High CENPM Gene Expression Predict Poor Survival Outcome in Lung Adenocarcinoma

Zhe Huang
University of South China
Nong Yang
University of South China
Yongchang Zhang (zhangyongchang@csu.edu.cn)
University of South China

Research Article

Keywords: CENPM, TGCA database, GEO database, Lung adenocarcinoma, Prognosis

Posted Date: December 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1109250/v1

Abstract

Introduction: Lung adenocarcinoma is a disease with high morbidity and mortality. The aim of our study was to investigate the relationship between the gene expression of centromere protein M (CENPM) and its prognostic impact in lung adenocarcinoma.

Method: By analyzing the data of lung adenocarcinoma in database, the *CENPM* gene expression in lung adenocarcinoma and its relationship with clinical stage and survival time were analyzed using datasets from The Cancer Genome Atlas (TCGA)and Gene Expression Omnibus (GEO) datasets. Genes associated with *CENPM* expression were analyzed and subjected to functional and pathway enrichment analysis. Finally, the genetic results and treatment and survival outcomes of 20 patients with lung adenocarcinoma from our hospital were analyzed.

Result: *CENPM* transcripts were found to be highly expressed in lung adenocarcinoma as compared with normal tissues (3.628 VS. 2.227, P < 0.001). *CENPM* expression was positively associated with tumor stage (3.803 vs. 3.444, p < 0.001) and nodal stage (3.992 vs. 3.573, p < 0.001). Patients with low CENPM expression achieved better progression-free survival (45.9 months vs. 25.7months, p < 0.001) and overall survival (57.5 months vs. 47.5 months, p=0.001). The *CENPM* expression was negatively correlated with the infiltration of most immune cells in lung adenocarcinoma tissues and positively correlated with PD1 (r = 0.231, p < 0.001) and PD-L1 (r = 0.116, p < 0.007). CENPM-related genes were enriched in the set of genes with poor prognosis as well as the set of cell cycle-related genes in lung adenocarcinoma. *CENPM* expression was also negatively correlated with T lymphocyte and B lymphocyte signaling pathways. Finally, *CENPM*-related genes were related in Rho GTPases and ATR signaling pathways.

Conclusion: Our findings demonstrate that *CENPM* gene is highly expressed and is associated with poor prognosis in lung adenocarcinoma.

Introduction

Lung cancer is one of the most common malignancies threatening human health (1, 2). Among them, the incidence of lung adenocarcinoma accounts for more than half of the total population of lung cancer. Although remarkable progress has been achieved in the diagnosis and treatment of lung cancer, it remains one of the most fatal and poor prognosis diseases. Nowadays, the discovery of new targets, the study of new anti-tumor mechanisms, and the establishment of new predictive models have become one of the main works in anti-tumor exploration.

Centromere protein M (CENPM) is a component of the CENPA-NAC (nucleosome-associated) complex, a complex that plays a central role in the assembly of kinetochore proteins, mitotic progression, and chromosome segregation. The CENPA-NAC complex recruits the CENPA-CAD (nucleosome distal) complex and may be involved in incorporating newly synthesized CENPA into centromeres (3). In recent years, *CENPM* gene has been studied in multiple cancer types (4, 5); however, no study has explored the correlation between *CENPM* and lung cancer prognosis. In order to deeply understand the prognostic

impact of *CENPM* in lung adenocarcinoma, we used various tumor databases to analyze the gene expression level of *CENPM* in normal and tumor tissue samples, survival results, and various biomarkers of the tumor microenvironment. We also constructed a preliminary prediction model and validated it based on gene sequencing of 20 patients with lung adenocarcinoma from our hospital.

Method

Data collection

The gene expression data for *CENPM* was obtained as level 3 open access RNA-sequencing data (normalized fragments per kilobase of transcript per million mapped reads [FPKM]) from the Cancer Genome Atlas (TCGA) (https://genomecancer.ucsc.edu/), including the lung adenocarcinoma (LUAD) cohort. Clinical factors include gender, stage, age, grade, tumor (T) stage, metastasis (M) stage, node (N) stage, survival status, and survival time. Initial validation was performed using gene expression data derived from normal tissues from the genotype-tissue expression (GTEx) database combined with the TCGA-LUAD dataset. The differential gene expression data were obtained from three independent lung adenocarcinoma datasets, including GSE43458 (6), GSE32863 (7), and GSE140797(8), from the gene expression omnibus (GEO) database.

CENPM **expression in lung adenocarcinoma**

Training set: The gene expression of *CENPM* from the TCGA-LUAD cohort was compared with that in tissue messenger RNA (mRNA), and the means of the two groups were compared by independent sample t-test and paired sample t-test. *CENPM* expression of normal tissues from the GTEx database and lung adenocarcinoma tissues from the TCGA database were compared. *CENPM* expression in lung adenocarcinoma tissues and normal tissues from 3 independent datasets from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) were further analyzed for validation.

Survival Analysis

The *CENPM* gene expression in lung adenocarcinoma tissues from the TCGA database was compared in groups. Kaplan Meier survival curves comparing the overall survival (OS), progression-free survival (PFS) of patients with lung adenocarcinoma according to level of *CENPM* gene expression. A p value of < 0.05 indicated statistical difference.

Immune Infiltrates Analysis

The *CENPM* expression inform the TCGA-LUAD dataset was correlated with 24 immune cell markers. Mann-Whitney U test (Wilcoxon rank sum test) was used to analyze the relationship between high and low *CENPM* expression and immune cell infiltration in tumor tissues (9).

Gene ontology (GO)/Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Set Enrichment Analysis (GSEA)

The top 20 genes identified to have either positive or negative correlation with *CENPM* gene expression were selected for GO/KEGG enrichment analysis.

GSEA created a list of all gene permutations related to *CENPM* expression. The samples were then divided according to level of *CENPM* expression as high and low to distinguish potential functions and use GSEA to clarify significant survival differences. Genome replacement is performed multiple times with each exam. The level of *CENPM* expression was used as a phenotypic marker. Normalized enrichment scores (NES) and nominal P-values were used to classify the signaling pathways for the enrichment of each phenotype.

Prognostic Model

The clinical characteristics of the patients, including gender, age, smoking history, clinical stage, and TNM stage, as well as the survival time of the patients, were downloaded from the TCGA database. Univariate analysis was performed using. The receiver operating characteristic (ROC) curve was plotted using *CENPM* expression as a single factor to calculate the area under the curve (AUC). Multivariate analysis was performed using Cox proportional hazards regression method to understand the prognostic impact of various clinical factors, *CENPM* gene expression level, and overall survival. Hazard ratio (HR) and corresponding 95% confidence intervals (CI) were calculated. A nomogram model was constructed for predicting overall survival probabilities based on the results of the multivariate Cox regression analysis.

Statistical analysis

Statistical analysis was performed using R version 3.6.3 software for statistical analysis, and GraphPadPrism8 software for mapping. Normally distributed data were expressed as mean ± standard deviation (x ± s), and compared using t-test. Paired sample t-test was used for comparing the same individual value at different time points. The rank sum test was used to compare groups of data that did not conform to a normal distribution. $P < 0.05$ was considered as statistically significant.

Results

Expression of CENPM in lung adenocarcinoma

Using the TCGA database, we analyzed the *CENPM* gene expression in various tumors and normal tissue samples. We found that *CENPM* gene was highly expressed in almost all tumor types, including lung adenocarcinoma (n=515) (Figure 1A). *CENPM* was found to be highly expressed in lung adenocarcinoma by analyzing all the differentially expressed genes between lung adenocarcinoma and normal tissues from the database (Figure 1B). From the TCGA database comprised of 535 lung adenocarcinoma

samples and 59 normal tissue samples, we could see that *CENPM* expression was significantly higher in lung adenocarcinoma tissues (3.628 VS. 2.227, P < 0.001; Figure 1C, Table 1). Consistently, *CENPM* was highly expressed in lung adenocarcinoma samples from the TCGA database compared with normal tissue samples from the GTEx database (3.628 vs. 2.139, p < 0.001; Figure 1D). Next, we analyzed 57 matched lung adenocarcinomas and normal tissues from the TCGA database and found that the *CENPM* expression in tumor tissues was significantly higher than that in paired normal tissues (3.709 VS. 2.234, p < 0.001; Figure 1E, Table 1).

These findings were further validated using combined data from three independent datasets derived from GSE43458, GSE32863, and GSE140797 datasets. *CENPM* expression was significantly higher in lung adenocarcinoma (n=145) as compared with normal tissue samples (n=95) (3.012 vs. 2.901, p<0.001; Figure 1F, Table S1). Next, we compared 65 paired tumor and normal samples from the same dataset and saw consistent results that lung adenocarcinoma tissues had higher *CENPM* expression than their paired adjacent normal lung tissues (3.089 vs 2.921, p<0.001; Figure 1G, Table S1).

Next, we analyzed the diagnostic value of *CENPM* expression for lung adenocarcinoma. The area under the ROC curve was 0.874 (95% CI: 0.834 – 0.913). From the results, it can be seen that the predictive ability of *CENPM* expression was accurate in distinguishing between tumor and normal tissues. Using an optimal *CENPM* gene expression cut-off of 2.919, the sensitivity was 86.4% and specificity was 78.9% in distinguishing between tumor and normal tissues (Figure 1H). Taken together, our analysis showed that *CENPM* was highly expressed in lung adenocarcinoma than normal tissues.

Table 1
CENPM expression in TCGA database

	Unpaired Samples			**Paired samples**		
	Normal(n=59)	Tumor(n=535)	p	Normal(n=57)	Tumor(n=57)	p
Mean	2.227	3.682	<0.001	2.234	3.709	<0.001
SD	0.749	0.999		0.761	0.968	
Max	4.005	6.722		4.005	6.16	
Min	0.663	0.721		0.663	1.901	

Cenpm Expression And Clinical Characteristics

Table 2 shows the baseline clinical characteristics of patients with lung adenocarcinoma from the TCGA database. According to 50% of the population, 267 patients had low *CENPM* expression and 268 patients had high *CENPM* expression. No statistical difference was found in age, gender, and smoking history between the two groups. However, TNM and clinical stage were significantly different between the two groups.

We also analyzed the mRNA expression of bronchial epithelial cells from smokers with and without lung cancer using the GSE4115 dataset (10). We found that *CENPM* expression was significantly higher in bronchial epithelial cells from patients diagnosed with lung tumors than patients without lung cancer (2.99 VS. 2.52, $p < 0.001$; (Figure S1), suggesting that *CENPM* gene expression is cancer-specific and not due to smoking.

Next, we analyzed the *CENPM* expression with the baseline clinicopathological status of the patients. The gene expression of *CENPM* in pathological stage II-IV was significantly higher than that in stage I (3.919 vs. 3.622, $p = 0.015$, Figure 2A). Similarly, patients with T2-4 had significantly higher *CENPM* expression than patients with T1 (3.803 vs. 3.444, $p < 0.001$, Figure 2B), and patients with N1-3 had significantly higher *CENPM* expression than patients with N0 (3.992 vs. 3.573, $p < 0.001$, Figure 2C). However, *CENPM* expression was not statistically different in patients with or without metastasis (4.098 versus 3.716, $p = 0.060$, Figure 2D). Taken together, these data demonstrate that CENPM expression is positively correlated with advanced stage and is cancer-specific.

Table 2
Characteristic of Baseline

Characteristic	CENPM Low (n=267)	CENPM High(n=268)	p
Age, median (range)	67 (60, 73)	65 (57, 71)	0.057
Gender, n (%)			0.068
Female	157 (58.8%)	129 (48.1%)	
Male	110 (41.2%)	139 (51.9%)	
Smoker, n (%)			1.000
No	38 (14.2%)	37 (13.8%)	
Yes	224 (85.8%)	222 (86.2%)	
T stage, n (%)			0.002
T1	108 (20.3%)	67 (12.6%)	
T2	126 (23.7%)	163 (30.6%)	
T3	24 (4.5%)	25 (4.7%)	
T4	8 (1.5%)	11 (2.1%)	
N stage, n (%)			< 0.001
N0	191 (36.8%)	157 (30.3%)	
N1	33 (6.4%)	62 (11.9%)	
N2	30 (5.8%)	44 (8.5%)	
N3	0 (0%)	2 (0.4%)	
M stage, n (%)			0.022
M0	180 (46.6%)	181 (46.9%)	
M1	6 (1.6%)	19 (4.9%)	
Pathologic stage, n (%)			< 0.001
Stage I	171 (32.4%)	123 (23.3%)	
Stage II	50 (9.5%)	73 (13.9%)	
Stage III	34 (6.5%)	50 (9.5%)	
Stage IV	7 (1.3%)	19 (3.6%)	

According to *CENPM* expression from the TCGA database, we analyzed the patient survival time between the two groups. Patients with low *CENPM* expression achieved significantly longer overall survival (OS) time than those with high *CENPM* expression (57.5 months vs. 47.5 months, p = 0.001, Figure 3A). The median progression free survival (PFS) was 45.9 months (95%CI: 33.6-83.9) for patients with low *CENPM* expression and 25.7 months (21.2-36.0) for patients with high expression, which showed that the survival of those with low *CENPM* expression group would be longer (p<0.001, Figure 3B). The survival analysis was further validated using an independent dataset derived from GSE68465 (11). The data were divided into 2 groups according to *CENPM* expression, with 148 patients in each group. The OS of *CENPM* low expression group was 78 months (95%CI: 65-105.4), while that of high expression group was 48 months (95%CI: 40-130). Patients with low *CENPM* expression achieved a better survival time than those with high *CENPM* expression (p= 0.033, Figure 3C).

Table 3 shows the treatment response of the two groups. In *CENPM* low expression group, 183 patients achieved complete response (CR), 4 patients had partial response (PR), 19 patients had stable disease (SD), and 19 patients had progressive disease (PD). The objective response rate (ORR) was 83.2% and disease control rate (DCR) was 91.6% for the low *CENPM* expression group. In *CENPM* high expression group, 149 patients were CR, 2 patients were PR, 18 patients were SD, and 52 patients were PD. The ORR was 68.3% and DCR was 76.5% for the high *CENPM* expression group. Patients with low *CENPM* expression achieved significantly better ORR and DCR than those with high *CENPM* expression (p<0.001 and p< 0.001).

Table 4 summarizes the Cox regression analyses of the basic clinical characteristics of patients from the TCGA database for OS. We found that the patient's age, gender, and smoking history did not directly affect the patient's OS in this cohort. T, M, N stage as well as pathological stage were identified as prognostic factors for OS, wherein the more advanced the stage, the worse the OS. *CENPM* expression was identified as an independent prognostic factor for lung adenocarcinoma (HR=1.665, 95% CI: 1.242-2.233, p< 0.001, Figure 3D). The statistically significant factors based on the Cox regression analyses, including T, M, N stage, and the expression of *CENPM*, were selected as variables to construct an OS prediction model (Figure 3E). Taken together, our data indicate that high *CENPM* expression is associated with poor prognosis in lung adenocarcinoma.

Table 3
Treatment outcome

	Low expression(n=225)	High expression(n=221)	p
CR	183(81.4%)	149(67.4%)	
PR	4(1.8%)	2(0.9%)	
SD	19(8.4%)	18(8.2%)	
PD	19(8.4%)	52(23.5%)	
ORR	187(83.2%)	151(68.3%)	<0.001
DCR	206(91.6%)	169(76.5%)	<0.001

Table 4
Cox Regression of Clinical Characteristics

Characteristics	Total(N)	HR(95% CI)	P value
Age	516		
<=65	255	Reference	
>65	261	1.223 (0.916-1.635)	0.172
Gender	526		
Female	286	Reference	
Male	249	1.070 (0.803-1.426)	0.642
Smoker	512		
No	75	Reference	
Yes	446	0.894 (0.592-1.348)	0.591
T stage	523		
T1	175	Reference	
T2	289	1.521 (1.068-2.166)	0.020
T3&T4	68	3.066 (1.950-4.823)	<0.001
N stage	510		
N0	348	Reference	
N1	95	2.382 (1.695-3.346)	<0.001
N2&N3	76	2.968 (2.040-4.318)	<0.001
M stage	377		
M0	361	Reference	
M1	25	2.136 (1.248-3.653)	0.006
Pathologic stage	518		
Stage I&Stage II	417	Reference	
Stage III&Stage IV	110	2.664 (1.960-3.621)	<0.001
CENPM	526		
Low	268	Reference	
High	267	1.665 (1.242-2.233)	<0.001

Expression Of Cenpm And Immune Infiltration

Next, we analyzed the immune infiltration data for the lung adenocarcinoma samples from the TCGA database. We found that the *CENPM* expression of CENPM was negatively correlated with the infiltration level of most immune cells (Figure 4A), such as macrophages (r = -0.120, p = 0.006, Figure 4B), B lymphocytes (r = -0.110, p = 0.015 Figure 4C), T lymphocytes (r = -0.096, p = 0.027, Figure 4D), and CD8 + T cells (r = -0.120, p = 0.004, Figure 4E). These results suggest an inhibitory status of the tumor immune microenvironment of lung adenocarcinomas with high *CENPM* expression.

We further analyzed the current possible correlation of immune checkpoint inhibition. High *CENPM* expression was positively correlated with PD-1 (*PDCD1*, r = 0.231, p < 0.001), and PD-L1 (*CD274*, r = 0.116, p = 0.007), but not correlated with *TGFB1*, and *CTLA4* (Figure 4F, Table S2). These associations may serve as potential biomarkers for response with immune checkpoint inhibitors.

Cenpm-related Genes And Enrichment

We next analyzed the genes correlated with *CENPM* expression. The top 10 positively correlated genes and the top 10 negatively correlated genes were selected (Figure 5A). Further GO/KEGG enrichment of genes associated with *CENPM* expression identified the genes involved in the cell cycle-related pathways (Table S4), and genes associated with platinum resistance (Figure 5B).

CESA enrichment was further performed for genes associated with *CENPM* expression in all TCGA databases. We found that *CENPM* expression was enriched mainly in cyclins, and most genes were associated with poor prognosis in lung cancer. Furthermore, the *CENPM* expression was negatively correlated with T cell and B cell signaling pathways (Figure 5C).

Cenpm Gene Expression And Real-world Lung Adenocarcinoma

In our center, 20 patients with lung adenocarcinoma who submitted for gene expression profiling were analyzed. Among them, 6 patients had stage III and 14 patients had stage IV lung adenocarcinoma (Table S3). Four of these patients were treated with surgery and one with concurrent chemoradiotherapy. We found that *CENPM* expression was significantly higher in tumor tissues (n = 20) compared with adjacent non-cancerous tissues (n = 8) (2.767 vs. 1.337, p = 0.002; Figure 6A-B). We next analyzed the survival outcomes of these patient received chemotherapy. Patients with high expression of CENPM achieved poor PFS (12.5 months VS. 5.8 months, p=0.02, Figure 6C). However, OS was not reached and was not included in the analysis.

Discussion

Our work evaluated the expression of *CENPM* in lung adenocarcinoma form the TCGA database, and its relationship with clinical features and prognosis. We also validated these findings using multiple independent dataset form the GEO database and a small cohort from our hospital. Moreover, the mechanism of the prognostic impact of high *CENPM* expression in lung adenocarcinoma was elucidated using differential expression analysis. Our findings revealed that the high gene expression of *CENPM* was significantly correlated with the development and prognosis of lung adenocarcinoma, and may reveal potential biomarkers or prognostic indicators.

We found that *CENPM* gene expression was significantly higher in lung adenocarcinoma than that in normal tissues from the TCGA database, GEO database, and our real-world cohort. We further investigated the ability of *CENPM* expression in distinguishing between lung adenocarcinoma tissues and normal tissues, which yielded an AUC of 0.874, indicating that *CENPM* expression could be a biomarker for screening of lung adenocarcinoma. In addition, the expression of CENPM is positively associated with more advanced stages of the disease, which also suggests that *CENPM* expression may be associated with the development or progression of lung adenocarcinoma. We further demonstrated the poor prognosis associated with high *CENPM* expression in lung adenocarcinoma using independent datasets from the TCGA and GEO datasets. Cox regression results showed that clinical stage and *CENPM* expression were independent predictors of poor prognosis in lung adenocarcinoma. We also preliminarily established a OS prediction model based on the results of these analyses. From the small real-world cohort, we have also observed a trend of poorer prognosis in patients with high *CENPM* expression (>3 TPM). However, the small sample size limits our conclusion.

The high *CENPM* expression was negatively correlated with the infiltration of most immune cells, suggesting that the poor prognosis of lung adenocarcinoma with high *CENPM* expression might be due to an altered tumor microenvironment. Further analysis revealed that *CENPM* expression was correlated with immune checkpoint-related molecules PD-1 and PD-L1. In the current treatment of lung cancer, the treatment with immune checkpoint inhibitors improves the prognosis of many patients (12–14). Multiple monoclonal antibodies against PD-1 and PD-L1 are also widely used as first-line treatment of lung cancer in clinical practice (15–18). In view of this, *CENPM* gene expression may serve as a biomarker for immunotherapy response in lung adenocarcinoma.

Furthermore, the *CENPM* expression was negatively correlated with the signaling pathways of B lymphocytes and T lymphocytes, which was consistent with the reduced T and B lymphocyte infiltration level we previously described, and this inhibition may have a prognostic impact in lung cancer patients (19–21). In addition, high *CENPM* expression in lung adenocarcinoma was correlated with enrichement in the RHO GTPases pathway and ATR signaling pathway. Rho GTPases are highly expressed in many malignancies and are closely associated with tumor invasion and metastasis (22–24). Similarly, the ATR signaling pathway has an important role for maintaining DNA stability in tumor cells (25–27). In lung adenocarcinoma with high *CENPM* expression, targeted therapy targeting these two signaling pathways may be a good therapeutic option (28–30).

Since a majority of our data were derived from publicly available databases, bias may occur in practice. Our study also included a small cohort from our hospital, which severely limited the clinical validation of our findings gathered from data mining. In addition, our study only analyzed the gene expression data and other clinical and survival correlations. The lack of other molecular data, including DNA level and protein level data, also severely limited our conclusion. The clinical impact of high CENPM expression in lung adenocarcinoma and their associated mechanisms still need further validation.

Conclusion

Our study found that the mRNA expression of CENPM was significantly increased in patients with lung adenocarcinoma. High *CENPM* expression was found to be associated with a variety of clinical and molecular characteristics, including alteration in immune infiltration level and poor prognosis. Hence, *CENPM* may be a useful prognostic biomarker for patients with lung adenocarcinoma.

Declarations

Acknowledgements

The authors would like to thank Dr. Analyn Lizaso for her editing assistance.

Authors' contributions

Yongchang Zhang: Responsible for conceptualization, organization, data collection, auditing, supervision, project management, funding acquisition, writing review and editing.

Zhe Huang: Responsible for data curation, methodology, formal analysis, original draft preparation, writing review and editing.

Nong Yang: Responsible for critical comments and suggestions, writing review and editing.

All authors approved the final version of the manuscript.

Funding

This work received financial support from the Natural Science Foundation of Hunan Province (2020SK2031, kq1801102, 2020SK2030, 2018RS3106, 2019-TJ-N04, 2019SK4010, and 2020JJ3025), and CAS “Light of West China Program”. The funding agencies had no role in the study design, data collection, analysis, interpretation, manuscript writing, and decision to submit the article for publication.

Competing interests

All authors declare no Competing Financial or Non-Financial Interests.

Consent for publication

All patients provided written informed consent to take part in the study.

Ethics approval and consent to participate

Approval was obtained from the Hunan Cancer Hospital Institutional Review Board Committee (2017YYQ-SSB-274).

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7-30.
2. Thai AA, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung cancer. Lancet. 2021;398(10299):535-54.
3. Foltz DR, Jansen LE, Black BE, Bailey AO, Yates JR, 3rd, Cleveland DW. The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol. 2006;8(5):458-69.
4. Zheng C, Zhang T, Li D, Huang C, Tang H, Ni XF, et al. Upregulation of CENPM facilitates tumor metastasis via the mTOR/p70S6K signaling pathway in pancreatic cancer. Oncol Rep. 2020;44(3):1003-12.
5. Zou Y, Sun Z, Sun S. LncRNA HCG18 contributes to the progression of hepatocellular carcinoma via miR-214-3p/CENPM axis. J Biochem. 2020;168(5):535-46.
6. Kabbout M, Garcia MM, Fujimoto J, Liu DD, Woods D, Chow CW, et al. ETS2 mediated tumor suppressive function and MET oncogene inhibition in human non-small cell lung cancer. Clin Cancer Res. 2013;19(13):3383-95.
7. Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res. 2012;22(7):1197-211.
8. Liang J, Li H, Han J, Jiang J, Wang J, Li Y, et al. Mex3a interacts with LAMA2 to promote lung adenocarcinoma metastasis via PI3K/AKT pathway. Cell Death Dis. 2020;11(8):614.
9. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39(4):782-95.
10. Gustafson AM, Soldi R, Anderlind C, Scholand MB, Qian J, Zhang X, et al. Airway PI3K pathway activation is an early and reversible event in lung cancer development. Sci Transl Med. 2010;2(26):26ra5.
11. Director's Challenge Consortium for the Molecular Classification of Lung A, Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. Nat Med. 2008;14(8):822-7.
12. Brueckl WM, Ficker JH, Zeitler G. Clinically relevant prognostic and predictive markers for immune-checkpoint-inhibitor (ICI) therapy in non-small cell lung cancer (NSCLC). BMC Cancer.

2020;20(1):1185.
13. Wang C, Qiao W, Jiang Y, Zhu M, Shao J, Wang T, et al. The landscape of immune checkpoint inhibitor plus chemotherapy versus immunotherapy for advanced non-small-cell lung cancer: A systematic review and meta-analysis. J Cell Physiol. 2020;235(5):4913-27.
14. Nabet BY, Esfahani MS, Moding EJ, Hamilton EG, Chabon JJ, Rizvi H, et al. Noninvasive Early Identification of Therapeutic Benefit from Immune Checkpoint Inhibition. Cell. 2020;183(2):363-76 e13.
15. Paz-Ares L, Vicente D, Tafreshi A, Robinson A, Soto Parra H, Mazieres J, et al. A Randomized, Placebo-Controlled Trial of Pembrolizumab Plus Chemotherapy in Patients With Metastatic Squamous NSCLC: Protocol-Specified Final Analysis of KEYNOTE-407. J Thorac Oncol. 2020;15(10):1657-69.
16. Herbst RS, Giaccone G, de Marinis F, Reinmuth N, Vergnenegre A, Barrios CH, et al. Atezolizumab for First-Line Treatment of PD-L1-Selected Patients with NSCLC. N Engl J Med. 2020;383(14):1328-39.
17. Hellmann MD, Paz-Ares L, Bernabe Caro R, Zurawski B, Kim SW, Carcereny Costa E, et al. Nivolumab plus Ipilimumab in Advanced Non-Small-Cell Lung Cancer. N Engl J Med. 2019;381(21):2020-31.
18. Yang Y, Wang Z, Fang J, Yu Q, Han B, Cang S, et al. Efficacy and Safety of Sintilimab Plus Pemetrexed and Platinum as First-Line Treatment for Locally Advanced or Metastatic Nonsquamous NSCLC: a Randomized, Double-Blind, Phase 3 Study (Oncology pRogram by InnovENT anti-PD-1-11). J Thorac Oncol. 2020;15(10):1636-46.
19. Stankovic B, Bjorhovde HAK, Skarshaug R, Aamodt H, Frafjord A, Muller E, et al. Immune Cell Composition in Human Non-small Cell Lung Cancer. Front Immunol. 2018;9:3101.
20. Muenst S, Laubli H, Soysal SD, Zippelius A, Tzankov A, Hoeller S. The immune system and cancer evasion strategies: therapeutic concepts. J Intern Med. 2016;279(6):541-62.
21. Gaissmaier L, Christopoulos P. Immune Modulation in Lung Cancer: Current Concepts and Future Strategies. Respiration. 2020:1-27.
22. Haga RB, Ridley AJ. Rho GTPases: Regulation and roles in cancer cell biology. Small GTPases. 2016;7(4):207-21.
23. Svensmark JH, Brakebusch C. Rho GTPases in cancer: friend or foe? Oncogene. 2019;38(50):7447-56.
24. Bustelo XR. RHO GTPases in cancer: known facts, open questions, and therapeutic challenges. Biochem Soc Trans. 2018;46(3):741-60.
25. Yazinski SA, Zou L. Functions, Regulation, and Therapeutic Implications of the ATR Checkpoint Pathway. Annu Rev Genet. 2016;50:155-73.
26. Smith J, Tho LM, Xu N, Gillespie DA. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. Adv Cancer Res. 2010;108:73-112.
27. Lecona E, Fernandez-Capetillo O. Targeting ATR in cancer. Nat Rev Cancer. 2018;18(9):586-95.

28. Zheng CW, Zeng RJ, Xu LY, Li EM. Rho GTPases: Promising candidates for overcoming chemotherapeutic resistance. Cancer Lett. 2020;475:65-78.
29. Qiu Z, Oleinick NL, Zhang J. ATR/CHK1 inhibitors and cancer therapy. Radiother Oncol. 2018;126(3):450-64.
30. Havill AM, Riggs D, Pitt BR, Gillis CN. Resolution of impaired pulmonary function and pulmonary hypertension after phorbol ester administration in rabbits. Am Rev Respir Dis. 1989;140(3):782-8.

Figures

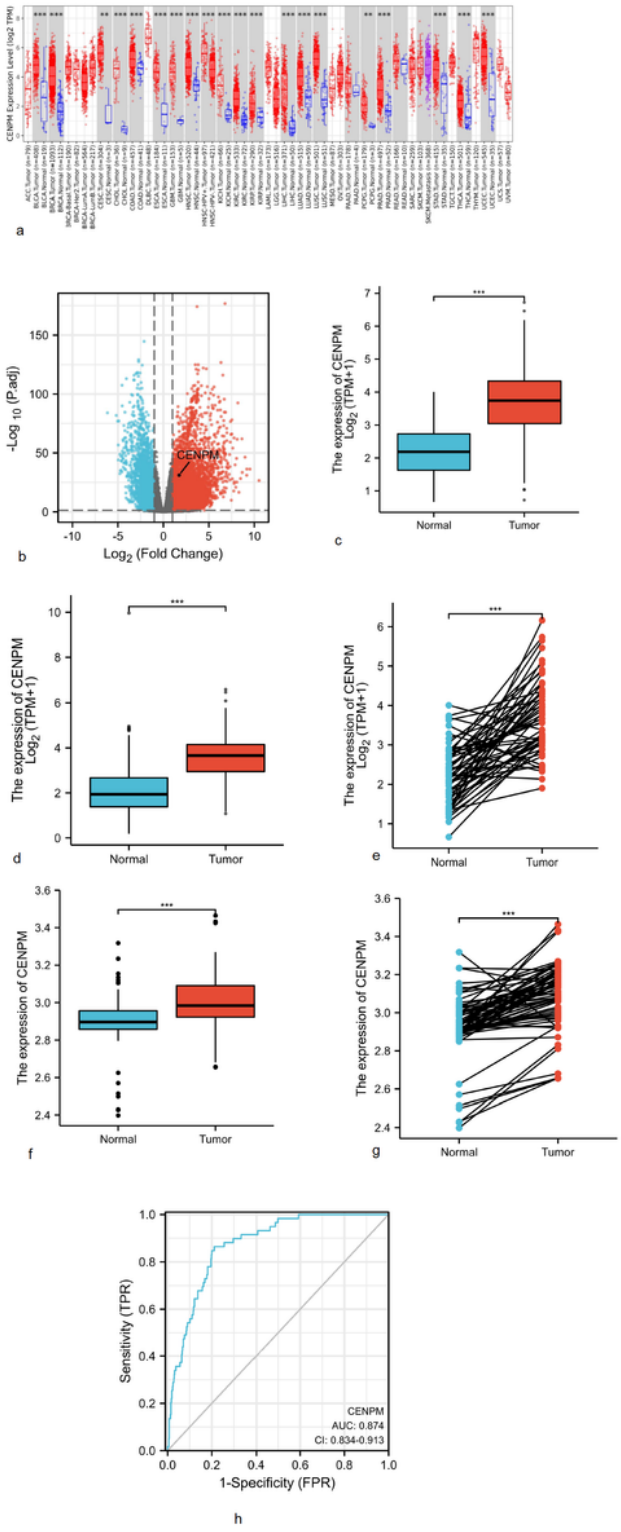

Figure 1

CENPM gene is highly expressed in lung adenocarcinoma. A. *CENPM* gene expression across tumor types and normal tissue, including lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). B. volcano map of genes associated with prognosis of lung adenocarcinoma. C. *CENPM* expression in lung adenocarcinoma (n=535) and normal tissues (n=59) in the TCGA database. D. *CENPM* expression in lung adenocarcinoma (TCGA, n=535) and normal tissues (GTEx, n=347). E. *CENPM* expression in paired

samples from the TCGA database (n= 57). F. *CENPM* expression in lung adenocarcinoma (n=145) and normal tissues (n=95) in the GEO database. G. *CENPM* expression in paired samples from the TCGA database (n= 65). H. Receiver Operative Characteristic (ROC) curve of *CENPM* gene expression for screening lung adenocarcinoma. Abbreviations: AUC, area under the curve; FPR, false positive rate; TPR, true positive rate; TPM, transcript per million. Asterisks denote the degree of statistical significance; ***, p<0.001 and **, p<0.01.

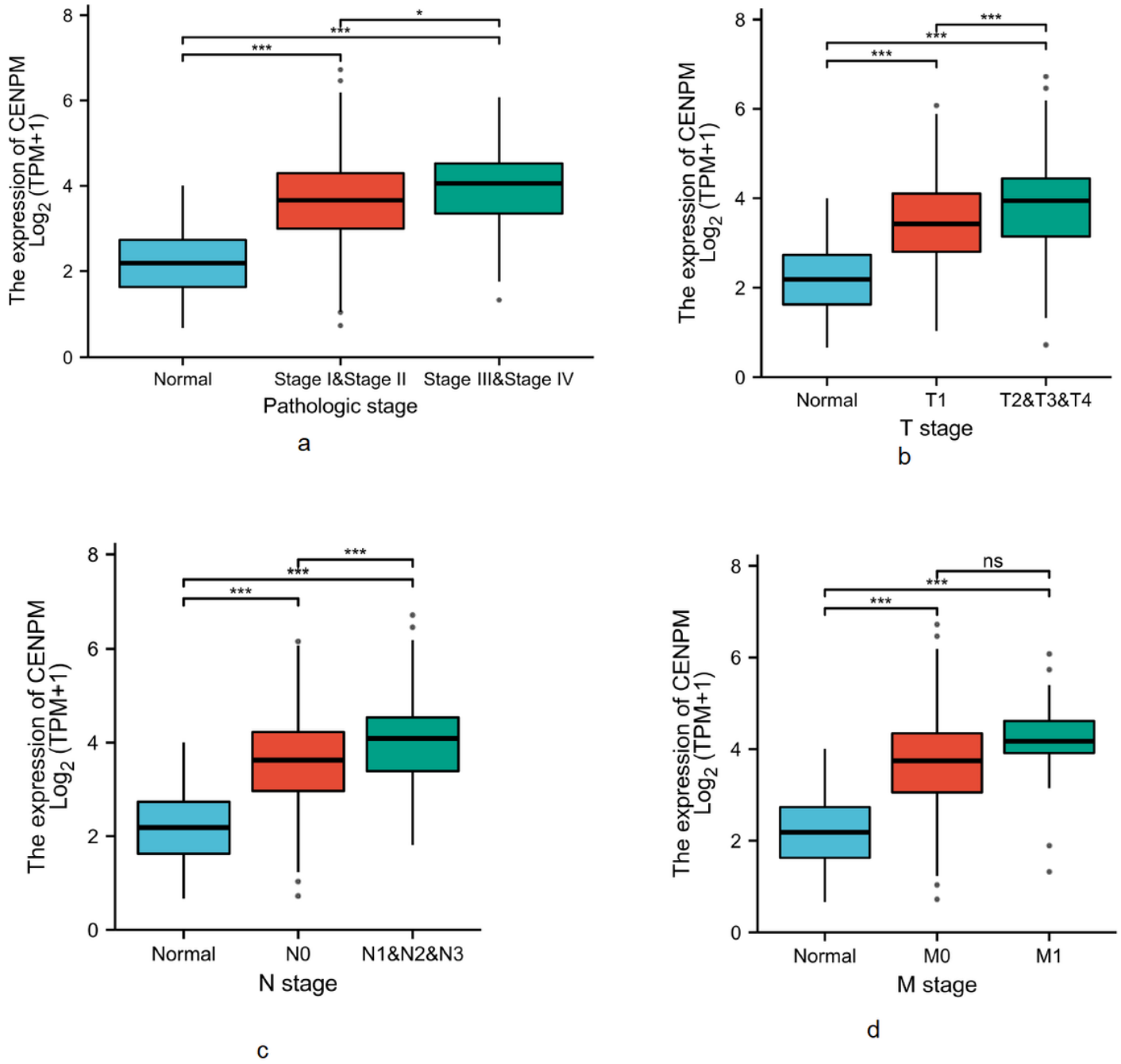

Figure 2

Positive association between *CENPM* gene expression and clinical stage. A-D. Comparison of *CENPM* expression between normal tissue and tumor across various pathologic stage (A); tumor (T) stage (B); lymph node (N) stage (C); and metastatic (M) stage (D). Abbreviations: TPM, transcript per million. Asterisks denote the degree of statistical significance; ***, p<0.001; *, p<0.05; ns, not statistically significant (p≥0.05).

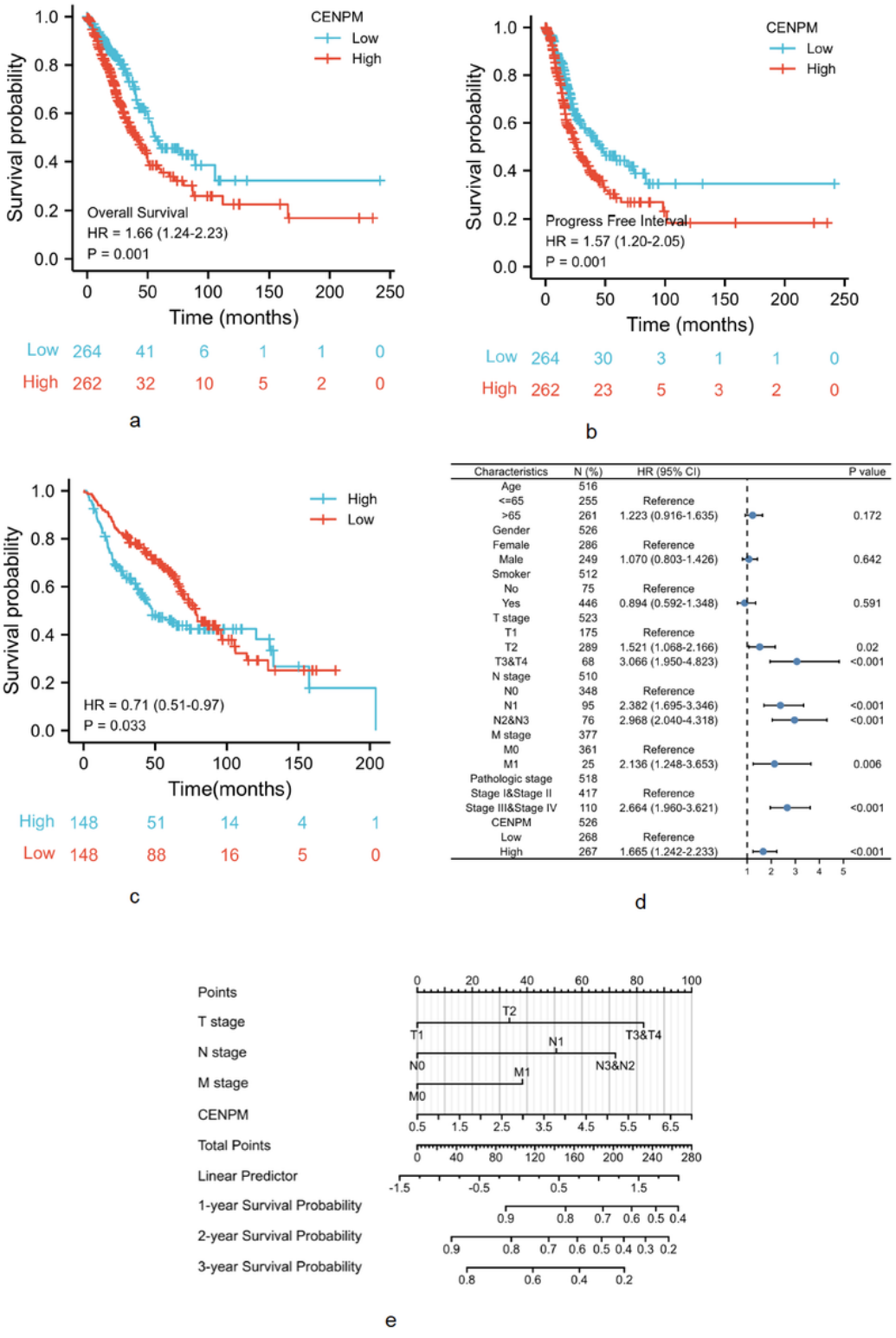

Figure 3

High *CENPM* gene expression with poor prognosis in lung adenocarcinoma. A. Kaplan-Meier survival curves for overall survival (OS) (A, C), and progression-free survival (PFS) of patients from the TCGA Database (A, B) and GEO Database (C) with high and low *CENPM* gene expression. D. Forest plot demonstrating the association between various clinical characteristics and OS. Abbreviations: CI, confidence intervals; HR, hazard ratio. E. Nomogram for predicting 1-year, 2-year, and 3-year OS probability for lung adenocarcinoma using Tumor (T), Node (N), Metastasis (M) stage, and *CENPM* gene expression as variables.

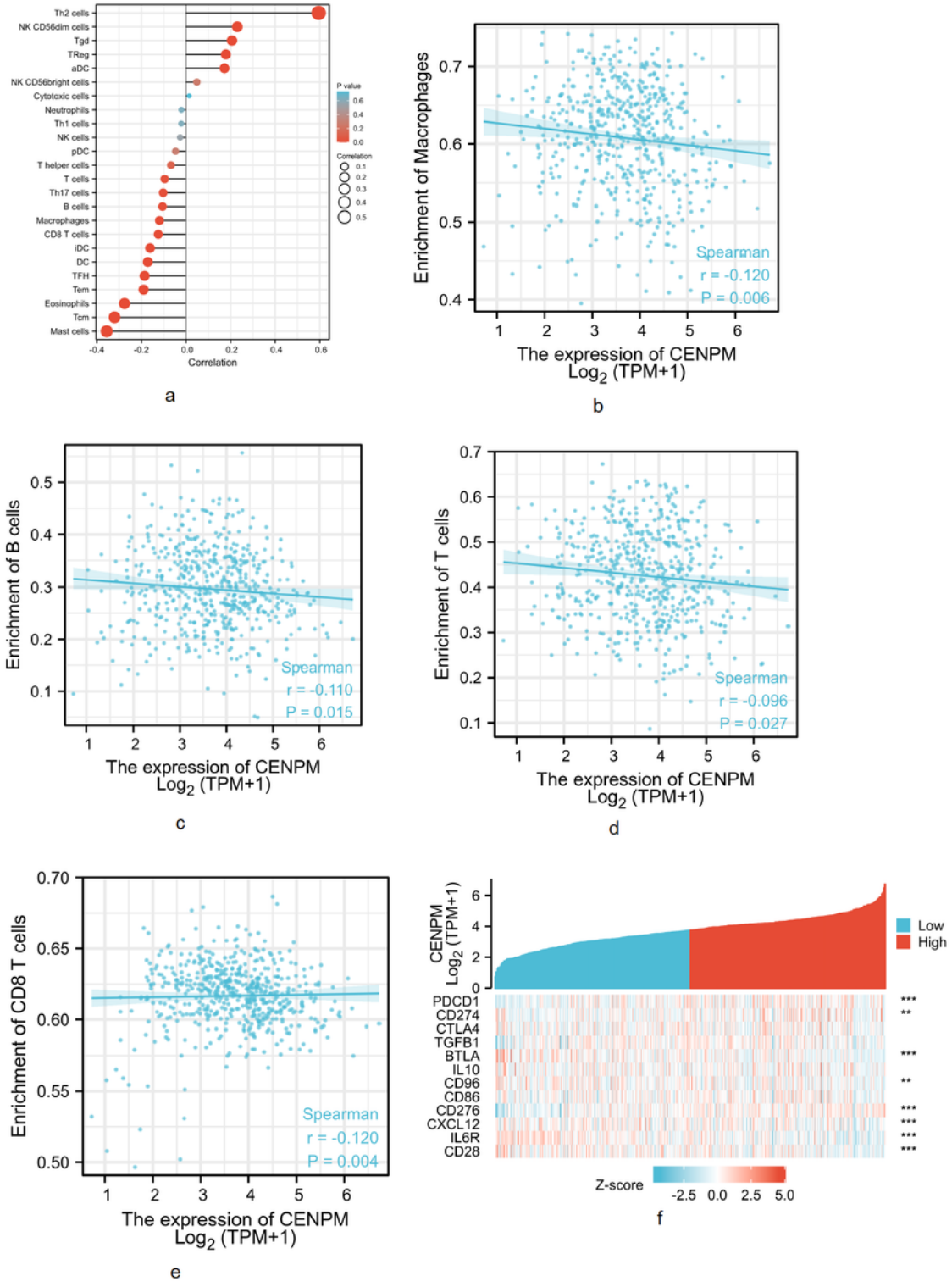

Figure 4

The association between *CENPM* gene expression and immune infiltration in lung adenocarcinoma. A-E: Relationship between *CENPM* expression and various immune cells in immune microenvironment (A), including macrophages (B), B cells (C), T cells (D), and CD8+ T-cells (E). F. Association between *CENPM* Expression and immune checkpoint genes. Abbreviations: TPM, transcript per million.

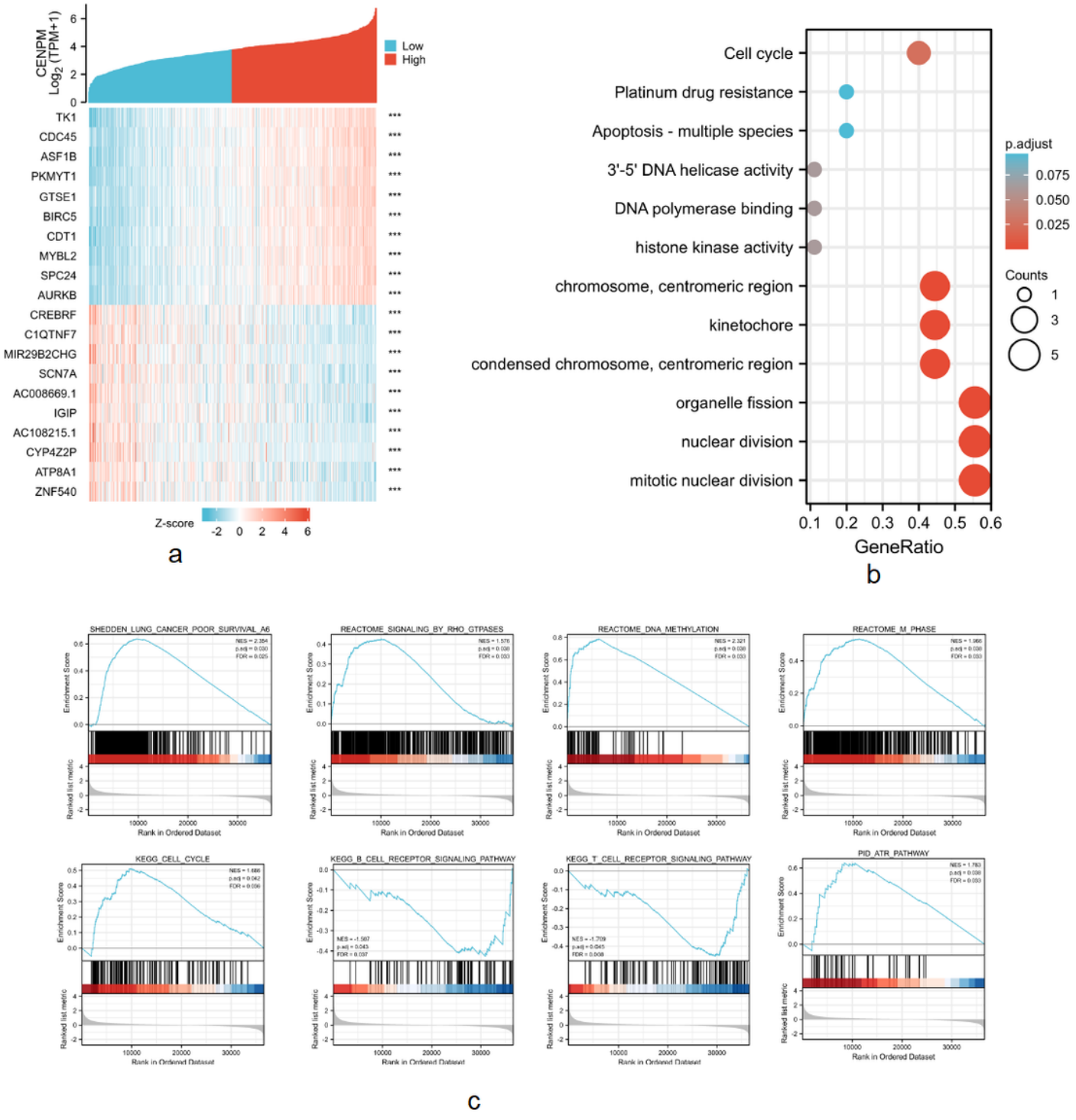

Figure 5

Differential expression and enrichment of *CENPM*. A-C. Differentially expressed genes (A) and signaling pathways (B) associated with *CENPM* gene expression. C. Gene Set Enrichment Analysis of various differentially expressed signaling pathways associated with *CENPM* expression. Abbreviations: FDR, false discovery rate; NES, normalized enrichment score; p.adj, adjusted p-value; TPM, transcript per million. Asterisks denote the degree of statistical significance; ***, p<0.001.

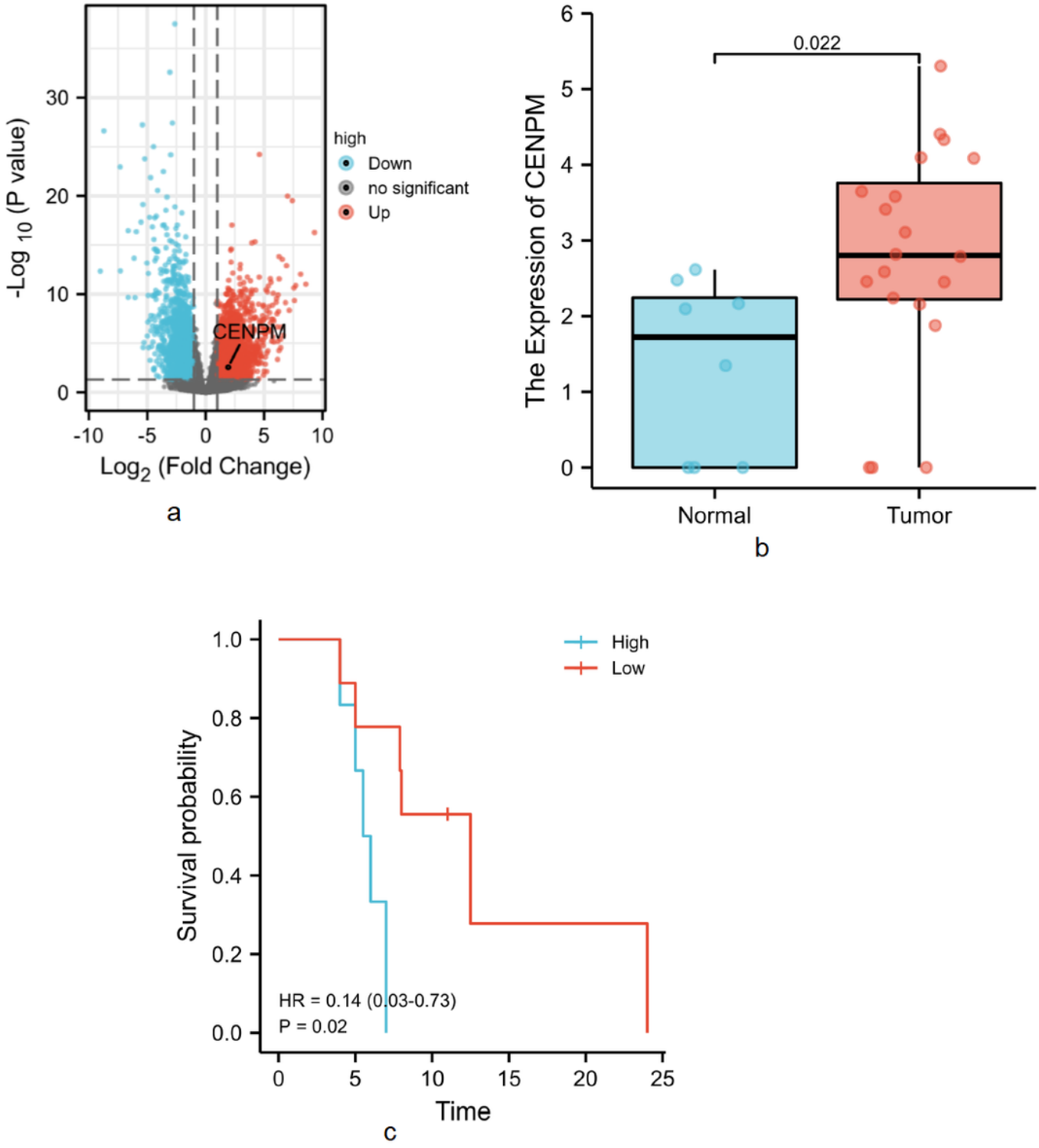

Figure 6

CENPM expression in real-world patients. A volcano map of genes different expression of lung adenocarcinoma and adjacent non-cancerous tissues. B. *CENPM* expression in lung adenocarcinoma (n=20) and normal tissues (n=) C. Kaplan-Meier survival curve comparing the overall survival (PFS) of patients with high (n=13) and low (n=5) *CENPM* expression.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.pptx
- TableS14.pdf