

Knockdown of CENPF inhibits the progression of lung adenocarcinoma mediated by ER β 2/5 pathway

Tang Hexiao^{1&*}MD, Ph.D, Bai Yuquan^{1&} MD, Xiong Lecai^{1&} MD, Wei Yanhong² MD, Ph.D, Hu Weidong¹ MD, Ph.D, Xu Ming¹ MD, Ph.D, Zhou Xuefeng¹ MD, Ph.D, Pan Gaofeng¹ MD, Ph.D, Zhang Li¹ MD, Ph.D, Zhu Minglin¹ MD, Ph.D, Tang Zheng¹ MD, Ph.D, Yang Zetian¹ MD, Ph.D, Zhou Xiao¹ MD, Cai Yi¹ MD, Michael Lanuti³ MD, Zhao Jinping¹MD, Ph.D.

¹ Department of Thoracic and Cardiovascular Surgery, Zhongnan Hospital of Wuhan University, Wuhan, 430071, China

² Department of Geriatrics, Zhongnan Hospital of Wuhan University, Wuhan, 430071, China

³ Division of Thoracic Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[&]Co-first author. They make the same contribution to this article.

* Correspondence to. **Tang Hexiao**, MD, Ph.D, the Department of Thoracic and Cardiovascular Surgery, Zhongnan Hospital of Wuhan University, Wuhan 430071, China. Telephone number: 086-15102723563. E-mail: thx1245@sina.com.

Word count: 4640 words.

Abstract

The signal transduction pathways of estrogen receptors (ER) mainly includes gene pathway and non-gene pathway. Studies have shown that the gene pathway of ER is related with the expression of nuclear proteins, and this is the key issue for our current research. With the GEO database analysis, Human centromere protein F (CENPF) is highly expressed in adenocarcinoma of lung (LUAD), and the co-expression of CENPF and ER β was found in the nucleus of LUAD cells. Meanwhile, CENPF and ER β 2/5 were related with T stage and poor prognosis ($P < 0.05$). Knockdown of CENPF gene significantly inhibited the biological effects of LUAD cells, the tumor growth of mice and the expression of ER β 2/5 ($P < 0.05$). Further, group experiments showed that knockdown CENPF inhibits biological effects of LUAD cells mediated by ER β pathway. All the results indicated that both CENPF and ER β 2/5 play important roles in the progression of LUAD, and knockdown of CENPF can inhibit the progression of LUAD by inhibiting the expression of ER2/5. Thus, the development of inhibitors against ER β 2/5 subtype and CENPF remained more effective in improving the therapeutic effect of LUAD.

Keywords: centromere protein F (CENPF); Estrogen Receptor beta; Lung Adenocarcinoma (LUAD); WGCNA package; non-small cell lung cancer (NSCLC)

Abbreviations:

ER: Estrogen receptors

AP-1: Activator protein 1

SP-1: Specificity protein 1

IL-6: Interleukin 6

EGF: Epidermal growth factor

CENPF: centromere protein F

LUAD: lung adenocarcinoma

BPL: benign pulmonary lesions

FBS: fetal bovine serum

FDR: false discovery rate

FC: fold change

WGCNA: weighted gene co-expression network analysis

GEPIA: Gene Expression Profiling Interactive Analysis

TCGA: The Cancer Genome Atlas

KD: knockdown

NC: negative control

ERE: estrogen response element

1. Introduction

Lung cancer is one of the most common and a leading cause of death in patients with malignant tumors¹. The incidence and mortality of lung cancer is on a significant increase worldwide². In addition to the risk factors including smoke and environment, growth factor pathway and hormonal regulation also play important roles in the pathogenesis of lung cancer³.

Estrogen receptors (ER) belong to the nuclear receptor steroid superfamily, and

are closely related with the imbalances of hormone regulation. ER is classified into two subtypes, ER α and ER β , which have different tissue distributions and biological effects in different types of tumors⁴. Previous studies have revealed that estrogen activates the transcription of target genes by binding to ER α and ER β directly, which is referred to as the classical gene signaling pathway⁵. Numerous studies have found that ER also interacts with other transcription factors such as activator protein 1 (AP-1), specificity protein 1 (SP-1), interleukin 6 (IL-6) and epidermal growth factor (EGF) through protein-protein interactions in the gene signaling pathway⁶. By analyzing four lung cancer gene chips, we found that the nuclear protein gene, human centromere protein F (CENPF), is highly expressed in lung adenocarcinoma (LUAD). Furthermore, the expression of CENPF and ER β 2/5 in the LUAD patients showed association with TNM staging, providing a basis for exploring the interaction between CENPF and ER β 2/5.

Among them, CENPF protein, discovered by Rattner et al., is involved in mitosis and infinite proliferation of tumor cells⁷. The full-length molecular weight of CENPF is 367 KDa and contains 3,210 amino acids⁸. It is involved in a variety of cancer diseases, especially in the development of prostate cancer, predicting the survival rate and tumor metastasis⁹. CENPF is directly associated with the disease outcomes after undergoing gene amplification¹⁰. However, the role of CENPF in the progression of LUAD still remained unclear.

In addition, a large number of studies including ours have shown that E2 promotes the progression of lung cancer by binding to nuclear ER¹¹. Our previous

related studies have found that only ER β 1/2/5 are expressed in the lung cancer tissues among five subtypes^{12, 13}. As a full-length fragment of ER β subtype, ER β 1 forms the basis for the action of other subtypes¹⁴, and so the role of ER β 2/5 in the progression of LUAD remained the focus of our current research.

Based on estrogen gene signaling pathway, a high co-expression of CENPF and ER β 2/5 that is associated with the clinicopathological features and prognosis of LUAD patients. CENPF is probably one of the key nuclear proteins in estrogen gene signaling pathway. Therefore, this study mainly explored the relationship between CENPF and ER β 2/5, and also explored their role alone in the development of LUAD. This study likely provides a better understanding of ER gene signaling pathway and improves the prognosis of LUAD patients.

2. Materials and Methods

2.1 Tissue specimens of patient and cell culture

This study was approved by the Ethics Review Committee of Wuhan University (no. 2018096). Fifty-six cases of LUAD and ten cases of benign pulmonary lesions (BPL) tissue specimens who underwent surgery from April 2014 to July 2017 for tissue chip, one pair of LUAD and paracancerous tissue for RNA-Seq and three pairs of tissues for profiling protein were obtained from the Department of Thoracic and Cardiovascular Surgery, Zhongnan Hospital of Wuhan University. The tissue chip was customized by Shanghai Core Biotech Co., Ltd.¹¹. The isolated tissue samples were immediately stored in liquid nitrogen and sent to Huada Gene (Beijing) for RNA-Seq

analysis⁴.

Human LUAD cell lines (A549, H1975, H1299 and PC-9) cultured in RPMI-1640 medium and normal lung bronchial cells BEAS-2B cultured in DMEM medium were purchased from the Chinese Academy of Sciences cell bank. The medium contained 10% fetal bovine serum (FBS) and double antibody (Gibco, 15140-122).

2.2 Lung cancer patient data set

Training data sets (GSE19804, GSE30219, GSE32863, GSE63459) based on the Affymetrix platform (Affymetrix HG-U133 Plus 2.0 array and HG-U133A array) and corresponding clinical information were retrieved from the Gene Expression Omnibus. Two NSCLC genome-wide expression profiles were extracted from GSE19804 (including 60 paired tumors and normal tissues) and GSE30219 (including 293 tumors and 14 non-tumor tissues). Two LUAD genome-wide expression profiles were extracted from GSE32863 (including 58 paired tumor and normal tissues) and GSE63459 (including 65 tumor tissues).

2.3 Analysis and verification of hub gene

The data sets of GSE19804, GSE30219, GSE32863 and GSE63459 were used to construct co-expression networks and clinical function related modules. The criteria for screening genes was performed according to the false discovery rate (FDR) <0.05 and $|\log_2 \text{fold change (FC)}| > 1.5$. Next, a weighted gene co-expression network analysis (WGCNA) package was used to construct a co-expression network^{15 16}.

Finally, the hub gene was selected based on the degree of centrality using the Venn diagram to obtain key genes.

The raw data of RNA-Seq was subjected to quality control, and then mapped with STAR¹⁷ to obtain differential genes. The screening criteria for differential genes were $\text{abs}(\log_2\text{FC}) > 1$ and $p \text{ value} < 0.05$.

The Oncomine, Gene Expression Profiling Interactive Analysis (GEPIA) and clinical data from The Cancer Genome Atlas (TCGA) database were used to verify the expression, progression and prognosis of hub gene.

2.4 Immunohistochemistry

Previously published articles included detailed steps for immunohistochemistry¹¹. CENPF (ab5) and ER β (ab3576) were purchased from Abcam. ER β 1 (MCA1974ST), ER β 2 (MCA2279GT) and ER β 5 (MCA4676T) were purchased from AbD Serotec. The specificity of the above antibodies were confirmed by numerous laboratories including ours^{11, 18}. Immunohistochemical method to analyze the optical density was calculated by Image-Pro Plus software.

2.5 Western blotting

Detailed western blotting analysis has been described previously¹⁹. E-cadherin (3195), N-cadherin (13116), MMP2 (40994), CDK2 (2546), CDK4 (12790) and β -actin (4970) were purchased from Cell Signaling Technology. CCND1 (60186-1-Ig) and GADPH (1E6D9) were obtained from Proteintech. The specificity of the above

antibodies was verified by numerous laboratories including ours^{11, 20, 21}.

2.6 Reverse transcription and quantitative real-time PCR (RT-qPCR)

Specific experimental methods were found in our previous published article⁶. Primers were designed based on CENPF mRNA sequence in GenBank. The primers used were as follows: CENPF, 3'-CTCTCCCGTCAACAGCGTTC; CENPF, 5'-GTTGTGCATATTCTTGGCTTGC. Data was analyzed using $2^{-\Delta\Delta C_t}$ method.

2.7 Cell culture experiment

The sense sequence of CENPF knockdown (KD) and negative control (NC) were integrated into the pWSLV-sh08-GFP vector for transfecting of A549 and H1299 cells. The stable CENPF-deficient A549 and H1299 cells were immunofluorescently labeled with anti-CENPF and anti-ER β ²². Cell apoptosis assay was measured using TdT-mediated dUTP-biotin nick end labeling test (TUNEL, Roche Applied Science, Germany) according to the manufacturer's instructions²³. MTT²⁴, invasion, migration, scratch and cell cycle experiments were used to evaluate the effect of LV-CENPF sh and LV-NC on the biological function of A549 and H1299 cells^{25, 26}.

2.8 Xenograft mouse model

Male nude mice were obtained from Beijing HFK Bioscience Co., Ltd., China. Mice were housed in specific-pathogen free (SPF) environment for one week, and then were subcutaneously injected with 100 μ L of 8×10^6 LV-CENPF sh or LV-NC cells. Tumor size was measured with an interval of three days (tumor size = length \times

width $^2 \times 0.5 \text{ mm}^3$). When the tumor size has reached 80-120 mm³, the mice were injected with E2 (0.036 mg/ml, purity 98%, Sigma) or Fulvestrant (Ful, 0.800 mg/ml, Sigma) subcutaneously twice a week (6 weeks)¹¹. All mice were then sacrificed on day 45, wherein the xenograft tumors were harvested and the tumor weight and size were measured²⁷. Finally, the tumor tissues were divided into two parts, one part was fixed with 4% paraformaldehyde for undergoing immunohistochemistry, and another part was quickly frozen with liquid nitrogen and stored at -80°C to perform western blotting.

2.9 Data Analysis

Data are expressed as mean \pm SD. All analyses were performed at least thrice and the representative data were obtained from three independent experiments. Two-tailed Student's t-test was used to assess significance differences between the groups. The effect of LV-CENPF sh and LV-NC on the biological function of A549 and H1299 cells, the expression of key signaling molecules, immunohistochemistry results in tissue microarray and *in vivo* experiments were analyzed by one-way analysis of variance. Statistical analysis was performed using SPSS 22.0 software (SPSS Inc., Chicago, IL). $P < 0.05$ was considered to be statistically significant.

3 Results

3.1 Bioinformatics analysis of lung cancer datasets and the determination of CENPF

Differential genes with similar expression pattern in each dataset (GSE19804,

GSE30219, GSE32863, GSE63459) were used (Figure 1A-C, Figure S2A-I, $\beta=5, 6, 5, 7$)²⁸. The four datasets obtained the module genes that are significantly associated with TNM staging of lung cancer, which included the brown module of GSE19804 (n=185), the green module of GSE30219 (n=413), the yellow module of GSE32863 (n=63) and the yellow module of GSE63459 (n=160) (Figure 1B; Figure S2B, E, H; Table S1). By overlapping the module genes, five key genes including CENPF, CDC20, TOP2A, CCNB2 and BIRC5 were obtained (Figure 1D-G). Finally, based on the central degree of the five key genes in different datasets and relevant literature^{9, 29}, the hub gene was identified as CENPF.

3.2 CENPF is highly expressed in LUAD and negatively correlated with the prognosis of LUAD patients

The results revealed that CENPF was highly expressed in LUAD when compared to the normal lung tissues (Figure 2A-H). The expression of CENPF showed a positive correlation with TNM staging of LUAD ($P<0.01$, Figure 2I-M). Additionally, high expression of CENPF was negatively correlated with overall survival and disease-free survival in patients with LUAD ($P=0.01, 0.027, 0.0267$, Figure 2N-P). The resulting RNA-Seq included 1515 high expressed genes and 1370 low expressed genes in LUAD patients (Figure 2Q). CENPF was included in high expressed genes (Figure 2R, $P < 0.05$) and was associated with poor prognosis of LUAD patients (Figure 2S, T).

3.3 Expression of CENPF, ER β , ER β 1, ER β 2 and ER β 5, and its

relationship with TNM staging and prognosis of LUAD patients

The expression of CENPF, ER β , ER β 1, ER β 2 and ER β 5 in different TNM stages of LUAD and benign primary lesions (BPL) was examined (Figure S1A, Figure S3A-D). The results revealed that CENPF, ER β , ER β 2 and ER β 5 were highly expressed in LUAD and showed a positive correlation with TNM staging and T grade of LUAD patients ($P < 0.001$, Figure S1B, C), but showed no association with N grade (Figure S1D). Meanwhile, high expression of CENPF was significantly associated with shorter survival by analyzing the high expression of CENPF in LUAD patients (Table 1).

3.4 Knockdown CENPF inhibits the biological effects of LUAD cells and tumor growth *in vivo*

Compared with BEAS-2B and other LUAD cells, CENPF showed high expression in A549 and H1299 cells ($P < 0.05$, Figure 3A; Figure S4A, B). RT-qPCR and cellular immunofluorescence confirmed that the efficiency of CENPF knockdown was above 70% in A549 and H1299 cells (Figure 3B; Figure S4C). Cell proliferation of CENPF-KD group was significantly weaker than NC group from day 3 in stable CENPF-deficient cell lines A549 and H1299 ($P = 0.007$, 0.000 , Figure 3C, D). At the same time, cells in CENPF-KD group demonstrated less Ki67 staining than NC group (Figure S4D). In addition, the invasion and migration of A549 cells in the CENPF-KD group were significantly decreased ($P = 0.000$, 0.000 , Figure 3E; Figure S4E), the related protein E-cadherin was significantly increased ($P = 0.009$, Figure 3F; Figure

S4G), and N-cadherin and MMP2 were significantly decreased when compared with NC group ($P=0.004$, 0.012 , Figure 3G; Figure S4H). In the stable CENPF-deficient cell line H1299, a similar trend was obtained ($P<0.01$, Figure S4E-H). Scratch experiment also demonstrated similar results ($P=0.000$, 0.000 , Figure 3H; Figure S4I). In A549 cells, cell percentage and DNA content were significantly increased in the G1 phase in the CENPF-KD group when compared with NC group ($P=0.011$, Figure 3I; Figure S4J). At the same time, the expression of CCND1, CDK2 and CDK4 was significantly lowered in CENPF-KD group ($P=0.022$, 0.001 , 0.002 , Figure 3J; Figure S4M). In stable CENPF-deficient cell line H1299, similar results were obtained ($P<0.05$, Figure S4K-M). The CENPF-KD group also showed a significant increase in apoptosis when compared with NC group ($P=0.001$, 0.001 , Figure 3K; Figure S4N).

The tumor weight and size in CENPF-KD group were lower than NC group after sacrificing in *in vivo* experiments ($P=0.001$, 0.001 , Figure 3L-N). The results of immunohistochemistry also showed that the expression of CENPF and ER β 2/5 was significantly decreased in CENPF-KD group when compared with NC group ($P=0.000$, 0.000 , 0.000 , Figure S4O, P).

3.5 Knockdown CENPF inhibits biological effects of LUAD cells mediated by ER β pathway and tumor growth *in vivo*

CENPF and ER β were co-localized in the nucleus of LUAD cells (Figure 4A). To investigate the effect of CENPF knockdown on the biological effects of LUAD cells mediated by ER β pathway, the cells were divided into the following groups: NC

group, CENPF-NC+E2 group, CENPF-NC+Ful group, CENPF-KD+E2 group and CENPF-KD+Ful group. Cell proliferation in CENPF-KD+Ful group was significantly lower than CENPF-NC+Ful group at 48 hours ($P=0.000$, 0.000 , Figure 4B). In A549 cells, invasion and migration of cells in CENPF-KD+E2 group were significantly reduced when compared with CENPF-NC+E2 group ($P=0.000$, 0.000 , Figure 4C; Figure S5A). Similarly, the expression of MMP2 and N-cadherin were significantly decreased in CENPF-KD+E2 group when compared with the CENPF-NC+E2 group ($P=0.002$, 0.016 , Figure 4E, F; Figure S5B, D). Similar results were obtained in stable CENPF-deficient cell line H1299 ($P<0.01$, Figure 4D-F, Figure S5A, C, D). Scratch experiment also showed that the migration of CENPF-KD+E2 group was significantly lower than CENPF-NC+E2 group ($P=0.000$, 0.000 , Figure 4G; Figure S5E). Flow cytometry analysis showed that cell percentage in CENPF-KD+E2 group was significantly reduced in A549 and H1299 cells in G2/M phase than CENPF-NC+E2 group ($P=0.001$, 0.021 , Figure 4H, I; Figure S5F, G). At the same time, the protein expression of CCND1, CDK2 and CDK4 in CENPF-KD+E2 group demonstrated a significant decrease ($P=0.003$, 0.008 , 0.006 , $P=0.043$, 0.004 , 0.005 , Figure 4J, Figure S5H).

In the lung cancer model of nude mice, tumor size and weight of mice in the CENPF-KD+Ful group were significantly lowered than CENPF-NC+Ful group ($P=0.001$, 0.039 , Figure 5A-C; Figure S6E). Immunohistochemistry of mice tumor tissues showed that the expression of CENPF and ER β 2/5 in CENPF-KD+E2 group was significantly lower than that in CENPF-NC+E2 group ($P=0.000$, 0.000 , 0.000 ,

Figure 5D, E).

3.6 Knockdown of CENPF inhibits the expression of ERβ2/5 both *in vitro* and *in vivo*

The effect of CENPF knockdown on the expression of ERβ2/5 was examined in CENPF-NC and CENPF-KD groups both *in vitro* and *in vivo*. The protein expression of ERβ2/5 in CENPF-KD group was significantly lower than that in NC group ($P_{A549}=0.013, 0.000$; $P_{H1299}=0.006, 0.002$, Figure 5F; Figure S6A). The protein expression of CENPF and ERβ2/5 in CENPF-KD group was lower than NC group *in vivo* ($P=0.024, 0.020, 0.003$, Figure 5G; Figure S6B).

The effect of CENPF knockdown on the expression ERβ2/5 under the action of E2 and Ful was examined both *in vitro* and *in vivo*. The results revealed that the protein expression of CENPF and ERβ2/5 in CENPF-KD+Ful group was significantly lower than CENPF-NC+Ful group ($P=0.002, 0.004, 0.001$, Figure 5H; Figure S6C) in A549 cells. Similar results were obtained in stable CENPF-deficient cell line H1299 ($P < 0.01$, Figure 5H; Figure S6C). The protein expression of CENPF and ERβ2/5 was significantly decreased in CENPF-KD+Ful group when compared with CENPF-NC+Ful group in mice tumor tissues ($P=0.020, 0.004, 0.002$, Figure 5I; Figure S6D).

4 Discussion

Lung cancer remains one of the most common malignancies, posing a serious health problem worldwide. Targeted therapy for cell proliferation-related pathways

and hormone therapy for lung cancer are considered to be important treatment strategies³⁰. ER signaling pathways mainly included gene signaling pathways and non-gene signaling pathways³¹. Among them, the gene signaling pathway involves integration of estrogen with ER in the nucleus and then correlates with the estrogen response element (ERE) of the target gene, finally forming a complex to promote the recruitment of RNA polymerase II on the promoter, and regulating the transcription of genes. During this process, different combinations of synergistic activators determine the specificity of ER for the activation of target genes⁵. Therefore, the key genes involved in TNM staging of LUAD in the four lung cancer datasets (GSE19804, GSE30219, GSE32863, GSE63459) were analyzed by the "WGCNA" R package (Figure 1A-C; Figure S2, Table S1). Finally, CENPF was found by overlapping the key genes (Figure 1D-G).

High expression of CENPF is associated with malignant progression and poor prognosis in patients with LUAD. Studies have shown that CENPF has been up-regulated in a variety of malignancies, including nasopharyngeal carcinoma, esophageal squamous cell carcinoma and prostate cancer^{20, 32, 33}. The expression of CENPF in LUAD was detected by the Oncomine database and RNA-Seq data, and showed high expression of CENPF in LUAD (Figure 2A-F, Q, R). This result was further confirmed by analyzing the expression of CENPF according to LUAD staging in four data sets and in *in vitro* experiment (Figure 2I-L; Figure 5F, G), indicating that high expression of CENPF in LUAD might be related with its occurrence. At the same time, studies have showed that CENPF is involved in protein-mediated mitosis

and cell proliferation³⁴. The expression of CENPF and DNA content were significantly reduced, (Figure 3I), and the cell proliferation was significantly decreased in stably CENPF-deficient LUAD cells (Figure 3C, D). These results indicated that low expression of CENPF inhibited proliferation of tumor cells in LUAD. Other studies showed that forkhead box M1 (FOXM1) and CENPF synergistically promoted malignant progression and poor prognosis of prostate cancer⁹. It has been reported that chicken ovalbumin upstream promoter-transcription factor 2 (COUP-TFII) promoted metastasis of prostate cancer through signal transduction of FOXM1 and CENPF²⁰. Therefore, we speculated that dysregulation of miRNA-COUP-TFII-FOXM1-CENPF axis might also be associated with malignant progression, poor prognosis and metastasis of LUAD. Our study revealed that LUAD cells showed a significant reduction in invasion and metastasis after CENPF knockdown (Figure 3E-H), and the expression of CENPF was significantly related to TNM staging of LUAD (Figure 2I-L; Figure S1B, C). In addition, based on the clinical data from LUAD patients in GEPIA and TCGA, LUAD patients with high expression of CENPF had a poor prognosis (Figure 2N-P). These results indicated that abnormal expression of CENPF was significantly associated with TNM staging, poor prognosis, and malignant metastasis of LUAD. Although the characteristics and tissue sources of prostate cancer and LUAD are different, CENPF might play an important role in the development of LUAD through a similar mechanism.

In addition, according to our previous study, E2 promoted the progression of lung cancer by binding to ER β ³⁵. Significant differences in the biological functions were

caused due to different subtypes of ER β in E2 target organ tumors such as breast cancer, cervical cancer and prostate cancer^{13, 36}. Our previous findings indicated that ER β 2/5 are expressed in lung cancer¹². Our study results revealed that ER β 2/5 are highly expressed in LUAD patients (Figure S3B, D), and high expression of ER β 2/5 was positively associated with TNM staging of LUAD (Figure S1B, C). It can be observed that ER β 2/5 also plays an important role in the occurrence of LUAD.

CENPF knockdown inhibited the progression of LUAD mediated by ER β 2/5 pathway. High expression of CENPF and ER β 2/5 are associated with the development of LUAD, and so we focused on the relationship between CENPF and ER β 2/5 in patients with LUAD. Invasion, migration and proliferation of LUAD cells in CENPF-KD+E2 group showed significant reduction *in vitro* when compared to CENPF-NC+E2 group (Figure 4C-I, Figure S5A-G). The protein expression of ER β 2/5 in CENPF-KD+E2 group was significantly lower *in vivo* than CENPF-NC+E2 group (Figure 5H, I; Figure S6C, D). From these results, we confirmed that the knockdown of CENPF inhibited the progression of LUAD mediated by ER β 2/5 pathway both *in vitro* and *in vivo*. This is another important point regarding the mechanism of CENPF, except for FOXM1⁹ and COUP-TFII²⁰.

Taken together, these findings indicated that both CENPF and ER β 2/5 are highly expressed in LUAD cells and their expression is associated with TNM staging and prognosis of LUAD patients. CENPF knockdown inhibited proliferation, invasion and metastasis of LUAD cells mediated by ER β 2/5 pathway. Thus, the development of inhibitors against ER β 2/5 subtype and CENPF remained more effective in improving

the therapeutic effect of LUAD. Of course, there are some shortcomings in this study. Firstly, as the molecular weight of CENPF is too large and plasmid construction is easy to off target, no LUAD cell lines with high expression of CENPF can be constructed. Secondly, the number of specimens used for RNA-Seq is limited, as only specimens were used to validate the results of big database. These issues are key points that requires focus in our future research.

Acknowledgements

This project was supported by the Program of Excellent Doctoral (Postdoctoral) of Zhongnan Hospital of Wuhan University (Grant No. ZNYB2019002). The authors thank MedSci staff for her great help in providing language help, and thank all authors for their help in writing assistance and proof reading the article.

Disclosure Statement: The authors declare no potential conflicts of interest.

Authors are accountable for all aspects of the work (including full data access, integrity of the data and the accuracy of the data analysis). And we declare that the views expressed in the submitted article are our own and not an official position of the institution or funder.

References

1. Funahashi S, Okazaki Y, Nagai H, et al. Twist1 was detected in mesenchymal 1
Funahashi S, Okazaki Y, Nagai H, et al. Twist1 was detected in mesenchymal
cells of mammary fibroadenoma and invasive components of breast carcinoma in rats.
Journal of toxicologic pathology. 2019; 32: 19-26.

- 2 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*. 2018; 68: 394-424.
- 3 Chang GC, Tseng CH, Hsu KH, et al. Predictive factors for EGFR-tyrosine kinase inhibitor retreatment in patients with EGFR-mutated non-small-cell lung cancer - A multicenter retrospective SEQUENCE study. *Lung Cancer*. 2017; 104: 58-64.
- 4 Anraku M, Waddell TK, de Perrot M, et al. Induction chemoradiotherapy facilitates radical resection of T4 non-small cell lung cancer invading the spine. *J Thorac Cardiovasc Surg*. 2009; 137: 441-447.e441.
- 5 Ciuleanu TE, Ahmed S, Kim JH, et al. Randomised Phase 2 study of maintenance linsitinib (OSI-906) in combination with erlotinib compared with placebo plus erlotinib after platinum-based chemotherapy in patients with advanced non-small cell lung cancer. *British Journal of Cancer*. 2017; 117: 757-766.
- 6 Bai Y, Shen W, Zhu M, et al. Combined detection of estrogen and tumor markers is an important reference factor in the diagnosis and prognosis of lung cancer. *Journal of cellular biochemistry*. 2019; 120: 105-114.
- 7 Rattner JB, Rao A, Fritzler MJ, Valencia DW, Yen TJ. CENP-F is a .ca 400 kDa kinetochore protein that exhibits a cell-cycle dependent localization. *Cell motility and the cytoskeleton*. 1993; 26: 214-226.
- 8 Yerushalmi R, Woods R, Kennecke H, Speers C, Knowling M, Gelmon K. Patterns of relapse in breast cancer: changes over time. *Breast cancer research and treatment*. 2010; 120: 753-759.

- 9 Aytes A, Mitrofanova A, Lefebvre C, et al. Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer cell*. 2014; 25: 638-651.
- 10 Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A. Reverse engineering cellular networks. *Nature protocols*. 2006; 1: 662-671.
- 11 Antoon JW, Martin EC, Lai R, et al. MEK5/ERK5 signaling suppresses estrogen receptor expression and promotes hormone-independent tumorigenesis. *PLoS One*. 2013; 8: e69291.
- 12 Butts C, Socinski MA, Mitchell PL, et al. Tecemotide (L-BLP25) versus placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. *The Lancet Oncology*. 2014; 15: 59-68.
- 13 Leung YK, Lam HM, Wu S, et al. Estrogen receptor beta2 and beta5 are associated with poor prognosis in prostate cancer, and promote cancer cell migration and invasion. *Endocrine-related cancer*. 2010; 17: 675-689.
- 14 Shaaban AM, Green AR, Karthik S, et al. Nuclear and cytoplasmic expression of ERbeta1, ERbeta2, and ERbeta5 identifies distinct prognostic outcome for breast cancer patients. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2008; 14: 5228-5235.
- 15 Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS computational biology*. 2008; 4: e1000117.
- 16 Mason MJ, Fan G, Plath K, Zhou Q, Horvath S. Signed weighted gene co-expression network analysis of transcriptional regulation in murine embryonic

stem cells. BMC Genomics. 2009; 10: 327.

17 Dobin A, Gingeras TR. Optimizing RNA-Seq Mapping with STAR. Methods in molecular biology (Clifton, NJ). 2016; 1415: 245-262.

18 Gong Y, Xu Z, Jin C, et al. Treatment of Advanced Non-small-Cell Lung Cancer with Qi-Nourishing Essence-Replenishing Chinese Herbal Medicine Combined with Chemotherapy. Biol Proced Online. 2018; 20.

19 Antonia SJ, Villegas A, Daniel D, et al. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. New England Journal of Medicine. 2017; 377: 1919-1929.

20 Abe T, Shirai K, Saitoh J, et al. Incidence, risk factors, and dose-volume relationship of radiation-induced rib fracture after carbon ion radiotherapy for lung cancer. Acta oncologica (Stockholm, Sweden). 2016; 55: 163-166.

21 Jing X, Peng J, Dou Y, et al. Macrophage ERalpha promoted invasion of endometrial cancer cell by mTOR/KIF5B -mediated epithelial to mesenchymal transition. Immunology and cell biology. 2019.

22 Taylor CE, Heimer GV, Lea DJ, Tomlinson AJ. A COMPARISON OF A FLUORESCENT ANTIBODY TECHNIQUE WITH A CULTURAL METHOD IN THE DETECTION OF INFECTIONS WITH SHIGELLA SONNEI. Journal of clinical pathology. 1964; 17: 225-230.

23 Kyrylkova K, Kyryachenko S, Leid M, Kioussi C. Detection of apoptosis by TUNEL assay. Methods in molecular biology (Clifton, NJ). 2012; 887: 41-47.

24 Weiss-Steider B, Cordova Y, Aguiniga-Sanchez I, Ledesma-Martinez E,

Dominguez-Melendez V, Santiago-Osorio E. Sodium caseinate and alpha-casein inhibit proliferation of the mouse myeloid cell line 32D cl3 via TNF-alpha.

Biomedica. 2019; 39.

25 Agostini P, Naidu B, Cieslik H, et al. Effectiveness of incentive spirometry in patients following thoracotomy and lung resection including those at high risk for developing pulmonary complications. *Thorax*. 2013; 68: 580-585.

26 Sabran A, Kumolosasi E, Jantan I. Effects of annexin A1 on apoptosis and cell cycle arrest in human leukemic cell lines. *Acta pharmaceutica (Zagreb, Croatia)*. 2019; 69: 75-86.

27 Guilmain W, Colin S, Legrand E, et al. CD9P-1 expression correlates with the metastatic status of lung cancer, and a truncated form of CD9P-1, GS-168AT2, inhibits in vivo tumour growth. *Br J Cancer*. 2011; 104: 496-504.

28 Ayoub NM, Al-Shami KM, Yaghan RJ. Immunotherapy for HER2-positive breast cancer: recent advances and combination therapeutic approaches. *Breast cancer (Dove Medical Press)*. 2019; 11: 53-69.

29 An C, Zhang J, Chu H, et al. Study of Gefitinib and Pemetrexed as First-Line Treatment in Patients with Advanced Non-Small Cell Lung Cancer Harboring EGFR Mutation. *Pathology and Oncology Research*. 2016; 22: 763-768.

30 Bao P, Zhao W, Li Y, Liu Y, Zhou Y, Liu C. Protective effect of ulinastatin in patients with non-small cell lung cancer after radiation therapy: a randomized, placebo-controlled study. *Medical oncology (Northwood, London, England)*. 2015; 32: 405.

- 31 Niikawa H, Suzuki T, Miki Y, et al. Intratumoral estrogens and estrogen receptors in human non-small cell lung carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2008; 14: 4417-4426.
- 32 Pimkhaokham A, Shimada Y, Fukuda Y, et al. Nonrandom chromosomal imbalances in esophageal squamous cell carcinoma cell lines: possible involvement of the ATF3 and CENPF genes in the 1q32 amplicon. *Japanese journal of cancer research : Gann*. 2000; 91: 1126-1133.
- 33 Cheng H, Sun N, Sun X, et al. Polymorphisms in hMSH2 and hMLH1 and response to platinum-based chemotherapy in advanced non-small-cell lung cancer patients. *Acta Biochimica et Biophysica Sinica*. 2010; 42: 311-317.
- 34 Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: an evidence-based and economic analysis. *Ontario health technology assessment series*. 2010; 10: 1-57.
- 35 Gomez DR, Blumenschein GR, Lee JJ, et al. Local consolidative therapy versus maintenance therapy or observation for patients with oligometastatic non-small-cell lung cancer without progression after first-line systemic therapy: a multicentre, randomised, controlled, phase 2 study. *The Lancet Oncology*. 2016; 17: 1672-1682.
- 36 Cagney DN, Thirion PG, Dunne MT, et al. A Phase II Toxicity End Point Trial (ICORG 99-09) of Accelerated Dose-escalated Hypofractionated Radiation in Non-small Cell Lung Cancer. *Clin Oncol (R Coll Radiol)*. 2018; 30: 30-38.

Table

Variables	case (n=56)	high	low	P value
Gender				0.389
Male	37	14	23	
Female	19	5	14	
Age(year)				0.942
>65	21	7	14	
≤65	35	12	23	
Tumor size(cm)				0.007
≥5	12	8	4	
<5	44	11	33	
Smoking				0.757
Smoking	22	8	14	
No-smoking	34	11	23	
TNM stage				0.019
I	9	0	9	
II-III	47	19	28	
Tgrade				0.044
T1-T3	54	17	37	
T4	2	2	0	
N grade				0.372
N0	40	15	25	
N1-N2	16	4	12	
Relapse				0.041
Relapse	5	4	1	
No-relapse	51	15	36	
Survival months		27.39±7.54	33.41±13.86	0.048

Table 1. CENPF expression in the lung adenocarcinoma. Bold numbers represent statistical significance.

Figures

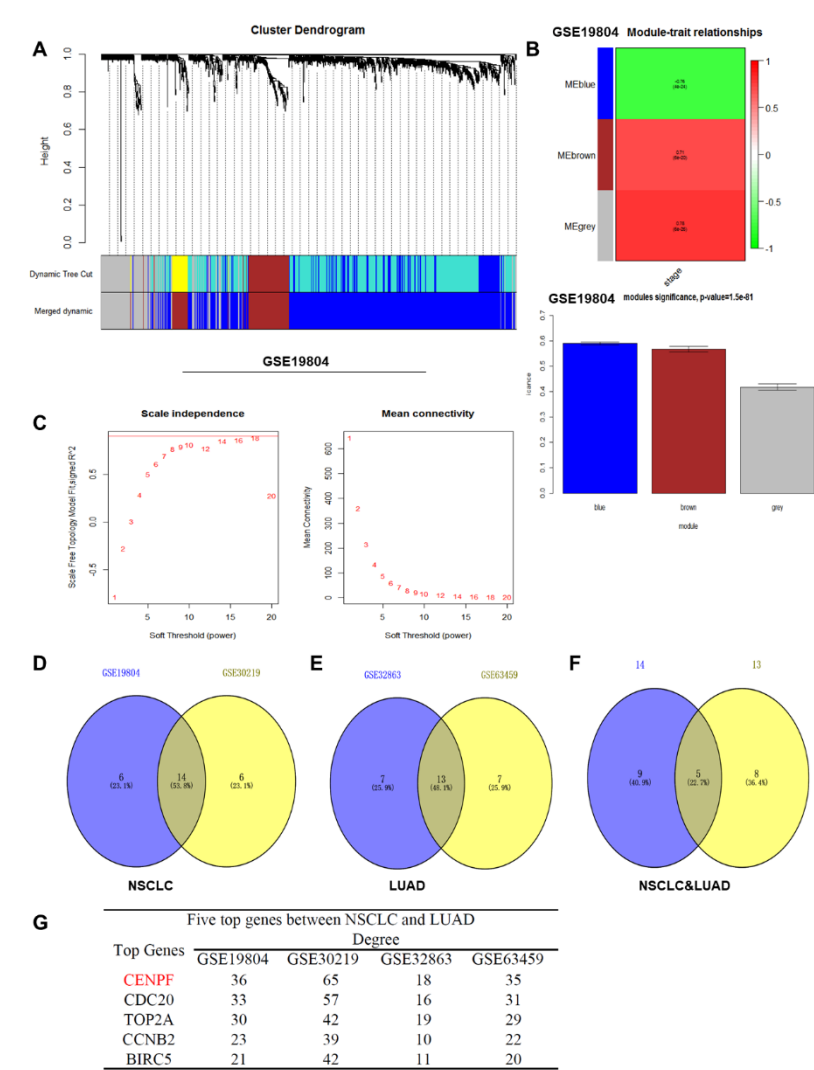


Figure 1. WGCNA analysis and determination of the CENPF gene.

(A)Dendrogram of all differentially expressed genes clustered based on a dissimilarity measure (1-TOM). (B) Heat maps and distribution of differential genes for different modules related to NSCLC staging (GSE19804). (C) Analysis of the scale-free fit index for various soft-thresholding power (β) and Analysis of the mean connectivity

for various soft-thresholding power (GSE19804). (D) There are 14 gene differential expressions in the NSCLC staging modules. (E) There are 13 gene differential expressions in the LUAD staging modules. (F) In the four datasets, there were 5 overlapping genes that were significantly differentially expressed between NSCLC and LUAD. (G) The degree values of the five key genes in different datasets.

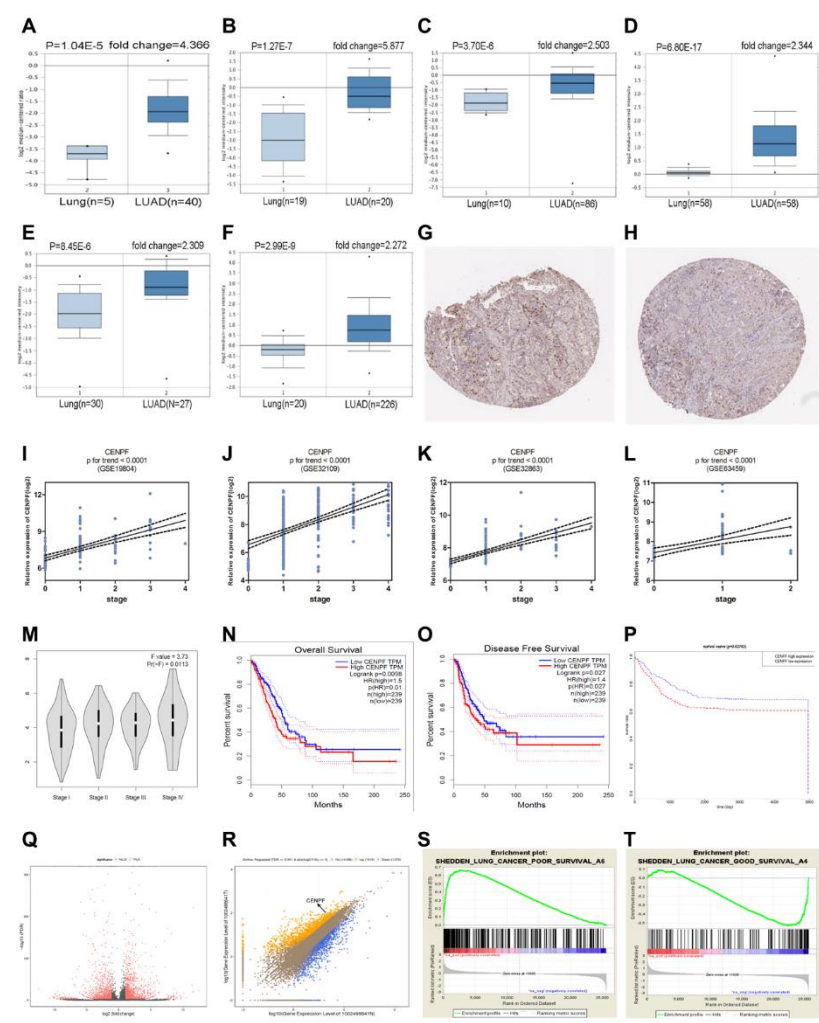


Figure 2. CENPF is upregulated in LUAD and is related with TNM staging and prognosis of LUAD patients. (A-F) Oncomine database results show that CENPF expression is significantly up-regulated in LUAD. (The corresponding P value and fold change are given above the picture). (G, H) The Human Protein Atlas database

indicates that CENPF is strongly expressed in LUAD. G: LUAD (patient ID.4923, male, 57); H: normal lung tissue (patient ID. 4208, male, 75). (I-L) Analyzes the relationship between CENPF and LUAD staging based on four datasets. (M) Verify the correlation between the expression of CENPF and the pathological stage of LUAD (based on TCGA data in GEPIA). N-P: Survival analysis. (N) Overall survival analysis. (O) Disease-free survival analysis. (P) Survival curves of CENPF gene in LUAD patients based on TCGA database. (Q) RNA sequencing analysis of volcano maps. (R) RNA sequencing results indicate that CENPF is highly expressed in LUAD tissues. Orange represents a high expression of the gene in LUAD, and blue represents a low expression of the gene in LUAD. (S, T) The CENPF gene was analyzed using Gene Set Enrichment Analysis (GSEA). The positive expression of the CENPF is related with a low prognosis in LUAD patients.

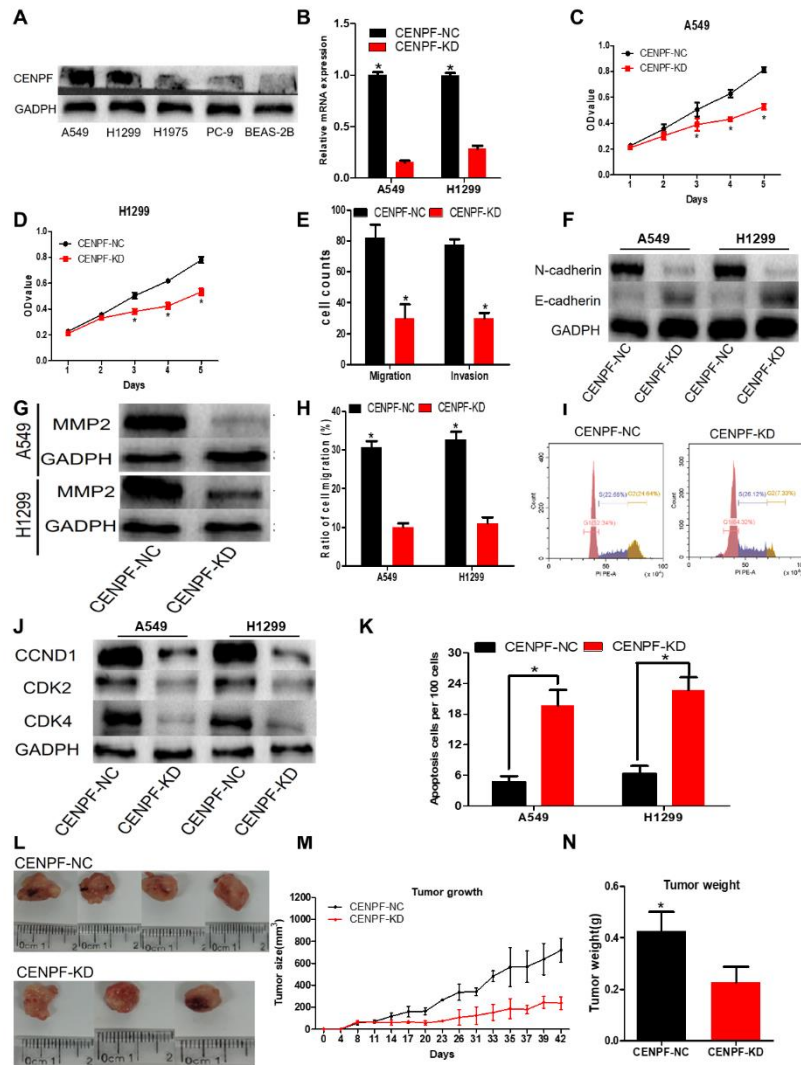


Figure 3. Knockdown of CENPF inhibits the biological effects of LUAD cells and the growth of LUAD in vivo. (A) Protein expression of CENPF in different cell lines. (B) The knockdown efficiency of LV-CENPF sh or LV-NC transfected with A549 and H1299 cells was verified by RT-qPCR. *P < 0.05 vs CENPF-KD. (C, D) MTT measures the proliferative viability of cells at different times. *P < 0.05 vs CENPF-KD (E) Quantified histograms of migration and invasion in A549 cells. (F, G) Protein expression of N-cadherin, E-cadherin and MMP2 in A549 and H1299 cells. (H) Quantified histograms of scratch experiment of A549 and H1299. (I) Percentage of CENPF-deficient cells A549 at different stages of the cell cycle (G1, S and G2/M).

(J) Expression of CCND1, CDK2 and CDK4 in A549 and H1299. (K) Quantified histograms of representative TUNEL staining of CENPF-deficient cell lines A549 and H1299. (L) Pictures of mice tumor tissues after dissection. (M, N) Analysis of tumor size and tumor weight. *P < 0.05.

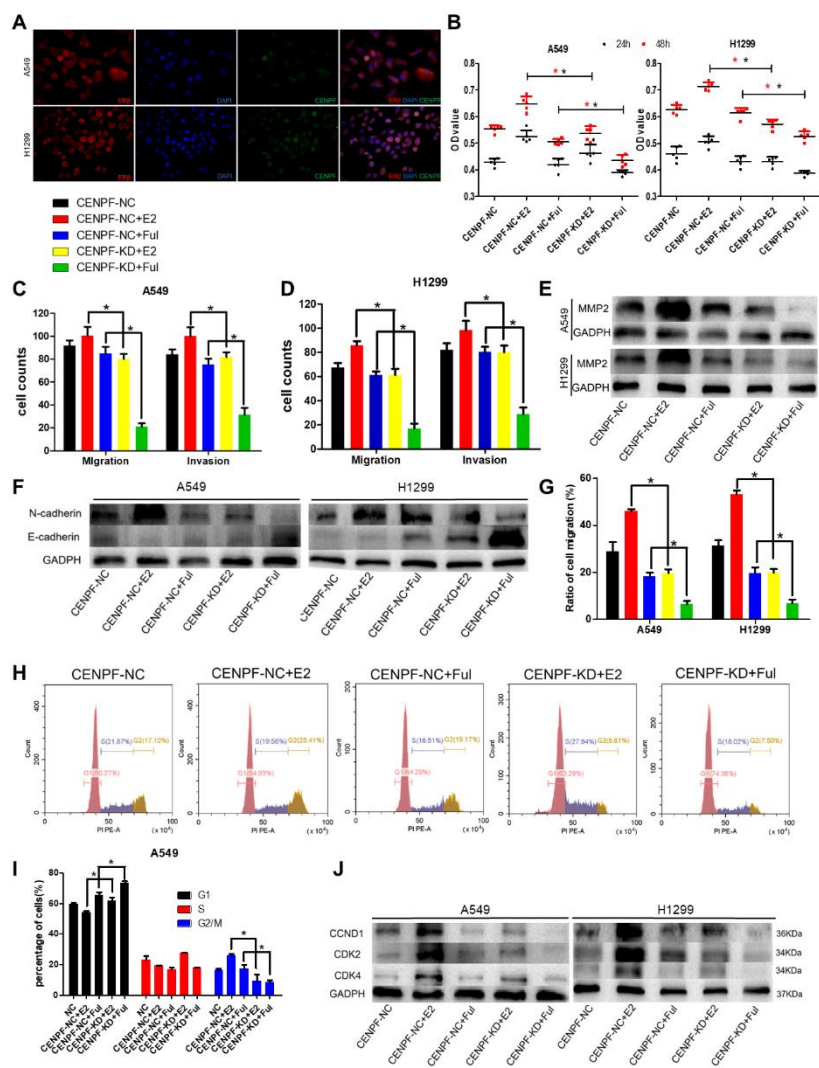


Figure 4. Knockdown of CENPF inhibits proliferation, invasion and migration of LUAD cells via the ERβ pathway. (A) Immunofluorescence showed the co-localization of CENPF and ERβ in A549 and H1299 cells (400 x). (B) Cell proliferation assays of differently grouped cells at specific times in A549 and H1299 cells. (C, D) Corresponding quantified histograms of migration and invasion in

A549 and H1299 cells. (E, F) Protein expression of MMP2, N-cadherin and E-cadherin in A549 and H1299 cells. (G) Corresponding quantified histograms of scratch experiment in A549 and H1299 cells. (H, I) Percentage of the A549 cells at different stages of the cell cycle (G1, S and G2/M) and corresponding quantified histograms. (J) Protein expression of CCND1, CDK2 and CDK4 in A549 and H1299 cells. *P < 0.05.

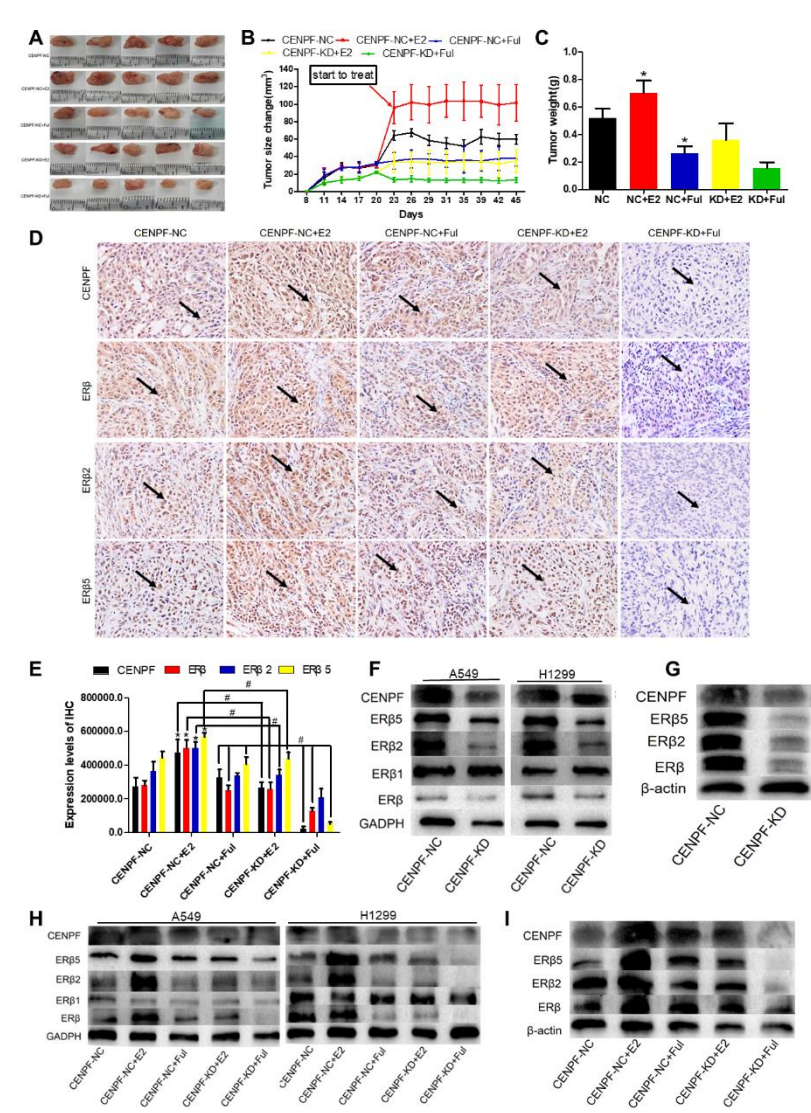


Figure 5. Knockdown of CENPF can inhibit ERβ2/5 pathway-mediated tumor tissue growth in vivo and in vitro. (A) Tumor images after dissection of nude mice.

(B, C) Statistical analysis of tumor size and tumor weight. (D, E) Immunohistochemical analysis of the expression of CENPF, ER β , ER β 2 and ER β 5 in nude mice tumor tissues and corresponding quantified histograms. (F, G) Protein expression of CENPF, ER β , ER β 1, ER β 2 and ER β 5 in vitro and in vivo experiment. (H, I) Protein expression of CENPF, ER β , ER β 1, ER β 2 and ER β 5 in vitro and in vivo experiment after treated with E2 and Ful treatment. *P < 0.05.