Journal of Biosciences and Medicines, 2025, 13(10), 76-95
https://www.scirp.org/journal/jbm
ISSN Online: 2327-509X
ISSN Print: 2327-5081

Comprehensive Analysis of Prognostic and Diagnostic Value, Immune Infiltration, and Co-Regulation Networks of Centromere Protein Family Members in Pancreatic Ductal Adenocarcinoma

Guohao Yin[1], Tao Liu[2*], Wei Qin[2*]

[1]Clinical Medical College (Affiliated Hospital), Jining Medical University, Jining, China
[2]Hepatobiliary Surgery Department, Jining Medical University Affiliated Hospital, Jining, China
Email: *liutaogdwk@163.com, *qinw0218@163.com

How to cite this paper: Yin, G.H., Liu, T. and Qin, W. (2025) Comprehensive Analysis of Prognostic and Diagnostic Value, Immune Infiltration, and Co-Regulation Networks of Centromere Protein Family Members in Pancreatic Ductal Adenocarcinoma. *Journal of Biosciences and Medicines*, **13**, 76-95.
https://doi.org/10.4236/jbm.2025.1310007

Received: August 20, 2025
Accepted: October 10, 2025
Published: October 13, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Aim: The centromere protein family members (CENPs) are a mitosis-associated protein family. Accumulative evidence indicates that CENPs are abnormally expressed and affect the occurrence and development of a variety of tumors, but their functions in pancreatic ductal adenocarcinoma (PDAC) remain unclear. In this study, we systematically analyzed CENPs in PDAC. **Methods:** Gene Expression Profiling Interactive Analysis (GEPIA), Kaplan-Meier plotter, cBioPortal, Oncomine, TIMER, Metascape, and NetworkAnalyst were used to analyze differential expression, prognostic and diagnostic value, genetic alteration, immune cell infiltration, functional enrichment, and co-regulation networks of CENPs in PDAC patients. The higher expression levels and prognostic values of CENPs in PDAC were determined by GEPIA and Kaplan-Meier Plotter. **Results:** The 7 hub CENPs (CENPA, CENPE, CENPF, CENPK, CENPL, CENPM, and CENPN) that predict worse prognosis caused by genetic alteration were identified by cBioPortal. Overexpression of these genes in PDAC was further confirmed by the Oncomine and GEO databases. The diagnostic values of CENPs were evaluated by ROC curves. The expression of CENPs was also significantly correlated with the infiltration of diverse immune cells in PDAC. Finally, we constructed interaction and co-regulation networks of CENPs and analyzed their functions. **Conclusion:** Taken together, these results indicated that 7 hub CENPs (CENPA, CENPE, CENPF, CENPK, CENPL, CENPM, and CENPN) could be prognostic and diagnostic biomarkers as well as possible synergistic targets of immunotherapy for PDAC patients.

Keywords

Centromere Protein Family Members, Pancreatic Ductal Adenocarcinoma, Prognosis, Diagnostic Value, Immune Infiltration

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC), which accounts for approximately 90% of all pancreatic cancers, is a leading cause of cancer-related deaths worldwide. [1] [2] Although the treatment of pancreatic cancer has made great progress in recent years, the therapeutic effects of PDAC are still unsatisfactory due to a lack of effective early diagnostic markers, a high post-surgical recurrence rate, chemoresistance, and immunotherapy resistance. [3] [4] Therefore, we extended the research field to PDAC based on a variety of databases, with the purpose of determining the potential prognostic indicators and therapeutic targets.

The centromere is the junction of two sister chromatids of metaphase chromosomes that specifies the mitotic behavior of chromosomes. Centromere protein family members (CENPs) are a family of genes encoding centromeric proteins, which are involved in multiple biological functions. [5]-[7] Several CENPs were reported to play an important role in a variety of tumors, such as CENPA in lung cancer and liver cancer, CENPF and laryngeal carcinoma, CENPU and ovarian cancer, etc. [8]-[12] There are only a few articles that have preliminarily explored the relationship between some CENPs and PDAC. For example, some articles point out that the compound UA62784 is a novel fluorenone with inhibitory activity against the centromere protein E (CENP-E) kinesin-like protein, and its activity suggests a potential role for antimitotic drugs in treating pancreatic carcinomas. Some studies have also indicated that CENP-F and CENP-M are associated with PADC. [13]-[16] The relationship between them remains to be further systematically elucidated so as to elaborate the value of CENPs in the prognosis evaluation and treatment of PDAC.

With the development of high-throughput sequencing technology, the specific mechanisms of tumorigenesis and the development of various tumors have been systematically elucidated by integrated bioinformatics analysis. [17]-[19] In this study, we extended the knowledge of the role of CENPs in PDAC based on various large databases for conducting a comprehensive analysis of the pathogenesis and progression of PDAC.

2. Methods

2.1. Gene Expression Profiling Interactive Analysis (GEPIA)

GEPIA is an interactive web server that compares mRNA expression data based on thousands of tumor and normal tissue samples from the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTX) dataset. [20] More than 35 members of the CENPs family are currently known. In this study, a differential expres-

sion analysis was conducted on 22 major centriole proteins in 179 pancreatic cancer tissues and 171 normal pancreatic tissues. [21]

2.2. Kaplan-Meier Plotter

Kaplan-Meier plotter is a commonly used web-based tool for assessing the effect of genes on survival based on the TCGA database. [22] The effects of the differentially expressed genes on overall survival (OS) and relapse-free survival (RFS) were determined by Kaplan-Meier survival analysis. The log-rank P value and hazard ratio (HR) with 95% confidence intervals are shown on the plot.

2.3. cBioPortal

cBioPortal is a comprehensive web resource that provides us with visual and multidimensional cancer genomics data based on the TCGA database. In our study, genetic alterations of CENPs and their effects on mRNA expression, protein expression, and prognosis were obtained from cBioPortal.

2.4. Oncomine

The mRNA levels of differentially expressed CENPs in PDAC were further determined through the Oncomine database, which is a publicly accessible online database providing powerful, genome-wide expression analysis with cancer gene expression information. [23]

2.5. GEO Dataset

The GEO dataset (GDS) is the expression data that NCBI staff organized according to the data uploaded by users. There are only two GEO datasets comparing the differentially expressed mRNA between pancreatic cancer and normal pancreas, GDS4102 and GDS4336, [24] [25] which were enrolled in this study.

2.6. ROC Curve

The ROC curve is a common method to evaluate diagnostic value. In this study, the clinical data of 167 patients (normal pancreas) in the GTX database and 179 patients (PDAC) in the TCGA database were used to draw ROC curves.

2.7. TIMER2.0

TIMER2.0 is a comprehensive resource for systematic analysis of immune infiltrates across diverse cancer types. [26] Associations between immune infiltrates and CENPs were identified via the "Gene" module, and associations between immune infiltrates and clinical outcome were identified via the "Outcome" module.

2.8. Metascape

Metascape is a user-friendly gene analysis tool for gene annotation and analysis. [27] In this study, Metascape was used to conduct enrichment analysis and network construction. Specifically, for each given gene list, pathway, and process enrichment

analysis has been carried out with the following ontology sources: KEGG Pathway, GO Biological Processes, Reactome Gene Sets, Canonical Pathways, CORUM, TRRUST, DisGeNET, PaGenBase, Transcription Factor Targets, and COVID. All genes in the genome have been used as the enrichment background. Terms with a p-value < 0.01, a minimum count of 3, and an enrichment factor > 1.5 are collected and grouped into clusters. Kappa scores are used as the similarity metric when performing hierarchical clustering on the enriched terms, and sub-trees with a similarity of >0.3 are considered a cluster. The most statistically significant term within a cluster is chosen to represent the cluster. PPI enrichment analysis was performed using the following databases: BioGrid, InWeb-IM, and OmniPath. Further, the MCODE algorithm was applied to identify densely connected network components.

2.9. NetworkAnalyst

NetworkAnalyst is a comprehensive network visual analytics platform for gene expression analysis. [28] PPI and transcriptional co-regulatory network of CENPs were constructed using the online NetworkAnalyst database.

2.10. Patient and Public Involvement

None.

3. Results

3.1. Identification of Differentially Expressed CENPs in PDAC

In order to explore the distinct expression profile of the 22 CENPs in PDAC patients, the GEPIA database was used. Fifteen CENPs were identified as being significantly differentially expressed between pancreatic ductal adenocarcinoma and normal pancreatic tissues, and all of them were highly expressed in cancer (**Figure 1**). The significantly differentiated members of the CENP family were as follows: CENPA, CENPB, CENPE, CENPF, CENPH, CENPK, CENPL, CENPM, CENPN, CENPO, CENPR, CENPU, CENPV, CENPW, and CENPX.

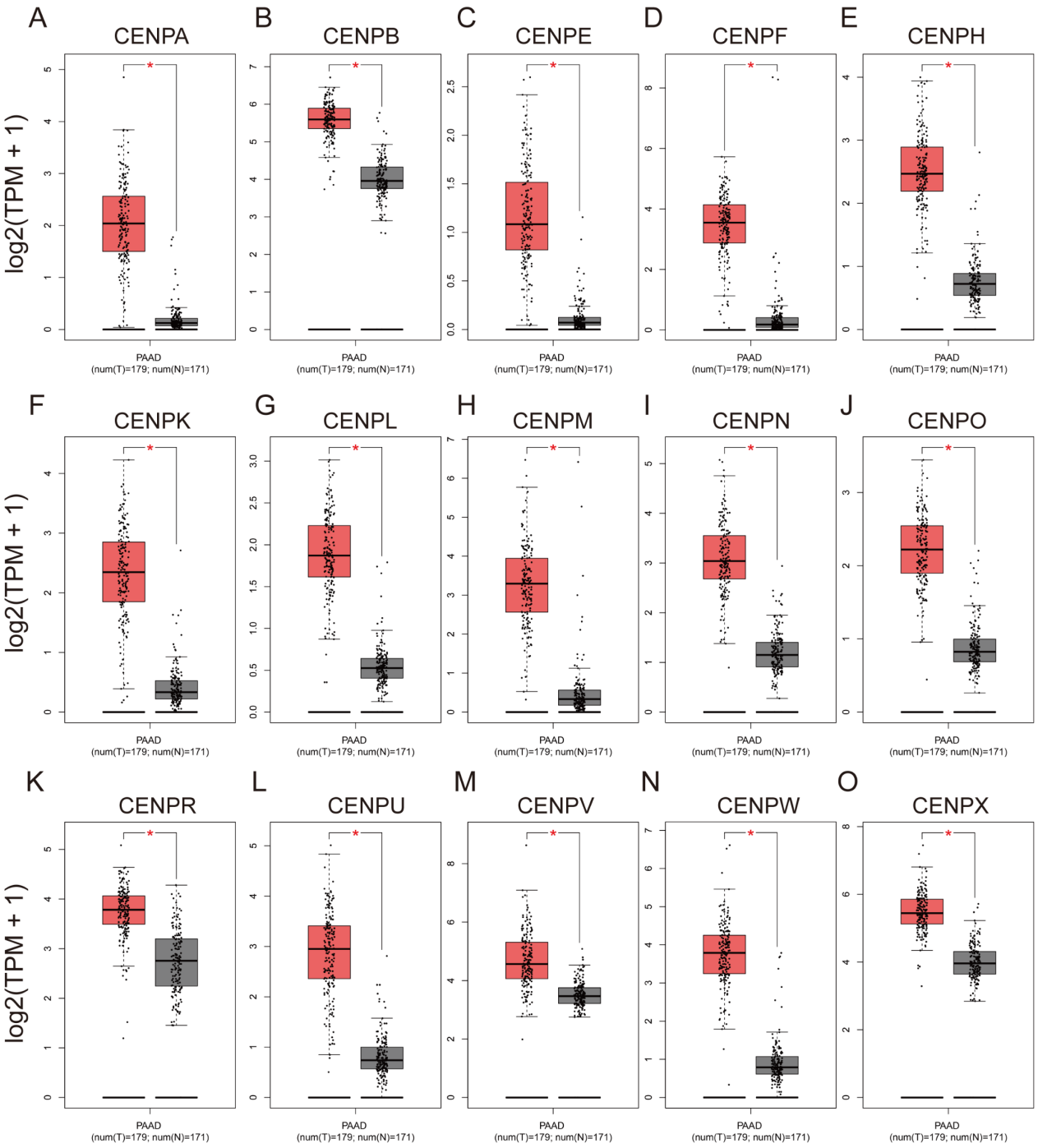
3.2. Prognostic Value of CENPs

To evaluate the value of differentially expressed CENPs in the progression of PDAC, correlations between CENPs and clinical outcomes were analyzed by Kaplan-Meier Plotter. Fourteen CENPs were enrolled, except CENPR, which had no prognostic information in Kaplan-Meier Plotter. As shown in **Figure 2**, twelve CENPs were significantly associated with OS, and ten CENPs were significantly associated with RFS (**Figure 3**). There are nine genes in their intersection, including CENPA, CENPB, CENPE, CENPF, CENPK, CENPL, CENPM, CENPN, and CENPW, which are used for further study.

3.3. Genetic Variations and the Prognostic Values of Differentially Expressed CENPs

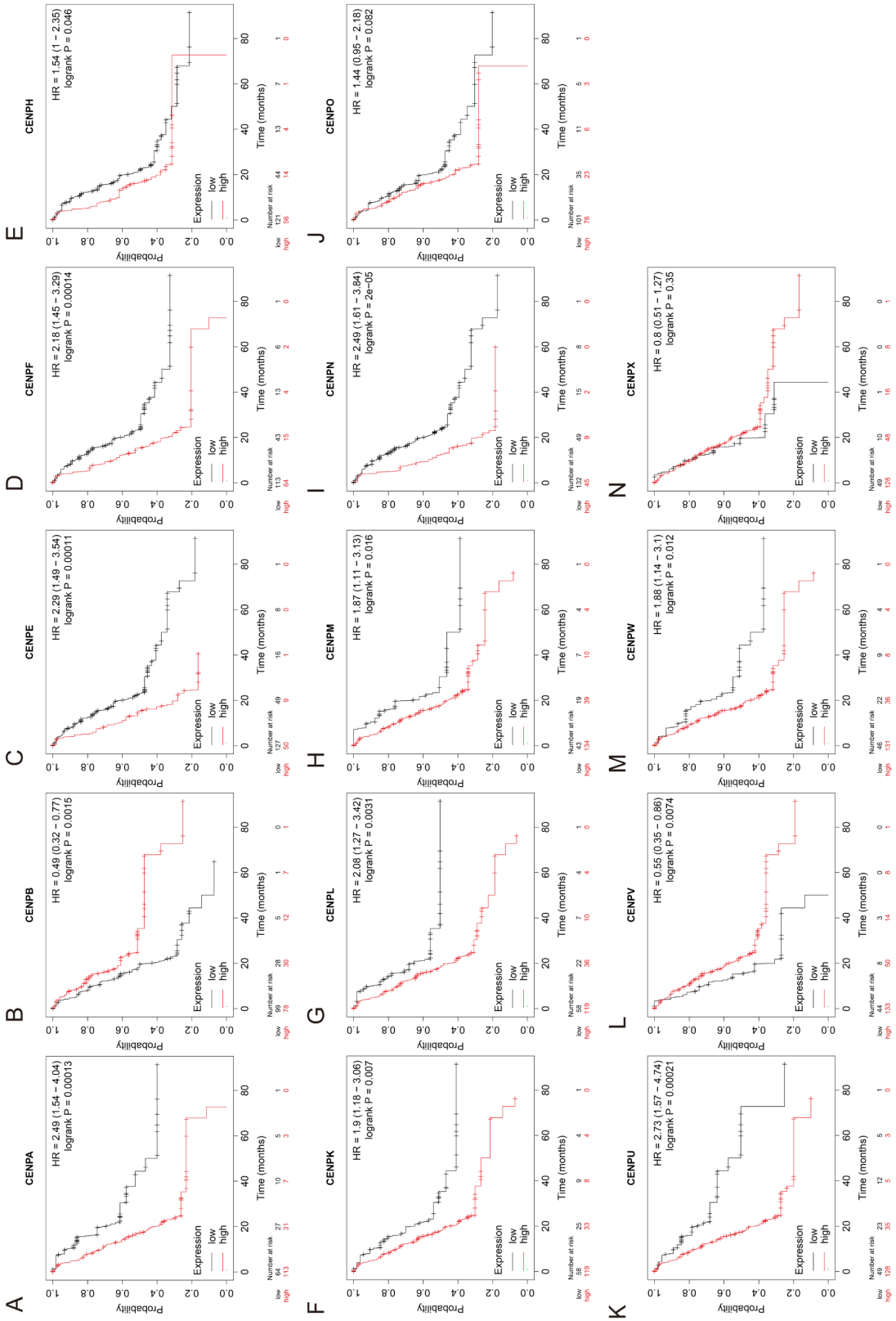
The cBioPortal database was used to determine if the upregulation of CENPs in

PDAC was caused by genetic alteration. As shown in Figure 4(A), a high mutation rate of CENPs was observed in PDAC patients, and the specific mutation information for each gene is shown in Figure 4(B). Additionally, the relationship between overall mutation and prognosis of these CENPs was further explored. We



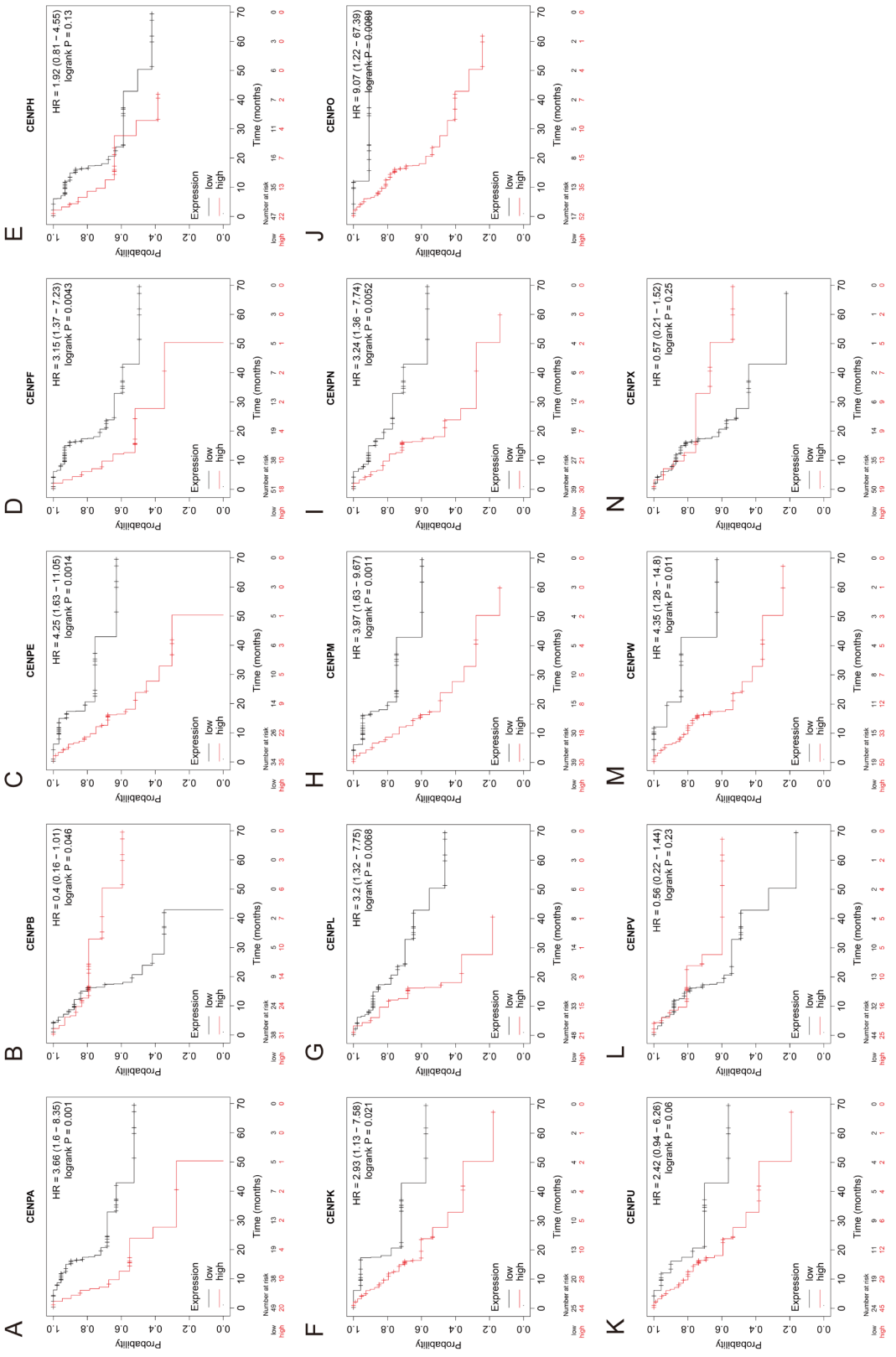
Note: *p < 0.01. CENPs: Centromere Protein Family Members; PDAC: Pancreatic Ductal Adenocarcinoma; GEPIA: Gene Expression Profiling Interactive Analysis.

Figure 1. Expression levels of CENPs in PDAC (GEPIA). (A) - (O) mRNA expression levels of CENPs in pancreatic cancer and normal pancreatic tissues.



Note: CENPs: Centromere Protein Family Members; OS: Overall Survival; PDAC: Pancreatic Ductal Adenocarcinoma.

Figure 2. Prognostic values (OS) of CENPs in PDAC patients (Kaplan-Meier plotter). (A) - (N) Effect of CENPs mRNA level on OS rate in PDAC patients.



CENPs: Centromere Protein Family Members; RFS: Relapse-Free Survival; PDAC: Pancreatic Ductal Adenocarcinoma.

Figure 3. Prognostic values (RFS) of CENPs in PDAC patients (Kaplan-Meier plotter). (A) - (N) Effect of CENPs mRNA levels on RFS rate in PDAC patients.

found that there was no significant difference between the overall mutation of these 9 CENPs and prognosis (data not shown). We selected 7 genes (CENPA, CENPE, CENPF, CENPK, CENPL, CENPM, and CENPN) that have a significant correlation with the survival rate of patients in terms of their overall genetic variations. But it is noteworthy that when combined with 7 genes for analysis, a high genetic alteration rate of CENPs was associated with shorter OS (**Figure 4(C)**) and DFS (**Figure 4(D)**) of PDAC patients. Therefore, all the patients were divided into an altered group and a non-altered group according to whether there was a gene mutation of the 7 hub CENPs, which was used for further analysis. Gene mutation often causes functional changes by altering the expression levels of RNA and protein molecules. The changes of RNA and protein expression caused by the seven CENP gene mutations are shown in **Figure 4(E)** and **Figure 4(F)**. Next, we analyzed the enrichment of significantly upregulated RNA and protein molecules by Metascape (**Figure 4(G)**, **Figure 4(H)**). The enrichment terms networks were visualized in **Figure 4(I)** and **Figure 4(J)**. These results were helpful in further understanding and exploring the mechanisms of CENP mutation causing poor prognosis in patients with PDAC.

3.4. Validation of the Expression Levels of the 7 Hub CENPs in PDAC

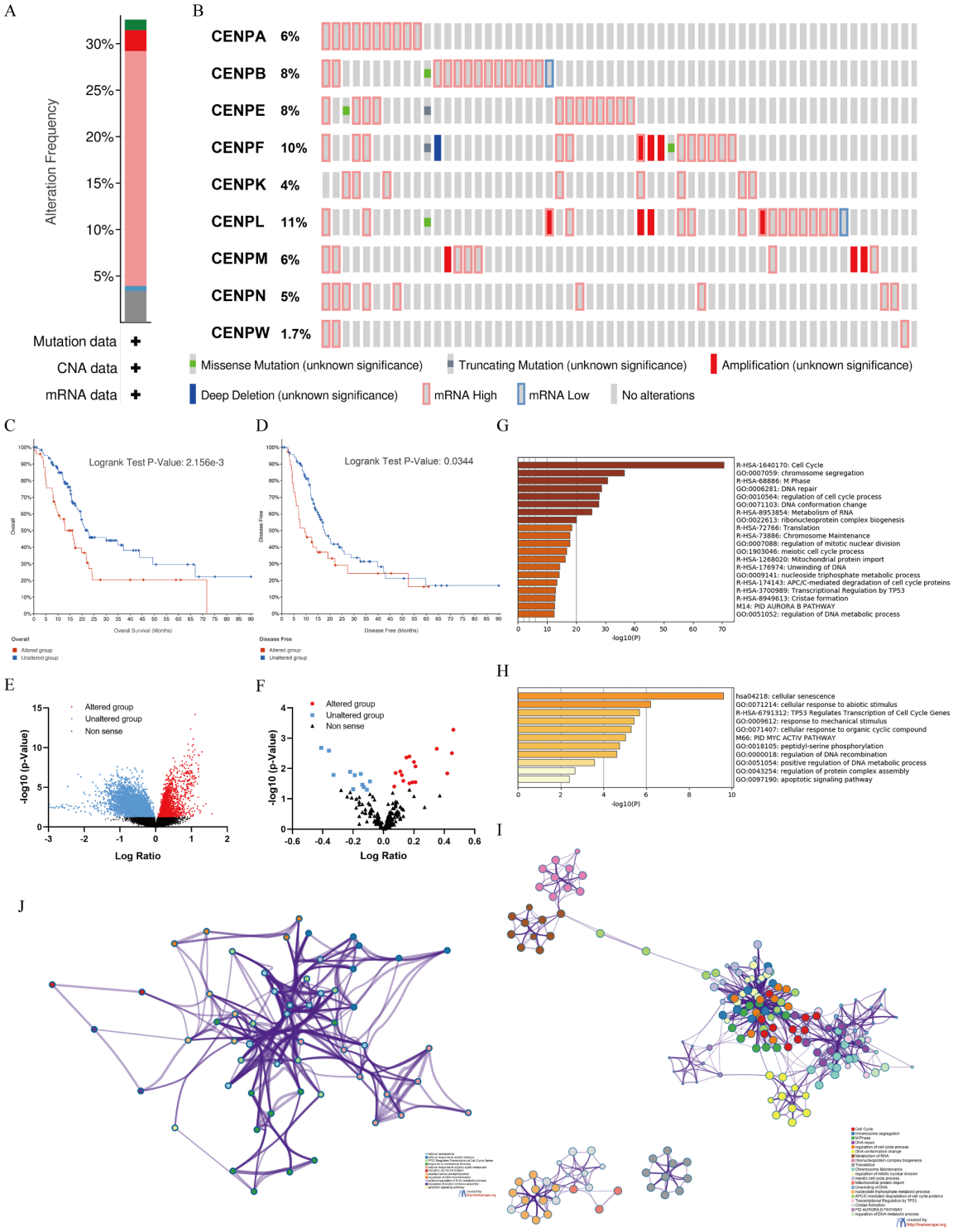
To further validate the mRNA expression levels of the 7 hub CENPs between PDAC and normal pancreatic tissues, Oncomine and GEO databases were used. As displayed in **Figure 5**, all 7 CENPs were significantly highly expressed in PDAC tissues, consistent with previous results. As shown in **Figure 6**, 6 CENPs were significantly highly expressed in PDAC tissues except for CENPA, which has no expression data in GDS4336.

3.5. Diagnostic Values of CENPs in PDAC

The clinical data of 167 patients (normal pancreas) in the GTX database and 179 patients (PDAC) in the TCGA database were used to draw ROC curves for evaluating the diagnostic values of CENPs in PDAC. As shown in **Figure 7**, all seven CENPs showed good diagnostic value for PDAC.

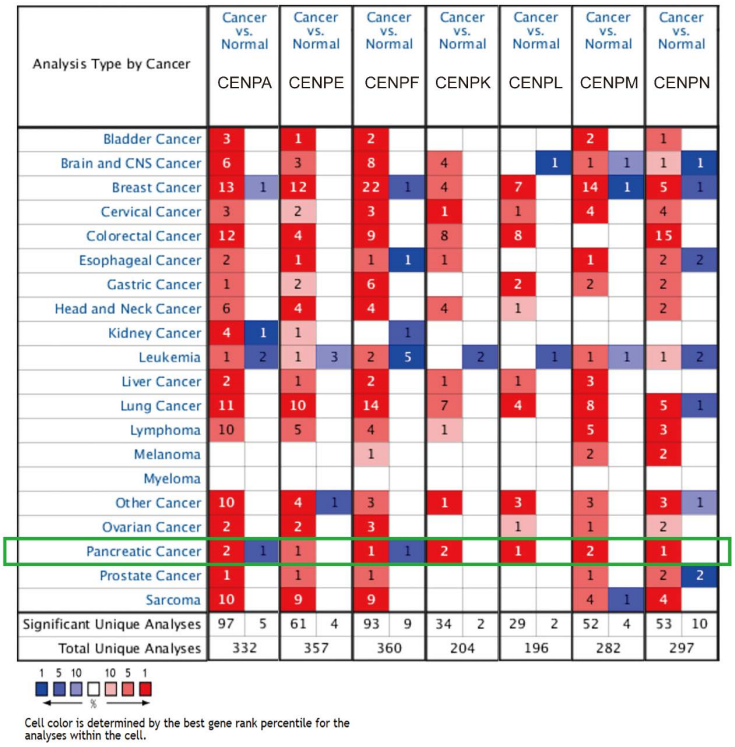
3.6. Immune Cell Infiltration of CENPs in Patients with PDAC

Tumorigenesis is often accompanied by immune dysfunction, and immunotherapy has become an increasingly important part of the treatment of a variety of tumors. [29]-[31] However, the response of patients with different tumor types to immunotherapy varies greatly, so it is urgent to explore the immunological characteristics of different tumors and the corresponding targets of immunotherapy in order to improve their immunotherapeutic response. In this study, we explored the relationship between CENPs and the main infiltrating immune cells in PDAC using the TIMER database. As shown in **Figure 8(A)**, the expression of CENPA was significantly related to the infiltration of $CD8^+$ T cells, $CD4^+$ T cells, regulatory T cells, monocytes, and macrophages. CENPE was related to the infiltration of $CD8^+$ T cells, B cells, and



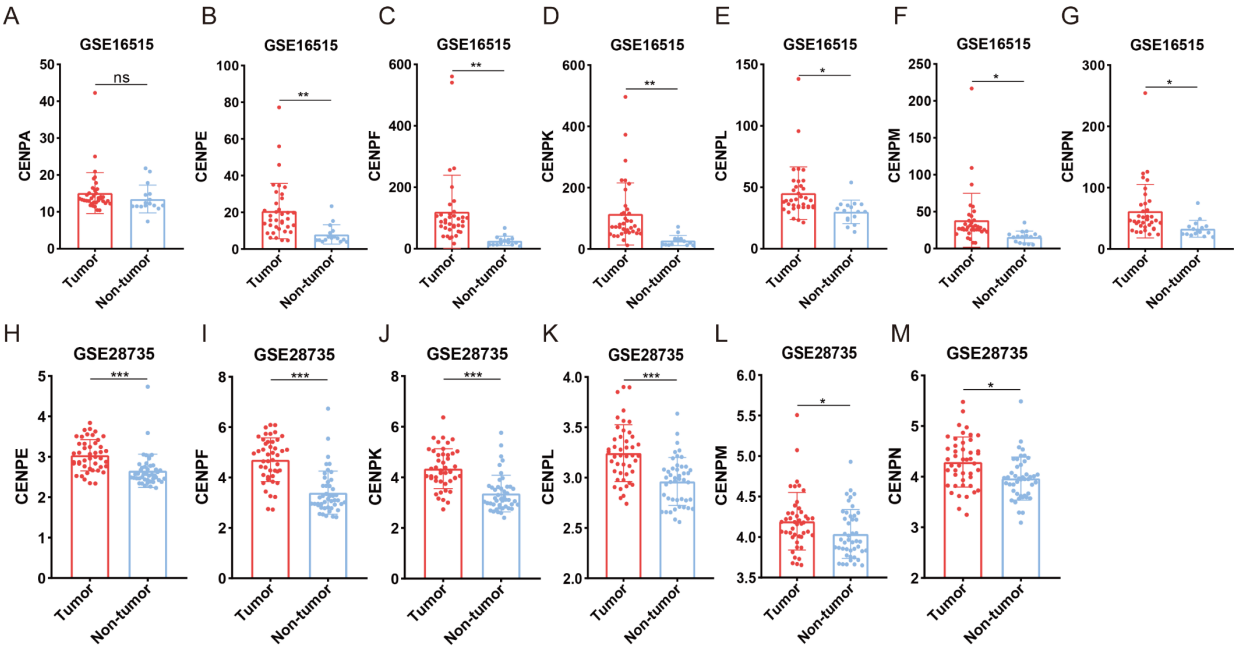
Note: CENPs: Centromere Protein Family Members; PDAC: Pancreatic Ductal Adenocarcinoma; OS: Overall Survival; DFS: Disease-Free Survival.

Figure 4. Genetic alteration frequency of CENPs in PDAC (cBioPortal). (A) Summary of genetic alterations in CENPs. (B) OncoPrint visual summary of alterations in a query of CENPs. (C), (D) Kaplan-Meier plots comparing OS/DFS in cases with/without CENPs alterations. (E), (F) The changes in RNA/protein expression profiles caused by CENPs gene mutations. (G), (H) Enrichment analysis of significantly upregulated RNA/protein molecules. (I), (J) Network of enriched terms of upregulated RNA/protein molecules, colored by term types.



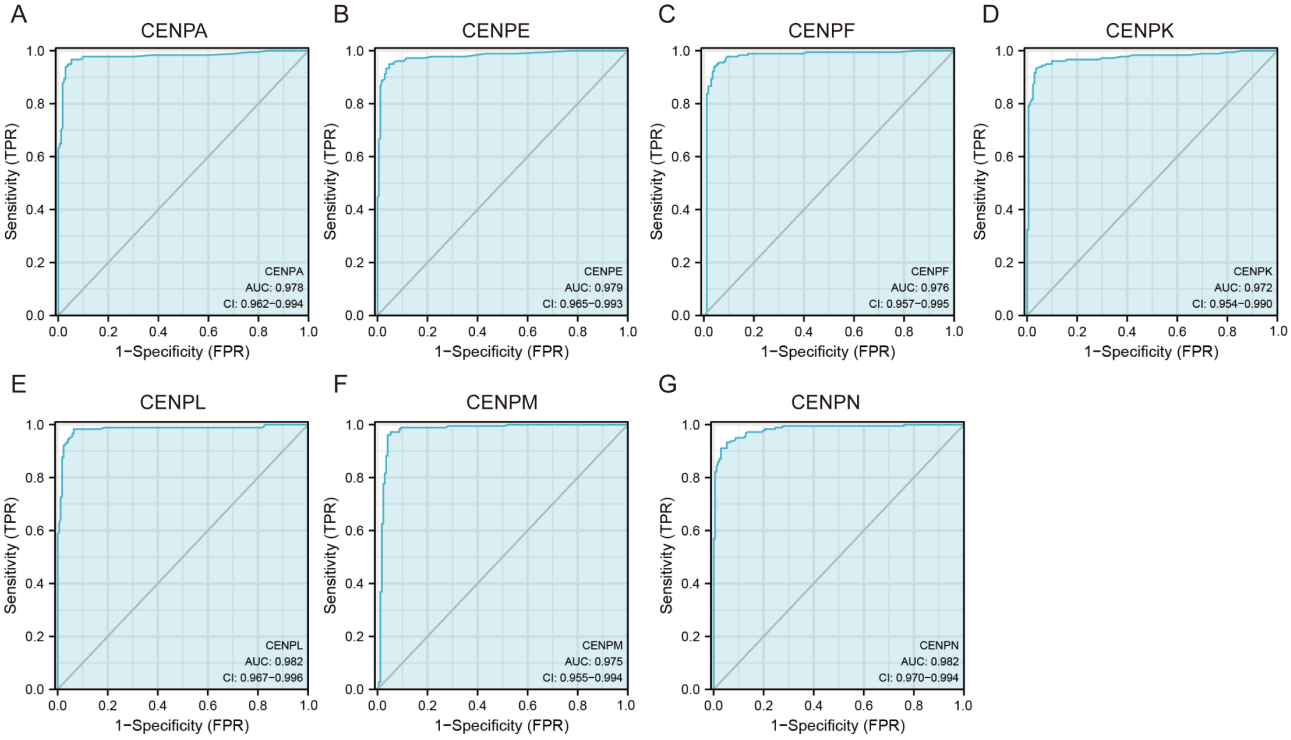
Note: CENPs: Centromere Protein Family Members.

Figure 5. Transcriptional expression of CENPs in 20 different types of cancer (ONCOMINE). Differences in mRNA expression were compared by the students' t-test. Cut-offs for p-value and fold change were as follows: p-value: 0.01, fold change: 2, gene rank: 10%, data type: mRNA.



Note: *p < 0.05, **p < 0.01. CENPs: Centromere Protein Family Members; GEO: Gene Expression Omnibus; PDAC: Pancreatic Ductal Adenocarcinoma.

Figure 6. Validation of CENPs mRNA expression levels in GEO datasets. CENPs expression levels between tumor and nontumor tissues in PDAC patients in GEO datasets, including GDS4102 (A) - (G) and GDS4336 (H) - (M).



Note: CENPs: Centromere Protein Family Members; TCGA: The Cancer Genome Atlas; GTX: Genotype-Tissue Expression.

Figure 7. Diagnostic values of CENPs. TCGA and GTX databases were used to evaluate the diagnostic values of CENPS (A) - (G).

monocytes (**Figure 8(B)**). CENPF was related to the infiltration of $CD8^+$ T cells and monocytes (**Figure 8(C)**). CENPK was related to the infiltration of B cells (**Figure 8(D)**). Additionally, CENPM was related to the infiltration of regulatory T cells (**Figure 8(F)**), while CENPL and CENPN were not significantly related to any of the eight immune cells (**Figure 8(E)**, **Figure 8(G)**). Notably, except for the expression of CENPA and monocyte infiltration, CENPK and B cell infiltration, other CENPs were negatively correlated with immune cell infiltration.

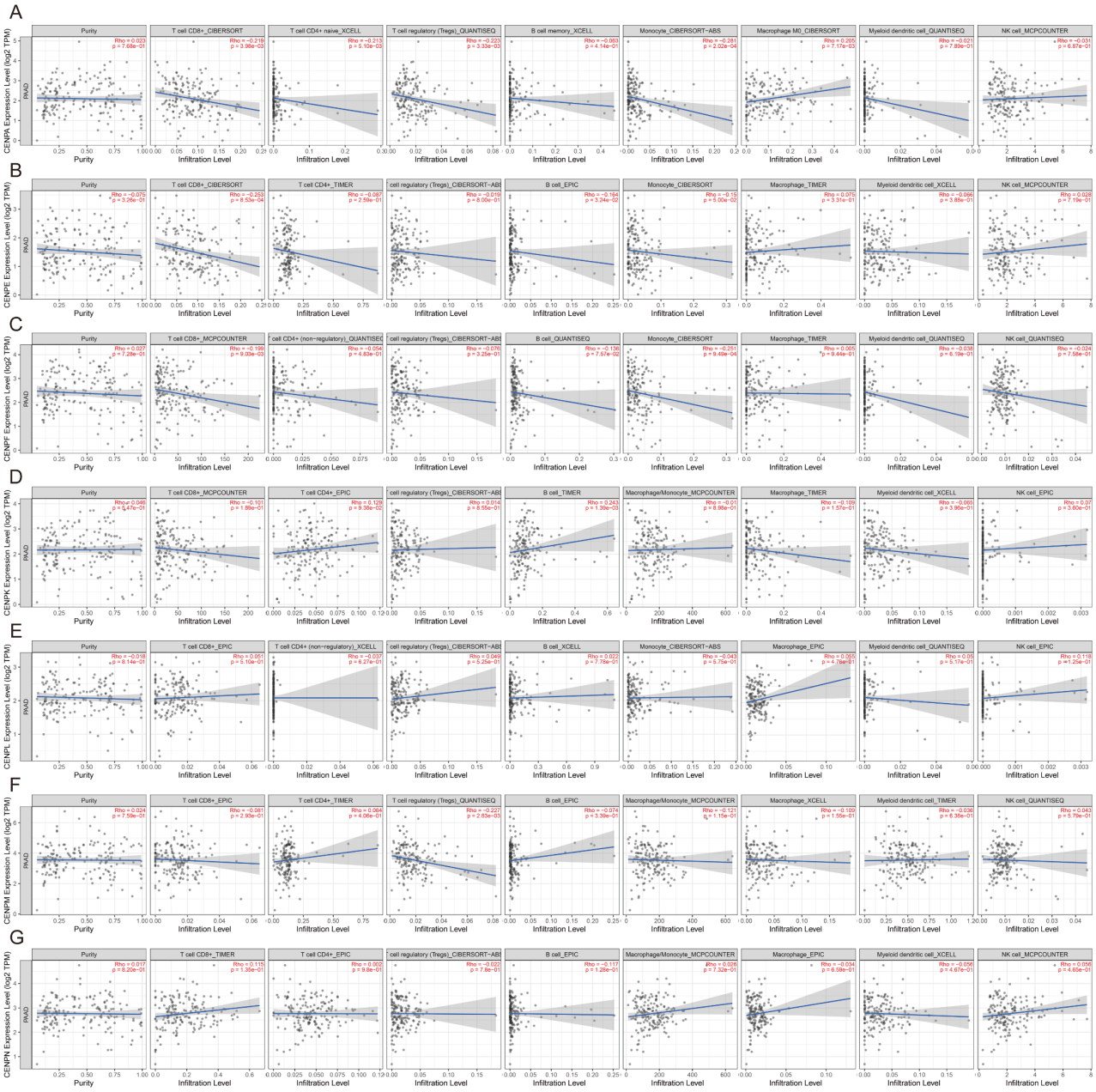
3.7. Analysis of Immune Infiltration and Prognosis

Next, the prognostic value of the combination analysis of the expression levels of CENPs and the infiltration of immune cells in PDAC was evaluated using the TIMER database. All positive results of the combined analyses are shown in **Figure 9**. Specific information is summarized below: 1) Monocyte infiltration predicts better prognosis in CENPA and CENPE with lower expression of PDAC (**Figure 9(A)**, **Figure 9(C)**). 2) $CD4^+$ T cell infiltration predicts better prognosis in CENPE and CENPF with higher expression of PDAC (**Figure 9(B)**, **Figure 9(D)**). 3) B cell infiltration predicts better prognosis in CENPL with higher expression of PDAC (**Figure 9(F)**), while their infiltration predicts worse prognosis in CENPK and CENPM with lower expression of PDAC (**Figure 9(E)**, **Figure 9(G)**). 4) Macrophage infiltration predicts better prognosis in CENPM higher expression PDAC (**Figure 9(H)**). 5) NK cells predict worse prognosis in CENPF, CENPK, CENPL, and CENPN with higher expression of PDAC (**Figures 9(I)-**

(L)). The co-expression patterns of these CENPs and immune cells may provide new co-targets for immunotherapy.

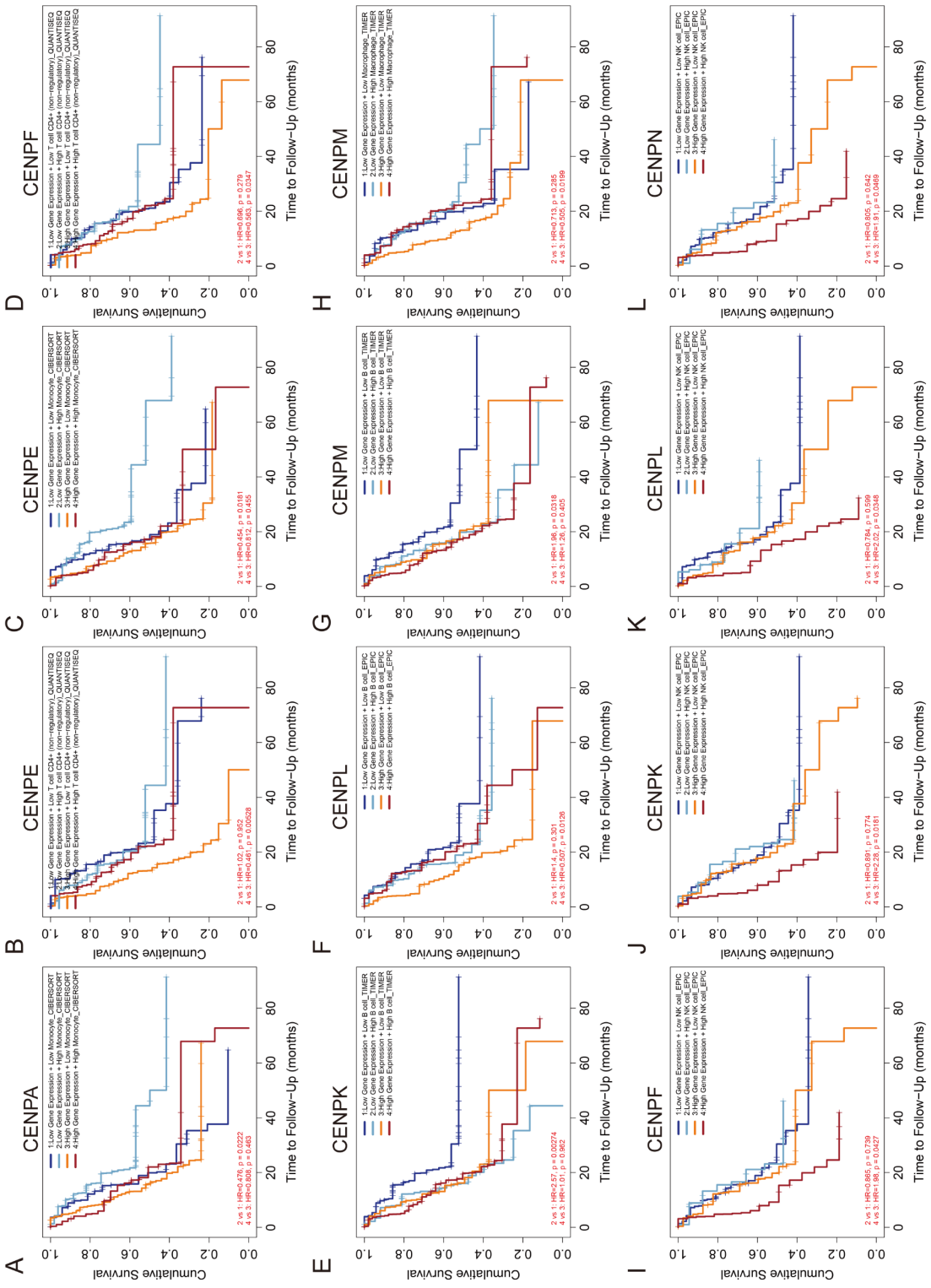
3.8. Construction of the PPI and Transcriptional Co-Regulatory Network

Many proteins perform their functions by interacting with other proteins. [32] Therefore, the PPI network of the 7 hub CENPs and their possible interacting



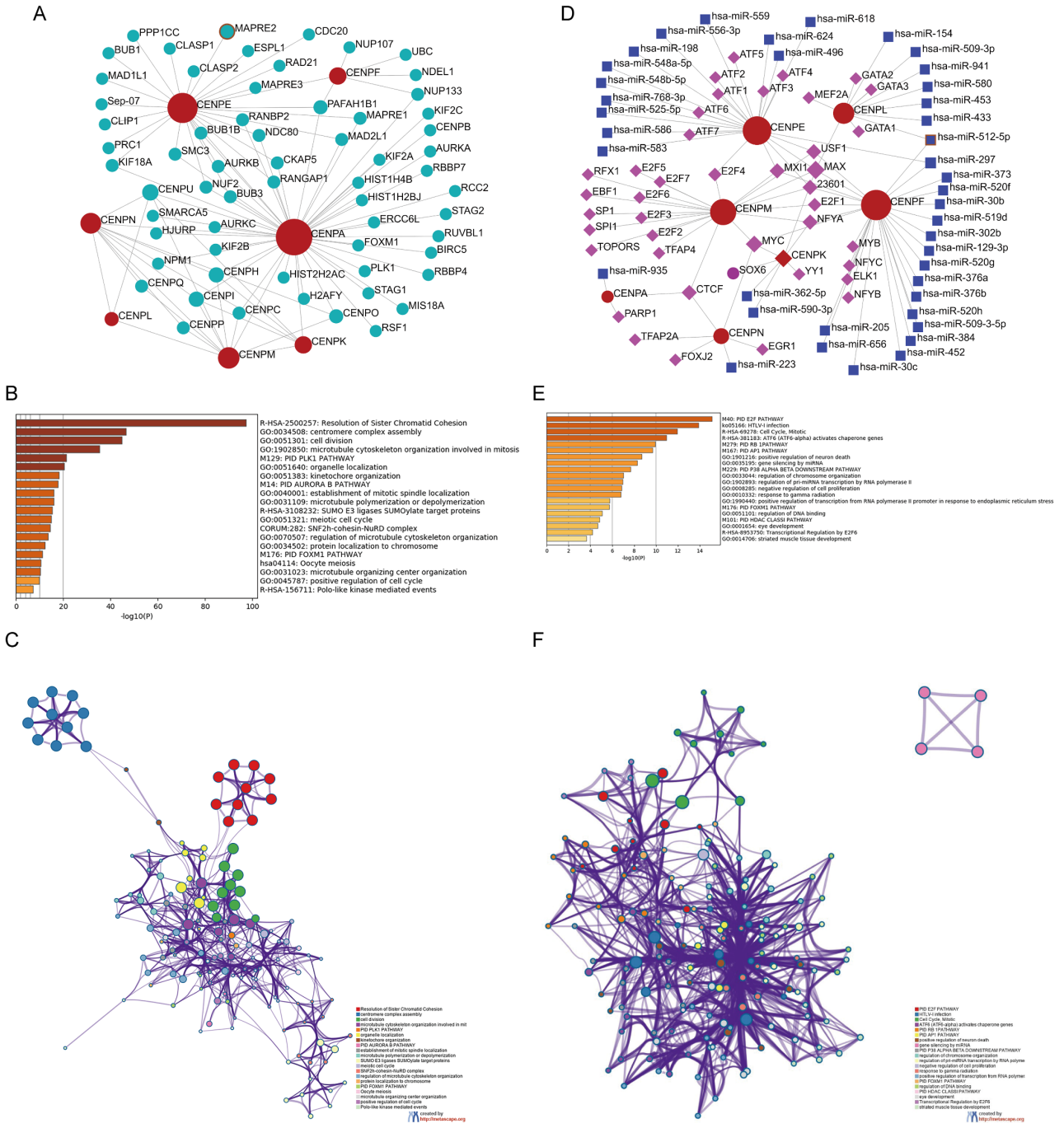
CENPs: Centromere Protein Family Members.

Figure 8. Correlations between differentially expressed CENPs and immune cell infiltration (TIMER). Correlations between the abundance of immune cells and the expression of (A) CENPA, (B) CENPE, (C) CENPF, (D) CENPK, (E) CENPL, (F) CENPM, and (G) CENPN.



Note: CENPs: Centromere Protein Family Members.

Figure 9. Analysis of immune infiltration and prognosis. Prognostic analysis between (A) CENPA and monocyte infiltration, (B) CENPE and $CD4^+$ T cell infiltration, (C) CENPE and monocyte infiltration, (D) CENPF and $CD4^+$ T cell infiltration, (E) CENPK and B cell infiltration, (F) CENPL and B cell infiltration, (G) CENPM and B cell infiltration, (H) CENPM and macrophage infiltration, (I) CENPF and NK cell infiltration, (J) CENPK and NK cell infiltration, (K) CENPL and NK cell infiltration, (L) CENPN and NK cell infiltration were conducted by TIMER.



PPI: Protein-Protein Interaction; CENPs: Centromere Protein Family Members.

Figure 10. Construction of the PPI network and transcriptional co-regulatory network based on Metascape. (A) PPI network of the 7 hub CENPs and their possible interacting proteins. (B) Transcription factor and microRNA co-regulatory network of CENPs. (C), (D) Enrichment analysis of all the genes in the PPI network or transcriptional co-regulatory network. (E), (F) Enrichment term networks of the PPI network or transcriptional co-regulatory network were visualized by Metascape.

proteins was constructed using the NetworkAnalyst database, as shown in **Figure 10(A)**. Next, we conducted enrichment analysis of all the genes in this interaction network (**Figure 10(C)**). Finally, the enrichment term networks were also visualized in **Figure 10(E)**. Besides gene mutation, transcriptional regulation is also an im-

portant reason for the increase in gene expression. In order to further explore the reasons for the increased expression of CENPs in PDAC and to find new therapeutic targets, we constructed the transcription factor and microRNA co-regulatory network of CENPs by NetworkAnalyst (**Figure 10(B)**). Subsequently, enrichment analysis and enrichment term networks were conducted as shown in **Figure 10(D)** and **Figure 10(F)**. These results were helpful in finding the key molecules of CENPs that cause poor prognosis and in understanding their transcriptional regulatory mechanisms in PDAC, which will cooperate with CENPs to become a potential therapeutic target for pancreatic cancer.

4. Discussion

The centromere is a specialized region on the chromosome that directs equal chromosome segregation. Accumulative evidence indicates that CENPs are aberrantly expressed and affect the occurrence and development of a variety of tumors. [8]-[12] [33] However, the patterns of expression and the exact roles of distinct CENPs in PDAC remain unclear. In this study, we systematically explored the role of CENPs in PDAC through integrated bioinformatic analysis.

At first, 15 significantly upregulated CENPs in PDAC were identified using the GEPIA database. Subsequently, 9 CENPs that predict worse OS and RFS in PDAC were further determined by Kaplan-Meier Plotter. All of the 9 CENPs were mutated, and their mRNA expression was increased, as detected by the cBioPortal database. In order to further clarify the effect of CENPs as a whole, we studied the effect of different combinations of them on prognosis. It is noteworthy that when combined with 7 genes (CENPA, CENPE, CENPF, CENPK, CENPL, CENPM, and CENPN) for analysis, a high genetic alteration rate of CENPs (the altered group) was associated with shorter OS and DFS in PDAC. The changes in RNA and protein expression levels caused by these 7 gene mutations were further explored to clarify the mechanism of their oncogenic function in PDAC. In the altered group, the genes with significantly increased mRNA expression were enriched in the cell cycle, chromosome segregation, M phase, DNA repair, regulation of the cell cycle process, DNA conformation change, metabolism of RNA, ribonucleoprotein complex biogenesis, translation, and chromosome maintenance. Moreover, the genes with significantly increased protein expression were enriched in cellular senescence, cellular response to abiotic stimulus, TP53 regulation of transcription of the cell cycle, response to mechanical stimulus, cellular response to organic cyclic compound, PID MYC ACTIV pathway, peptidyl-serine phosphorylation, regulation of DNA recombination, positive regulation of DNA metabolic process, and regulation of protein complex assembly. These results, combined with the PPI network of CENPs, will help to further clarify the oncogenic role of CENPs in PDAC. After that, the prognostic and diagnostic values of CENPs in PDAC were further validated by ONCOMINE, GEO databases, and ROC curves. In order to explore the possibility of CENPs as synergistic targets of immunotherapy, the relationship between CENPs and the main infiltrating immune cells in PDAC was explored by

the TIMER database. We found that CENPs were significantly associated with the infiltration of various immune cells, including $CD8^+$ T cells, $CD4^+$ T cells, regulatory T cells, B cells, monocytes, and macrophages in PDAC. Notably, except for the expression of CENPA and monocyte infiltration, as well as CENPK and B cell infiltration, other CENPs were negatively correlated with immune cell infiltration. These results suggest that we may be able to influence the infiltration of different immune cells in PDAC by interfering with the expression of related CENPs, so as to enhance the efficacy of immunotherapy.

In addition, we also found some interesting results of prognostic value through the combination analysis of the expression of CENPs and infiltration of immune cells in PDAC. Specifically, $CD4^+$ T cell, monocyte, and macrophage infiltration predict a better prognosis in PDAC under different expression of CENPs. However, the increase of $CD8^+$ T cells induced by high expression of CENPA, CENPE, and CENPF did not significantly improve the prognosis of patients with PDAC. In addition, B-cell infiltration shows a better or worse prognosis; its diverse role was consistent with previous reports. [34] It is noteworthy that NK cell infiltration indicates poor prognosis in PDAC patients with high expression of CENPF, CENPK, CENPL, and CENPN. In the tumor microenvironment, PDAC cells with high expression of CENPF, CENPK, CENPL, and CENPN may induce dysfunction of NK cells or inhibit the anti-tumor effect of NK cells through certain pathways. These results are not completely consistent with the existing understanding of immunotherapy, and it is worth further exploring the specific mechanism of CENPs in tumor immunity. [35]-[37] At last, we constructed the transcriptional factor and microRNA co-regulatory network of CENPs. Enrichment analysis was also conducted. The main 10 enrichment items include the PID E2F pathway, HTLV-I infection, cell cycle, ATF6 activates chaperone genes, PID RB pathway, PID AP1 pathway, positive regulation of neuron death, gene silencing by miRNA, PID P38 ALPHA BETA downstream pathway, and regulation of chromosome organization. These results were helpful in understanding the transcriptional regulatory mechanism of CENPs in PDAC so as to better study the role of CENPs in PDAC.

There were some limitations in our study. First, although high mRNA expressions of CENPA, CENPE, CENPF, CENPK, CENPL, CENPM, and CENPN were independent prognostic and diagnostic factors of PDAC patients, all the data analyzed in our study were retrieved from online databases; further studies consisting of larger sample sizes are required to validate our findings. Second, we did not explore the potential mechanisms of distinct CENPs in PDAC through in vivo and in vitro experiments; this will be one of our main research priorities in the future. Thirdly, the analysis relies on bulk transcriptomic data, which may mask cell-type-specific expression patterns and interactions within the complex tumor microenvironment.

In summary, our results showed that all 7 hub CENPs were effective prognostic and diagnostic markers in PDAC patients. Genetic alteration is an important reason for the increased expression of CENPs in PDAC. The expression of CENPs

was also significantly correlated with the infiltration of diverse immune cells and predicted prognosis in PDAC. Finally, PPI and co-regulation networks were constructed to explore the specific mechanisms of CENPs in PDAC. Taken together, these results indicated that the 7 hub CENPs (CENPA, CENPE, CENPF, CENPK, CENPL, CENPM, and CENPN) could be prognostic and diagnostic biomarkers and possible synergistic targets for immunotherapy in PDAC patients.

Authors' Contributions

All authors contributed to the preparation of the manuscript. HGY and QW conceived this study. QW and TL obtained the datasets from an online database. HGY analyzed the data and prepared the manuscript. All authors approved the final version of the manuscript.

Availability of Data and Materials

The datasets generated and analyzed during the current study are available in:

- GEPIA (http://gepia.cancer-pku.cn/index.html);
- Kaplan-Meier plotter (http://kmplot.com/analysis/);
- cBioPortal (https://www.cbioportal.org/);
- ONCOMINE (https://www.oncomine.com/);
- GEO (https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS4102, https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS4336);
- TCGA (https://cancergenome.nih.gov/);
- TIMER (https://cistrome.shinyapps.io/timer/);
- Metascape (http://metascape.org);
- NetworkAnalyst (https://www.networkanalyst.ca/NetworkAnalyst/Secure/AnalysisOverview.xhtml).

Ethical Approval and Consent to Participate

This research is based on public databases. As researchers did not access any information that could lead to the identification of an individual patient, no ethical issue was raised in this research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. and Jemal, A. (2018) Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA*: *A Cancer Journal for Clinicians*, **68**, 394-424. https://doi.org/10.3322/caac.21492

[2] Siegel, R.L., Miller, K.D. and Jemal, A. (2017) Cancer Statistics, 2017. *CA*: *A Cancer Journal for Clinicians*, **67**, 7-30. https://doi.org/10.3322/caac.21387

[3] Wang, X., Li, X., Wei, X., Jiang, H., Lan, C., Yang, S., *et al.* (2020) PD-L1 Is a Direct

Target of Cancer-FOXP3 in Pancreatic Ductal Adenocarcinoma (PDAC), and Combined Immunotherapy with Antibodies against PD-L1 and CCL5 Is Effective in the Treatment of PDAC. *Signal Transduction and Targeted Therapy*, **5**, Article No. 38. https://doi.org/10.1038/s41392-020-0144-8

[4] Kuehn, B.M. (2020) Looking to Long-Term Survivors for Improved Pancreatic Cancer Treatment. *JAMA*, **324**, 2242-2244. https://doi.org/10.1001/jama.2020.21717

[5] Wong, C.Y.Y., Lee, B.C.H. and Yuen, K.W.Y. (2020) Epigenetic Regulation of Centromere Function. *Cellular and Molecular Life Sciences*, **77**, 2899-2917. https://doi.org/10.1007/s00018-020-03460-8

[6] Ranjan, R., Snedeker, J. and Chen, X. (2019) Asymmetric Centromeres Differentially Coordinate with Mitotic Machinery to Ensure Biased Sister Chromatid Segregation in Germline Stem Cells. *Cell Stem Cell*, **25**, 666-681.e5. https://doi.org/10.1016/j.stem.2019.08.014

[7] Kabeche, L., Nguyen, H.D., Buisson, R. and Zou, L. (2018) A Mitosis-Specific and R Loop-Driven ATR Pathway Promotes Faithful Chromosome Segregation. *Science*, **359**, 108-114. https://doi.org/10.1126/science.aan6490

[8] Welner, S., Trier, N.H., Frisch, M., Locht, H., Hansen, P.R. and Houen, G. (2013) Correlation between Centromere Protein-F Autoantibodies and Cancer Analyzed by Enzyme-Linked Immunosorbent Assay. *Molecular Cancer*, **12**, Article No. 95. https://doi.org/10.1186/1476-4598-12-95

[9] Li, H., Zhang, H. and Wang, Y. (2018) Centromere Protein U Facilitates Metastasis of Ovarian Cancer Cells by Targeting High Mobility Group Box 2 Expression. *American Journal of Cancer Research*, **8**, 835-851.

[10] Sharma, A.B., Dimitrov, S., Hamiche, A. and Van Dyck, E. (2018) Centromeric and Ectopic Assembly of CENP-A Chromatin in Health and Cancer: Old Marks and New Tracks. *Nucleic Acids Research*, **47**, 1051-1069. https://doi.org/10.1093/nar/gky1298

[11] Bayo, J., Fiore, E.J., Dominguez, L.M., Real, A., Malvicini, M., Rizzo, M., *et al*. (2019) A Comprehensive Study of Epigenetic Alterations in Hepatocellular Carcinoma Identifies Potential Therapeutic Targets. *Journal of Hepatology*, **71**, 78-90. https://doi.org/10.1016/j.jhep.2019.03.007

[12] Fei, X., Zhu, L., Zhou, H., Qi, C. and Wang, C. (2019) A Novel Intergenic Region between CENPA and DPYSL5-ALK Exon 20 Fusion Variant Responding to Crizotinib Treatment in a Patient with Lung Adenocarcinoma. *Journal of Thoracic Oncology*, **14**, e191-e193. https://doi.org/10.1016/j.jtho.2019.04.012

[13] Henderson, M.C., Shaw, Y.Y., Wang, H., Han, H., Hurley, L.H., Flynn, G., *et al*. (2009) UA62784, a Novel Inhibitor of Centromere Protein E Kinesin-Like Protein. *Molecular Cancer Therapeutics*, **8**, 36-44. https://doi.org/10.1158/1535-7163.mct-08-0789

[14] Shaw, A.Y., Henderson, M.C., Flynn, G., Samulitis, B., Han, H., Stratton, S.P., *et al*. (2009) Characterization of Novel Diaryl Oxazole-Based Compounds as Potential Agents to Treat Pancreatic Cancer. *The Journal of Pharmacology and Experimental Therapeutics*, **331**, 636-647. https://doi.org/10.1124/jpet.109.156406

[15] Zheng, C., Zhang, T., Li, D., Huang, C., Tang, H., Ni, X., *et al*. (2020) Upregulation of CENPM Facilitates Tumor Metastasis via the mTOR/p70s6K Signaling Pathway in Pancreatic Cancer. *Oncology Reports*, **44**, 1003-1012. https://doi.org/10.3892/or.2020.7673

[16] Chen, H., Wang, X., Wu, F., Mo, X., Hu, C., Wang, M., *et al*. (2021) Centromere Protein F Is Identified as a Novel Therapeutic Target by Genomics Profile and Contributing to the Progression of Pancreatic Cancer. *Genomics*, **113**, 1087-1095.

https://doi.org/10.1016/j.ygeno.2020.10.039

[17] Zhang, Q., Lou, Y., Yang, J., Wang, J., Feng, J., Zhao, Y., *et al*. (2019) Integrated Multiomic Analysis Reveals Comprehensive Tumour Heterogeneity and Novel Immunophenotypic Classification in Hepatocellular Carcinomas. *Gut*, **68**, 2019-2031. https://doi.org/10.1136/gutjnl-2019-318912

[18] Gao, Q., Zhu, H., Dong, L., Shi, W., Chen, R., Song, Z., *et al*. (2019) Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma. *Cell*, **179**, 1240. https://doi.org/10.1016/j.cell.2019.10.038

[19] Li, C., Sun, Y., Yu, G., Cui, J., Lou, Z., Zhang, H., *et al*. (2020) Integrated Omics of Metastatic Colorectal Cancer. *Cancer Cell*, **38**, 734-747.e9. https://doi.org/10.1016/j.ccell.2020.08.002

[20] Tang, Z., Kang, B., Li, C., Chen, T. and Zhang, Z. (2019) GEPIA2: An Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Research*, **47**, W556-W560. https://doi.org/10.1093/nar/gkz430

[21] Perpelescu, M. and Fukagawa, T. (2011) The ABCs of CENPs. *Chromosoma*, **120**, 425-446. https://doi.org/10.1007/s00412-011-0330-0

[22] Menyhárt, O., Nagy, Á. and Győrffy, B. (2018) Determining Consistent Prognostic Biomarkers of Overall Survival and Vascular Invasion in Hepatocellular Carcinoma. *Royal Society Open Science*, **5**, Article ID: 181006. https://doi.org/10.1098/rsos.181006

[23] Rhodes, D.R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B.B., *et al*. (2007) Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Gene Expression Profiles. *Neoplasia*, **9**, 166-180. https://doi.org/10.1593/neo.07112

[24] Pei, H., Li, L., Fridley, B.L., Jenkins, G.D., Kalari, K.R., Lingle, W., *et al*. (2009) FKBP51 Affects Cancer Cell Response to Chemotherapy by Negatively Regulating Akt. *Cancer Cell*, **16**, 259-266. https://doi.org/10.1016/j.ccr.2009.07.016

[25] Zhang, G., Schetter, A., He, P., Funamizu, N., Gaedcke, J., Ghadimi, B.M., *et al*. (2012) DPEP1 Inhibits Tumor Cell Invasiveness, Enhances Chemosensitivity and Predicts Clinical Outcome in Pancreatic Ductal Adenocarcinoma. *PLOS ONE*, **7**, e31507. https://doi.org/10.1371/journal.pone.0031507

[26] Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., *et al*. (2020) TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic Acids Research*, **48**, W509-W514. https://doi.org/10.1093/nar/gkaa407

[27] Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., *et al*. (2019) MetaScape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nature Communications*, **10**, Article No. 1523. https://doi.org/10.1038/s41467-019-09234-6

[28] Zhou, G., Soufan, O., Ewald, J., Hancock, R.E.W., Basu, N. and Xia, J. (2019) NetworkAnalyst 3.0: A Visual Analytics Platform for Comprehensive Gene Expression Profiling and Meta-analysis. *Nucleic Acids Research*, **47**, W234-W241. https://doi.org/10.1093/nar/gkz240

[29] Borghaei, H., Gettinger, S., Vokes, E.E., Chow, L.Q.M., Burgio, M.A., de Castro Carpeno, J., *et al*. (2021) Five-Year Outcomes from the Randomized, Phase III Trials Checkmate 017 and 057: Nivolumab versus Docetaxel in Previously Treated Non-Small-Cell Lung Cancer. *Journal of Clinical Oncology*, **39**, 723-733. https://doi.org/10.1200/jco.20.01605

[30] Zhang, Y., Ma, S., Liu, X., Xu, Y., Zhao, J., Si, X., *et al*. (2021) Supramolecular Assembled Programmable Nanomedicine as *in Situ* Cancer Vaccine for Cancer Immuno-

therapy. *Advanced Materials*, **33**, Article ID: 2007293. https://doi.org/10.1002/adma.202007293

[31] Jin, Q., Zhu, W., Zhu, J., *et al*. (2021) Nanoparticle-Mediated Delivery of Inhaled Immuno-Therapeutics for Treating Lung Metastasis. *Advanced Materials*, **33**, e2007557

[32] Pillarsetty, N., Jhaveri, K., Taldone, T., Caldas-Lopes, E., Punzalan, B., Joshi, S., *et al*. (2019) Paradigms for Precision Medicine in Epichaperome Cancer Therapy. *Cancer Cell*, **36**, 559-573.e7. https://doi.org/10.1016/j.ccell.2019.09.007

[33] Cheng, Z., Yu, C., Cui, S., Wang, H., Jin, H., Wang, C., *et al*. (2019) Circtp63 Functions as a Cerna to Promote Lung Squamous Cell Carcinoma Progression by Upregulating FOXM1. *Nature Communications*, **10**, Article No. 3200. https://doi.org/10.1038/s41467-019-11162-4

[34] Lu, Y., Zhao, Q., Liao, J., Song, E., Xia, Q., Pan, J., *et al*. (2020) Complement Signals Determine Opposite Effects of B Cells in Chemotherapy-Induced Immunity. *Cell*, **180**, 1081-1097.e24. https://doi.org/10.1016/j.cell.2020.02.015

[35] Dong, W., Wu, X., Ma, S., Wang, Y., Nalin, A.P., Zhu, Z., *et al*. (2019) The Mechanism of Anti-PD-L1 Antibody Efficacy against PD-L1-Negative Tumors Identifies NK Cells Expressing PD-L1 as a Cytolytic Effector. *Cancer Discovery*, **9**, 1422-1437. https://doi.org/10.1158/2159-8290.cd-18-1259

[36] Jiménez-Sánchez, A., Cybulska, P., Mager, K.L., Koplev, S., Cast, O., Couturier, D., *et al*. (2020) Unraveling Tumor-Immune Heterogeneity in Advanced Ovarian Cancer Uncovers Immunogenic Effect of Chemotherapy. *Nature Genetics*, **52**, 582-593. https://doi.org/10.1038/s41588-020-0630-5

[37] Huntington, N.D., Cursons, J. and Rautela, J. (2020) The Cancer-Natural Killer Cell Immunity Cycle. *Nature Reviews Cancer*, **20**, 437-454. https://doi.org/10.1038/s41568-020-0272-z

Abbreviations

CENPs	Centromere Protein Family Members
PDAC	Pancreatic Ductal Adenocarcinoma
GEPIA	Gene Expression Profiling Interactive Analysis
TCGA	The Cancer Genome Atlas
GTEx	Genotype-Tissue Expression
OS	Overall Survival
RFS	Relapse-Free Survival
HR	Hazard Ratio
DFS	Disease-Free Survival
PPI	Protein-Protein Interaction
GEO	Gene Expression Omnibus
GDS	Geo Dataset
ROC Curve	Receiver Operating Characteristic Curve