Research Paper

Knockdown of CENPF inhibits the progression of lung adenocarcinoma mediated by ERβ2/5 pathway

Tang Hexiao[1,*], Bai Yuquan[1,*], Xiong Lecai[1,*], Wei Yanhong[2], Shen Li[3], Hu Weidong[1], Xu Ming[1], Zhou Xuefeng[1], Pan Gaofeng[1], Zhang Li[1], Zhu Minglin[1], Tang Zheng[1], Yang Zetian[1], Zhou Xiao[1], Cai Yi[1], Michael Lanuti[3], Zhao Jinping[1]

[1]Department of Thoracic Surgery, Zhongnan Hospital of Wuhan University, Wuhan, China
[2]Department of Geriatrics, Zhongnan Hospital of Wuhan University, Wuhan, China
[3]Division of Thoracic Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA
*Co-first author

Correspondence to: Tang Hexiao; **email:** thx1245@sina.com
Keywords: centromere protein F (CENPF), estrogen receptor beta, lung adenocarcinoma (LUAD), WGCNA package, non-small cell lung cancer (NSCLC)
Received: May 9, 2020 **Accepted:** September 5, 2020 **Published:** January 10, 2021

ABSTRACT

Many studies have reported that estrogen (E2) promotes lung cancer by binding to nuclear estrogen receptors (ER), and altering ER related nuclear protein expressions. With the GEO database analysis, Human centromere protein F (CENPF) is highly expressed in lung adenocarcinoma (LUAD), and the co-expression of CENPF and ERβ was found in the nucleus of LUAD cells through immunofluorescence. We identified the nuclear protein CENPF and explored its relationship with the ER pathway. CENPF and ERβ2/5 were related with T stage and poor prognosis (P<0.05). CENPF knockout significantly inhibited LUAD cell growth, the tumor growth of mice and the expression of ERβ2/5 (P<0.05). The protein expression of CENPF and ERβ2/5 in the CENPF-Knockdown+Fulvestrant group was lower than CENPF- Negative Control +Fulvestrant group (P=0.002, 0.004, 0.001) in A549 cells. The tumor size and weight of the CENPF-Knockdown+Fulvestrant group were significantly lower than CENPF- Negative Control +Fulvestrant group (P=0.001, 0.039) in nude mice. All the results indicated that both CENPF and ERβ2/5 play important roles in the progression of LUAD, and knockdown CENPF can inhibit the progression of LUAD by inhibiting the expression of ER2/5. Thus, the development of inhibitors against ERβ2/5 and CENPF remained more effective in improving the therapeutic effect of LUAD.

INTRODUCTION

Lung cancer is one of the most commonly diagnosed cancers and a leading cause of cancer related mortality [1]. The global incidence and mortality of lung cancer is increasing significantly [2]. In addition to environmental exposures such as smoking, the growth factor pathway and hormonal regulation also play critical roles in the carcinogenesis of lung cancer [3].

Estrogen receptors (ERs) belong to the nuclear receptor steroid superfamily, and are closely linked to hormonal regulations. ERs are classified into two subtypes, ERα and ERβ, and these have different tissue distributions and biological effects in various tumor types [4]. Previous studies have revealed that estrogen activates the transcription of target genes by binding directly to ERα and ERβ [5]. Numerous studies have found that ER 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 [6]. Analysis of four lung cancer gene chips revealed that the nuclear protein gene, human centromere protein F (CENPF), is

highly expressed in lung adenocarcinoma (LUAD). Furthermore, the expression of CENPF and ERβ2/5 in LUAD patients have been shown to be correlated to TNM staging, providing a basis for exploring the interactions between CENPF and ERβ2/5. Rattner et al. demonstrated that CENPF is involved in mitosis and tumor proliferation [7]. The full-length molecular weight of CENPF is 367 KDa and contains 3,210 amino acids [8]. In prostate cancer, CENPF has been shown to predict survival and tumor metastasis [9]. CENPF is directly associated with disease outcomes after undergoing gene amplification [10]. However, the role of CENPF in the progression of LUAD still remained unclear.

A large number of studies have shown that estrogen (E2) promotes the progression of lung cancer by binding to nuclear ERs [11]. Our previous studies have shown that among the five types of ERs, lung cancer tissues express found the ERβ1/2/5 [12, 13]. As a full-length fragment of ERβ subtype, ERβ1 is responsible for the action of other subtypes [14], and so the role of ERβ2/5 in the progression of LUAD remained the main focus of our current research.

Based on estrogen gene signaling pathway, a high co-expression of CENPF and ERβ2/5 showed association with the clinicopathological features and prognosis of LUAD patients. CENPF is hypothesized as one of the key nuclear proteins in estrogen gene signaling pathway. Therefore, this study mainly explored the relationship between CENPF and ERβ2/5, and also explored their role alone in the development of LUAD. This study will provide a better understanding of ERs gene signaling pathway and improve the prognosis of LUAD patients.

RESULTS

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–1C, Supplementary Figure 1A–1I, β=5, 6, 5, 7) [15]. The module genes obtained by the four datasets showed significant association with TNM staging of lung cancer, which included the brown module of GSE19804 (n=185), the turquoise module of GSE30219 (n=413), the yellow module of GSE32863 (n=63) and the yellow module of GSE63459 (n=160) (Figure 1B; Supplementary Figure 1B, 1E, 1H; Supplementary Table 1). By overlapping the module genes, five key genes including CENPF, CDC20, TOP2A, CCNB2 and BIRC5 were obtained (Figure 1D–1G). Finally, based on the central degree of the five key genes in different datasets and relevant literature [9, 16], the hub gene was identified as CENPF.

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

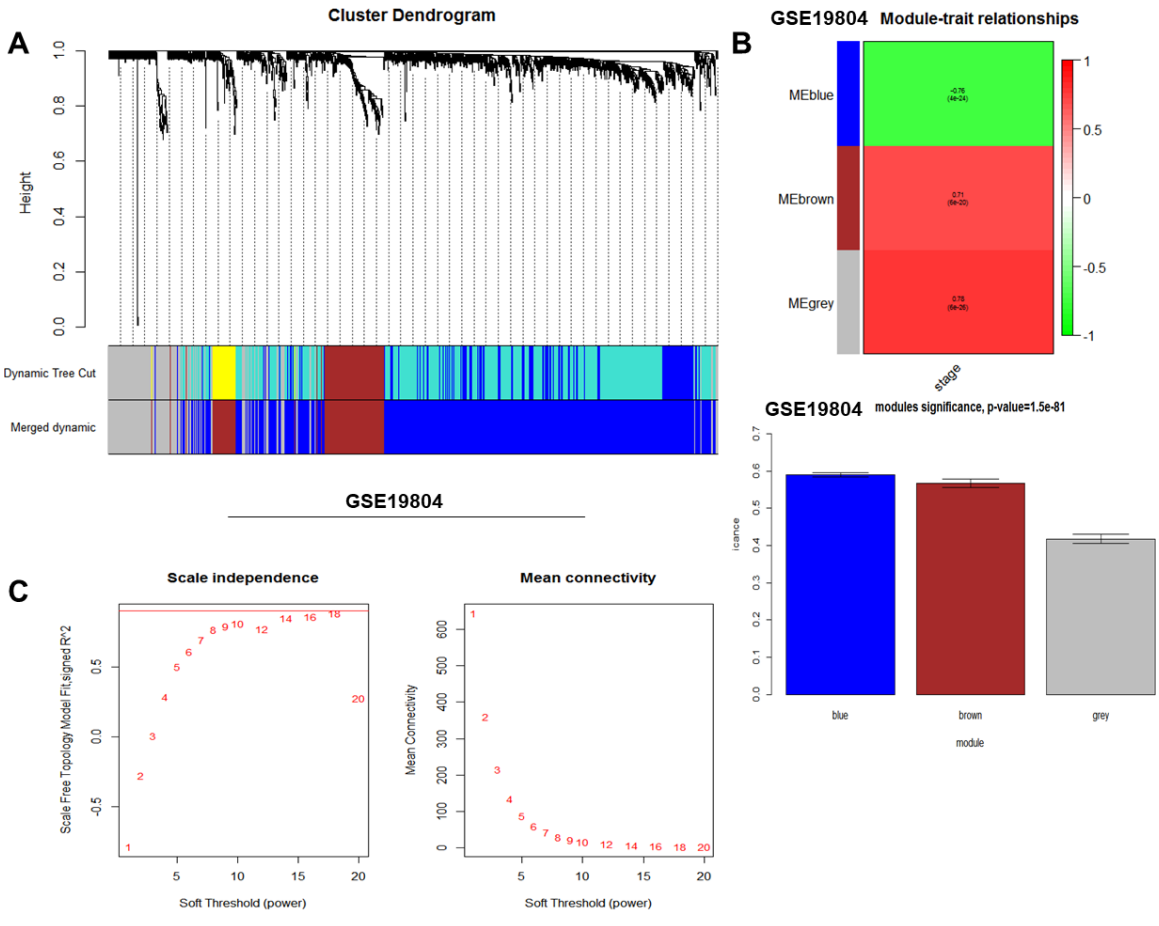
CENPF was highly expressed in LUAD when compared to normal lung tissues (Figure 2A–2H). The expression of CENPF was positively correlated to the TNM staging of LUAD (P<0.01, Figure 2I–2M). Additionally, high expression of CENPF was negatively correlated with overall survival and disease-free survival in LUAD patients (P=0.01, 0.027, 0.0267, Figure 2N–2P). The result of 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 in LUAD patients (Figure 2S, 2T).

Expression of CENPF, ERβ, ERβ1, ERβ2 and ERβ5, and its relationship with TNM staging and prognosis of LUAD patients

The expressions of CENPF, ERβ, ERβ1, ERβ2 and ERβ5 in different TNM stages of LUAD and benign primary lesions (BPL) were examined (Figure 3A, Supplementary Figure 2A–2D). The results revealed that CENPF, ERβ, ERβ2 and ERβ5 were highly expressed in LUAD and showed a positive correlation with TNM staging and T grade of LUAD patients (P<0.001, Figure 3B, 3C), but showed no association with nodal involvements (Figure 3D). Moreover, analysis of high expression of CENPF in LUAD patients showed significant association of CENPF with shorter survival rate (Table 1).

CENPF knockdown inhibits biological effects of LUAD cells

Compared with BEAS-2B and other LUAD cells, CENPF was highly expressed in A549 and H1299 cells (P<0.05, Figure 4A; The corresponding gray value are shown in Supplementary Figure 3A, 3B). RT-PCR and cellular immunofluorescence confirmed that CENPF knockdown (KD) was seen in 70% in A549 and H1299 cells (Figure 4B; Supplementary Figure 3C). Cell proliferation of CENPF-KD group was significantly weaker than the control (NC) group from day 3 in stable CENPF-deficient cell lines A549 and H1299 (P=0.007, 0.000, Figure 4C, 4D). At the same time, cells in CENPF-KD group demonstrated less Ki67 staining than NC group (Supplementary Figure 3D). In addition, the invasion and migration of A549 cells in the CENPF-KD group were significantly decreased (P=0.000, 0.000, Figure 4E; Supplementary Figure 3E, 3F) while



G

Five top genes between NSCLC and LUAD

Top Genes	Degree			
	GSE19804	GSE30219	GSE32863	GSE63459
CENPF	36	65	18	35
CDC20	33	57	16	31
TOP2A	30	42	19	29
CCNB2	23	39	10	22
BIRC5	21	42	11	20

Figure 1. WGCNA analysis and determination of the CENPF gene. (**A**) Dendrogram of differentially expressed genes clustered based on a dissimilarity measure (1-TOM). (**B**) Heat map distribution histogram of differential genes for modules related to NSCLC staging in GSE19804 (The same results in the GSE30219, GSE32863, GSE63459 databases are shown in Supplementary Figure 1). (**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.

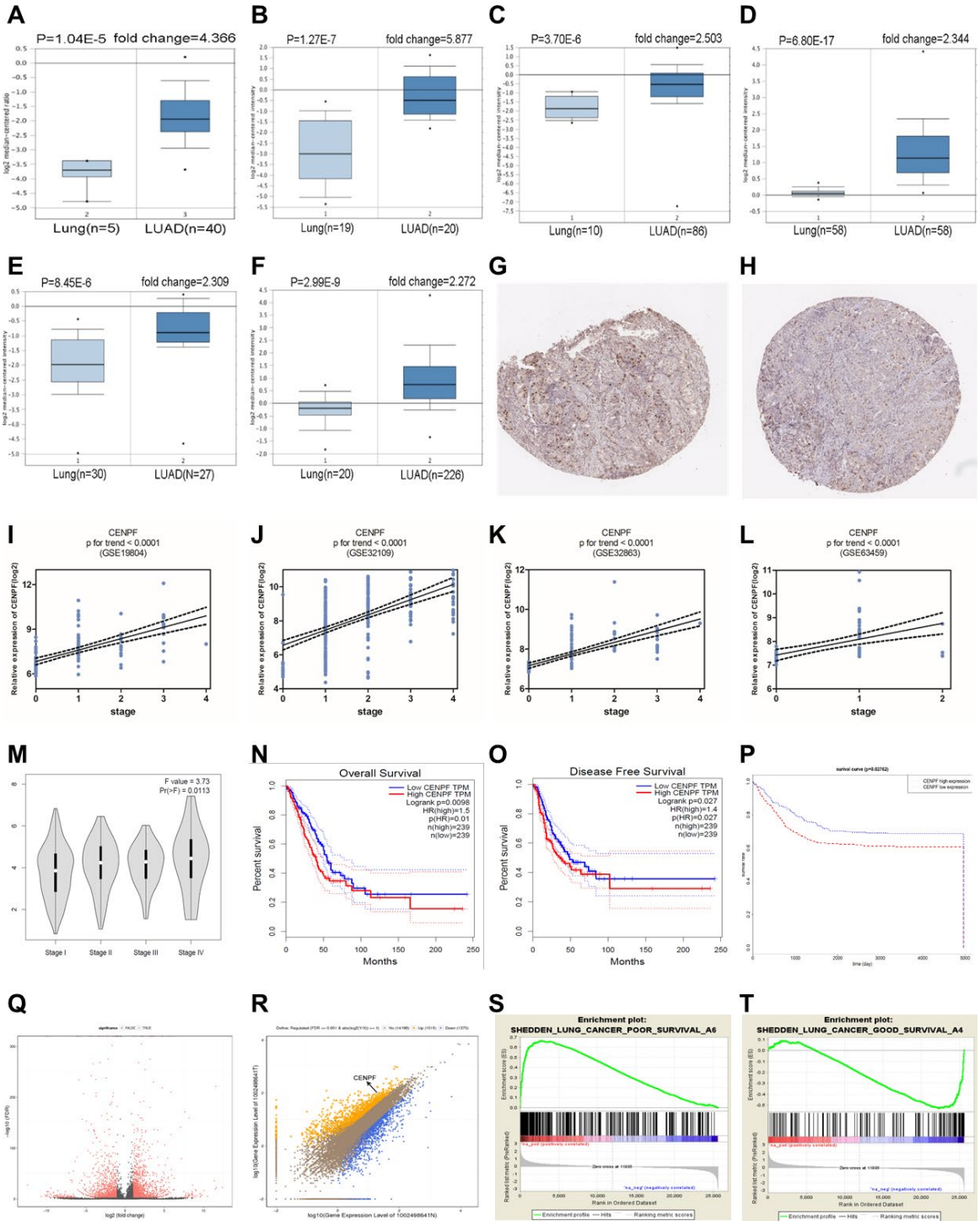


Figure 2. CENPF is upregulated in LUAD and is related with TNM staging and prognosis of LUAD patients. (**A**–**F**) Oncomine database results show that CENPF expression is significantly up-regulated in LUAD. (The corresponding P value and fold change are given above the picture). (**G**, **H**) The 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**, **O**) Kaplan Meier curves of OS (overall survival), DFS (Disease-free survival) in a cohort of LUAD stratified by CENPF expression. (**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.

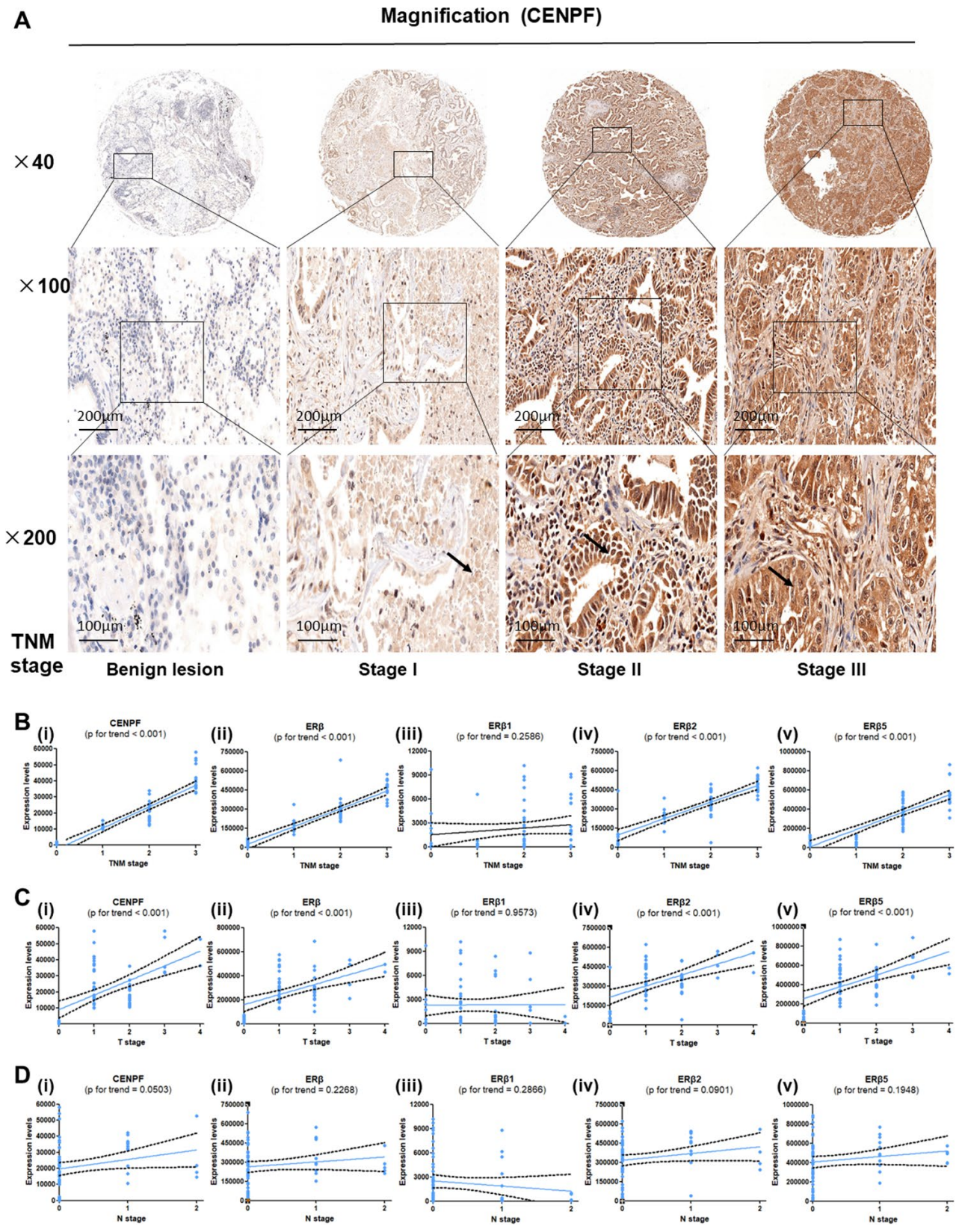


Figure 3. Expression of CENPF, ERβ, ERβ2 and ERβ5 are associated with T stage and TNM stage in LUAD patients. (**A**) Tissue microarray (TMA) was used to analyze the expression of CENPF in benign lung lesions and different TNM staging tissues of LUAD. The magnification of each slice is 40×, 100×, 200× in order. (**B**–**D**) Analysis of the relationship between the expression of CENPF, ERβ, ERβ1, ERβ2 and ERβ5 and the TNM staging or T stage or N stage of LUAD. The corresponding P value is marked above the picture.

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

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

E-cadherin expression was significantly increased (P=0.009, Figure 4F; The corresponding gray value are shown in Supplementary Figure 3G), and N-cadherin and MMP2 were significantly decreased when compared with NC group (P=0.004, 0.012, Figure 4F, 4G; Supplementary Figure 3G, 3H). A similar trend was obtained in the stable CENPF-deficient cell line H1299 (P<0.01, Supplementary Figure 3E–3H). Scratch experiment also demonstrated similar results (P=0.000, 0.000, Figure 4H; Supplementary Figure 3I). In A549 cells, the 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 4I; Supplementary Figure 3J). 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 4J; The corresponding gray value are shown in Supplementary Figure 3M). Similar results were obtained in the stable CENPF-deficient cell line H1299 (P<0.05, Supplementary Figure 3K–3M). The CENPF-KD group also showed a significant increase in apoptosis when compared with NC group (P=0.001, 0.001, Figure 4K; Supplementary Figure 3N).

CENPF Knockdown inhibits biological effects of LUAD cells mediated by ERβ2/5 pathway

CENPF and ERβ were co-localized in the nucleus of LUAD cells (Figure 5A). To investigate the biological effects of CENPF knockdown in LUAD cells mediated by ERβ signaling pathway, the cells were divided into CENPF-NC, CENPF-NC+E2, CENPF-NC+Ful, CENPF-KD+E2, and CENPF-KD+Ful. Cell proliferation in CENPF-KD+Ful group was significantly lower than CENPF-NC+Ful group at 48 hours (P=0.000, 0.000, Figure 5B). In A549 cells, the invasion and migration in CENPF-KD+E2 group were significantly reduced when compared with CENPF-NC+E2 group (P=0.000, 0.000, Figure 5C; Supplementary Figure 4A). Similarly, the expression of MMP2 and N-cadherin were significantly decreased in CENPF-KD+E2 group when compared with CENPF-NC+E2 group (P=0.002, 0.016, Figure 5E, 5F; The corresponding gray value are shown in Supplementary Figure 4B–4D). Similar results were obtained in stable CENPF-deficient cell line H1299 (P<0.01, Figure 5D–5F, Supplementary Figure 4A–4D). Scratch experiment also showed that the migration of CENPF-KD+E2 group was significantly

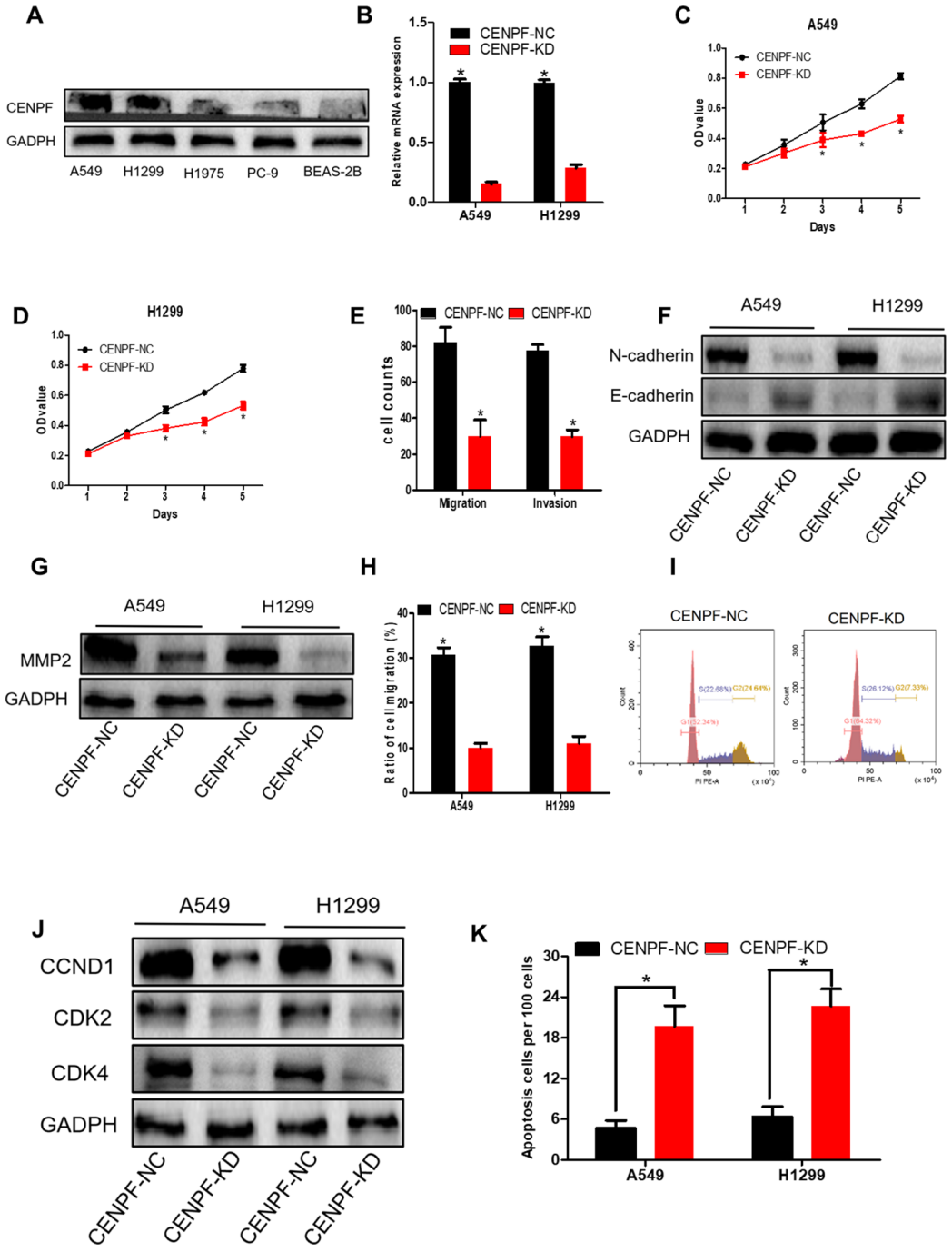


Figure 4. Knockdown of CENPF inhibits cell proliferation, migration, invasion and increases apoptosis of LUAD cells. (**A**) The protein level of CENPF in A549 and H1299 cell lines were higher than in normal cell lines BEAS-2B and other LUAD cells. GAPDH served as the internal control. (The corresponding gray value are shown in Supplementary Figure 3). (**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 showed that CENPF knockdown suppressed the proliferative viability of cells in A549 and H1299 cells. *P < 0.05 vs CENPF-KD. (**E**) Migration assays and invasion assays revealed that CENPF-KD decreased cell migration and invasion abilities of A549. (**F**, **G**) The related protein E-cadherin was significantly increased (P=0.009, Figure 4F; The corresponding gray value are shown in Supplementary Figure 3G), and N-cadherin and MMP2 were significantly decreased when compared with NC group (P=0.004, 0.012; The corresponding gray value are shown in Supplementary Figure 3H). (**H**) Quantified histograms of scratch experiment of A549 and H1299. (**I**) The cell percentage and DNA content were significantly increased in the G1 phase in the CENPF-KD group(P=0.011). (**J**) The expression of CCND1, CDK2 and CDK4 was significantly lowered in CENPF-KD group (P=0.022, 0.001, 0.002; The corresponding gray value are shown in Supplementary Figure 3M). (**K**) CENPF knockdown increased apoptosis of A549 and H1299 cell lines (P=0.001, 0.001). Each experiment was performed in triplicate and repeated three times. P values were calculated with two-tailed unpaired Student's t test.

lower than the CENPF-NC+E2 group (P=0.000, 0.000, Figure 5G; Supplementary Figure 4E). Flow cytometry analysis showed that the percentage of cells in G2/M phase in CENPF-KD+E2 group was significantly reduced than CENPF-NC+E2 group in A549 and H1299 cells (P=0.001, 0.021, Figure 5H; Supplementary Figure 4G–4I). 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 5I. The corresponding gray value are shown in Supplementary Figure 4F).

The effect of CENPF knockdown on the expression of ERβ2/5 was examined *in vitro* in CENPF-NC and CENPF-KD groups. The protein expression of ERβ2/5 in CENPF-KD group was significantly lower than the CENPF-NC group (P_{A549}=0.013, 0.000; P_{H1299}=0.006, 0.002, Figure 5J; The corresponding gray value are shown in Supplementary Figure 5A). We further explored the effect of CENPF knockdown on the expression ERβ2/5 under the action of E2 and Ful *in vitro*. 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 5K; The corresponding gray value are shown in Supplementary Figure 5C) in A549 cells. Similar results were obtained in stable CENPF-deficient cell line H1299 (P < 0.01, Figure 5J, 5K; Supplementary Figure 5C).

Knockdown of CENPF can inhibit ERβ2/5 pathway-mediated tumor growth *in vivo*

The tumor weight and size in the CENPF-KD group were lower than NC group (P<0.001, 0.001, Figure 6A–6C). The results of immunohistochemistry also showed that the expression of CENPF and ERβ2/5 was significantly decreased in CENPF-KD group when compared to the NC group (P=0.000, 0.000, 0.000, Supplementary Figure 5D, 5E).

In the lung cancer model of nude mice, tumor size and weight in the CENPF-KD+Ful group were significantly lowered than CENPF-NC+Ful group (P=0.001, 0.039, Figure 6D–6F; Supplementary Figure 5G). Immunohistochemistry staining demonstrated that the expression of CENPF and ERβ2/5 in CENPF-KD+E2 group was significantly lower than that in CENPF-NC+E2 group (P=0.000, 0.000, 0.000, Figure 6G, 6H).

Similar to the *in vitro* experiments, the effect of CENPF knockdown on the expression of ERβ2/5 was examined in CENPF-NC and CENPF-KD groups, and examined under the action of E2 and Ful *in vivo*. The protein expression of CENPF and ERβ2/5 in CENPF-KD group was lower than CENPF-NC group *in vivo* (P=0.024, 0.020, 0.003, Figure 6I; The corresponding gray value are shown in Supplementary Figure 5B). 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 6J; The corresponding gray value are shown in Supplementary Figure 5F).

DISCUSSION

Lung cancer is one of the most common malignancies and poses as a major health crisis globally. Targeted therapy for cell proliferation-related pathways and hormone therapy for lung cancer are considered important treatment modalities [17]. ER signaling pathways mainly included gene signaling and non-gene signaling pathways [18]. Among them, the gene signaling pathway involves the integration of estrogen with ER and the association with estrogen response element (ERE) to promote the recruitment of RNA polymerase II and regulate gene transcription. During this process, different combinations of synergistic activators determine the specificity of ER for activating the target genes [5]. 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–1C; Supplementary Figure 1, Supplementary Table 1). Ultimately, CENPF was found by overlapping the key genes (Figure 1D–1G).

High expression of CENPF was shown to correlate with the 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 [19–21]. The expression of CENPF in LUAD was detected by the Oncomine database. Similarly, RNA-Seq data showed high expression of CENPF in LUAD (Figure 2A–2F, 2Q, 2R). This result was further confirmed by analyzing the expression of CENPF according to LUAD staging in the four datasets and in *in vitro* experiment (Figure 2I–2L; Figure 6G, 6H), indicating that high expression of CENPF in LUAD might be related with its occurrence. At the same time, studies have shown that CENPF mediates mitosis and cell proliferation [22]. The expression of CENPF and DNA content were significantly reduced (Figure 4I), and the cell proliferation was significantly decreased in stably CENPF-deficient LUAD cells (Figure 4C, 4D). 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

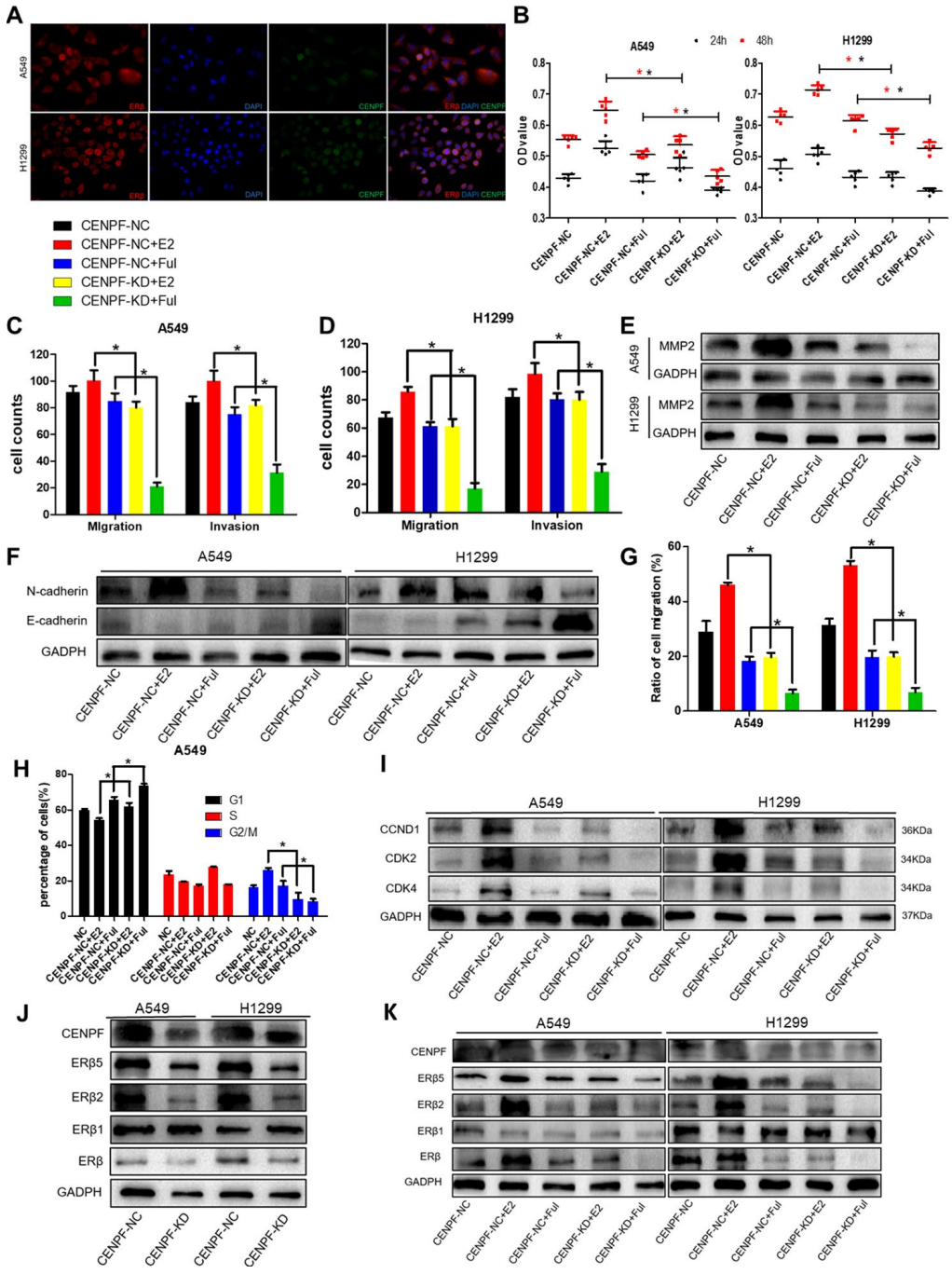


Figure 5. Knockdown of CENPF inhibits proliferation, invasion and migration of LUAD cells via the ERβ2/5 pathway. (**A**) Immunofluorescence showed the co-localization of CENPF and ERβ in A549 and H1299 cells (400 x). (**B**) Cell proliferation assays of different grouped cells at specific times in A549 and H1299 cells. (**C**, **D**) Corresponding quantified histograms of migration and invasion in A549 and H1299 cells. The invasion and migration of cells in CENPF-KD+E2 group were significantly reduced when compared with CENPF-NC+E2 group. (**E**, **F**) Protein expression of MMP2, N-cadherin and E-cadherin in A549 and H1299 cells (The corresponding gray value are shown in Supplementary Figure 4B–4D): The expression of MMP2 and N-cadherin were significantly decreased in CENPF-KD+E2 group when compared with CENPF-NC+E2 group. (**G**) Scratch experiment showed that the migration of CENPF-KD+E2 group was significantly lower than CENPF-NC+E2 group in A549 and H1299 cells (P=0.000, 0.000). (**H**) Corresponding quantified histograms of the A549 cells at different stages of the cell cycle (G1, S and G2/M). (**I**) Protein expression of CCND1, CDK2 and CDK4 in A549 and H1299 cells (The corresponding gray value are shown in Supplementary Figure 4F). *P < 0.05. (**J**) Knockdown of CENPF inhibited the expression of ERβ2/5 *in vitro*. (The corresponding gray value are shown in Supplementary Figure 5A). (**K**) Protein expression of CENPF, ERβ, ERβ1, ERβ2 and ERβ5 *in vitro* experiment after treated with E2 and Ful treatment (The corresponding gray value are shown in Supplementary Figure 5C). *P < 0.05. P values were calculated with two-tailed unpaired Student's t-test, or one-way analysis of variance.

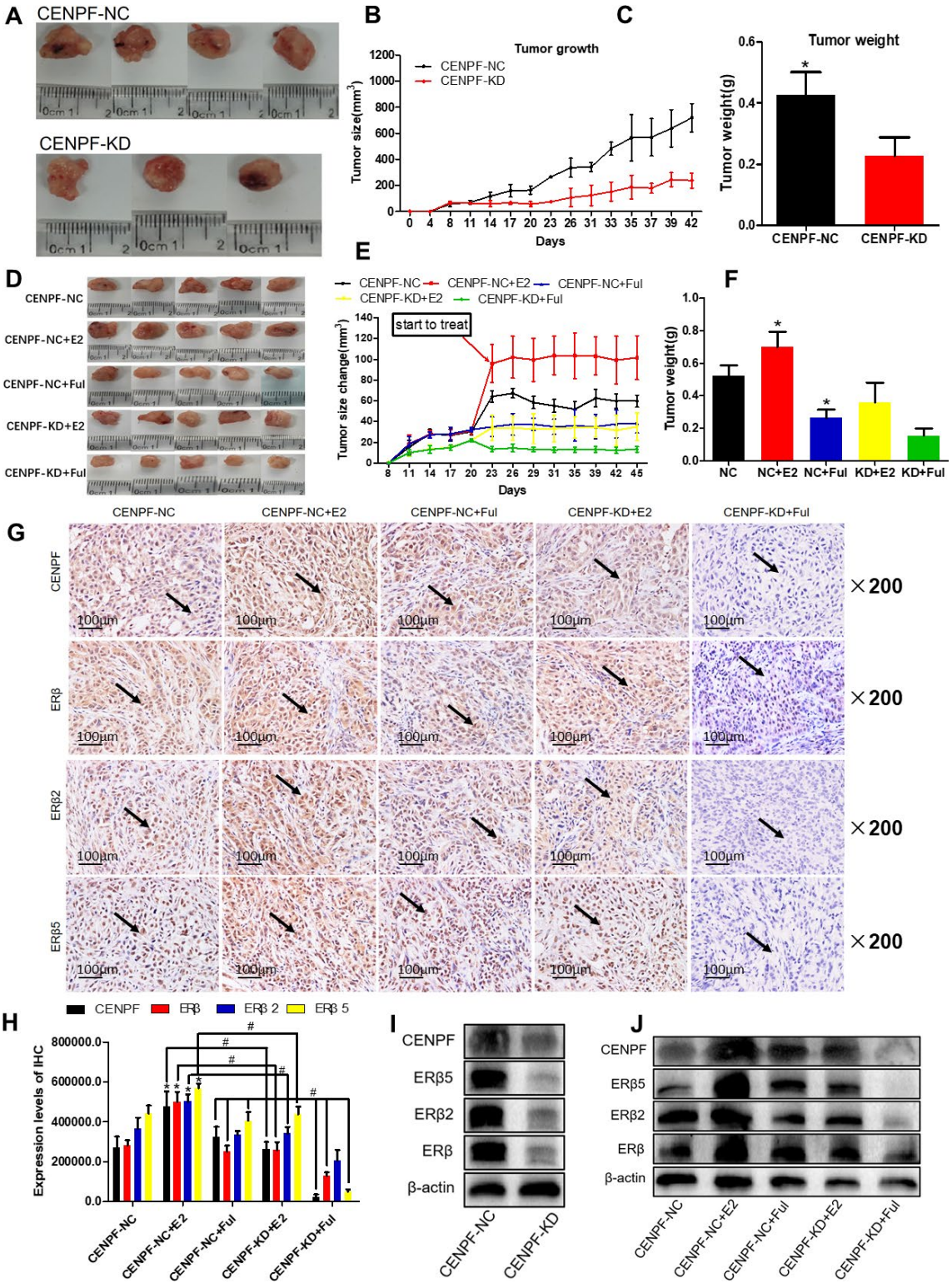


Figure 6. Knockdown of CENPF can inhibit ERβ2/5 pathway-mediated tumor tissue growth *in vivo*. (**A**) Pictures of mice tumor tissues after resection. (**B**, **C**) Analysis of tumor size and tumor weight. *P < 0.05. (**D**) Tumor images of nude mice. (**E**, **F**) Statistical analysis of tumor size and tumor weight. (**G**, **H**) Immunohistochemical analysis of the expression of CENPF, ERβ, ERβ2 and ERβ5 in nude mice tumor tissues and corresponding quantified histograms. (**I**) Knockdown of CENPF inhibited the expression of ERβ2/5. (The corresponding gray value are shown in Supplementary Figure 5B). (**J**) Protein expression of CENPF, ERβ, ERβ1, ERβ2 and ERβ5 *in vivo* experiment after treated with E2 and Ful treatment (The corresponding gray value are shown in Supplementary Figure 5F). *P < 0.05. P values were calculated with two-tailed unpaired Student's t-test, or one-way analysis of variance.

and poor prognosis of prostate cancer [9]. 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 CENPF [19]. We speculated that the dysregulation of miRNA-COUP-TFII-FOXM1-CENPF axis can be associated with malignant progression, poor prognosis and metastasis in LUAD. Our study revealed that LUAD cells showed a significant reduction in invasion and metastasis after CENPF knockdown (Figure 4E–4H). The expression of CENPF was significantly related to TNM staging of LUAD (Figure 2I–2L; Figure 3B, 3C). In addition, based on the clinical data of LUAD patients obtained from GEPIA and TCGA, LUAD patients with high expression of CENPF was correlated to a poor prognosis (Figure 2N–2P). These results indicated that abnormal expression of CENPF was significantly associated with TNM staging, poor prognosis, and malignant metastasis of LUAD.

In addition, our previous study reported that E2 promoted the progression of lung cancer by binding to ERβ [23]. The biological effects of ERβ in E2 varies based on the targeted organ tumors, including breast, cervical, and prostate cancers [13, 24]. Our previous findings indicated that ERβ2/5 are expressed in lung cancer [12]. Our study results revealed that ERβ2/5 is also highly expressed in LUAD patients (Supplementary Figure 2B, 2D). In addition, ERβ2/5 showed high positive association with the TNM staging of LUAD (Figure 3B, 3C).

CENPF knockdown inhibited the progression of LUAD mediated by the ERβ2/5 pathway. High expression of CENPF and ERβ2/5 is associated with the development of LUAD. Invasion, migration and proliferation of LUAD cells in the CENPF-KD+E2 group showed significant reduction *in vitro* when compared to the CENPF-NC+E2 group (Figure 5C–5H, Supplementary Figure 4A–4F). In the *in vitro* system, the protein expression of ERβ2/5 in CENPF-KD+E2 group was significantly lower than CENPF-NC+E2 group (Figure 5K; Supplementary Figure 5C). In order to eliminate the effects of endogenous estrogen due to gender differences, *in vivo* experiments were only conducted in male mice. Consistent with the *in vitro* experiment results, the expression of ERβ2/5 protein in CENPF-KD+E2 group was significantly lower than that in CENPF-NC+E2 group (Figure 6J; Supplementary Figure 5F). From these results, we confirmed that the knockdown of CENPF inhibited the progression of LUAD mediated by ERβ2/5 pathway both *in vitro* and *in vivo*. This is another important point regarding the mechanism of CENPF, except for FOXM1 [9] and COUP-TFII [19].

Taken together, these findings indicated that both CENPF and ERβ2/5 are highly expressed in LUAD cells and their expression is associated with TNM staging and prognosis in LUAD patients. CENPF knockdown inhibited proliferation, invasion and metastasis of LUAD cells mediated by ERβ2/5 pathway. Thus, the development of inhibitors against the ERβ2/5 subtypes and CENPF can have great therapeutic impact in LUAD. However, there are some shortcomings in this study that should be acknowledged. First, due to the large molecular weight of CENPF, plasmid construction can easily be created off target, thus LUAD cell lines expressing CENPF can be difficult to construct. Second, the number of specimen used for RNA-Seq is limited. These issues are key points that should be targeted for future studies.

MATERIALS AND METHODS

Tissue specimens of patient and cell culture

This study was approved by the Ethics Review Committee of Wuhan University. Tissue specimens from 56 LUAD cases and 10 benign pulmonary lesions (BPL) cases who underwent surgery from April 2014 to July 2017 were collected for tissue chip. One pair of LUAD and peri-cancerous tissues were collected 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 analysis [4].

Human LUAD cell lines (A549, H1975, H1299 and PC-9) were cultured in RPMI-1640 medium. 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).

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 non-small cell lung cancer (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 tumors and normal tissues) and GSE63459 (including 65 tumor tissues).

Analysis and verification of hub gene

The data sets of GSE19804, GSE30219, GSE32863 and GSE63459 were used to construct co-expression networks and clinical functioning related modules. The genes were screened according to the false discovery rate (FDR) <0.05 and | log2 fold change (FC) |> 1.5. Next, a weighted gene co-expression network analysis (WGCNA) package was used to construct a co-expression network [25, 26]. 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 STAR [27] to obtain differential genes. The screening criteria for differential genes were abs(log2FC) > 1 and p 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.

Immunohistochemistry

The detailed steps for conducting immunohistochemistry was described previously [11]. CENPF (ab5) and ERβ (ab3576) were purchased from Abcam. ERβ1 (MCA1974ST), ERβ2 (MCA2279GT) and ERβ5 (MCA4676T) were purchased from AbDSerotec. The specificity of the above antibodies was confirmed by numerous laboratories including ours [11, 28]. Immunohistochemical method to analyze the optical density was calculated by Image-Pro Plus software.

Western blotting

Detailed western blotting analysis has been done as described previously [29]. 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, 19, 30].

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

Specific experimental methods were shown in our previously published study [6]. Primers were designed based on CENPF mRNA sequence in GenBank. The primers used were as follows: CENPF, 3- CTCTCCC GTCAACAGCGTTC; CENPF, 5- GTTGTGCATATT CTTGGCTTGC. Data was analyzed using $2^{-\Delta\Delta Ct}$ method. GAPDH, Forward Primer: GGTGA AGGTC GGAGT CAACG; GAPDH, Reverse Primer: CAAAG TTGTC ATGGA TGHACC.

Cell culture experiment

The sense sequence of CENPF knockdown (KD) and negative control (NC) were integrated into the pWSLV-sh08-GFP vector for transfection of A549 and H1299 cells. The stable CENPF-deficient A549 and H1299 cells were immunofluorescently labeled with anti-CENPF and anti-ERβ [31]. Cell apoptosis assay was conducted using TdT-mediated dUTP-biotin nick end labeling test (TUNEL, Roche Applied Science, Germany) according to the manufacturer's instructions [32]. MTT [33], invasion, migration, scratch and cell cycle experiments were used to evaluate the effects of LV-CENPF sh and LV-NC on the biological function of A549 and H1299 cells [34, 35].

Xenograft mouse model

Male nude mice were obtained from Beijing HFK Bioscience Co., Ltd., China. Mice were housed in specific-pathogen free environment for one week, and then were subcutaneously injected with 100 μL of $8x10^6$ LV-CENPF sh or LV-NC cells. Tumor size was measured every three days (tumor size = length × $width^2$ × 0.5 mm^3). When the tumor size has reached 80-120 mm^3, 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 measured [36]. Tumor tissues were fixed in 4% paraformaldehyde or frozen with liquid nitrogen and stored at -80° C.

Data analysis

Data are expressed as means ± 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 significant 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.

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.

AUTHOR CONTRIBUTIONS

(I) Conception and design: Tang Hexiao, Bai Yuquan and Xiong Lecai. (II) Administrative support: Zhao Jinping. (III) Provision of study materials or patients: Zhao Jinping, Hu Weidong. (IV) Collection and assembly of data: Tang Hexiao, Bai Yuquan and Xiong Lecai. (V) Data analysis and interpretation: Bai Yuquan and Xiong Lecai. (VI) Manuscript writing: All authors. (VII) Final approval of manuscript: All authors.

ACKNOWLEDGMENTS

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.

CONFLICTS OF INTEREST

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.

FUNDING

This project was supported by the Program of Excellent Doctoral (Postdoctoral) of Zhongnan Hospital of Wuhan University (Grant No. ZNYB2019002).

REFERENCES

1. Funahashi S, Okazaki Y, Nagai H, Chew SH, Ogawa K, Toyoda T, Cho YM, Toyokuni S. Twist1 was detected in mesenchymal cells of mammary fibroadenoma and invasive components of breast carcinoma in rats. J Toxicol Pathol. 2019; 32:19–26. https://doi.org/10.1293/tox.2018-0029 PMID:30739992

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 Cancer J Clin. 2018; 68:394–424. https://doi.org/10.3322/caac.21492 PMID:30207593

3. Chang GC, Tseng CH, Hsu KH, Yu CJ, Yang CT, Chen KC, Yang TY, Tseng JS, Liu CY, Liao WY, Hsia TC, Tu CY, Lin MC, 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. https://doi.org/10.1016/j.lungcan.2016.12.002 PMID:28213001

4. Anraku M, Waddell TK, de Perrot M, Lewis SJ, Pierre AF, Darling GE, Johnston MR, Zener RE, Rampersaud YR, Shepherd FA, Leighl N, Bezjak A, Sun AY, et al. Induction chemoradiotherapy facilitates radical resection of T4 non-small cell lung cancer invading the spine. J Thorac Cardiovasc Surg. 2009; 137:441–47.e1. https://doi.org/10.1016/j.jtcvs.2008.09.035 PMID:19185167

5. Ciuleanu TE, Ahmed S, Kim JH, Mezger J, Park K, Thomas M, Chen J, Poondru S, VanTornout JM, Whitcomb D, Blackhall F. 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. Br J Cancer. 2017; 117:757–66. https://doi.org/10.1038/bjc.2017.226 PMID:28772281

6. Bai Y, Shen W, Zhu M, Zhang L, Wei Y, Tang H, Zhao J. Combined detection of estrogen and tumor markers is an important reference factor in the diagnosis and prognosis of lung cancer. J Cell Biochem. 2019; 120:105–14. https://doi.org/10.1002/jcb.27130 PMID:30216488

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 Motil Cytoskeleton. 1993; 26:214–26. https://doi.org/10.1002/cm.970260305 PMID:7904902

8. Yerushalmi R, Woods R, Kennecke H, Speers C, Knowling M, Gelmon K. Patterns of relapse in breast cancer: changes over time. Breast Cancer Res Treat. 2010; 120:753–59. https://doi.org/10.1007/s10549-009-0510-2 PMID:19701704

9. Aytes A, Mitrofanova A, Lefebvre C, Alvarez MJ, Castillo-Martin M, Zheng T, Eastham JA, Gopalan A, Pienta KJ, Shen MM, Califano A, Abate-Shen C. Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. Cancer Cell. 2014; 25:638–51. https://doi.org/10.1016/j.ccr.2014.03.017 PMID:24823640

10. Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A. Reverse engineering cellular networks. Nat Protoc. 2006; 1:662–71. https://doi.org/10.1038/nprot.2006.106 PMID:17406294

11. Antoon JW, Martin EC, Lai R, Salvo VA, Tang Y, Nitzchke AM, Elliott S, Nam SY, Xiong W, Rhodes LV, Collins-Burow B, David O, Wang G, et al. MEK5/ERK5 signaling suppresses estrogen receptor expression and promotes hormone-independent tumorigenesis. PLoS One. 2013; 8:e69291. https://doi.org/10.1371/journal.pone.0069291 PMID:23950888

12. Butts C, Socinski MA, Mitchell PL, Thatcher N, Havel L, Krzakowski M, Nawrocki S, Ciuleanu TE, Bosquée L, Trigo JM, Spira A, Tremblay L, Nyman J, et al, and START Trial Team. Tecemotide (L-BLP25) versus placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. Lancet Oncol. 2014; 15:59–68. https://doi.org/10.1016/S1470-2045(13)70510-2 PMID:24331154

13. Leung YK, Lam HM, Wu S, Song D, Levin L, Cheng L, Wu CL, Ho SM. Estrogen receptor beta2 and beta5 are associated with poor prognosis in prostate cancer, and promote cancer cell migration and invasion. Endocr Relat Cancer. 2010; 17:675–89. https://doi.org/10.1677/ERC-09-0294 PMID:20501637

14. Shaaban AM, Green AR, Karthik S, Alizadeh Y, Hughes TA, Harkins L, Ellis IO, Robertson JF, Paish EC, Saunders PT, Groome NP, Speirs V. Nuclear and cytoplasmic expression of ERbeta1, ERbeta2, and ERbeta5 identifies distinct prognostic outcome for breast cancer patients. Clin Cancer Res. 2008; 14:5228–35. https://doi.org/10.1158/1078-0432.CCR-07-4528 PMID:18698041

15. Ayoub NM, Al-Shami KM, Yaghan RJ. Immunotherapy for HER2-positive breast cancer: recent advances and combination therapeutic approaches. Breast Cancer (Dove Med Press). 2019; 11:53–69. https://doi.org/10.2147/BCTT.S175360 PMID:30697064

16. An C, Zhang J, Chu H, Gu C, Xiao F, Zhu F, Lu R, Shi H, Zhang H, Yi X. Study of gefitinib and pemetrexed as first-line treatment in patients with advanced non-small cell lung cancer harboring EGFR mutation. Pathol Oncol Res. 2016; 22:763–68. https://doi.org/10.1007/s12253-016-0067-4 PMID:27126186

17. 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. Med Oncol. 2015; 32:405. https://doi.org/10.1007/s12032-014-0405-x PMID:25502081

18. Niikawa H, Suzuki T, Miki Y, Suzuki S, Nagasaki S, Akahira J, Honma S, Evans DB, Hayashi S, Kondo T, Sasano H. Intratumoral estrogens and estrogen receptors in human non-small cell lung carcinoma. Clin Cancer Res. 2008; 14:4417–26. https://doi.org/10.1158/1078-0432.CCR-07-1950 PMID:18579664

19. Abe T, Shirai K, Saitoh J, Ebara T, Shimada H, Tashiro M, Okano N, Ohno T, Nakano T. Incidence, risk factors, and dose-volume relationship of radiation-induced rib fracture after carbon ion radiotherapy for lung cancer. Acta Oncol. 2016; 55:163–66. https://doi.org/10.3109/0284186X.2015.1088169 PMID:26399488

20. Pimkhaokham A, Shimada Y, Fukuda Y, Kurihara N, Imoto I, Yang ZQ, Imamura M, Nakamura Y, Amagasa T, Inazawa J. Nonrandom chromosomal imbalances in esophageal squamous cell carcinoma cell lines: possible involvement of the ATF3 and CENPF genes in the 1q32 amplicon. Jpn J Cancer Res. 2000; 91:1126–33. https://doi.org/10.1111/j.1349-7006.2000.tb00895.x PMID:11092977

21. Cheng H, Sun N, Sun X, Chen B, Li F, Feng J, Cheng L, Cao Y. Polymorphisms in hMSH2 and hMLH1 and response to platinum-based chemotherapy in advanced non-small-cell lung cancer patients. Acta Biochim Biophys Sin (Shanghai). 2010; 42:311–17. https://doi.org/10.1093/abbs/gmq023 PMID:20458443

22. Medical Advisory Secretariat. Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: an evidence-based and economic analysis. Ont Health Technol Assess Ser. 2010; 10:1–57. PMID:23074401

23. Gomez DR, Blumenschein GR Jr, Lee JJ, Hernandez M, Ye R, Camidge DR, Doebele RC, Skoulidis F, Gaspar LE, Gibbons DL, Karam JA, Kavanagh BD, Tang C, 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. Lancet Oncol. 2016; 17:1672–82. https://doi.org/10.1016/S1470-2045(16)30532-0 PMID:27789196

24. Weiss-Steider B, Córdova Y, Aguiñiga-Sánchez I, Ledesma-Martínez E, Domínguez-Meléndez V, Santiago-Osorio E. El sodium caseinate and alfa-casein inhibit proliferation of the mouse myeloid cell line 32D clone 3 (32Dcl3) via TNF-α. Biomedica. 2019; 39:291–99.

https://doi.org/10.7705/biomedica.v39i3.4094 PMID:31529816

25. Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. PLoS Comput Biol. 2008; 4:e1000117. https://doi.org/10.1371/journal.pcbi.1000117 PMID:18704157

26. 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. https://doi.org/10.1186/1471-2164-10-327 PMID:19619308

27. Dobin A, Gingeras TR. Optimizing RNA-seq mapping with STAR. Methods Mol Biol. 2016; 1415:245–62. https://doi.org/10.1007/978-1-4939-3572-7_13 PMID:27115637

28. Gong Y, Xu Z, Jin C, Deng H, Wang Z, Zhou W, Zhang M, Zhao X, Wang L. Treatment of advanced non-small-cell lung cancer with qi-nourishing essence-replenishing Chinese herbal medicine combined with chemotherapy. Biol Proced Online. 2018; 20:9. https://doi.org/10.1186/s12575-018-0074-9 PMID:29618954

29. Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, Yokoi T, Chiappori A, Lee KH, de Wit M, Cho BC, Bourhaba M, Quantin X, et al, and PACIFIC Investigators. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. N Engl J Med. 2017; 377:1919–29. https://doi.org/10.1056/NEJMoa1709937 PMID:28885881

30. Jing X, Peng J, Dou Y, Sun J, Ma C, Wang Q, Zhang L, Luo X, Kong B, Zhang Y, Wang L, Qu X. Macrophage ERα promoted invasion of endometrial cancer cell by mTOR/KIF5B-mediated epithelial to mesenchymal transition. Immunol Cell Biol. 2019; 97:563–76. https://doi.org/10.1111/imcb.12245 PMID:30779215

31. 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. J Clin Pathol. 1964; 17:225–30. https://doi.org/10.1136/jcp.17.3.225 PMID:14159448

32. Kyrylkova K, Kyryachenko S, Leid M, Kioussi C. Detection of apoptosis by TUNEL assay. Methods Mol Biol. 2012; 887:41–47. https://doi.org/10.1007/978-1-61779-860-3_5 PMID:22566045

33. Cagney DN, Thirion PG, Dunne MT, Fleming C, Fitzpatrick D, O'Shea CM, Finn MA, O'Sullivan S, Booth C, Collins CD, Buckney SJ, Shannon A, Armstrong JG. 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. https://doi.org/10.1016/j.clon.2017.10.010 PMID:29097074

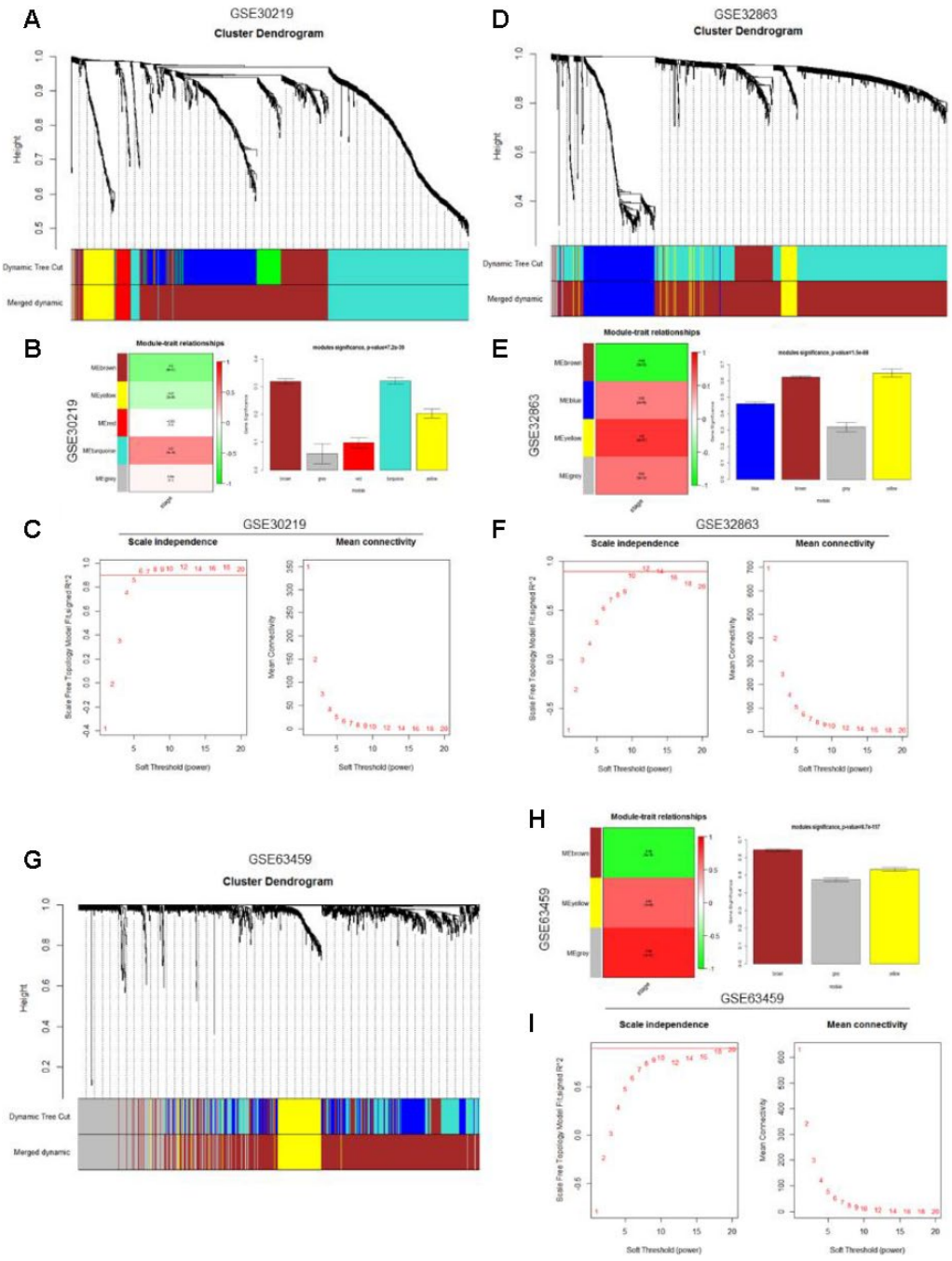
34. Agostini P, Naidu B, Cieslik H, Steyn R, Rajesh PB, Bishay E, Kalkat MS, Singh S. Effectiveness of incentive spirometry in patients following thoracotomy and lung resection including those at high risk for developing pulmonary complications. Thorax. 2013; 68:580–85. https://doi.org/10.1136/thoraxjnl-2012-202785 PMID:23429831

35. Sabran A, Kumolosasi E, Jantan I. Effects of annexin A1 on apoptosis and cell cycle arrest in human leukemic cell lines. Acta Pharm. 2019; 69:75–86. https://doi.org/10.2478/acph-2019-0005 PMID:31259717

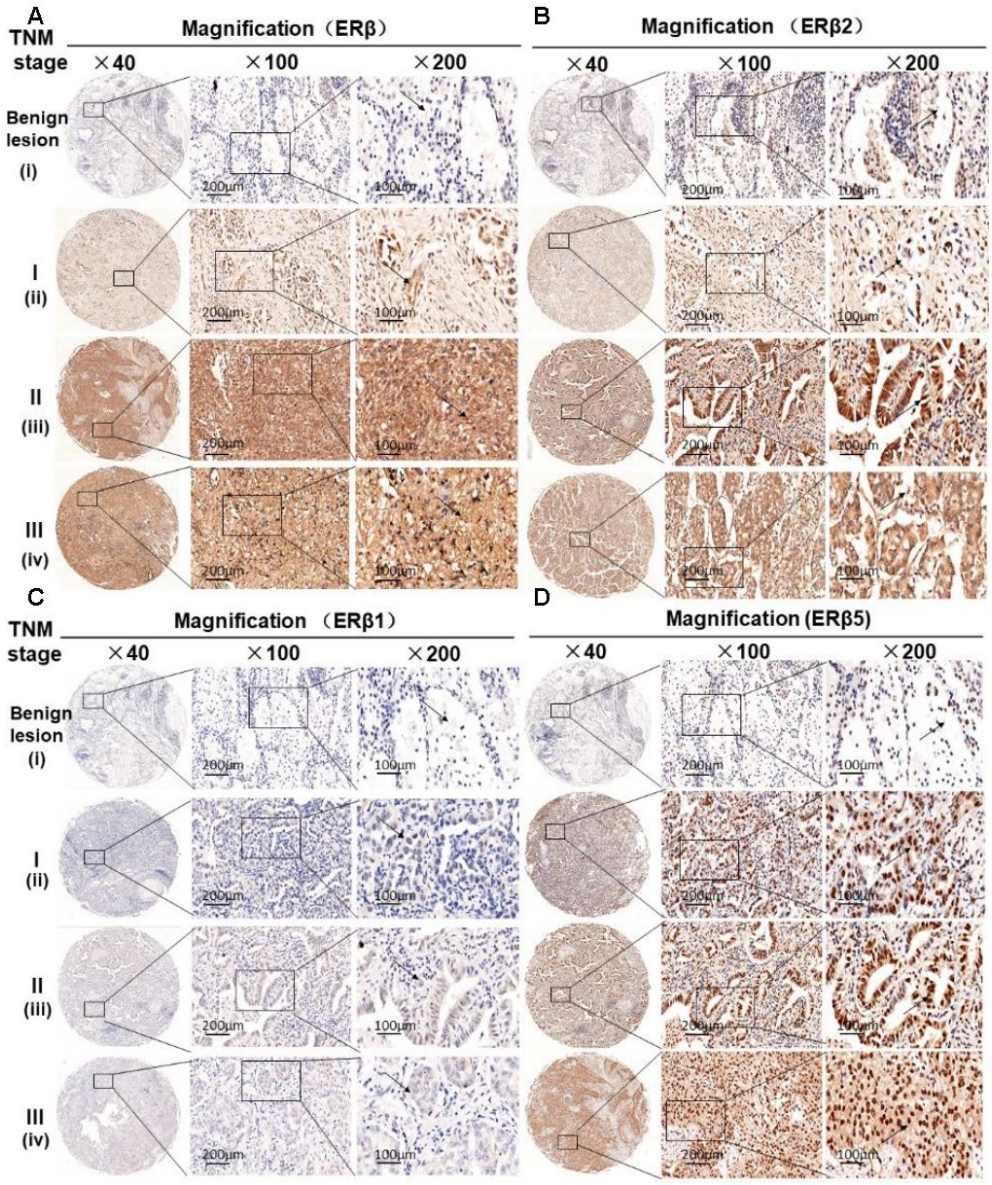
36. Guilmain W, Colin S, Legrand E, Vannier JP, Steverlynck C, Bongaerts M, Vasse M, Al-Mahmood S. CD9P-1 expression correlates with the metastatic status of lung cancer, and a truncated form of CD9P-1, GS-168AT2, inhibits in vivo tumour growth. Br J Cancer. 2011; 104:496–504. https://doi.org/10.1038/sj.bjc.6606033 PMID:21206492

SUPPLEMENTARY MATERIALS

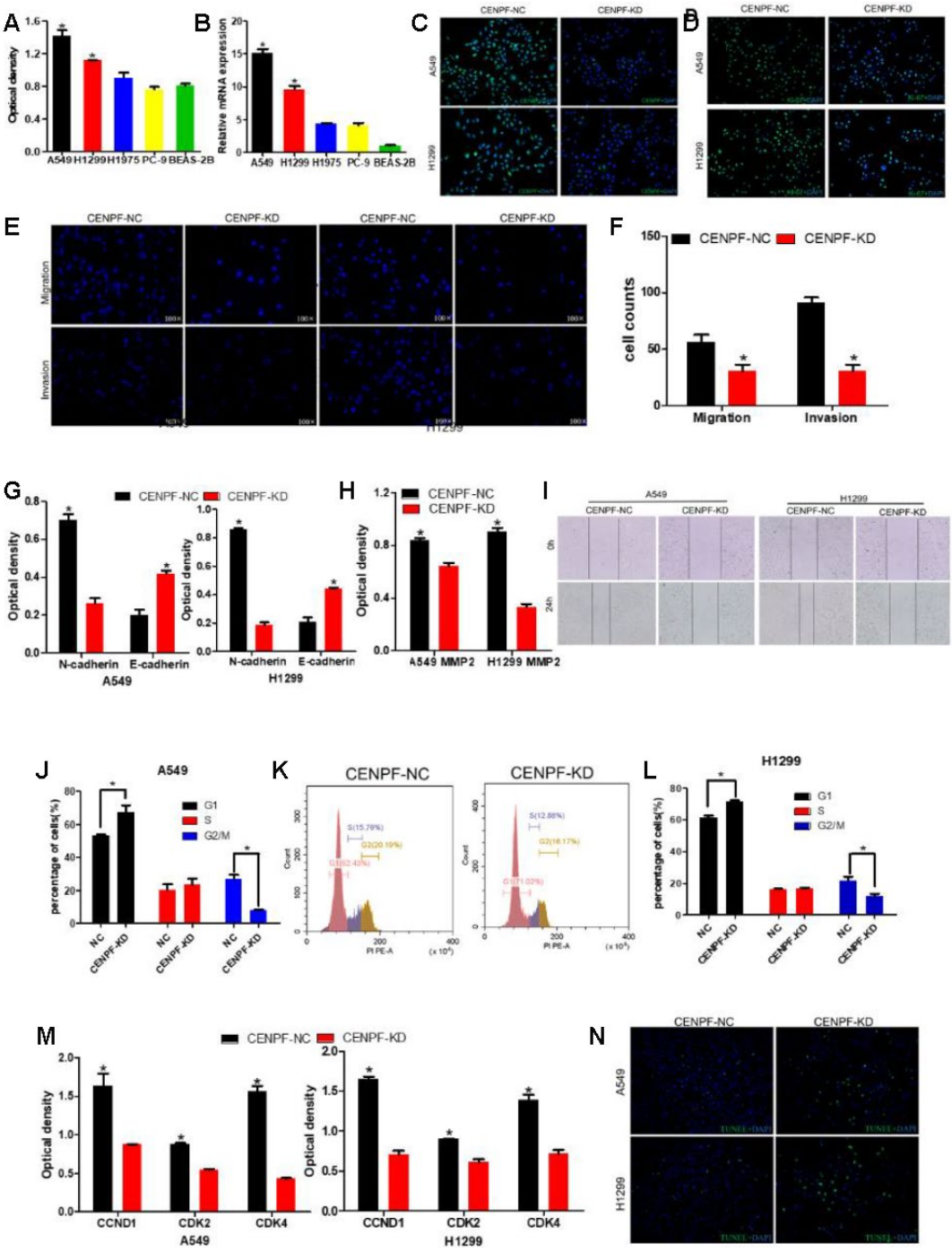
Supplementary Figures



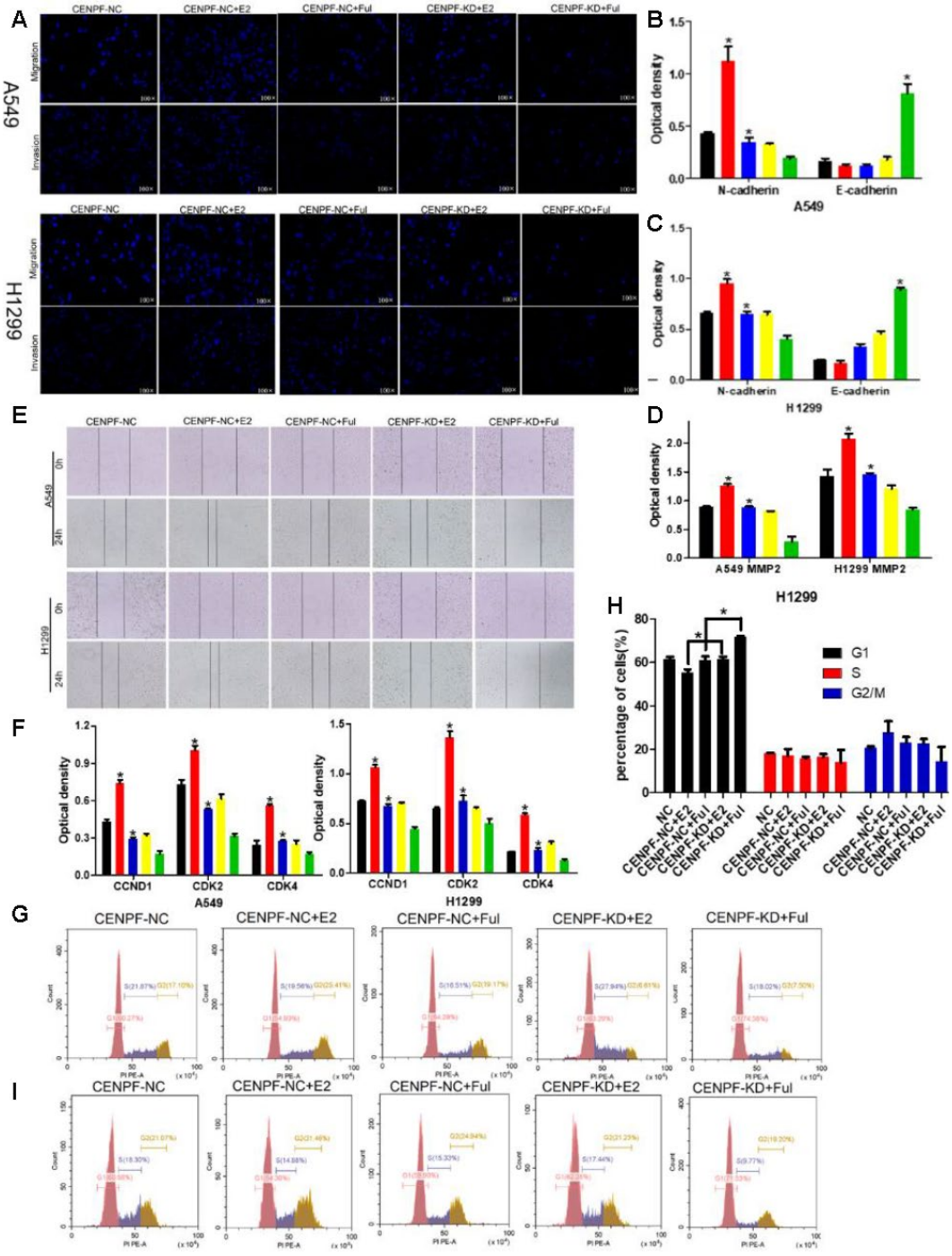
Supplementary Figure 1. WGCNA analysis and determination of the CENPF gene. (**A**, **D**, **G**) Dendrogram of all differentially expressed genes clustered based on a dissimilarity measure (1-TOM) (GSE30219, GSE32863, GSE63459). (**B**, **E**, **H**) Heat maps and distribution of differential genes for different modules related to NSCLC staging (GSE30219) and LUAD staging (GSE32863, GSE63459). (**C**, **F**, **I**) Analysis of the scale-free fit index for various soft-thresholding power (β) and analysis of the mean connectivity for various soft-thresholding power (GSE30219, GSE32863, GSE63459).



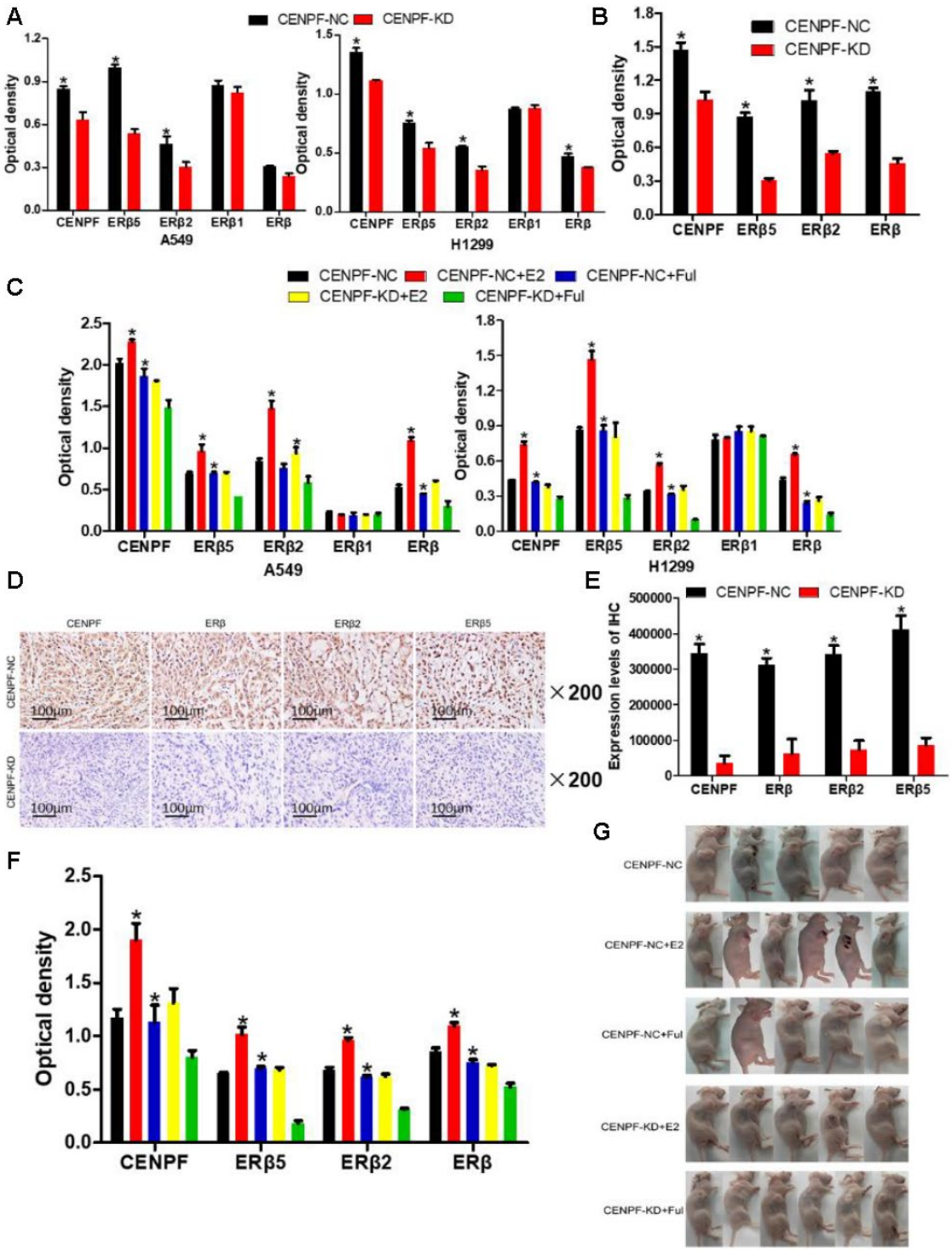
Supplementary Figure 2. (**A**–**D**) Expression of ERβ, ERβ1, ERβ2 and ERβ5 in benign lung lesions (i) and different TNM staging of LUAD (ii=I stage, iii=II stage, iv=III stage). The magnification of each slice is 40×, 100×, 200× in order.



Supplementary Figure 3. Knockdown of CENPF inhibits the biological effects of LUAD cells. (**A**) The protein level of CENPF in A549 and H1299 cell lines were higher than in normal cell lines BEAS-2B and other LUAD cells. GAPDH served as the internal control. *P < 0.05 vs other cells. (**B**) mRNA expression of CENPF in different cell lines. *P < 0.05 vs other cells. (**C**) Representative cellular immunofluorescence images after transfection of CENPF (200×). Green stands for CENPF and blue stands for DAPI. (**D**) Representative Ki67 staining (green) shows cell proliferation of LUAD cells after CENPF-NC or KD treatment (200x). The nuclei were counterstained with DAPI (blue). (**E**, **F**) Migration assays and invasion assays revealed that CENPF-KD decreased cell migration and invasion abilities of A549 and H1299. (**G**, **H**) Corresponding gray value analysis showed that the related protein E-cadherin was significantly increased (P=0.009, Figure 4F), and N-cadherin and MMP2 were significantly decreased when compared with NC group (P=0.004, 0.012) of N-cadherin, E-cadherin and MMP2 in A549 and H1299 cells. (**I**) Representative scratched pictures of A549 and H1299 cells. (**J**–**L**) Percentage of CENPF-KD cells H1299 at different stages of the cell cycle (G1, S and G2/M) and corresponding quantified histograms of A549 and H1299. (**M**) Corresponding gray value analysis showed that the expression of CCND1, CDK2 and CDK4 was significantly lowered in CENPF-KD group (P=0.022, 0.001, 0.002). (**N**) Representative TUNEL staining (green) shows (200x). P values were calculated with two-tailed unpaired Student's t-test, or one-way analysis of variance.



Supplementary Figure 4. Knockdown of CENPF inhibits proliferation, invasion and migration of LUAD cells via the ERβ2/5 pathway. (**A**) Migration and invasion pictures of A549 and H1299 cells. (**B**–**D**) Corresponding quantified histograms of MMP2, N-cadherin and E-cadherin in A549 and H1299 cells. (**E**) Representative scratched images of A549 and H1299 cells. (**F**) Corresponding quantified histograms of CCND1, CDK2 and CDK4 in A549 and H1299. *P < 0.05. (**G**–**I**) Percentage of the A549 cells and H1299 cells at different stages of the cell cycle (G1, S and G2/M) and corresponding quantified histograms. P values were calculated with two-tailed unpaired Student's t-test, or one-way analysis of variance.



Supplementary Figure 5. Knockdown of CENPF can inhibit the expression of ERβ2/5 *in vitro* and *in vivo*. (**A**, **B**) Corresponding gray value analysis showed that knockdown of CENPF inhibited the expression of ERβ2/5 *in vitro* and *in vivo* experiment. (**C**) Corresponding gray value analysis of CENPF, ERβ, ERβ1, ERβ2 and ERβ5 *in vitro* experiment after treated with E2 and Ful treatment. (**D**, **E**) Immunohistochemical analysis of CENPF, ERβ, ERβ2 and ERβ5 expression in nude mice tumor tissues and corresponding quantitative histograms. *P < 0.05. (**F**) Corresponding gray value analysis of CENPF, ERβ, ERβ1, ERβ2 and ERβ5 *in vivo* experiment after treated with E2 and Ful treatment. (**G**) Pictures of nude mice sacrificed at 45 days. *P < 0.05. P values were calculated with two-tailed unpaired Student's t-test, or one-way analysis of variance.

Supplementary Table

Supplementary Table 1. Genes in key modules.

Hub module	Genes							
Brown Module (GSE19804) (n=185)	CTHRC1	ARHGAP31	CST1	SUGCT	SORD	PCLAF	TNS1	TNPO1
	PAICS	CDCA7	LAMP3	MDK	SFXN1	COMP	TFAP2A	SULF1
	IGSF9	GJB2	P3H4	UHRF1	CST2	CTTN	MMP12	ATOH8
	DNAH14	NMU	PGM2L1	SLC39A8	AQP4	GPX3	CDCA3	SLC46A2
	PEBP4	MS4A15	SOX4	GREM1	KLF9	HSPB8	JPT1	ANP32E
	TENM4	SRPX2	RUNX2	CST4	CPB2	SCG5	E2F8	CENPU
	ACACB	VEPH1	SLC2A1	LRRC15	PCDH7	FIGNL1	MXRA5	MND1
	THY1	ADAM12	ADGRD1	FAM199X	CCDC34	SHCBP1	MKI67	MFAP2
	TRIM59	LRRK2	IQGAP3	FHL2	CILP2	DEPDC1	ABCA3	LAMP5
	FAP	SLC2A5	PLAU	PCP4	PLPP4	SFTA1P	IGF2BP3	DIO2
	KNL1	GPX8	FNDC1	CACNA2D2	DEPDC1B	DEPDC7	PRR11	VCAN
	PLLP	CYP4B1	HOXB7	HLF	WISP1	HOXA10	HSD17B6	FUT9
	GPR87	CILP	CEMIP	SUSD2	SCGB3A2	CDK1	BUB1	CCNB1
	CCNB2	MAD2L1	CDC20	TOP2A	CCNA2	KIF11	BUB1B	DLGAP5
	KIF2C	KIF20A	AURKA	NDC80	CENPF	NUSAP1	BIRC5	NCAPG
	PRC1	TPX2	UBE2C	NUF2	TTK	ZWINT	RRM2	MELK
	CDKN3	PTTG1	CENPK	CHEK1	NEK2	MCM4	ASPM	MCM2
	PBK	KIF4A	CEP55	FOXM1	KIF15	COL1A1	ANLN	COL1A2
	COL10A1	COL11A1	COL3A1	COL5A2	COL5A1	COL8A2	PLOD2	TYMS
	KIF14	ORC6	HMMR	KIF26B	GMNN	FEN1	MMP1	FBXO32
	GINS1	GINS2	GTSE1	ZBTB16	RMI2	UBE2T	ITGA11	BRIP1
	EZH2	CCNE2	PSAT1	PAFAH1B3	MMP11	THBS2	NME1	STIL
	PSPH	FANCI	MMP13	ADAMTS12	ECT2	IGFBP3	SFTPD	PLA2G1B
	SFTPB							
Turquoise module (GSE30219) (n=413)	GPD1	BTNL9	SCARA5	PLAC9	GDF10	HSPB6	FHL1	DAPK2
	TNPO1	FAM189A2	PGR	ANKRD29	DLC1	EMCN	ABCA8	PDE2A
	CSRNP1	ARHGAP6	LOC100506990	PSMD6-AS2	FHL5	PGM5	NOSTRIN	FILIP1
	LTBP4	SPATA13	STARD13	RUFY2	TNS2	PREX2	CFAP70	SYNPO2
	GPX3	IFT57	PCLAF	JCAD	VEPH1	ATAD2	CX3CR1	PALMD
	ADIRF	RBP4	PLPP3	HBB	VAPA	CDCA3	FRY	ATOH8
	TGFBR3	SELENOI	VSIG2	CTHRC1	FAM199X	SRD5A1	HIGD1B	CAB39L
	CDH19	IL33	HMGB3	NR4A1	DIXDC1	NR4A3	MAGI2-AS3	
	ARHGAP44	PLSCR4	ATP5S	COX7A1	UHRF1	SHC3	HLF	2-Mar
	DEPDC1	RNF125	CDO1	HMGB3P1	CKAP2	MMP12	SPARCL1	NDRG2
	FAM13C	NETO2	AUNIP	NTN4	KNSTRN	MKI67	DNAH14	PSAT1
	THY1	E2F8	KNL1	DSP	SPAG5	CRIM1	SIX4	
	PRICKLE2	FIGNL1	RAI2	CCDC34	C1orf112	PIMREG	WFDC1	TMPO
	CENPU	CDCA2	DEPDC1B	SPP1	METTL7A	HELLS	CDCA7	FAM162B
	ECE2	CNRIP1	SAMD4A	HYAL1	ESRP1	NEBL	AFF3	PDK1
	GRAMD2A	DONSON	SGO2	IQGAP3	MYRF	CACNA2D2	RAB11FIP1	CYS1
	CKAP2L	PARPBP	ADAM12	PRR11	PDK4	KL	ABCA3	MLLT11
	COCH	NDC1	DUXAP10	MT1M	PPP1R3C	EPB41L5	SULF1	RERG
	TRIM59	MDK	CGNL1	ARNTL2	ZNF367	NEIL3	RNASE4	8-Sep
	GGH	ADGRD1	KCNE1	SLC12A8	IGSF9	TMEM125	SLC7A5	NEGR1
	CPB2	NR4A2	ID4	IGF2BP3	SLC2A5	CA2	FAXDC2	
	CAMK2N1	PFN2	WASF1	LOXL2	RPL39L	CTSV	PID1	MEST
	CABYR	MTFR2	BCL11A	CDK5R1	CDHR3	C12orf56	MACROD2	CHRNA5
	FAM83D	CRABP2	WISP1	NR3C2	HOXA10	ADAMDEC1	PI15	APOD
	SLC22A3	PRAME	WDR72	TMEM158	FAP	ZIC2	APOBEC3B	FNDC1
	DNALI1	HOXD10	ATP8A1	DLX6	MLPH	CNTNAP2	MAGEA6	CAPN8
	HOXC6	DLX5	HES6	HORMAD1	AQP3	MAGEA12	MAGEA1	TSPAN8
	MS4A8	LGSN	CALB1	CDK1	CCNB1	CCNB2	BUB1	CDC20
	MAD2L1	PLK1	AURKB	CCNA2	CENPE	CDCA8	KIF2C	BUB1B
	NDC80	TOP2A	BIRC5	KIF11	CENPF	KIF18A	CENPL	CENPH
	CENPN	CENPI	CENPK	KIF20A	NUF2	ESPL1	DLGAP5	KIF23
	AURKA	ZWINT	CDCA5	ZWILCH	KNTC1	SPC25	SKA1	PRC1
	UBE2C	CHEK1	NCAPG	TPX2	HIST1H2BD	HIST1H2BH	RACGAP1	TTK
	MCM4	NUSAP1	CDC45	CDC6	MCM2	RFC4	RRM2	POLE2
	PTTG1	CDKN3	ASPM	KIF4A	MELK	SMC2	CDT1	FOXM1
	ORC6	SMC4	NEK2	KIF15	MCM8	CDC25A	SKP2	PBK
	CEP55	MCM10	RAD51	ANLN	PIK3R1	HJURP	OIP5	CENPW

	KIF18B	GMNN	KIF26B	DBF4	TYMS	NCAPG2	EXO1	CCNE1
	ITGA1	NUP155	KIF14	RMI2	FEN1	NCAPH	BRIP1	CDC25C
	AGTR1	HMMR	TIMELESS	COL1A1	CKS1B	UBE2S	MYBL2	COL1A2
	CKS2	COL5A2	PLOD2	LMO7	ADRB2	ZBTB16	ARRB1	GINS2
	CCNE2	RAD51AP1	RNF144B	EZH2	FOS	NR3C1	LMNB1	COL3A1
	CBX2	COL10A1	COL11A1	FBXO5	MMP1	NMU	CTTN	DTL
	EPAS1	GINS1	ECT2	GTSE1	AR	SYT1	SSTR1	GAL
	UBE2T	MND1	P2RY14	CITED2	LPL	SERPINA1	VIPR1	RAMP2
	NME1	RAP1A	CLU	AOX1	MYH11	KPNA2	DUSP1	SORBS1
	FANCI	CBX7	CERS6	ASF1B	NME5	SHCBP1	ALDH2	E2F7
	SGMS2	TRIP13	MAOB	TFAP2A	MMP9	MAOA	SCNN1B	FXYD1
	LRRFIP1	SHANK3	GRIA1	ATP1A2	PAICS	PPAT	PRKCE	
	PAFAH1B3	SUV39H2	FGFR4	GATA2	SCN4B	TPD52	PLA2G1B	
	ADAMTSL3	TSC22D3	MMP11	TK1	GREM1	RAD54B	RNASEH2A	LMNB2
	SLC2A1	IGFBP3	DSCC1	STIL	PGF	THBS2	EGLN3	BMP2
	CDKN2A	HMGA2						
Yellow module (GSE32863) (n=79)	CACNA2D2	UHRF1	UBE2T	Pfs2	C9orf140	CDC45L	PGC	
	MGC24665	C17orf53	ECT2	C1QTNF6	SLPI	CDCA7	SUSD2	WDR51A
	A2M	HES6	PGCP	PTTG3	HDC	LEPREL1	FLRT3	TBX2
	NMU	SCG5	RPL39L	SELENBP1	ZNF533	SCGB3A2	RNASE1	VSIG2
	TRIP13	FOLR1	ANG	KCNK12	C4BPA	IGJ	PCP4	TUBB2B
	AGR3	SCGB3A1	TOP2A	CCNB2	CDC20	AURKB	AURKA	CDCA8
	UBE2C	KIF20A	TPX2	BIRC5	PRC1	NUSAP1	CENPF	PTTG1
	MCM2	NEK2	CDCA5	TYMS	MCM4	MELK	ASPM	NUP155
	TUBG1	FEN1	TIMELESS	CCNF	KIFC1	KIF1A	STIL	TK1
	H2AFX	KIAA0101	ALDH2	E2F2	SFTPD	CCNE1	MAOA	SFTPB
Yellow module (GSE63459) (n=160)	ADRB2	MARCKSL1	CES1	IDH2	CSTF2	PSAT1	C6orf125	NOSTRIN
	KIAA0859	ST6GALNAC6		STOM	PLA2G4F	WFS1	MAL	XPO5
	SAP18	TNFSF12	MTA3	MGC13170	NARF	SLC2A1	C17orf53	CALM1
	GSN	CDC45L	MOCS1	ALG8	TSPAN4	C1orf112	TLE4	EDNRA
	C6orf129	IFT57	WDR51A	CIRBP	AHCY	TMEM132A	C20orf20	
	MGC24665	MS4A2	GDDR	TPSAB1	HMGA1	BOLA2	ECRG4	FLJ40629
	HCAP-G	CABYR	PCCB	NUDT1	GPR116	PLOD1	C9orf140	SUSD2
	NCALD	SLPI	C20orf24	ABCA3	VSIG2	PTTG3	KLF2	DLG7
	TROAP	PGCP	TMEM106C	CPA3	SFTPD	MSN	DLC1	
	DKFZp762E1312	SELENBP1	CDC2	NR3C2	HES6	C16orf60	RAFTLIN	C18orf56
	C1orf116	RPS6KA2	HLA-DRB6	CTSH	RNASE1	IGF2BP3	FOXA2	BUB1
	CCNB2	CCNA2	CDC20	AURKB	TOP2A	CDCA8	AURKA	KIF11
	CENPF	KIF2C	KIF20A	TPX2	BIRC5	PRC1	CHEK1	NUSAP1
	UBE2C	CENPM	MCM3	MCM4	MCM2	RFC4	KNTC1	CDCA5
	SPC24	PTTG1	RCC2	TTK	CDKN3	NEK2	PLK4	MCM7
	MELK	CEP55	FOXM1	ASPM	TYMS	MCM6	E2F3	FEN1
	HMMR	TIMELESS	MCM10	ANLN	CDK4	E2F2	EXO1	TUBG1
	CKS1B	CCNE1	CCNF	EZH2	RAB3IP	KIAA0101	RAD51AP1	POLQ
	OIP5	SHMT2	UBE2T	TK1	STIL	CBX2	ALDH2	PAICS
	CCT3	GAPDH	APOA1BP	MRPL12	CD59	RPL34	MRPS17	RPL39L
	TRIP13	MAOA	FOLR1					