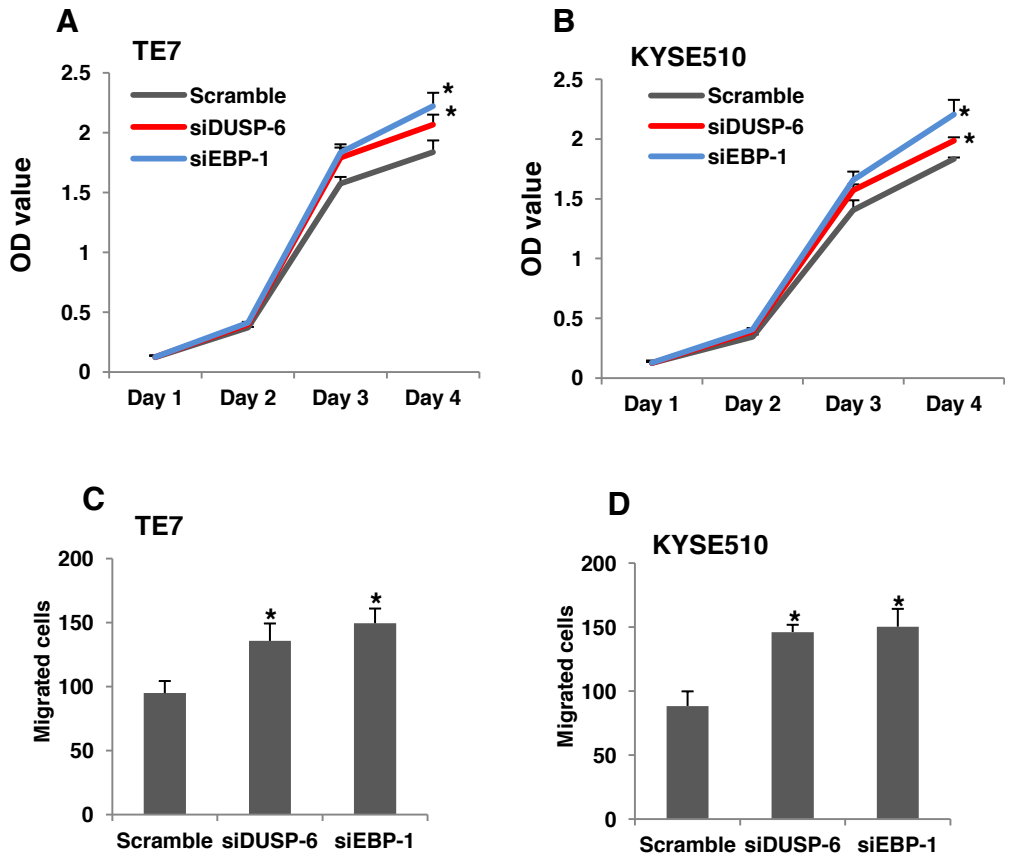
Supplementary Figure 19. TE7 or KYSE510 cells were transfected with either control siRNA (Scramble) or LINC01503 siRNA (siLINC01503). Cell lysates were analyzed by Western blotting for the expressions of the important molecules of ERK/MAPK and PI3K/Akt pathways. Signal intensity for the levels of phosphoproteins was quantified by densitometric scanning and normalized by the total amounts of each protein.



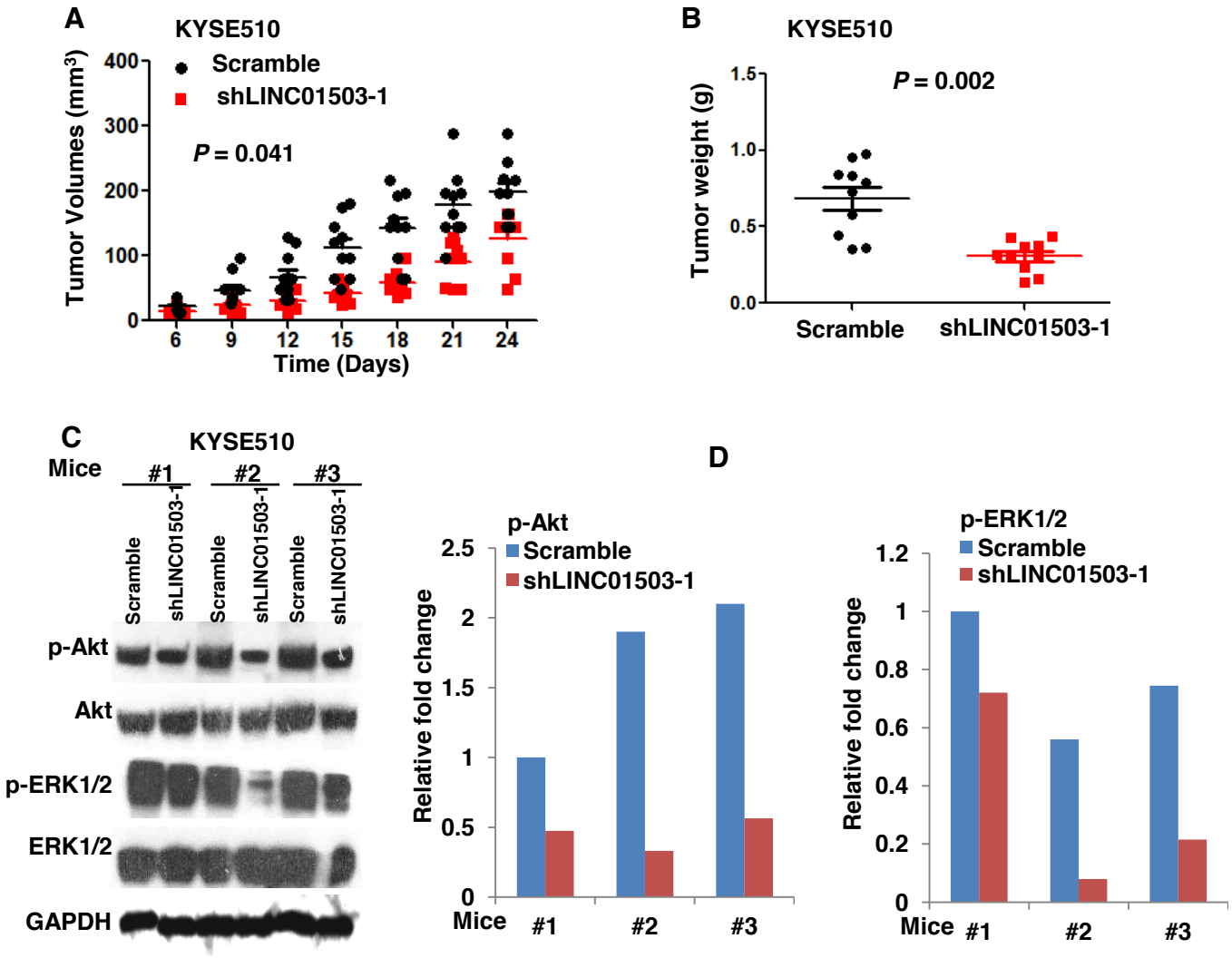
Supplementary Figure 20. LINC01503 activates ERK/MAPK and PI3K/Akt pathways. TE7 or KYSE510 cells transfected with either control siRNA (Scramble) or LINC01503 siRNA (siLINC01503) were cultured in medium with EGF for different duration. Cell lysates were analyzed by Western blotting (GAPDH was served as loading control).



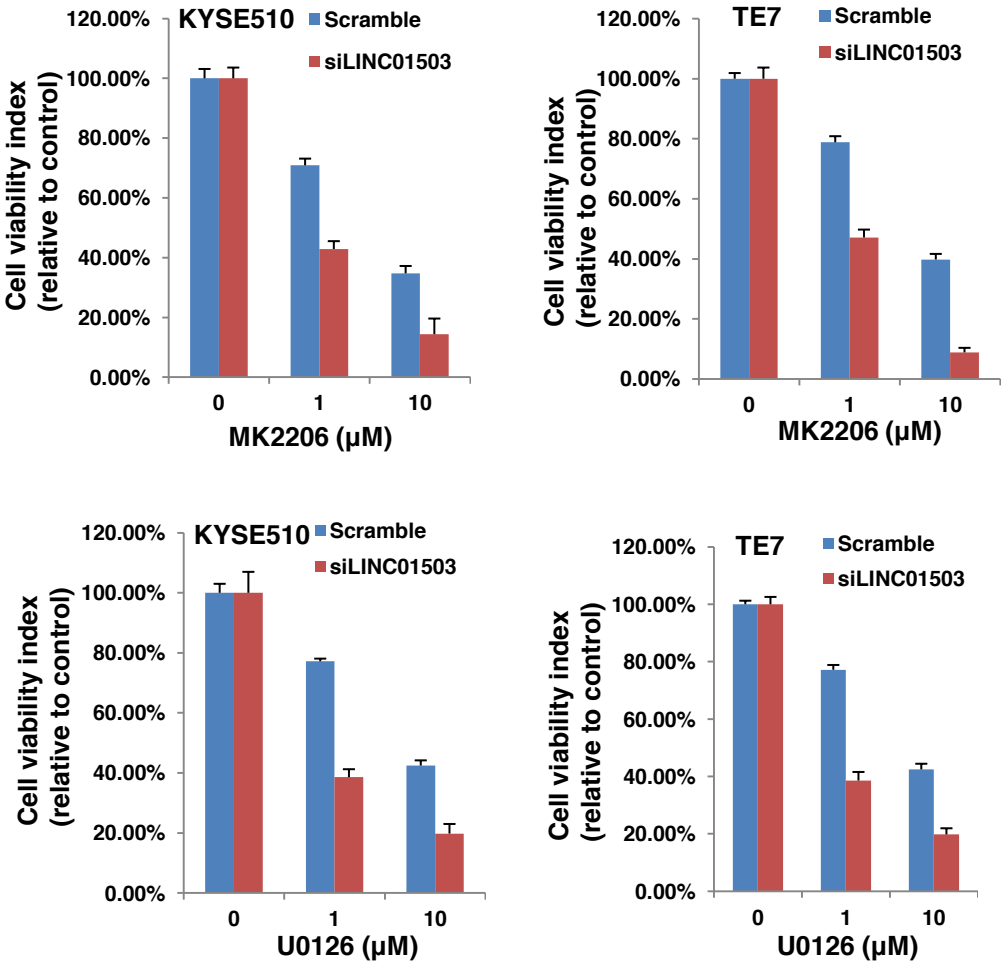
Supplementary Figure 21. Immunoprecipitation assay was performed with cell extracts from TE7 cells transfected with LINC01503 siRNA by using antibodies against ERK2, EBP-1 or normal IgG. Immunoprecipitates (IP) or cell lysates (Input) were analyzed by Western blotting.



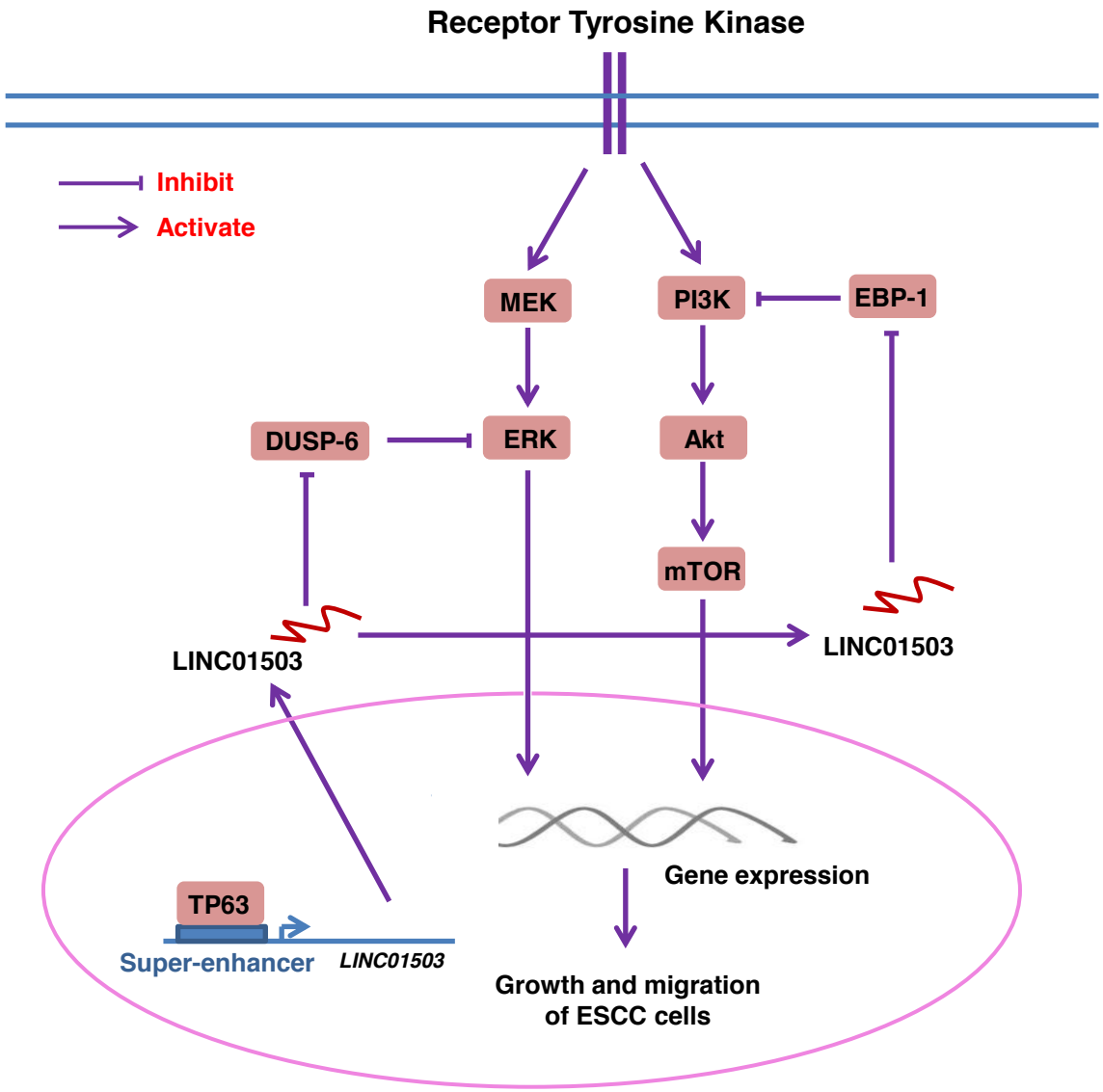
Supplementary Figure 22. Effect of DUSP-6 or EBP-1 on the growth and migration of ESCC cells. (A and B) Growth curves of cells transfected with control siRNA (Scramble), DUSP-6 siRNA (siDUSP-6) or EBP-1 siRNA (siEBP-1). (C and D) Migration assays. Mean ± s.d. are shown, n = 3. *, $P \leq 0.05$.



Supplementary Figure 23. KYSE510 cells with stable depletion of LINC01503 (shLINC01503-1) were injected subcutaneously into nude mice (10 in each group). Tumor volumes (A) and weights (B) were determined. Data are means ± s.d.. (C, D) Activation of PI3K/Akt and ERK/MAPK pathways in paired ESCC xenografts was determined by Western blotting (C). Signal intensity for the levels of p-Akt and p-ERK1/2 was quantified by densitometric scanning and normalized by the total amounts of each protein (D).



Supplementary Figure 24. Colony formation assay was employed to determine the effect of Akt inhibitor (MK2206) or MEK inhibitor (U0126) on the LINC01503 silencing ESCC cells. Indicated cells were treated with the inhibitors for 12 days and OD value was measured.



Supplementary Figure 25. Proposed working model of LINC01503 in ESCC cells.

Supplementary Table 1. Cloning primers, shRNA and siRNA-pool sequences

Cloning primers	
LINC01503	Primer-F: CCGCTCGAGGGGCTCGGAATACCCACC
	Primer-R: CGCGGATCCTGCTGAAAGAAACTCAT
LINC01503-E1	PRIMER-F: CGACGCGTCGCCCCGAAGAGCACTTCACTC
	PRIMER-R: CCGCTCGAGAGGTTTATCGCTCAAGGCCA
LINC01503-E2	PRIMER-F: CGACGCGTCGGCTACGTTTGCTGTCCCAAC
	PRIMER-R: CCGCTCGAGGGTGAGCCTGAGCACAAACTA
LINC01503-E2A	PRIMER-F: CGACGCGTCGGCCCAGCACAAGGCCTGTAA
	PRIMER-R: CCGCTCGAGAAGTGCCACCAGCAGGACAC
LINC01503-E2B	PRIMER-F: CGACGCGTCGCGCTCAGCCCAAGGTTCTGT
	PRIMER-R: CCGCTCGAGCGGGTCTCATGAGCAGACGTG
LINC01503-E3	PRIMER-F: CGACGCGTCGGGTCAAGAGTGAAAGAGAGGGG
	PRIMER-R: CCGCTCGAGAGGGGAGAGAGGGTTTTCAA
shRNA target sequences	
TP63-1	CCGTTTCGTCAGAACACACAT
TP63-3	GGACAGCAGCATTGATCAA
siRNA-pool sequences	
LINC01503	SEQUENCE-1: GGACGAAUGCAGAGCCCUA
	SEQUENCE-2: GGAUGGGUGUUUAAGGCAU
	SEQUENCE-3: GGAGACAAAUGACGGCCUU
	SEQUENCE-4: UUCCAAAGCUCUGUUAUUA
PVT1	SEQUENCE-1: ACCUAUGAGCUUUGAAUAA
	SEQUENCE-2: GAGAACUGUCCUUACGUGA
	SEQUENCE-3: CUUCAACCCAUUACGAUUU
	SEQUENCE-4: GUACGAACUUCAUCGCCCA
FAM83H-AS1	SEQUENCE-1: GGACAGAGUAGGAGCGUAA
	SEQUENCE-2: CCAAGAUGCUAGUCACACA
	SEQUENCE-3: GCUGAAUCACGUCAAGUAU
	SEQUENCE-4: CAAUGUCAGUGUCCGGCUA
MIR205HG	SEQUENCE-1: CAGAGAGAAAUGUGAUCAA
	SEQUENCE-2: GGUGAUGGGCAGAUUGGAA
	SEQUENCE-3: CCGCAGAUGGGAACCUUAA
	SEQUENCE-4: GCUAUGAAAAUUUGCGUUU
EBP-1	SEQUENCE-1: GAACAAAGUUGCCCACUCA
	SEQUENCE-2: UCUAUGAGAAGGAGGGUGA
	SEQUENCE-3: UAAACAGUAUGGACUGAAA
	SEQUENCE-4: CGGGCGAGGACGAGCAACA
DUSP-6	SEQUENCE-1: GAACUGUGGUGUCUUGGUA
	SEQUENCE-2: UGGCUUACCUUAUGCAGAA
	SEQUENCE-3: GACUGUGGCUUACCUUAUG
	SEQUENCE-4: GCGACUGGAACGAGAAUAC

Supplementary Table 2. Probes used in this study

Name	Sequences
ChIRP	
LINC01503XIE_1	caaaggccgtcatttgtctc
LINC01503XIE_2	aagacaaaggcctgcaggaa
LINC01503XIE_3	gtattcagagaggggtagtg
LINC01503XIE_4	cggctggtaaaaatacctgc
LINC01503XIE_5	actcattgcatcgtgttttt
A3300-ChIRP_1	gtcctgaatgttccagaatc
A3300-ChIRP_2	tttgttcaatcctcttcaaa
A3300-ChIRP_3	caggggagccaagaattatt
A3300-ChIRP_4	tcaacgcagatctgttcact
A3300-ChIRP_5	tcttcccaagcacaaggtag
A3300-ChIRP_6	tggaagtatttaatcctccc
A3300-ChIRP_7	atcagggtgctcagattctc
A3300-ChIRP_8	gtgttgccaagaaattgtca
A3300-ChIRP_9	gaagggtggtcaagcaggtc
A3300-ChIRP_10	tccaagtctttccttctcta
RNA FISH	
LINC01503_1	caaaggccgtcatttgtctc
LINC01503_2	ctgacacgtaggtacacact
LINC01503_3	aagacaaaggcctgcaggaa
LINC01503_4	cactgttgctactgtggttg
LINC01503_5	tggcagcagtcacacatcag
LINC01503_6	gtattcagagaggggtagtg
LINC01503_7	ctctacaggatgctccagag
LINC01503_8	cggctggtaaaaatacctgc
LINC01503_9	gagtgcacagtgcaagagta
LINC01503_10	gggggactgatgatgaacac
LINC01503_11	tggaaatgccttaaacaccc
LINC01503_12	actcattgcatcgtgttttt

Supplementary Table 3. Clinicopathological characteristics of ESCC patients

Clinical and pathological indexes	Case No.	3-year OS (%)	5-year OS (%)	*P**	3-year DFS (%)	5-year DFS (%)	*P**
Specimens	113						
Mean age, 57.7							
Age (year)							
≤57	60	55.8	45.6	0.826	43.5	29.3	0.456
>57	53	57.5			44.0	37.8	
Gender							
Male	86	62.2	58.0	0.887	62.5	41.2	0.784
Female	27	53.2			36.1		
Tobacco use							
No	73	56.8	47.1	0.887	42.8	38.6	0.282
Yes	40	56.8			25.0		
Alcohol use							
No	101	66.8	57.8	0.613	66.8	43.0	0.367
Yes	12	48.6			25.0		
Therapies							
Preoperative Radiotherapy	0						
Preoperative Chemotherapy	3	66.7		0.804	66.7		0.364
Postoperative Radiotherapy	44	41.0	43.7	0.545	25.2		0.014
Postoperative Chemotherapy	45	53.5		0.672	36.1	31.0	0.548
Tumor size							
≤3cm	25	70.2		0.001	33.8		0.101
3-5cm	66	60.2	58.2		48.9	40.6	
>5cm	22	18.2			11.4		
Tumor location							
upper	6	53.5		0.396	0		0.679
middle	64	62.4	53.2		47.2	37.9	
lower	43	48.6			34.3	28.6	
Histologic grade							
G1	18	60.6		0.782	55.4	26.6	0.313

	G2	86	63.2	57.2		61.1	42.8	
	G3	9	44.4			22.2		
Invasive depth								
	T1	1	0		0.097	0		0.108
	T2	18	83.3			66.7		
	T3	93	63.9	52.0		65.1	38.0	
	T4	1	100			0		
Lymph node metastasis								
	N0	53	68.9		0.003	54.7	46.2	0.001
	N1	38	50.4			26.0		
	N2	18	27.8			16.7		
	N3	4	0			0		
pTNM-stage								
I	IA	0	25.0		0.011	21.7	0	0.001
	IB	4						
II	IIA	25	71.4			55.7	51.4	
	IIB	29						
III	IIIA	35	45.1	35.7		37.3	20.4	
	IIIB	15						
	IIIC	5						

*Log-rank test of Kaplan Meier method; P <0.05 was considered significant.
All patients underwent surgical treatment.
OS, overall survival; DFS, disease-free survival.

Supplementary Table 4. Univariate and multivariate analysis of factors associated with OS

Variables	Sig.*	HR	95% CI for HR	
			Lower	Upper
Univariate analyses				
Age (>57 vs ≤57)	0.827	0.939	0.535	1.649
Gender (Female vs Male)	0.888	0.954	0.479	1.830
Tobacco use (No vs Yes)	0.888	0.959	0.535	1.727
Alcohol use (No vs Yes)	0.617	1.244	0.529	2.923
Tumor Size	0.002			
Tumor Size (3-5cm vs ≤3cm)	0.003	0.255	0.105	0.623
Tumor Size (>5cm vs ≤3cm)	0.004	0.399	0.213	1.746
Tumor Location	0.407			
Tumor Location (Middle vs Upper)	0.503	0.609	0.143	2.593
Tumor Location (Lower vs Upper)	0.202	0.688	0.387	1.222
Histologic grade	0.407			
Histologic grade (G2 vs G1)	0.786	0.667	0.211	2.105
Histologic grade (G3 vs G1)	0.490	0.808	0.317	2.060
Invasive depth (T3+T4 vs T1+T2)	0.070	2.572	0.925	7.154
Lymph node metastasis (N1+N2+N3 vs N0)	0.009	2.219	1.221	4.034
pTNM-stage (III+IV vs I+II)	0.010	2.150	1.202	3.846
LINC01503	0.027	1.891	1.074	3.327
Multivariate analyses				
Histologic grade	0.002			
Histologic grade (G2 vs G1)	0.003	0.255	0.104	0.623
Histologic grade (G3 vs G1)	0.003	0.384	0.205	0.723
pTNM-stage (III+IV vs I+II)	0.055	1.707	0.989	2.945
LINC01503	0.020	1.967	1.112	3.479

*Multivariate analysis, Cox proportional hazards regression model. Variables were adopted for their prognostic significance by univariate analysis. OS, overall survival; CI, confidence interval; HR, hazard ratio.

Supplementary Table 5. Univariate and multivariate analysis of factors associated with DFS

Variables	Sig.*	HR	95% CI for HR	
			Lower	Upper
Univariate analyses				
Age (>58 vs ≤58)	0.464	0.837	0.520	1.348
Gender (Female vs Male)	0.787	1.080	0.617	1.890
Tobacco use (No vs Yes)	0.292	1.295	0.801	2.096
Alcohol use (No vs Yes)	0.377	1.374	0.679	2.781
Tumor Size	0.116			
Tumor Size (3-5cm vs ≤3cm)	0.068	0.513	0.251	1.051
Tumor Size (>5cm vs ≤3cm)	0.068	0.593	0.338	1.042
Tumor Location	0.692			
Tumor Location (Middle vs Upper)	0.437	1.516	0.531	4.331
Tumor Location (Lower vs Upper)	0.885	0.964	0.590	1.576
Histologic grade	0.337			
Histologic grade (G2 vs G1)	0.312	0.618	0.243	1.571
Histologic grade (G3 vs G1)	0.141	0.552	0.250	1.219
Invasive depth (T3+T4 vs T1+T2)	0.036	2.311	1.057	5.055
Lymph node metastasis (N1+N2+N3 vs N0)	0.001	2.303	1.405	3.776
pTNM-stage (III+IV vs I+II)	0.001	2.258	1.392	3.662
LINC01503	0.035	1.753	0.958	2.518
Multivariate analyses				
pTNM-stage (III+IV vs I+II)	0.001	2.281	1.406	3.701
LINC01503	0.063	1.583	0.975	2.570

*Multivariate analysis, Cox proportional hazards regression model. Variables were adopted for their prognostic significance by univariate analysis. DFS, disease-free survival; CI, confidence interval; HR, hazard ratio.

Supplementary Table 6. Outer primer pairs for selected regions of interest in 4C.

Bait Coordinates	Forward Primer	Reverse Primer
E1 chr9:132,085,090-132,085,770	ACTCAGCCTTGACTCGGATG	AGTAAAACGGGGCTGCCTTG
E2 chr9:132,093,945-132,094,278	CACCCTGACTCCTCATCTTGT	GTTTCACTGGAGCCTCCTGG
E3 chr9:132,097,172-132,097,559	GGTGGGGAGTTAATGGCAGG	TTCCTCGCTTCCGGGAGTAT

Supplementary Table 7. Nested primer pairs for selected baits in 4C.

Bait Coordinates	Forward Primer	Reverse Primer
E1 chr9:132,085,090-132,085,770	AATGATACGGCGACCACCGAG ATCTACACTATCCTCT**TCGTC GGCAGCGTCAGATGTGTATA AGAGACAG**TTCATTGCTGAGT CCCTGGC	CAAGCAGAAGACGGCATACGAG ATGTAGAGAG**GTCTCGTGGGC TCGGAGATGTGTATAAGAGAC AG**GGGAGGTCTCATGGTAGTTG C
E2 chr9:132,093,945-132,094,278	AATGATACGGCGACCACCGAG ATCTACACTATCCTCT**TCGTC GGCAGCGTCAGATGTGTATA AGAGACAG**TCGTGGGCATTTA TCCATCCTA	CAAGCAGAAGACGGCATACGAG ATCCTCTCTG**GTCTCGTGGGCT CGGAGATGTGTATAAGAGACA G**TGGGAGGTCTAGTGGTGACA
E3 chr9:132,097,172-132,097,559	AATGATACGGCGACCACCGAG ATCTACACTATCCTCT**TCGTC GGCAGCGTCAGATGTGTATA AGAGACAG**CTGTTGGGAGAG GGTCCCCT	CAAGCAGAAGACGGCATACGAG ATAGCGTAGC**GTCTCGTGGGCT CGGAGATGTGTATAAGAGACA G**TGAGCTGGTTCTTGAAAACCCT

Red font indicates Nextera® Index Kit – PCR primers; gold font indicates i5 indices (N503, for replicate 1); green font indicates i7 indices (from top to bottom, N707 to N709); blue font indicates Nextera® transposase sequences and black font indicates designed nested primer sequences. N504 index was used for replicate 2.

Supplementary Table 8. qRT-PCR primers used in this study

Genes	Primer sequences
LINC01503	Primer-F: CTCGGAATACCCACCTTTCTG
	Primer-R: ACCAGAGCGACTCCATTTTG
LINC01503 (ChIP-PCR)	Primer-F: AGGCCCTGTGAGAGGTAGGG
	Primer-R: CAGAGAGGAGGCCAGCACAC
TP63	Primer-F: TTCGGACAGTACAAAGAACGG
	Primer-R: GCATTTCATAAGTCTCACGGC
GAPDH	Primer-F: CCATGGGGAAGGTGAAGGTC
	Primer-R: GAAGGGGTCATTGATGGCAAC
U1	Primer-F: TCCCAGGGCGAGGCTTATCCATT
	Primer-R: GAACGCAGTCCCCCACTACCACAAAT
β-actin	Primer-F: ACCTTCTACAATGAGCTGCG
	Primer-R: CCTGGATAGCAACGTACATGG

Supplementary Materials and Methods

4C analysis

Preparation of 4C-template

4C templates were prepared using a protocol that has been modified from a previous paper.[1] In brief, TE7 cells were single-cell suspended, and their chromatin was cross-linked with 1% formaldehyde for 10 min at room temperature. Cells were then lysed and cross-linked DNA was digested with a primary restriction enzyme, HindIII-HF [R3104L, New England Biolabs (NEB)]. Next, HindIII-digested DNA was subjected to proximity ligation using T4 DNA ligase (EL0013, Thermo Scientific), followed by cross-link removal using Proteinase K (AM2546, Ambion), yielding 3C libraries. These 3C libraries were then subjected to a second restriction enzyme digestion using DpnII (R0543L, NEB), followed by another proximity ligation using T4 DNA ligase. For each viewpoint, a total of 3.2 µg of the resulting 4C templates was used to carry out a scale-up inverse, nested PCR, of which 32 reactions (100 ng in each) were pooled and purified using MinElute PCR Purification kit (Qiagen). Ten µg of the PCR products were then run on 4-20% TBE PAGE gel (5 µg per well). DNA smears from 200bp to 1000bp were extracted, and were subjected to Agencourt AMPure XP Bead clean-up (Beckman Coulter) using a bead-to-DNA ratio of 0.8 before high-throughput sequencing on the MiSeq system (Illumina).

4C-seq primer design

The inverse primers were designed based on a viewpoint region. The UCSC Genome Browser (build GRCh37/hg19) was used to locate the region of interest. Upon loading HindIII and DpnII tracks, two HindIII restriction sites flanking the region of interest were identified and the sequence between the nearest HindIII and DpnII restriction sites were selected as the viewpoint region. Based on this region, two pairs of primers

(outer and nested) were designed with the following settings: optimal primer melting temperature of 60 º C (minimum 57 º C and maximum 62 º C); GC content between 40 and 60%. Appropriate adaptors (Nextera® Index Kit - PCR primer, Nextera® transposase sequence) and index sequences were then added to the nested primer pair (Supplementary Table 6 and 7).

4C analysis

Long-range genomic interaction regions generated by 4C-Seq experiment were analyzed by using the R package r3CSeq.[2] Briefly, for each replicate the raw reads were aligned to the masked version of the reference human genome (masked for the gap, repetitive and ambiguous sequences) downloaded from the R Bioconductor repository (BSgenome.Hsapiens.UCSC.hg19. masked). Chromosome 9 was selected as the viewpoint chromosome and HindIII was used as the restriction enzyme to digest the genome. A non-overlapping window size of 2 kb was selected to identify interacted regions. The number of mapped reads for each window were counted and then normalized to obtain RPM (reads per million per window) values to perform statistical analysis. Interacted regions were then plotted and visualized on UCSC Genome Browser as custom tracks. The data has been deposited to Gene Expression Omnibus (GSE106432).

Antibodies and reagents

Antibodies: H3K27ac (Abcam, ab4729), TP63 (Santa Cruz, sc-8431), ERK2 (Santa Cruz, sc-1647), EBP1 (Santa Cruz, sc-292466), 14-3-3σ (Santa Cruz, sc-100638), 14-3-3ζ (Santa Cruz, sc-732), DUSP-6 (Santa Cruz, sc-137245), Ribosomal Protein LP0 (1B4) (Santa Cruz, sc-293260), EGFR (Cell Signaling Technology, #2085), Phospho-EGFR (Cell Signaling Technology, #3777), MEK1/2 (Cell Signaling Technology, #4694), Phospho-MEK1/2 (Cell Signaling Technology, #9154), p44/42

MAPK (Erk1/2) (Cell Signaling Technology, #4695), Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (Cell Signaling Technology, #8544), GAPDH (Cell Signaling Technology, #2118), Akt (Cell Signaling Technology, #4691), Phospho-Akt (Cell Signaling Technology, #4060), p70S6K (Cell Signaling Technology, #9208), Phospho-p70S6K (Cell Signaling Technology, #2708), mTOR (Cell Signaling Technology, #2972), Phospho-mTOR (Cell Signaling Technology, #5536), PI3K p85(Cell Signaling Technology, #4257), Anti-rabbit IgG (Cell Signaling Technology, #7074), Anti-mouse IgG (Cell Signaling Technology, #7076). Reagents: human EGF and IGF-1 (PROSPEC), MG132 (Sigma-Aldrich), U0126 (Sigma-Aldrich), MK2206 (Selleckchem).

PCR

Total RNA was extracted using RNeasy isolation Kit (Qiagen). For quantitative real-time PCR (qRT-PCR), cDNA was generated using qScript cDNA Synthesis Kit (Quanta Biosciences). qRT-PCR was performed on CFX96 qPCR System (Biorad Inc). Expression of each gene was normalized to β-actin as the reference. Primers for qRT-PCR are listed in Supplementary Table 8.

References

1. Splinter E, de Wit E, van de Werken HJG, et al. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. Methods 2012;58(3):221-230.
2. Thongjuea S, Stadhouders R, Grosveld FG, et al. r3Cseq: an R/Bioconductor package for the discovery of long-range genomic interactions from chromosome conformation capture and next-generation sequencing data. Nucleic Acids Res 2013;41(13):e132.