Original Article

Deregulated expression and subcellular localization of CPSF6, a circRNA-binding protein, promote malignant development of esophageal squamous cell carcinoma

Shichao Guo[1], Guangchao Wang[1], Zitong Zhao[1], Dan Li[1], Yongmei Song[1], Qimin Zhan[1,2]

[1]State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China; [2]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Laboratory of Molecular Oncology, Peking University Cancer Hospital & Institute, Beijing 100142, China

Correspondence to: Yongmei Song. State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. Email: symlh2006@163.com; Qimin Zhan. State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China; Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Laboratory of Molecular Oncology, Peking University Cancer Hospital & Institute, Beijing 100142, China. Email: zhanqimin@bjmu.edu.cn.

Abstract

Objective: Cleavage and polyadenylation specific factor 6 (CPSF6) has been documented as an oncoprotein in different types of cancer. However, functions of CPSF6 have not been investigated yet in esophageal squamous cell carcinoma (ESCC). Here, we aimed to investigate the potential clinical values and biological functions of CPSF6 in ESCC.

Methods: For determining the expression level of CPSF6 in ESCC patients, we analyzed published data, performed quantitative real-time polymerase chain reaction (RT-qPCR) and immunohistochemistry assays. Kaplan-Meier curves and log-rank tests were used for survival analyses. GO and KEGG analyses were done for CPSF6-related genes. Cell proliferation, colony formation and xenograft assays were conducted to verify the effects of CPSF6 on ESCC. In addition, cell cycle and apoptosis assays were also performed to manifest the functions of CPSF6 and circCPSF6. RNA pulldown and radioimmunoprecipitation (RIP) assays were used for confirming the interaction between circCPSF6 (hsa_circ_0000417) and CPSF6 protein. The regulatory relationship between CPSF6 protein and circCPSF6 was determined by RT-qPCR.

Results: We found that CPSF6 was upregulated in ESCC tissues and overexpression of cytoplasmic CPSF6 was associated with poor prognosis. GO and KEGG analyses suggested that CPSF6 could mainly affect cell division in ESCC. Further experiments manifested that CPSF6 promoted cell proliferation and colony formation *in vitro*. Xenograft assay showed that knockdown of CPSF6 significantly decreased tumor growth rate *in vivo*. Subsequently, we verified that depletion of CPSF6 led to cell cycle arrest and apoptosis. Finally, we validated that CPSF6, as a circRNA-binding protein, interacted with and regulated its circular isoform circCPSF6 (hsa_circ_0000417), of which depletion also resulted in cell cycle arrest and cell apoptosis in ESCC.

Conclusions: These findings gave us insight that overexpression of cytoplasmic CPSF6 protein is associated with poor prognosis in ESCC and CPSF6 may function as an oncoprotein, at least in part, through regulating circCPSF6 expression.

Keywords: CPSF6; esophageal squamous cell carcinoma; prognosis; cell proliferation; circRNA-binding protein; circCPSF6

Submitted Dec 03, 2021. Accepted for publication Feb 16, 2022.

doi: 10.21147/j.issn.1000-9604.2022.01.02
View this article at: https://doi.org/10.21147/j.issn.1000-9604.2022.01.02

Introduction

The incidence and mortality rate of esophageal cancer are the seventh and sixth worldwide, respectively (1). Esophageal cancer is a highly malignant tumor with a median survival of less than one year and a five-year survival rate of less than 30% for patients with intermediate to advanced disease (2-4). According to the pathological classification, esophageal cancer is divided into esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC). China has a high incidence of esophageal cancer. More than 50% of new patients worldwide are found in China, and more than 90% of them are ESCC (5). Despite the availability of more treatment strategies, the prognosis of ESCC patients is not promising due to its complexity. Biologically, ESCC shares some common pathogenic characteristics and mechanisms with head and neck squamous cell carcinoma (5,6). In addition, increasingly novel biological markers and potential therapeutic targets for ESCC were unveiled in recent years (7-9). However, the underlying mechanisms of tumorigenesis and progression of ESCC still need to be further unclosed.

Cleavage factor Im (CFIm) complex functions as a key regulator of alternative polyadenylation (APA), an RNA post-transcriptional mechanism for generating variable 3' untranslated regions (3' UTRs), raising transcriptome complexity of cells (10,11). Cleavage and polyadenylation specific factor 6 (CPSF6), also known as CFIm68, is one of CFIm complex subunits required for RNA cleavage and polyadenylation processing. Several independent research groups have found that CPSF6 is related to tumorigenesis and progression. For instance, CPSF6 is highly expressed in aggressive breast cancer and is a clinically relevant marker of poor patient outcomes (12). In hepatocellular carcinoma, overexpression of CPSF6 increases cell proliferation, migration and invasion both *in vitro* and *in vivo* (13). Whereas, there are no studies concerning the biological roles of CPSF6 in ESCC.

Here, we reported that CPSF6 was up-regulated in ESCC and the augmented expression level of cytoplasmic CPSF6 protein was associated with poor prognostic outcomes in ESCC. Functionally, we discovered that CPSF6 enhanced the malignant phenotypes of ESCC cells *in vitro* and *in vivo*. Moreover, CPSF6 protein could interact with circCPSF6 and impact its expression. We also manifested that depletion of circCPSF6 raised cell cycle arrest at G_0/G_1 phase and promoted cell apoptosis in ESCC, which as well arose in CPSF6-delpeted cells. Generally speaking, we conclude that CPSF6 functions as an oncoprotein and may be a potential prognostic indicator and therapeutic target for ESCC.

Materials and methods

Data collection and processing

We downloaded the GSE53624 dataset, which included CPSF6 mRNA expression data of paired tumor and normal tissues and relevant clinical information in 119 ESCC patients, from Gene Expression Omnibus (GEO) database (14). The probes were subsequently re-annotated using methods reported by our previous study (15). Additionally, we obtained the gene expression fragments per kilobase million (FPKM) values of The Cancer Genome Atlas (TCGA) ESCC cohort from TCGA Genomic Data Commons (GDC) portal, and then log2 transformed values were used to analyze CPSF6 expression. We divided the patients into "CPSF6 high" group and "CPSF6 low" group according to the Jorden index. To investigate the potential biological function of CPSF6, we identified CPSF6-related genes in GSE53624 dataset with adjusted $P<0.05$ and $|\text{Pearson correlation coefficient}| \geq 0.5$. The CPSF6-binding RNAs were grabbed from the Encyclopedia of RNA Interactomes (ENCORI) database (16) and organized artificially. R Bioconductor package "clusterProfiler" was used for GO and KEGG enrichment analyses (17).

ESCC tissue microarray and immunohistochemistry (IHC)

One microarray with ESCC tissues (n=55) and adjacent normal tissues (n=49) (HEsoS105Su01) was obtained from OUTDO Biotech (Shanghai, China), and the IHC detection of CPSF6 was performed by OUTDO Biotech according to standard procedures. The antibody of CPSF6 (#A5963, ABclonal, Wuhan, China) was diluted in 1:1,000 during this experiment. We defined the expression level (index) of CPSF6 by multiplying the intensity of the

staining and the frequency of positive cells. The CPSF6 expression indexes were used for paired-sample *t*-test analysis. In addition, the intensity of the staining was scored as follows: no staining, 0; light staining, 1; intermediate staining, 2; and dark staining, 3. The frequency of CPSF6-positive cells was scored as follows: less than 5%, 0; 5%−25%, 1; 26%−50%, 2; 51%−75%, 3; and 76%−100%, 4. So the CPSF6 staining strength was valued from 0 to 12 by multiplying intensity scores and frequency scores: 0−4 were considered as weak grades, 5−8 as moderate grades, and 9−12 as strong grades. Wilcoxon rank-sum test was used to analyze the frequency distribution of different IHC grades in ESCC tissues and adjacent normal tissues. We divided the patients into "CPSF6 high" group and "CPSF6 low" group according to the median value of CPSF6 expression. The details of clinical information of ESCC patients and IHC data are gathered in *Supplementary Table S1*.

Cell lines and cell culture

NE2 and NE3, the two immortalized human esophageal epithelial cell lines, were cultured in medium with 1:1 mixture of Defined Keratinocyte SFM and EpiLife added EpiLife Defined Growth Supplement (Gibco, California, USA). Human ESCC cell lines (YES2, KYSE30, KYSE70, KYSE140, KYSE150, KYSE180, KYSE410, KYSE450, KYSE510 and COLO680N) were cultured in RPMI 1640 (Gibco, California, USA) supplemented with 10% fetal bovine serum (FBS). All above cell lines were incubated at 37 °C in humidified air with 5% CO_2. Authentication was down for the identities of all cell lines and mycoplasma contamination was excluded.

Vector construction and siRNA synthesis

CPSF6 fragments (1,731 bp) were amplified from cDNA of ESCC cells by Q5 Hot Start High-Fidelity 2× Master Mix (NEB, Massachusetts, USA). The sequences of primers were as follows: forward primer: TGAACCGTCAGAT CCGCTAGCGCTAGATCCGCTGCTGCTG; reverse primer: GTCATCCTTATAATCACGATGACGATATT CGCGCTC. The pEX3-flag vector, reformed from the pEX3 vector (GenePharma, Shanghai, China), was cut by *Nhe* I enzyme (NEB, Massachusetts, USA) at 37 °C for 2 h. Then, the PCR products and linear vectors were both separated by electrophoresis within a 1% GelRed-stained agarose gel, and purified using gel extraction kits (TIANGEN, Beijing, China). After ligating with *pEASY*-Basic Seamless Cloning and Assembly Kit (TransGen Biotech, Beijing, China), plasmids were transformed into competent cells (*Escherichia coli*, *DH5α*) by heat shock. Positive colonies were identified by performing Sanger sequencing serviced by SinoGenoMax (Beijing, China). Correct plasmids were used for subsequent experiments or long-term storage at −80 °C.

siRNAs were designed and synthesized by RiboBio (Guangzhou, China). si-NC: GGCTCTAGAAAAGC CTATGC; si-CPSF6: GATGTCGGCGAAGAGTTCA; si-circCPSF6: GTTAGAAGGAAGCTGAATA.

Transfection

Before transfection, appropriate amounts of cells were seeded in 60-mm dishes and incubated for 24 h. siRNAs (a final concentration of about 100 nmol/L) and/or plasmids (4 μg) coated by lipofectamine 2000 reagent (Invitrogen, New York, USA) were transfected into cells following the protocol of manuals. The efficiency of transfection was determined by quantitative real-time polymerase chain reaction (RT-qPCR) and Western blot assays after 48 h of transfection.

Lentiviral infection

High titer lentiviruses were prepared by Genechem (Shanghai, China). Lentiviral infection procedure was referenced to manufacturer's manuals. One day before lentiviral infection, 1.0×10^5 cells were seeded into a well of a six-well plate so that about 20% confluence of cells was obtained after 24 h. On the second day, the spent medium was replaced by complete fresh medium with an appropriate volume of lentiviral particles to achieve a desired multiplicity of infection (MOI). The virus-contained medium was removed after infection for 16 h and new complete medium was added. The efficiency of CPSF6 depletion was observed and quantified with RT-qPCR and Western blot assays after 72 h of infection.

Extraction of nuclear and cytoplasmic fractions

The nuclear and cytoplasmic fractions of cells were extracted using the PARIS™ kit (AM1921, Invitrogen, New York, USA). Firstly, we collected 1.0×10^6 fresh cultured cells after washing with cold phosphate buffer saline (PBS). Secondly, cells were resuspended in 300 μL cell fractionation buffer by pipetting gently and centrifuged 5 min at 4 °C and 500× g after incubating on ice for 5−10

min. Supernatant with cytoplasmic fraction was collected carefully into a new nuclease-free microtube on ice for RNA and protein isolation later on. Thirdly, the same volume of cell fractionation buffer was added to wash the nuclear pellet to avoid contamination of the nuclear fraction with portions of the cytoplasmic fraction. Finally, the nuclei were re-pelleted and lysed in 300 μL cell disruption buffer by vortexing until the lysate was homogenous. The nuclear and cytoplasmic lysates were used to isolate RNA and protein.

RNA extraction, RNase R treatment and RT-qPCR analysis

Total RNAs were extracted according to the manual of the RNAExpress Total RNA Kit (NCM biotech, Suzhou, China), followed by first-strand cDNA synthesis using the M-MLV reverse transcription kit (Promega, Wisconsin, USA). For circRNA enrichment, 1 μg total RNAs were treated with 1 U RNase R (Geneseed, Guangzhou, China) at 37 °C for 15 min before reverse transcription. cDNA was amplified with *PerfectStart* Green qPCR SuperMix (TransGen Biotech, Beijing, China) and quantified by the CFX96 Real-Time System (Bio-Rad, California, USA). The relative expression of genes was normalized to ACTB (β-actin), and the fold changes of the target gene expression in relative to the control were calculated by the formula $2^{-\Delta\Delta Ct}$. Primers were listed as follow: β-actin: CTCCA TCCTGGCCTCGCTGT, GCTGTCACCTTCACCGT TCC; CPSF6: GGCGAAGAGTTCAACCAGGA, ATC TCGGTCTTCTGGGGCAT; circCPSF6: GAGTCGA CGTCATAAATCCCGTA, GTCTTCTGGGGCATCT CCATTAT.

Protein extraction and Western blot analysis

Cells were harvested and lysed for 10–20 min in cold Pierce IP Lysis Buffer (Thermo Scientific, New York, USA) which was supplemented with protease and phosphatase inhibitor cocktail (NCM biotech, Suzhou, China), and then the supernatant was collected into a new clear 1.5 mL centrifuge tube after centrifuging for 15 min (12,000× g, 4 °C). The protein concentration was quantified by the BCA protein assay kit (Applygen, Beijing, China). Denatured total proteins (30 μg) per sample were loaded into 10% SDS-PAGE gels and separated, followed by transferring to polyvinylidene difluoride (PVDF) membranes. After blocking in Tris buffered saline with Tween-20 (TBST) containing 5% skim milk for 1 h at room temperature, membranes were incubated with primary antibodies at 4 °C overnight. Then, the membranes were washed with TBST three times for 10 min each and incubated with conjugated secondary antibodies for 1 h at room temperature, followed by washing with TBST three times for 10 min each. The targets were developed with Enhanced Chemiluminescent HRP substrate (NCM biotech, Suzhou, China) and imaged on ImageQuant LAS 4000 (GE, Massachusetts, USA). The targeting protein levels were normalized to β-actin using ImageJ for gray-scale statistics.

Cell proliferation assay

Firstly, cells were digested by 0.25% trypsin into individual cells and cleaned by PBS twice. After that, the cells were resuspended with fresh medium and counted with the TC10 Automated Cell Counter (Bio-Rad, California, USA). Then, cells were diluted at a density of 3,000 cells per 100 μL for seeding. Each 100 μL suspension was plated into a well of 96 E-Plate after background-testing. Every sample made 6 parallels. The cell index was tested every 15 min lasting for at least 96 h by the xCELLigence Real-Time Cell Analyzer MP System (ACEA, Agilent Technologies, California, USA).

Colony formation assay

A total of 1,000 single cells were plated into a well of a six-well plate and incubated at 37 °C in humidified air with 5% CO_2. Colonies would format for about 2 weeks after seeding. During fostering, the medium would be replaced every 3–5 d for up to colony formation. Each group made three parallels. Colonies were stained with 2% crystal violet and then imaged. Colony number was counted using particle analysis of ImageJ.

In vivo xenograft assay

Twelve five-week-old female BALB/c nude mice purchasing from Vitalriver (Beijing, China) were housed under specific-pathogen-free (SPF) conditions for a one-week acclimation period and randomly divided into 2 groups with 6 mice in each. Each nude mouse was subcutaneously injected in the left dorsal flank with a total of 3×10^6 CPSF6-depleted KYSE450 cells or negative control KYSE450 cells, respectively. After 30 d, nude mice were sacrificed and the subcutaneous tumors were removed and weighted. The differences in tumor weight between "CPSF6-depleted KYSE450 cell injection" group and

"negative control KYSE450 cell injection" group were statistically analyzed. All animal cares and procedures were complied with the national policies for animal health and well-being, and the protocol was approved by Animal Care and Use Committee, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College.

Cell cycle assay

Firstly, the CPSF6-knockdown cells and negative control cells were respectively starved using serum-free medium for 24 h in order to synchronize most of the cells at G_0/G_1 stage. After that, cells were digested and re-cultured in complete medium for another 24 h. Then, adherent cells were rinsed twice with cold PBS. Eventually, cells were digested into single cells and fixed with 70% ethanol for 2 h at −20 °C prior to propidium iodide (PI) staining and flow cytometric analysis. Each group was conducted in triplets.

Cell apoptosis assay

The CPSF6-knockdown cells and negative control cells were respectively seeded into 60-mm dishes and cultured for 48 h so that cell confluence reached about 90%. Then, the cells were digested into individual cells, washed twice with pre-chilled PBS and resuspended with 1× binding buffer to achieve a density of 1.0×10^5 cells/mL. Annexin V-FITC (5 μL) were added into 195 μL of the resuspension, followed by adding 10 μL of PI and mixing thoroughly. Protecting from light, the resuspension was incubated for 20 min at room temperature, subsequently placed on ice pending for flow cytometric analysis. Each group was conducted in triplets.

Caspase 3 (#9662S, CST, Massachusetts, USA) and cleaved caspase 3 (#9661S, CST, Massachusetts, USA) were both detected by Western blot assay for indicating cell apoptosis. The level of β-actin (#A5316, Sigma, Missouri, USA) was regarded as an internal control.

RNA-protein pulldown assay

A total of 2.0×10^7 cells were harvested and lysed in 500 μL cold Pierce IP lysis buffer with protease and phosphatase inhibitor cocktail (NCM biotech, Suzhou, China) and RNase inhibitor (Promega, Wisconsin, USA). The cell suspension was incubated on ice for 10–20 min, followed by centrifuging for 15 min (12,000× g, 4 °C), and then the supernatant was collected into a new clear centrifuge tube (1.5 mL) for later use. The cell lysates were also compatible for radioimmunoprecipitation (RIP) assay.

Pre-cleared streptavidin agarose beads (25 μL) (GE, Massachusetts, USA) were mixed with 5 μg biotinylated single-strand oligonucleotides (Sense: CGCGAATAT CGTCATCGTTAGAAGGAAGCTGAATATGGTGG GCATGAT-biotin; Antisense: ATCATGCCCACCATAT TCAGCTTC CTTCTAACGATGACGATATTCGCG-biotin) in 500 μL RNA capture buffer (20 mmol/L pH 7.4 Tris-HCl, 1 mol/L NaCl, 1 mmol/L EDTA), and the mixture was incubated for 30 min at room temperature with agitation. After that, the beads were washed twice with 500 μL protein-RNA binding buffer (20 mmol/L pH 7.4 Tris-HCl, 50 mmol/L NaCl, 2 mmol/L $MgCl_2$, 5% glycerol, 0.1% Tween-20) and then collected for use. Next, cell lysates with 1 mg total protein were diluted in protein-RNA binding buffer to a final volume of 1 mL. The 1 mL diluted lysates were mixed with previously prepared probe-bound beads and incubated for 1–2 h at room temperature with rotation. Finally, the beads were washed three times with wash buffer (20 mmol/L pH 7.4 Tris-HCl, 150 mmol/L NaCl, 0.1% Tween-20), and subsequently boiled in 50 μL 2× sodium dodecyl sulfate (SDS) loading buffer. Western blot assay was used for further analysis. GAPDH (#10494-1-Ig, Proteintech, Wuhan, China) was considered as a negative control.

RNA-binding protein RIP assay

For the native RIP assay, cell lysates (1 mg total protein, containing protease and phosphatase inhibitor cocktail and RNase inhibitor) were incubated with 2 μg of targeting antibodies for 4 h at 4 °C on a rotator, and then 25 μL pre-cleared Protein A/G PLUS agarose beads (Santa Cruz, California, USA) were added into the mixture and incubated for 30 min at room temperature with rotation. Next, the slurry was washed 3 times with wash buffer (20 mmol/L pH 7.4 Tris-HCl, 150 mmol/L NaCl, 0.1% Tween-20). The captured RNAs were extracted using TRIzol followed by RT-qPCR analysis.

Statistical analysis and graphics production

IBM SPSS Statistics (Version 26.0; IBM Corp., New York, USA) and GraphPad Prism 8 (GraphPad Software, Inc., San Diego, CA, USA) were used for statistical analysis and graphics production. Survival analyses were performed according to the Kaplan-Meier method, and the statistical

differences were gathered by log-rank test. Wilcoxon rank-sum test and paired-sample *t*-test were both used for comparing the IHC staining strength of CPSF6 protein between ESCC tissues and adjacent normal tissues. The rest of statistical analyses were all gathered by Student's *t*-test. P<0.05 (two-tailed) was considered significantly difference.

Results

CPSF6 is high expressed in ESCC tissues and ESCC cell lines

Previous studies have demonstrated that CPSF6 plays a pro-oncogenic role in several human cancers (13,18,19). Herein, we investigated the expression and functions of CPSF6 in ESCC. Firstly, through the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) (20), we found that CPSF6 was universally up-regulated in tumor (statistically significant differences marked in red). Although GEPIA2 didn't distinguish subtypes of esophageal cancer, a trend of high expression of CPSF6 in esophageal cancer was observed (*Figure 1A*). In order to know the expression of CPSF6 in ESCC, we collected ESCC tissues (n=29) and adjacent non-tumor tissues (n=30) for RT-qPCR assay. The results showed the mRNA level of CPSF6 in tumors was on average about 1.8-fold higher than that in non-tumor tissues (P<0.001) (*Figure 1B*). In addition, both mRNA and protein level of CPSF6 were prevalently up-regulated in ESCC cell lines (YES2, KYSE30, KYSE70, KYSE140, KYSE150, KYSE180, KYSE410, KYSE450, KYSE510 and COLO680N), compared with two immortalized esophageal epithelial cell lines (NE2 and NE3) (*Figure 1C,D*). These results suggest that CPSF6 may positively contribute to tumorigenesis and progression in ESCC.

High level of CPSF6 mRNA is associated with poor prognosis in ESCC

Next, we retrieved the mRNA expression level of CPSF6 and clinical characteristic information of ESCC patients from public databases. Consistent with above results, the mRNA expression level of CPSF6 in ESCC was significantly higher in tumor tissues than in normal tissues (TCGA: P<0.001; GSE53624: P<0.001) (*Figure 2A*,*B*). In addition, overall survival analysis of all ESCC cases showed that high expression of CPSF6 was associated with a worse patient survival (*Figure 2C*). Furthermore, we conducted survival analyses of stage I+II cases, stage III cases and lymph node metastasis cases of ESCC, individually. Not only in the early-stage patients (stage I+II), but also in the late-stage patients (stage III and lymph node metastasis), all analyses showed that the "CPSF6 low" groups had better median overall survival time than the "CPSF6 high" groups (*Figure 2D*,*E*,*F*). Together, these results indicate that down-regulation of CPSF6 mRNA may have a survival benefit for ESCC patients and CPSF6 may play pro-cancer roles in both early and advanced stages of ESCC.

High level of cytoplasmic CPSF6 protein is detrimental to ESCC patient prognosis

Besides, we examined the protein level of CPSF6 in clinical ESCC tissues (n=55) and adjacent normal tissues (n=49) by immunohistochemistry (IHC) assay (*Figure 3A*). As expected, CPSF6 protein was significantly up-regulated in ESCC tissues comparing with adjacent normal tissues through Wilcoxon rank-sum test (P=0.005) and paired-sample *t*-test (49 pairs, P<0.001), respectively (*Figure 3B*,*C*). Unexpectedly, we found that the expression level of total CPSF6 protein was not associated with the median overall survival time of ESCC patients (P=0.934) (*Figure 3D*). By looking closely at the IHC images, we observed that the staining strength of cytoplasm-located CPSF6 protein was likely higher in ESCC cases than in normal cases. Moreover, we also confirmed that cytoplasm-located CPSF6 was expressed indeed higher in ESCC cell lines than in immortalized esophageal epithelial cell lines, but there were no such differences in the nucleus generally (*Figure 3E*). Thus, we rejudged the expression level of CPSF6 protein in the cytoplasm and nucleus, respectively, according to the IHC staining indexes. Wilcoxon rank-sum tests showed that, in the cytoplasm, there was statistically significant difference between ESCC tissues and adjacent normal tissues (P=0.001) (*Figure 3F*), while no statistically significant difference in the nucleus (P=0.051) (*Figure 3G*). Similarly, through paired-sample *t*-test, a significantly statistic difference of CPSF6 protein expression between ESCC tissues and adjacent normal tissues was only discovered in the cytoplasm (P<0.001), not in the nucleus (P=0.760) (*Figure 3H*,*I*). Given that, extending survival analyses were done based on CPSF6 protein level in the cytoplasm and nucleus, individually. Notably, we observed that the patients with augmented cytoplasmic CPSF6 protein had shorter median overall survival time (15 months) than those with low cytoplasmic CPSF6 protein

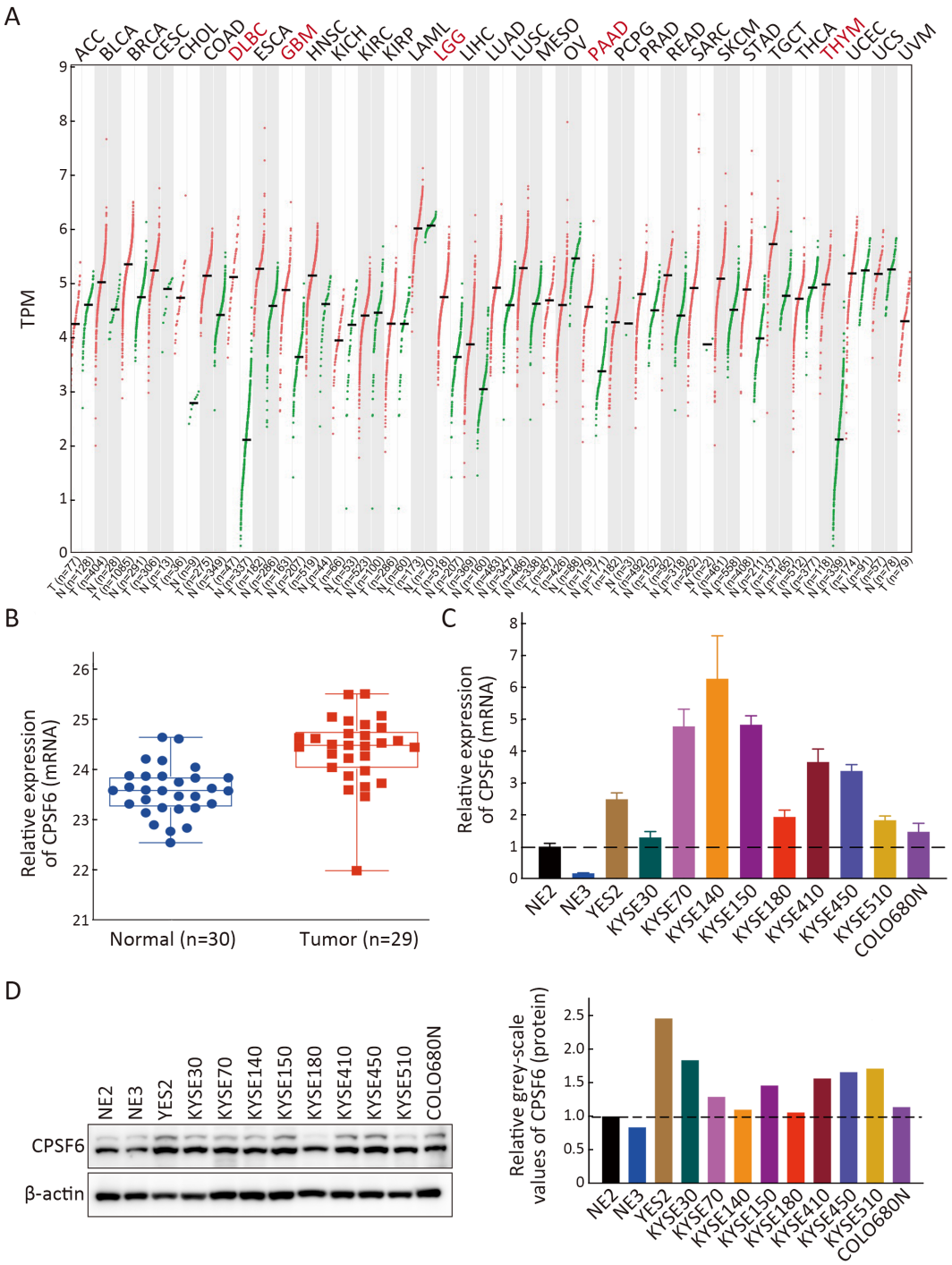

Figure 1 CPSF6 is highly expressed in ESCC. (A) CPSF6 mRNA expression profile across all types of tumor and normal tissues in GEPIA 2. Each dot represents CPSF6 expression (TPM) of one sample. Statistically significant differences between tumor and normal samples are marked in red; (B) RT-qPCR analysis for relative CPSF6 mRNA expression of ESCC tumor tissues (n=29) and normal tissues (n=30) (P<0.001); (C) RT-qPCR analysis for relative CPSF6 mRNA expression in 2 immortalized esophageal epithelial cell lines and 10 ESCC cell lines; (D) Western blot assay for CPSF6 protein expression in 2 immortalized esophageal epithelial cell lines and 10 ESCC cell lines. Bar plot shows the relative CPSF6 protein normalized to β-actin. ESCC, esophageal squamous cell carcinoma; TPM, transcripts per million; RT-qPCR, quantitative real-time polymerase chain reaction.

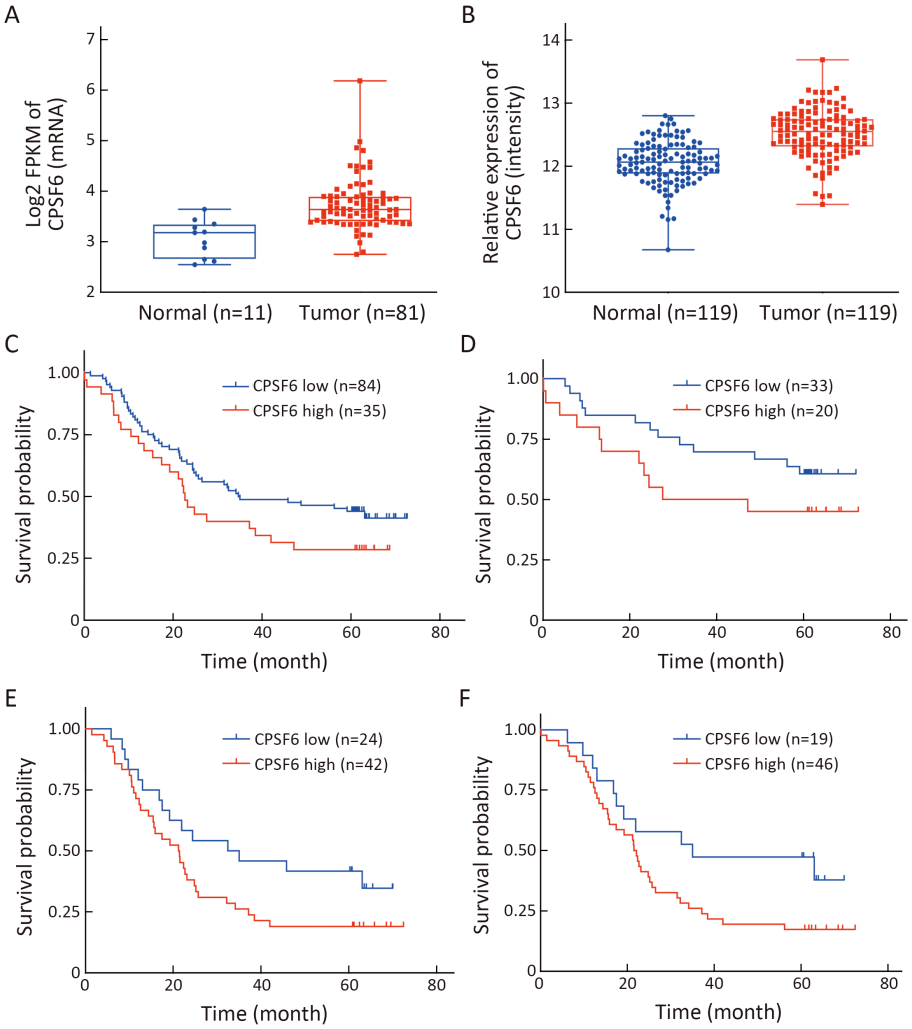

Figure 2 High mRNA level of CPSF6 is associated with poor prognosis in ESCC. (A) Relative CPSF6 mRNA expression (Log2 FPKM) of ESCC tumor samples (n=11) and normal samples (n=81) from the TCGA (P<0.001); (B) Relative CPSF6 mRNA expression (Log2 FPKM) of ESCC tumor samples (n=119) and normal samples (n=119) from the GSE53624 dataset (P<0.001); (C-F) Kaplan-Meier survival analyses for the correlation between CPSF6 mRNA expression and overall survival with all cases (C), stage I+II cases (D), stage III cases (E), and lymph node metastasis cases (F) using the data of GSE53624 dataset. ESCC, esophageal squamous cell carcinoma; TCGA, The Cancer Genome Atlas.

(51 months) (P=0.062) (*Figure 3J*). Otherwise, expression level of CPSF6 protein in the nucleus was not associated with overall survival time of ESCC patients (P=0.554) (*Figure 3K*). Above results reveal that it could be the cytoplasmic CPSF6 protein, rather than the total CPSF6 protein or the nucleoplasmic CPSF6 protein, that serves as a potential indicator for poor patient prognosis.

CPSF6 plays oncogenic roles in ESCC

We subsequently retrieved the expression information of CPSF6-related genes (adjusted P<0.05 and |Pearson correlation coefficient| ≥0.5) from the GSE53624 dataset for GO biological process (BP) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. As the results, CPSF6-related genes were mostly enriched in cell cycle- and apoptosis-related pathways (*Figure 4A*,*B*). So, we speculated that CPSF6 may affect cell growth and death in ESCC. For exploring these biological roles of CPSF6, we firstly knocked down and overexpressed CPSF6 in ESCC cells. Obviously, depletion of CPSF6 caused a sharp decrease in CPSF6 expression while overexpression of CPSF6 led to a large increase of CPSF6 expression (*Figure 5A*,*B*). Functionally, depletion of CPSF6

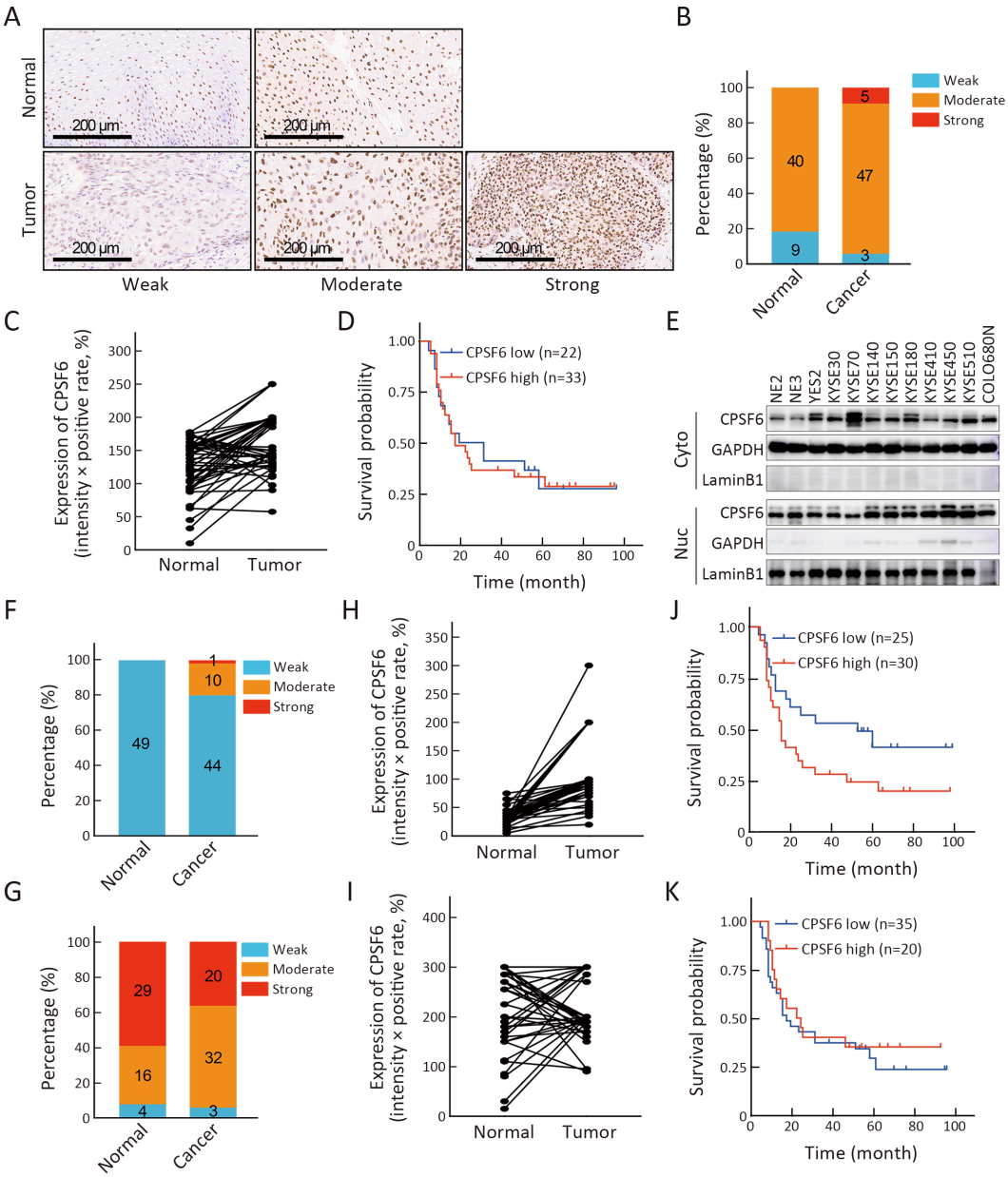

Figure 3 High cytoplasmic protein level of CPSF6 is detrimental to ESCC patient prognosis. (A) IHC staining of CPSF6 protein in ESCC and adjacent normal tissues; (B) Distribution of samples with different grades of total CPSF6 protein in ESCC (n=55) and adjacent normal tissues (n=49) (P=0.005); (C) Comparing total expression of CPSF6 protein in ESCC and adjacent normal tissues by IHC index [intensity × positive rate (%)] (P<0.001); (D) Kaplan-Meier survival analysis of cases according to the level of total expression of CPSF6 protein [HR=1.027 (95% CI: 0.637−1.965), P=0.934]; (E) Western blot assay for detecting the subcellular locations of CPSF6 protein in ESCC cell lines; (F) Distribution of samples with different grades of cytoplasmic CPSF6 protein in ESCC and adjacent normal tissues (P=0.001); (G) Comparing cytoplasmic expression of CPSF6 protein in ESCC and adjacent normal tissues by IHC index (P=0.051); (H) Kaplan-Meier survival analysis of cases according to the level of cytoplasmic expression of CPSF6 protein (P<0.001); (I) Distribution of samples with different grades of nucleoplasmic CPSF6 protein in ESCC and adjacent normal tissues (P=0.760); (J) Comparing nucleoplasmic expression of CPSF6 protein in ESCC and adjacent normal tissues by IHC index [HR=1.820 (95% CI: 0.964−3.439), P=0.062]; (K) Kaplan-Meier survival analysis of cases according to the level of nucleoplasmic expression of CPSF6 protein [HR=0.821 (95% CI: 0.427−1.576), P=0.554]. ESCC, esophageal squamous cell carcinoma; IHC, immunohistochemistry; HR, hazard ratio; 95% CI, 95% confidence interval.

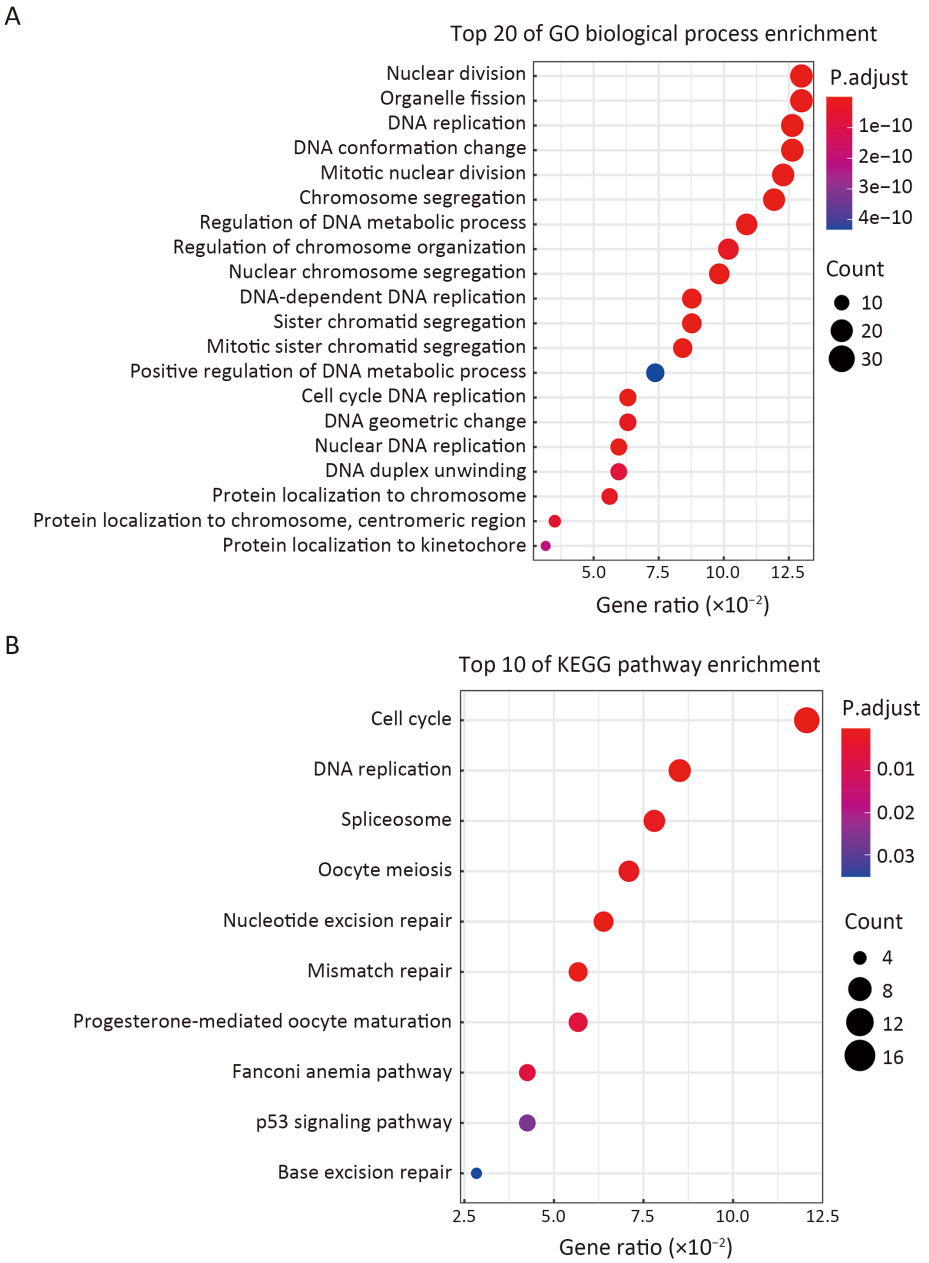

Figure 4 Gene enrichment analyses of CPSF6-related genes. (A) Top 20 biological process terms enriched in GO analysis; (B) Top 10 pathways enriched in KEGG analysis.

significantly attenuated cell proliferation and colony formation, while overexpression of CPSF6 significantly bolstered up these two malignant phenotypes in ESCC cells (*Figure 5C*,*D*). Otherwise, *in vivo* xenograft assay also demonstrated that the proliferation rate of subcutaneous tumor was significantly reduced after knocking down CPSF6 in ESCC cells *(Figure 5E*). Collectively, these observations indicate that CPSF6 plays an oncogenic role in ESCC.

Knockdown of CPSF6 induces cell-cycle arrest and apoptosis of ESCC

Recalling the results of signaling pathway analyses, we hypothesized that the pro-oncogenic effect of CPSF6 might be mediated by impacting cell cycle and apoptosis of ESCC cells. We thus conducted cell cycle and apoptosis assays with knockdown of CPSF6 in ESCC cell lines, respectively. As the results showed, the cell cycle was

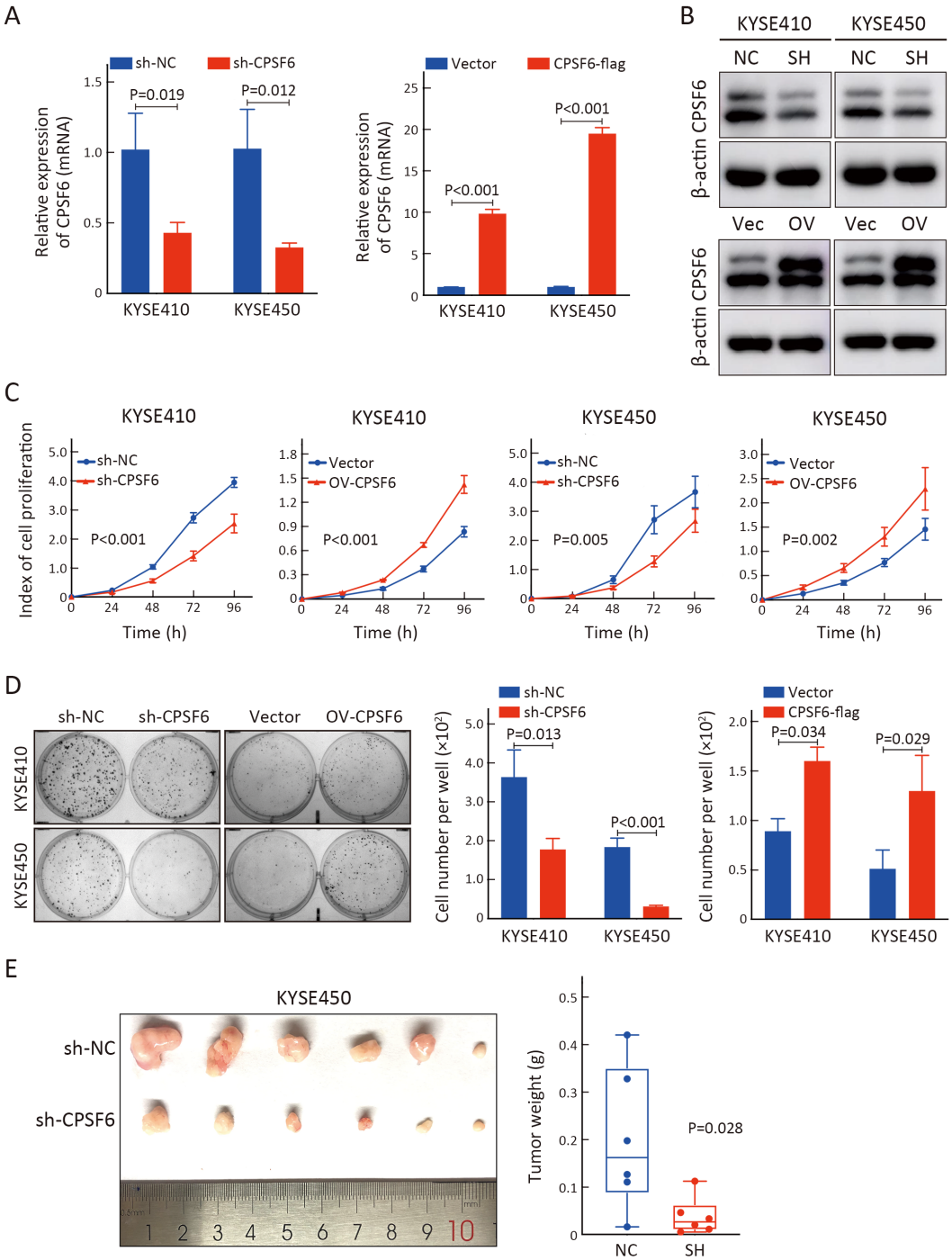

Figure 5 CPSF6 plays oncogenic roles *in vitro* and *in vivo*. (A,B) RT-qPCR assay (A) and Western blot assay (B) for detecting the mRNA and protein level of CPSF6 in ESCC cell lines (KYSE410 and KYSE450) after knockdown or overexpression of CPSF6; (C) Proliferation curves of ESCC cell lines (KYSE410 and KYSE450) after knockdown or overexpression of CPSF6; (D) Colony formation assay of ESCC cell lines (KYSE410 and KYSE450) after knockdown or overexpression of CPSF6. The number of colonies is presented in bar plots; (E) Subcutaneous tumorigenesis assay was performed after knockdown of CPSF6 in KYSE450 cells. Tumor weight was measured on d 30 after subcutaneous injection. A box-plot is presented for tumor weight. ESCC, esophageal squamous cell carcinoma.

significantly arrested at G_0/G_1 phase in CPSF6-depleted cells (*Figure 6A,B*). Further experiments proved the resistance to apoptosis of ESCC cells was also significantly diminished after knockdown of CPSF6 (*Figure 6C,D*). In parallel, we found that the level of cleaved caspase 3 was dramatically increased in CPSF6-depleted cells (*Figure 6E*),

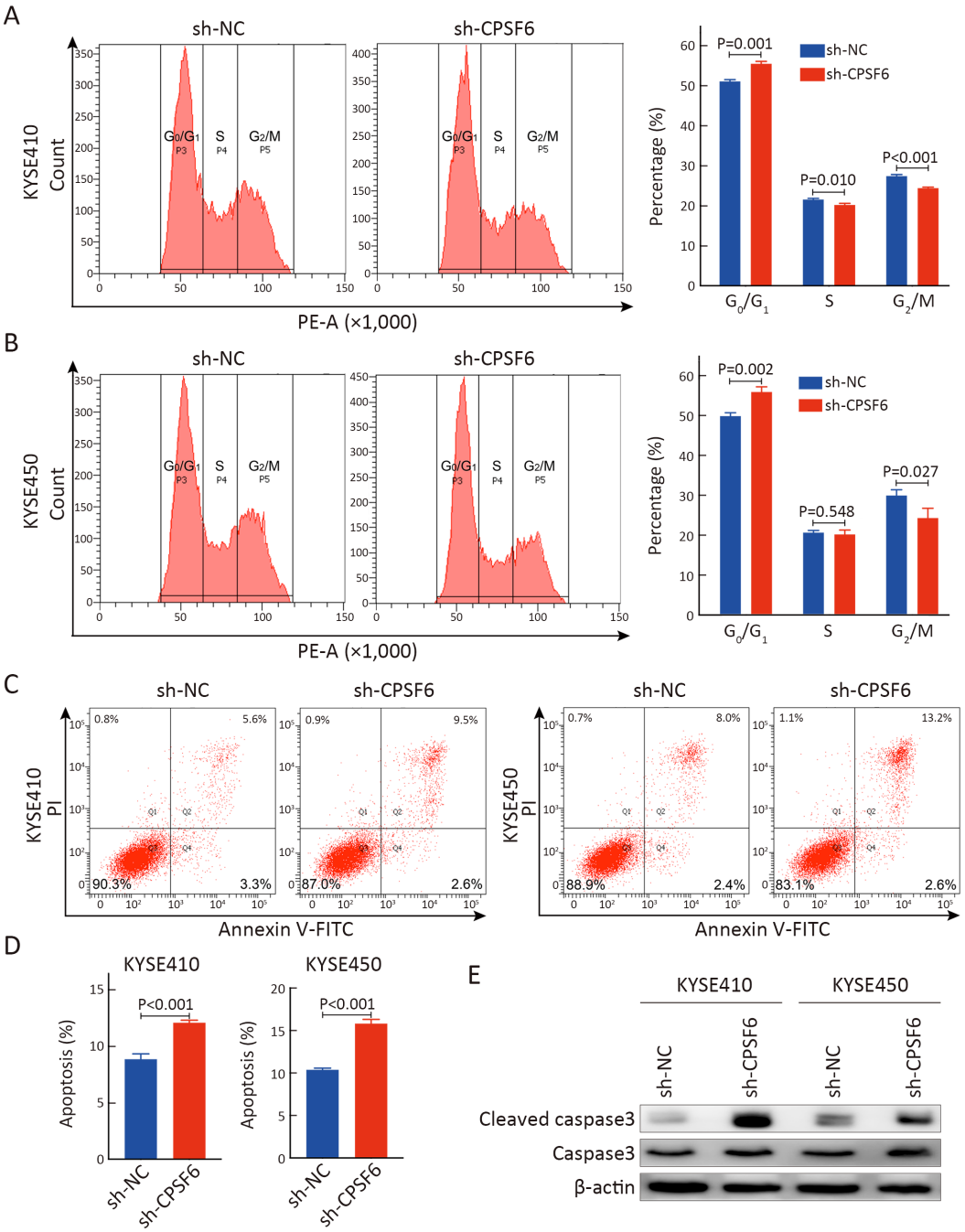

Figure 6 Knockdown of CPSF6 induces cell-cycle arrest and apoptosis of ESCC. (A,B) Cell cycle assay was performed after knockdown of CPSF6 in KYSE410 (A) and KYSE450 (B) cells, and the percentage of cells in each cell cycle phase is presented in bar plots; (C) Apoptosis assay was performed after knockdown of CPSF6 in ESCC cell lines (KYSE410 and KYSE45); (D) Bar plots are presented for the proportion of apoptotic cells; (E) Caspase 3 and cleaved caspase 3 were detected by Western blot after knockdown of CPSF6 in ESCC cell lines (KYSE410 and KYSE45). ESCC, esophageal squamous cell carcinoma.

which meant cell apoptosis was strongly activated after knockdown of CPSF6. Taken together, these results unveil that CPSF6 boosts the rate of cell division and the ability of cells to resist apoptosis, which may be the causes of promoting malignant phenotypes in ESCC.

CPSF6 protein binds to and regulates its circular isoform circCPSF6 (hsa_circ_0000417)

CPSF6 is an RNA-binding protein (RBP) which contains an N-terminal RNA recognition motif (21). For investigating the RNAs interacting with CPSF6 protein, we downloaded the genome-wide cross-linking immunoprecipitation sequencing (CLIP-seq) data of CPSF6 from the Encyclopedia of RNA Interactomes (ENCORI) database and the CLIP-seq data showed that CPSF6 protein can interact with different kinds of RNAs, including mRNA, circRNA, lncRNA and so on. Apart from a large number of mRNAs bound with CPSF6 protein, the number of CPSF6 protein-binding circRNA far exceeded that of other categories of RNA (*Figure 7A*). Accidentally, we noticed a circular isoform of CPSF6 bound to CPSF6 protein in CLIP-seq data. However, the circular isoform was not labeled with cirRNA ID in the CLIP-seq data. To identify the specific circular isoform of CPSF6, we queried our own RNA-seq data and determined that the major circular isoform of CPSF6 in ESCC was circCPSF6 (hsa_circ_0000417) (unpublished data). CircCPSF6 consists of exons 2-9 of *CPSF6* gene and is formed by back-splicing. We confirmed that circCPSF6 was resistant to RNase R and had a conventional back-splicing junction site consistent with annotation in Human circRNA Database (circbank) (*Figure 7B*). The circRNA pulldown assay and RIP assay showed that these two molecules could interact with each other indeed (*Figure 7 C*,*D*). Furthermore, we demonstrated the subcellular distribution of circCPSF6 and found it mainly located in the cytoplasm (79.7%) in ESCC cells (*Figure 7E*). Reminded that a portion of CPSF6 protein was also detected in the cytoplasm, we wondered if there was any regulatory relationship between these two molecules. For this purpose, CPSF6 and circCPSF6 were both detected in CPSF6- or circCPSF6-depleted cells. As the results, circCPSF6 underwent significant downregulation in CPSF6-depleted cells (*Figure 7F*). In contrast, no significant changes of CPSF6 were seen in circCPSF6-depleted cells (*Figure 7G*). As with the impacts of CPSF6, depletion of circCPSF6 arrested cell cycle in G_0/G_1 phase and increased apoptosis of ESCC cells (*Figure 7H*,*I*). Above results encourage us that CPSF6 may act as an upstream regulator of circCPSF6 and may play the oncogenic roles, at least in part, through circCPSF6.

Discussion

CPSF6 protein, as a regulator of APA, was mainly detected in nucleoplasm and nuclear speckles, and only small amount of that was located in the cytoplasm (22,23). Recent studies have reported that CPSF6 is dysregulated in hepatocellular carcinoma and gastric cancer and functions predominantly in the nucleus at the level of APA regulation to promote tumorigenesis and development (13,18). However, there were no studies reporting the clinical values and functions of cytoplasmic CPSF6 protein. In the present study, we found that cytoplasmic CPSF6 was up-regulated in ESCC and associated with poor prognosis of ESCC patients, which indicates that cytoplasmic CPSF6 could be a potential biological marker for prognostic assessment in ESCC. The high expression of cytoplasmic CPSF6 protein in ESCC also suggests that CPSF6 may provide previously unknown functions to sustain tumor cell survival.

The development of tumor is an extremely complex and systemic process in which a large number of genes are hyperactive or super-silenced. Dysregulation of gene expression is a key feature of ESCC (5,6,9,24). It has been reported that aberrant genes of ESCC mostly cluster in cell cycle, apoptosis and DNA damage control pathways (25), which affect malignant phenotypes, including cell proliferation, clone formation, migration and invasion, of cancer in different ways. In order to study the potential biological roles of CPSF6 in ESCC, we firstly retrieved the CPSF6-related genes in GSE53624 dataset and performed GO and KEGG enrichment analyses. Notably, CPSF6-related genes are mainly participated in the pathways associated with cell division and apoptosis, such as cell cycle, DNA replication, DNA repair and p53 signaling. These results prompted us that CPSF6 might affect growth and death of ESCC cells. Indeed, we revealed that overexpression of CPSF6 promoted cell proliferation and colony-formation capacity, while knockdown of CPSF6 diminished these two malignant phenotypes, inhibited subcutaneous tumor growth rate, as well as raised cell cycle arrest and apoptosis of ESCC cells. Disrupting cell cycle and promoting apoptosis of tumor cells are the keys to radiotherapy and chemotherapy for tumor treatment. For instance, radiotherapy can trigger cell cycle arrest at G_2/M

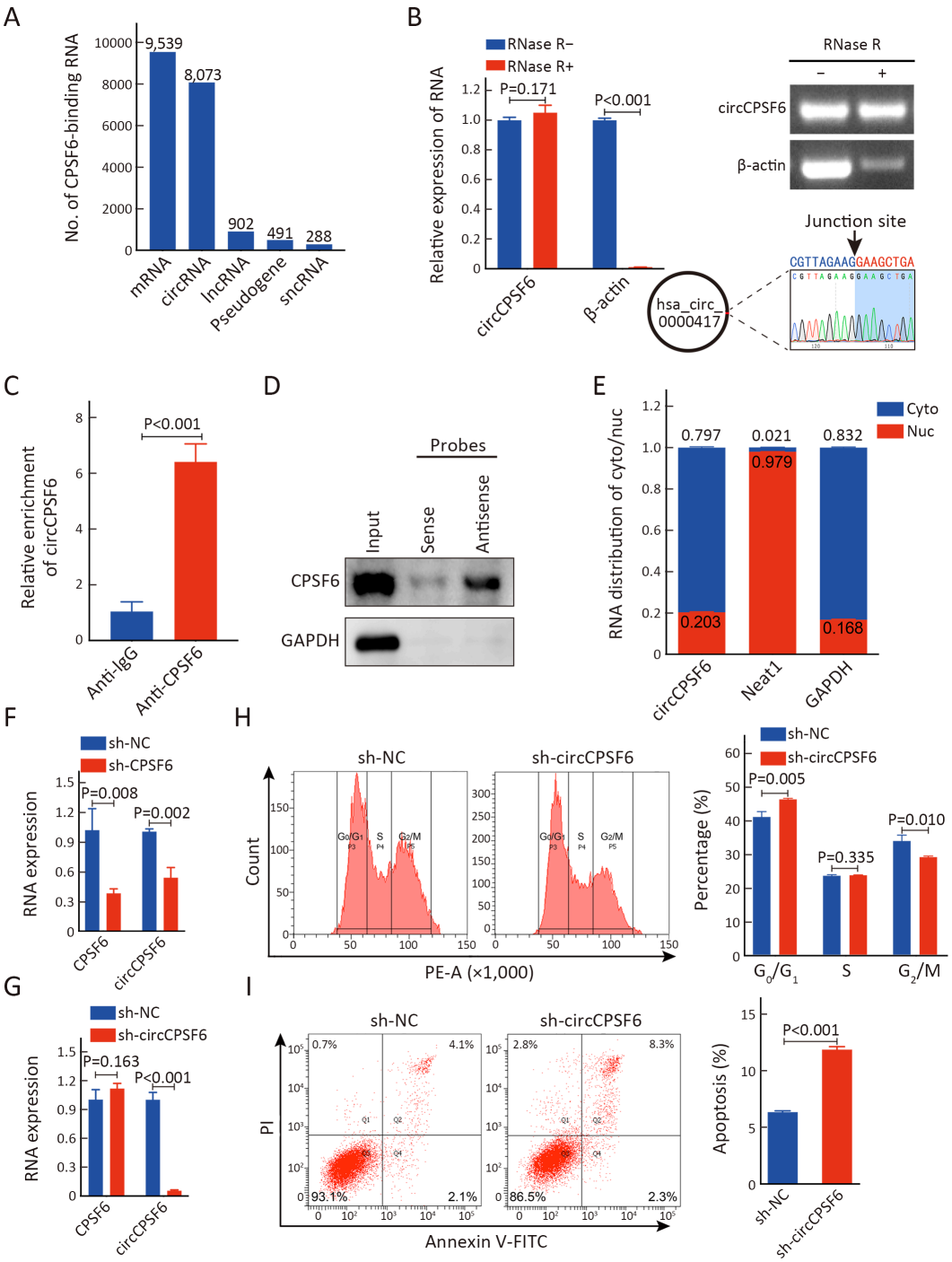

Figure 7 CPSF6 protein binds to and regulates its circular isoform circCPSF6 (hsa_circ_0000417). (A) A bar plot is presented for CPSF6 protein-binding RNA of CLIP-seq data from the ENCORI database; (B) Identification of circCPSF6 through RNase R treatment and Sanger sequencing of junction site; (C,D) Interaction between CPSF6 protein and circCPSF6 was detected by RIP assay (C) and RNA-protein pulldown assay (D); (E) Subcellular distribution of circCPSF6 in ESCC presented by a bar plot; (F,G) RT-qPCR assay for detecting the RNA level of CPSF6 and circCPSF6 after depletion of CPSF6 (F) and circCPSF6 (G), respectively; (H) Cell cycle assay was performed after knockdown of circCPSF6, and each cell cycle phase is presented by a bar plot; (I) Apoptosis assay was performed after knockdown of circCPSF6, and a bar plot is presented for the proportion of apoptotic cells. CLIP-seq, cross-linking immunoprecipitation sequencing; ESCC, esophageal squamous cell carcinoma.

phase and promote cell apoptosis through inducing DNA damage (26,27). In addition, cisplatin and paclitaxel, the first-line chemotherapeutic agents for the clinical treatment of esophageal cancer, exert their effects by inducing G_0/G_1 arrest and apoptosis of tumor cells (28,29). Considering the effects of CPSF6 to cell cycle and apoptosis, CPSF6 may be a potential therapeutic target to improve the outcomes of ESCC patients with radiotherapy and chemotherapy.

CPSF6, as a canonical RBP (21,23,30,31), plays oncogenic roles in human cancers (12,13,18). A few studies have reported that RBPs could dynamically regulate circRNA biogenesis, functions and metabolism (32-36). CircRNA, a novel identified type of RNA, has been reported to exhibit aberrant expression in multiform types of human cancers (37-39). RBP-circRNA interactions may play important roles in tumorigenesis and progression. Apart from the CLIP-seq data described above, no studies have reported that CPSF6 could bind with circRNA. Interestingly, we explored that CPSF6 protein was able to bind with its circular isoform circCPSF6 (hsa_circ_0000417) and perturbed circCPSF6 expression in ESCC cells. In particular, we found that circCPSF6 exhibited similar functions to CPSF6 protein in accelerating cell cycle and repressing apoptosis in ESCC cells, implying that CPSF6 might exert its tumor-promoting effects through circCPSF6. In further studies, we will investigate the roles of circCPSF6 in tumorigenesis and progression of ESCC and attempt to reveal the regulatory manner of CPSF6 protein to circCPSF6. Considering the fact of CPSF6 as a m*6*A reader (40), is it possible for CPSF6 to act as a broad-spectrum circRNA-binding protein that regulates circRNA metabolism in trafficking and decay? It is also an attractive issue that deserves investigating in the future.

Conclusions

Our investigations show that CPSF6 is significantly up-regulated in ESCC and high level of cytoplasmic CPSF6 is detrimental to ESCC patient prognosis. Biologically, CPSF6 may function as an oncoprotein in ESCC, at least in part, through regulating circCPSF6 expression.

Acknowledgements

This study was supported by National Natural Science Foundation of China (No. 81988101, 81830086 and 81872398) and CAMS Innovation Fund for Medical Sciences (No. 2021-1-I2M-014).

Footnote

Conflicts of Interest: These authors have no conflicts of interests to declare.

References

1. Li J, Xu J, Zheng Y, et al. Esophageal cancer: Epidemiology, risk factors and screening. Chin J Cancer Res 2021;33:535-47.
2. Lagergren J, Smyth E, Cunningham D, et al. Oesophageal cancer. Lancet 2017;390:2383-96.
3. Li M, Shao D, Zhou J, et al. Signatures within esophageal microbiota with progression of esophageal squamous cell carcinoma. Chin J Cancer Res 2020;32:755-67.
4. He Z, Ke Y. Precision screening for esophageal squamous cell carcinoma in China. Chin J Cancer Res 2020;32:673-82.
5. Cancer Genome Atlas Research Network, Analysis Working Group:Asan University, BC Cancer Agency, et al. Integrated genomic characterization of oesophageal carcinoma. Nature 2017;541:169-75.
6. Song Y, Li L, Ou Y, et al. Identification of genomic alterations in oesophageal squamous cell cancer. Nature 2014;509:91-5.
7. Chang J, Tan W, Ling Z, et al. Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. Nat Commun 2017;8:15290.
8. Cui Y, Chen H, Xi R, et al. Whole-genome sequencing of 508 patients identifies key molecular features associated with poor prognosis in esophageal squamous cell carcinoma. Cell Res 2020;30:902-13.
9. Yan T, Cui H, Zhou Y, et al. Multi-region sequencing unveils novel actionable targets and spatial heterogeneity in esophageal squamous cell carcinoma. Nat Commun 2019;10:1670.
10. Zhang Y, Liu L, Qiu Q, et al. Alternative polyadenylation: methods, mechanism, function, and role in cancer. J Exp Clin Cancer Res 2021;40:51.
11. Tian B, Manley JL. Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol 2017;18:18-30.
12. Binothman N, Hachim IY, Lebrun JJ, et al. CPSF6 is a clinically relevant breast cancer vulnerability target:

Role of CPSF6 in breast cancer. EBioMedicine 2017;21:65-78.
13. Tan S, Zhang M, Shi X, et al. CPSF6 links alternative polyadenylation to metabolism adaption in hepatocellular carcinoma progression. J Exp Clin Cancer Res 2021;40:85.
14. 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;63:1700-10.
15. Wang G, Guo S, Zhang W, et al. Co-expression network analysis identifies key modules and hub genes implicated in esophageal squamous cell cancer progression. Medicine in Omics 2021;1:100003.
16. Li JH, Liu S, Zhou H, et al. starBase v2. 0:decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014;42:D92-7.
17. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284-7.
18. Shi X, Ding K, Zhao Q, et al. Suppression of CPSF6 enhances apoptosis through alternative polyadenylation-mediated shortening of the *VHL* 3'UTR in gastric cancer cells. Front Genet 2021;12:707644.
19. Wang BJ, Liu DC, Guo QY, et al. NUDT21 suppresses breast cancer tumorigenesis through regulating CPSF6 expression. Cancer Manag Res 2020;12:3069-78.
20. Tang Z, Kang B, Li C, et al. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res 2019;47: W556-60.
21. Dettwiler S, Aringhieri C, Cardinale S, et al. Distinct sequence motifs within the 68-kDa subunit of cleavage factor I_m mediate RNA binding, protein-protein interactions, and subcellular localization. J Biol Chem 2004;279:35788-97.
22. Cardinale S, Cisterna B, Bonetti P, et al. Subnuclear localization and dynamics of the pre-mRNA 3' end processing factor mammalian cleavage factor I 68-kDa subunit. Mol Biol Cell 2007;18:1282-92.
23. Ruepp MD, Aringhieri C, Vivarelli S, et al. Mammalian pre-mRNA 3' end processing factor CF I_m68 functions in mRNA export. Mol Biol Cell 2009;20:5211-23.
24. Su H, Hu N, Yang HH, et al. Global gene expression profiling and validation in esophageal squamous cell carcinoma and its association with clinical phenotypes. Clin Cancer Res 2011;17:2955-66.
25. Gao YB, Chen ZL, Li JG, et al. Genetic landscape of esophageal squamous cell carcinoma. Nat Genet 2014;46:1097-102.
26. Dewey WC, Ling CC, Meyn RE. Radiation-induced apoptosis: relevance to radiotherapy. Int J Radiat Oncol Biol Phys 1995;33:781-96.
27. Pawlik TM, Keyomarsi K. Role of cell cycle in mediating sensitivity to radiotherapy. Int J Radiat Oncol Biol Phys 2004;59:928-42.
28. Wang TH, Wang HS, Soong YK. Paclitaxel-induced cell death: where the cell cycle and apoptosis come together. Cancer 2000;88:2619-28.
29. Zheng AW, Chen YQ, Fang J, et al. Xiaoaiping combined with cisplatin can inhibit proliferation and invasion and induce cell cycle arrest and apoptosis in human ovarian cancer cell lines. Biomed Pharmacother 2017;89:1172-7.
30. Awasthi S, Alwine JC. Association of polyadenylation cleavage factor I with U1 snRNP. RNA 2003;9: 1400-9.
31. Zhu Y, Wang X, Forouzmand E, et al. Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. Mol Cell 2018;69:62-74.e4.
32. Conn SJ, Pillman KA, Toubia J, et al. The RNA binding protein quaking regulates formation of circRNAs. Cell 2015;160:1125-34.
33. Li X, Liu CX, Xue W, et al. Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. Mol Cell 2017;67:214-27.e7.
34. Dong W, Dai ZH, Liu FC, et al. The RNA-binding protein RBM3 promotes cell proliferation in hepatocellular carcinoma by regulating circular RNA SCD-circRNA 2 production. EBioMedicine 2019;45:155-67.
35. Di Timoteo G, Dattilo D, Centrón-Broco A, et al. Modulation of circRNA metabolism by m^6A modification. Cell Rep 2020;31:107641.
36. Okholm TLH, Sathe S, Park SS, et al. Transcriptome-wide profiles of circular RNA and

RNA-binding protein interactions reveal effects on circular RNA biogenesis and cancer pathway expression. Genome Med 2020;12:112.
37. Patop IL, Kadener S. circRNAs in Cancer. Curr Opin Genet Dev 2018;48:121-7.
38. Chen S, Huang V, Xu X, et al. Widespread and functional RNA circularization in localized prostate cancer. Cell 2019;176:831-43.e22.
39. Vo JN, Cieslik M, Zhang Y, et al. The landscape of circular RNA in cancer. Cell 2019;176:869-81.e13.
40. Perez-Perri JI, Rogell B, Schwarzl T, et al. Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. Nat Commun 2018;9:4408.

Cite this article as: Guo S, Wang G, Zhao Z, Li D, Song Y, Zhan Q. Deregulated expression and subcellular localization of CPSF6, a circRNA-binding protein, promote malignant development of esophageal squamous cell carcinoma. Chin J Cancer Res 2022;34(1):11-27. doi: 10.21147/j.issn.1000-9604.2022.01.02

Table S1 CPSF6 protein expression data and relevant clinical information of ESCC microarray (HEsoS105Su01)

Sample code	Surgery time	Survival status	Follow-up time	Survival period (cut-off 2014/3) (month)	Remark	Gender	Age (year)	Distant metastasis	Primary foci	Pathological classification	Pathological grade	Tumor site	Pathological pattern	Surgical margin	Number of lymph nodes	Positive lymph nodes	T	N	M	AJCC cancer staging (7 edition)	Intensity (cytoplasm) (%)	Positivity rate (cytoplasm) (%)	Intensity (nucleus) (%)	Positivity rate (nucleus) (%)	Tissue type
D08A1704	2008/12/15	Death	2010/2/9	14		Female	62	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated		19	1	T3	N1	M0	3A	1	100	2	90	Cancer
																					1	45	3	100	Adjacent
D08A1716	2005/11/30	Death	2007/2/7	15		Male	51	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated		8	2	T3	N1	M0	3A	1	100	2	100	Cancer
																					1	35	1	30	Adjacent
D08A1720	2006/1/19	Death	2008/8/9	31		Male	65	No	Yes	Squamous cell carcinoma	I–II	Esophagus	Bulge	Negative	3	0	T3	N0	M0	2	1	90	2	45	Cancer
																					1	25	2	75	Adjacent
D08A1721	2006/2/3	Death	2006/7/7	5		Male	54	No	Yes	Squamous cell carcinoma	III	Esophagus	Bulge	Negative	10	5	T3	N2	M0	3B	2	100	2	85	Cancer
																					1	20	2	80	Adjacent
D08A1725	2006/3/13	Survival	2014/3	96		Male	51	No	Yes	Squamous cell carcinoma	II	Esophagus	Bulge	Negative	10	0	T2	N0	M0	2	1	90	2	75	Cancer
																					1	40	3	90	Adjacent
D08A1728	2006/4/3	Survival	2014/3	95		Female	37	No	Yes	Squamous cell carcinoma	II	Esophagus	Bulge	Negative	5	0	T3	N0	M0	2	2	100	2	100	Cancer
																					1	35	3	90	Adjacent
D08A1729	2006/4/21	Death	2007/7/9	15		Female	67	No	Yes	Squamous cell carcinoma	III	Esophagus	Ulcerated	Negative	17	4	T3	N2	M0	3B	1	100	2	100	Cancer
D08A1737	2006/6/16	Survival	2014/3	93		Female	62	No	Yes	Basal cell-like squamous cell carcinoma	II	Esophagus		Negative			T2		M0		1	40	3	90	Cancer
																					1	30	3	95	Adjacent
D08A1738	2006/6/29	Death	2011/4/5	58		Male	58	No	Yes	Squamous cell carcinoma	II	Middle and upper esophagus	Bulge	Negative	14	2	T3	N1	M0	3A	1	60	2	95	Cancer
																					1	65	2	100	Adjacent
D08A1742	2006/7/6	Death	2007/3/9	8		Male	57	Para-aortic lymph nodes (1/1)	Yes	Squamous cell carcinoma	II–III	Middle and lower esophagus	Ulcerated	Negative	6	2	T2	N1	M1	4	1	100	2	100	Cancer
																					– –	– –	– –	– –	Adjacent
D08A1747	2006/11/27	Death	2007/3/9	4		Male	73	No	Yes	Squamous cell carcinoma	II	Esophagus	Bulge	Negative	15	0	T2	N0	M0	2	1	75	2	95	Cancer
																					1	35	2/3	80	Adjacent
D08A1749	2007/1/23	Death	2007/9/2	8		Male	48	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated	Negative	6	2	T3	N1	M0	3A	1	100	3	95	Cancer

Table S1 (*continued*)

Table S1 (*continued*)

Sample code	Surgery time	Survival status	Follow-up time	Survival period (cut-off 2014/3) (month)	Remark	Gender	Age (year)	Distant metastasis	Primary foci	Pathological classification	Pathological grade	Tumor site	Pathological pattern	Surgical margin	Number of lymph nodes	Positive lymph nodes	T	N	M	AJCC cancer staging (7 edition)	Intensity (cytoplasm) (%)	Positivity rate (cytoplasm) (%)	Intensity (nucleus) (%)	Positivity rate (nucleus) (%)	Tissue type
																					1	40	1	85	Adjacent
D08A1750	2007/2/28	Death	2007/12/9	10		Male	59	No	Yes	Squamous cell carcinoma	II	Middle and lower esophagus	Ulcerated	Negative	3	3	T3	N2	M0	3B	1	100	2	90	Cancer
																					1	35	3	85	Adjacent
D08A1751	2007/3/8	Death	2008/10/27	19		Female	73	No	Yes	Squamous cell carcinoma	III	Middle and upper esophagus	Ulcerated	Squamous epithelial heterogeneity, carcinoma	10	0	T3	N0	M0	2	1	50	2	85	Cancer
																					1	45	2	100	Adjacent
D08A1754	2007/3/29	Death	2008/8/7	17		Female	58	No	Yes	Squamous cell carcinoma	II	Middle and lower esophagus	Infiltrating ulcer	Negative	23	3	T3	N2	M0	3B	2	100	2	100	Cancer
D08A1758	2007/5/28	Death	2008/1/9	8		Male	70	No	Yes	Squamous cell carcinoma	II	Middle esophagus	Ulcerated	Negative	1	0	T3	N0	M0	2	1	90	2	95	Cancer
																					1	40	3	90	Adjacent
D08A1760	2007/8/28	Death	2011/11/19	51		Female	57	No	Yes	Squamous cell carcinoma	II	Esophagus		Negative	11	0	T2	N0	M0	2	1	95	2	95	Cancer
																					1	75	2	100	Adjacent
D08A1761	2007/9/6	Death	2008/12/6	15		Male	78	No	Yes	Squamous cell carcinoma	II–III	Esophagus	Ulcerated	Negative	3	0	T3	N0	M0	2	1	100	2	80	Cancer
																					1	45	3	100	Adjacent
D08A1763	2007/9/24	Death	2008/4/1	7		Female	55	Supraclavicular lymph nodes (4/6)	Yes	Squamous cell carcinoma	I	Esophagus	Ulcerated	Negative	22	4	T3	N2	M1	4	1	100	2	95	Cancer
																					1	40	2	75	Adjacent
D08A1764	2007/10/16	Death	2009/9/20	23		Male	57	No	Yes	Squamous cell carcinoma	II	Esophagus	Narrowing	Negative	6	1	T3	N1	M0	3A	2	100	2	100	Cancer
																					1	30	3	95	Adjacent
D08A1772	2007/11/8	Death	2010/6/30	31		Male	64	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated	Negative	31	2	T3	N1	M0	3A	1	100	1	95	Cancer
																					1	55	3	100	Adjacent

Table S1 (*continued*)

Table S1 (*continued*)

Sample code	Surgery time	Survival status	Follow-up time	Survival period (cut-off 2014/3) (month)	Remark	Gender	Age (year)	Distant metastasis	Primary foci	Pathological classification	Pathological grade	Tumor site	Pathological pattern	Surgical margin	Number of lymph nodes	Positive lymph nodes	T	N	M	AJCC cancer staging (7 edition)	Intensity (cytoplasm) (%)	Positivity rate (cytoplasm) (%)	Intensity (nucleus) (%)	Positivity rate (nucleus) (%)	Tissue type
D08A1774	2007/11/28	Survival	2015/7	76	2015/7/7 (Death)	Female	70	No	Yes	Squamous cell carcinoma	II	Esophagus	Myxomatous	Negative	12	1	T2	N1	M0	2B	2	100	2	100	Cancer
																					1	25	3	95	Adjacent
D08A1785	2008/1/9	Death	2009/3/30	14		Male	53	No	Yes	Squamous cell carcinoma	I–II	Esophagus	Bulge	Negative	4	0	T2	N0	M0	2	2	100	3	100	Cancer
D08A1789	2008/1/24	Death	2008/8/17	7		Female	52	Supraclavicular lymph nodes (1/1)	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated	Negative	9	2	T3	N1	M1	4	1	70	2	90	Cancer
																					1	30	3	75	Adjacent
D08A1793	2008/2/15	Survival	2014/3	73		Male	57	No	Yes	Squamous cell carcinoma	II	Esophagus	Narrowing	Negative	17	0	T3	N0	M0	2	1	100	3	95	Cancer
																					1	35	2	85	Adjacent
D08A1795	2008/2/25	Death	2008/7/20	5		Male	44	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated	Infiltrative growth with focal point closer to the basal cut edge of the mass	12	0	T4a	N0	M0	3A	3	100	2	100	Cancer
																					1	30	3	90	Adjacent
D08A1803	2008/3/24	Death	2012/1/2	46		Male	54	Supraclavicular lymph nodes (1/1)	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated	Negative	27	3	T3	N2	M1	4	1	100	3	100	Cancer
																					1	25	2	80	Adjacent
D08A1809	2008/5/8	Death	2013/6/21	61		Female	54	No	Yes	Squamous cell carcinoma	II	Middle and upper esophagus	Bulge	Negative	8	0	T3	N0	M0	2	2	100	2	100	Cancer
																					1	35	3	85	Adjacent
D08A1810	2008/5/19	Survival	2014/3	70	2014/12/26 (Death)	Male	57	No	Yes	Squamous cell carcinoma	I	Esophagus	Infiltrating ulcer	Negative			T3		M0		1	85	2	95	Cancer
																					1	40	3	95	Adjacent
D08A1814	2008/6/11	Death	2009/11/1	17		Male	53	No	Yes	Squamous cell carcinoma	I–II	Esophagus	Ulcerated	Negative	30	0	T2	N0	M0	2	1	95	3	100	Cancer

Table S1 (*continued*)

Table S1 (*continued*)

Sample code	Surgery time	Survival status	Follow-up time	Survival period (cut-off 2014/3) (month)	Remark	Gender	Age (year)	Distant metastasis	Primary foci	Pathological classification	Pathological grade	Tumor site	Pathological pattern	Surgical margin	Number of lymph nodes	Positive lymph nodes	T	N	M	AJCC cancer staging (7 edition)	Intensity (cytoplasm) (%)	Positivity rate (cytoplasm) (%)	Intensity (nucleus) (%)	Positivity rate (nucleus) (%)	Tissue type
																					1	25	3	75	Adjacent
D08A1819	2008/7/4	Death	2009/3/9	8		Male	58	No	Yes	Squamous cell carcinoma	II	Esophagus		Negative	30	2	T2	N1	M0	2B	1	100	3	100	Cancer
																					1	35	2	95	Adjacent
D08A1820	2008/7/1	Death	2009/6/1	11		Male	62	No	Yes	Squamous cell carcinoma	II–III	Esophagus	Bulge	Negative	14	1	T3	N1	M0	3A	2	100	3	100	Cancer
																					1	30	3	95	Adjacent
D08A1821	2008/7/8	Death	2010/7/6	24		Male	72	No	Yes	Squamous cell carcinoma	II	Esophagus	Pelvic	Negative	23	0	T3	N0	M0	2	1	95	3	100	Cancer
																					1	65	3	95	Adjacent
D08A1822	2008/7/14	Death	2009/7/12	12		Male	59	No	Yes	Squamous cell carcinoma	II	Lower esophagus	Infiltrating ulcer	Negative	24	3	T3	N2	M0	3B	1	85	3	100	Cancer
																					1	40	3	75	Adjacent
D08A1823	2008/8/4	Death	2009/5/6	9		Male	59	No	Yes	Squamous cell carcinoma	I–II	Middle and lower esophagus		Negative	22	16	T3	N3	M0	3C	1	95	3	100	Cancer
																					1	35	3	100	Adjacent
D08A1824	2008/8/6	Death	2009/4/1	8		Male	59	No	Yes	Squamous cell carcinoma	II	Lower esophagus	Ulcerated	Negative	19	8	T3	N3	M0	3C	1	100	2	100	Cancer
																					1	40	3	90	Adjacent
D08A1825	2008/8/11	Survival	2014/3	67		Female	59	No	Yes	Squamous cell carcinoma	II	Middle esophagus	Ulcerated	Negative	32	0	T3	N0	M0	2	1	90	3	100	Cancer
																					1	35	2	95	Adjacent
D08A1843	2008/12/5	Survival	2014/3	63		Female	56	No	Yes	Squamous cell carcinoma	I–II	Middle and lower esophagus	Ulcerated	Squamous epithelial heterogeneous hyperplasia on the upper cut edge	21	0	T3	N0	M0	2	2	100	3	100	Cancer
D08A1861	2009/5/8	Survival	2014/3	58		Male	53	No	Yes	Squamous cell carcinoma	II	Middle esophagus	Ulcerated	Negative	22	0	T3	N0	M0	2	1	35	2	95	Cancer
																					1	5	1	15	Adjacent

Table S1 (*continued*)

Table S1 (*continued*)

Sample code	Surgery time	Survival status	Follow-up time	Survival period (cut-off 2014/3) (month)	Remark	Gender	Age (year)	Distant metastasis	Primary foci	Pathological classification	Pathological grade	Tumor site	Pathological pattern	Surgical margin	Number of lymph nodes	Positive lymph nodes	T	N	M	AJCC cancer staging (7 edition)	Intensity (cytoplasm) (%)	Positivity rate (cytoplasm) (%)	Intensity (nucleus) (%)	Positivity rate (nucleus) (%)	Tissue type
D08A1862	2009/5/11	Death	2010/1/3	8		Male	52	Aorta (1/1)	Yes	Squamous cell carcinoma	II	Middle esophagus	Ulcerated	Negative	16	0	T3	N0	M1	4	1	20	1	95	Cancer
																					1	15	2	55	Adjacent
D08A1866	2009/6/23	Death	2010/2/11	8		Female	50	No	Yes	Squamous cell carcinoma	II	Middle and lower esophagus	Ulcerated	Negative	35	1	T3	N1	M0	3A	1	100	2	100	Cancer
																					1	20	1/2	75	Adjacent
D08A1869	2009/7/2	Survival	2014/3	56		Male	75	No	Yes	Squamous cell carcinoma	II–III	Middle esophagus	Narrowing	Negative	18	1	T3	N1	M0	3A	1	45	2	100	Cancer
																					1	10	2	40	Adjacent
D08A1872	2009/8/25	Death	2010/6/15	10		Male	61	No	Yes	Squamous cell carcinoma	III	Esophagus	Ulcerated	Moderate to severe squamous epithelial heterogeneity	21	2	T3	N1	M0	3A	1	45	3	95	Cancer
D08A1874	2009/9/9	Survival	2014/11	54	2014/11/10 (Death)	Male	53	No	Yes	Squamous cell carcinoma	I–II	Middle and lower esophagus	Ulcerated	Negative	13	0	T2	N0	M0	2	1	95	3	100	Cancer
																					1	30	2	90	Adjacent
D08A1875	2009/9/7	Survival	2014/3	54		Male	60	No	Yes	Squamous cell carcinoma	I	Middle and lower esophagus	Ulcerated	Negative	24	0	T3	N0	M0	2	1	80	3	100	Cancer
																					1	25	2	80	Adjacent
D08A1878	2009/10/11	Survival	2014/3	53		Male	67	No	Yes	Squamous cell carcinoma	II	Middle and lower esophagus	Bulge	Negative	27	0	T3	N0	M0	2	1	75	2	90	Cancer
																					1	15	3	75	Adjacent
D08A1882	2009/12/9	Death	2012/1/2	25		Male	60	No	Yes	Squamous cell carcinoma	III	Esophagus	Bulge	Negative	21	0	T3	N0	M0	2	1	100	3	100	Cancer
																					1	30	3	85	Adjacent
D08A1883	2009/12/23	Survival	2014/3	51		Female	63	No	Yes	Squamous cell carcinoma	II–III	Esophagus	Ulcerated		34	0		N0	M0		1	75	2	100	Cancer
																					1	25	2	90	Adjacent

Table S1 (*continued*)

Table S1 (*continued*)

Sample code	Surgery time	Survival status	Follow-up time	Survival period (cut-off 2014/3) (month)	Remark	Gender	Age (year)	Distant metastasis	Primary foci	Pathological classification	Pathological grade	Tumor site	Pathological pattern	Surgical margin	Number of lymph nodes	Positive lymph nodes	T	N	M	AJCC cancer staging (7 edition)	Intensity (cytoplasm) (%)	Positivity rate (cytoplasm) (%)	Intensity (nucleus) (%)	Positivity rate (nucleus) (%)	Tissue type
D08A1884	2009/12/30	Survival	2014/3	51		Female	57	No	Yes	Squamous cell carcinoma	I–II	Esophagus	Ulcerated			0		N0	M0		1	55	2	100	Cancer
																					1	30	2	75	Adjacent
D08A1885	2010/3/8	Survival	2014/3	48		Male	62	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated		21	0		N0	M0		1	100	3	100	Cancer
																					1	35	3	85	Adjacent
D08A1886	2010/3/18	Death	2011/3/23	12		Female	69	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated		16	4		N2	M0	3	1	95	2	95	Cancer
																					1	30	2	90	Adjacent
D08A1887	2010/11/8	Death	2011/8/2	9		Male	70	No	Yes	Squamous cell carcinoma	II	Esophagus			20	7		N3	M0	3C	1	100	2	95	Cancer
																					1	45	3	100	Adjacent
D08A1888	2010/1/13	Death	2010/11/14	10		Male	73	No	Yes	Squamous cell carcinoma	II	Esophagus	Ulcerated			0		N0	M0		2	100	3	100	Cancer
																					1	55	3	100	Adjacent
D08A1890	2010/3/2	Death	2012/1/28	22		Male	48	No	Yes	Squamous cell carcinoma	II	Esophagus	Bulge		12	0		N0	M0		1	100	3	100	Cancer
																					1	30	3	95	Adjacent
D08A2224	2011/1/25	Survival	2014/3	38		Female	54	No	Yes	Squamous cell carcinoma, partial basal cell-like squamous cell carcinoma	III	Esophagus	Bulge		16	0	T1	N0	M0	1	1	100	2	100	Cancer
																					1	40	3	95	Adjacent

ESCC, esophageal squamous cell carcinoma.