Observational Study

Medicine®

OPEN

Integrated bioinformatics analysis reveals CDK1 and PLK1 as potential therapeutic targets of lung adenocarcinoma

Shuzhen Li, MSC[a], Hua Li, MSC[a], Yajie Cao, MSC[a], Haiying Geng, MSC[a], Fu Ren, PhD[b], Keyan Li, MD[c], Chunmei Dai, PhD[d], Ning Li, PhD[a,b,*]

Abstract

This study is to identify potential biomarkers and therapeutic targets for lung adenocarcinoma (LUAD).

GSE6044 and GSE118370 raw data from the Gene Expression Omnibus database were normalized with Robust Multichip Average. After merging these two datasets, the combat function of sva packages was used to eliminate batch effects. Then, limma packages were used to filtrate differentially expressed genes. We constructed protein–protein interaction relationships using STRING database and hub genes were identified based on connectivity degrees. The cBioportal database was used to explore the alterations of the hub genes. The promoter methylation of cyclin dependent kinase 1 (CDK1) and polo-like Kinase 1 (PLK1) and their association with tumor immune infiltration in patients with LUAD were investigated using DiseaseMeth version 2.0 and TIMER databases. The Cancer Genome Atlas-LUAD dataset was used to perform gene set enrichment analysis.

We identified 10 hub genes, which were upregulated in LUAD, among which 8 were successfully verified in the Cancer Genome Atlas and Oncomine databases. Kaplan–Meier analysis indicated that the expressions of CDK1 and PLK1 in LUAD patients were associated with overall survival and disease-free survival. The methylation levels in the promoter regions of these 2 genes in LUAD patients were lower than those in normal lung tissues. Their expressions in LUAD were associated with tumor stages and relative abundance of tumor infiltrating immune cells, such as B cells, CD4+ T cells, and macrophages. Moreover, cell cycle, DNA replication, homologous recombination, mismatch repair, P53 signaling pathway, and small cell lung cancer signaling were significantly enriched in CDK1 and PLK1 high expression phenotype.

CDK1 and PLK1 may be used as potential biomarkers and therapeutic targets for LUAD.

Abbreviations: CDK1 = cyclin dependent kinase 1, DEGs = differentially expressed genes, DFS = disease-free survival, GEO = Gene Expression Omnibus, GO = Gene Ontology, GSEA = Gene Set Enrichment Analysis, KEGG = Kyoto Encyclopedia of Genes and Genomes, LUAD = lung adenocarcinoma, NSCLC = non-small cell lung cancer, OS = overall survival, PLK1; polo-like Kinase 1, TCGA = the Cancer Genome Atlas, TKIs = tyrosine kinase inhibitors.

Keywords: bioinformatics analysis, biomarkers, differentially expressed genes, lung adenocarcinoma, prognosis

1. Introduction

Lung cancer, the leading cause of cancer-related death worldwide,[1] is classified into histological types of small cell lung cancer and non-small cell lung cancer (NSCLC).[2] Lung adenocarcinoma (LUAD), with the 5-year survival rate of only 15%, is the most common histologic subtype of NSCLC and one of the most malignant tumors that have poor prognosis.[3,4] Studies have shown the pathogenesis of LUAD is associated with

Editor: Mihai Dorin Vartolomei.

This work was supported by the Liaoning Province Natural Science Foundation (Grant No. 201602293), Liaoning Province Education Department Foundation of China (Grant No. JYTQN201727), Biological Anthropology Innovation Team Project of JZMU (Grant No. JYLJ201702), the Educational Department of Liaoning Province (Grant No. JYTFW201915), and, Innovative Talents in Colleges and Universities of Liaoning Province in 2019 of the Educational Department of Liaoning Province (Grant No. LR2019028).

The authors declare that they have no competing interests.

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

[a] Department of Biochemistry and Molecular Biology, College of Basic Medicine, Jinzhou Medical University, Jinzhou, Liaoning Province, PR China, [b] Liaoning Province Key Laboratory of Human Phenome Research, Jinzhou Medical University, Jinzhou, Liaoning Province, PR China, [c] Department of Cardiology, the First Affiliated Hospital of Jinzhou Medical University, Jinzhou Medical University, Jinzhou, Liaoning Province, PR China, [d] School of Pharmacy, Jinzhou Medical University, Jinzhou, Liaoning, PR China.

[*] Correspondence: Ning Li, Departments of Biochemistry and Molecular Biology and Biological Anthropology Institute, College of Basic Medicine, Jinzhou Medical University, No. 40, Section 3, Songpo Road, Linghe District, Jinzhou 121000, Liaoning, PR China (e-mail: 157026753@qq.com).

How to cite this article: Li S, Li H, Cao Y, Geng H, Ren F, Li K, Dai C, Li N. Integrated bioinformatics analysis reveals CDK1 and PLK1 as potential therapeutic targets of lung adenocarcinoma. Medicine 2021;100:32(e26474).

Received: 16 October 2020 / Received in final form: 15 April 2021 / Accepted: 8 June 2021

http://dx.doi.org/10.1097/MD.0000000000026474

many factors, such as epidermal growth factor receptor (EGFR) mutation, EML4-ALK gene fusion, and tumor suppressor RNA-binding motif protein 5.[5–7] Currently, first-generation EGFR tyrosine kinase inhibitors (TKIs), including gefitinib and erlotinib, have been used as therapeutic regimen for NSCLC, and for patients with resistance to the first-generation EGFR-TKIs, the second-generation EGFR-TKIs can be used.[8,9] However, secondary drug resistance still occurs, leading a failure in the course of treatment.[10] Hence, it is of great importance to develop more potential and effective diagnostic and therapeutic targets.

With the rapid development of genomics and sequencing techniques, the number of data available in public databases, such as Gene Expression omnibus (GEO) and the Cancer Genome Atlas (TCGA) databases, is increasing.[11–13] By mining the TCGA database, Liu et al[14] identified 5 competing endogenous RNAs interaction modules that play key roles in LUAD. According to the RNA-seq data in TCGA, the over expression of Keratin 8, which may function as an independent prognostic factor to predict overall survival (OS), is highly correlated with poor clinical outcome in patients with LUAD.[15] Using GEO database, Lu et al[16] identified many genes that may be used to distinguish LUAD from lung squamous cell cancer. However, the molecular mechanisms of LUAD have not been fully understood.

Herein, to explore the molecular mechanisms of LUAD further, the raw data of GSE6044 and GSE118370 were obtained from GEO and the differentially expressed genes (DEGs) between normal lung tissues and LUAD were identified. Subsequently, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment, and gene alteration analysis were carried out. Eventually, the cyclin dependent kinase 1 (CDK1) and polo-like Kinase 1 (PLK1) with $P < .05$ of both OS and disease-free survival (DFS) were further studied from different perspectives, including tumor stage, promoter methylation, tumor immune infiltration, and Gene Set Enrichment Analysis (GSEA). The findings of this study demonstrate that the overexpression of CDK1 and PLK1 are involved in tumor stage and poor prognosis of LUAD patients. This study may provide molecular targets and deep insight into the mechanisms of LUAD.

2. Materials and methods

2.1. Data sources

Raw data were downloaded from the GEO (www.ncbi.nlm.nih.gov/geo).[17] GSE6044 and GSE118370, 2 mRNA expression datasets of LUAD, were included. The dataset GSE6044 that was based on GPL201 (HG-Focus) Affymetrix Human HG-Focus Target Array platform included 5 normal lung tissues and 10 LUAD specimens; the microarray dataset of GSE118370 contained 6 normal lung tissues and 6 LUAD specimens on the basis of GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array platform.

2.2. Data processing

The GSE6044 and GSE118370 raw data were normalized with Robust Multichip Average of the Affy package in R (version3.5.2), respectively. And the mean values were used for those genes with several probes or a probe group. Then, K-Nearest Neighbors method was employed to deal with missing values in the 2 datasets. Afterward, the 2 datasets were integrated as 1. Subsequently, the combat function of the sva package was applied to eliminate batch effect.[18]

2.3. Identification of DEGs

The DEGs between normal lung tissues and LUAD specimens were screened by limma package, and statistically significant cut-off criteria of DEGs were defined as adjusted $P < .05$ and $|\log FC| > 1$.

2.4. GO annotation and KEGG pathway enrichment analysis

To understand the biological implications of DEGs, we analyzed the enrichment of functions and pathways using clusterprofiler, which is a package harboring an analysis and visualization function on the GO and KEGG analysis.[19] The $P < .05$ and adjusted $P < .05$ were regarded as the threshold values for remarkable enrichment.

2.5. Identification and alteration analysis of hub genes in cBioportal database

The STRING database (version 11.0; http://string-db.org/), which is a tool designed to analyze the interactions between proteins, was utilized to establish a global protein–protein interaction network of all DEGs. Among these genes, the top 10 genes ranked by connectivity degree were considered as hub genes. We used the pheatmap package for R (version 3.5.2) to draw the heatmap of the hub genes. We performed alteration analysis of 10 hub genes using LUAD dataset (The Cancer Genome Atlas, provisional), which included 586 cases in cBioportal (http://www.cbioportal.org/). The genetic alteration profiles consisted of mutations, copy number alterations (CNAs), and amplification.[20]

2.6. Validation of hub gene mRNA levels in TCGA and Oncomine databases

To validate hub gene expression in LUAD, we downloaded the RNA-seq data (HTSeq-Counts) available for LUAD from the TCGA database (https://cancergenome.nih.gov/). These data included 59 normal lung tissues and 535 LUAD specimens. We utilized edgeR[21] package to screen DEGs with significant cutoff criteria of adjusted $P < .05$ and $|\log FC| > 1$. Oncomine is an online data-mining platform that facilitates discovery from genome-wide expression analysis.[22] Su et al,[23] Hou et al,[24] Stearman et al,[25] and Yamagata et al[26] have searched lung cancer gene expression data in the Oncomine database for gene expression levels of the hub genes with $P < .05$ and $FC > 2$, however, the data type was confined to mRNA.

2.7. GEPIA database analysis

GEPIA (Gene Expression Profiling Interactive Analysis; http://gepia.cancer-pku.cn/) is an online interactive website that integrates gene expression and clinical data from TCGA and Genotype Tissue Expression (GTEx) projects. Numerous

sequencing data in GEPIA can be used for gene differential expression analysis, pathological stage analysis, survival analysis, and gene expression correlation analysis, etc.[27] Here, transcriptional and survival data of CDK1 and PLK1 were explored using GEPIA database.

2.8. Methylation level analysis of the 2 cell cycle-related hub genes

The human disease methylation database version 2.0 (DiseaseMeth 2.0; http://bioinfo.hrbmu.edu.cn/diseasemeth/) is an online database that contains abundant DNA methylation data associated with various human diseases, especially cancers.[28] Herein, the difference in methylation levels of the promoter regions of CDK1 and PLK1 between LUAD specimens and normal lung tissues were analyzed.

2.9. Correlation between the expressions of CDK1 and PLK1 and the tumor infiltrating immune cells

To assess the correlation between the expressions of CDK1 and PLK1 and the tumor infiltrating immune cells (B cells, $CD4^+$ T cells, $CD8^+$ T cells, macrophages, neutrophils and dendritic cells), we used the TIMER (https://cistrome.shinyapps.io/timer/) database, which provides 32 cancer types and 6 major analytic modules on the basis of TCGA datasets.[29]

2.10. GSEA of CDK1 and PLK1

LUAD RNA-seq (FPKM) data including 59 normal lung tissues and 535 LUAD specimens, were downloaded from TCGA database. We used software "gsea-3.0.jar" (Broad Institute) and gene set "c2.cp.kegg.v6.2.symbols.gmt" to perform the GSEA analysis of CDK1 and PLK1.

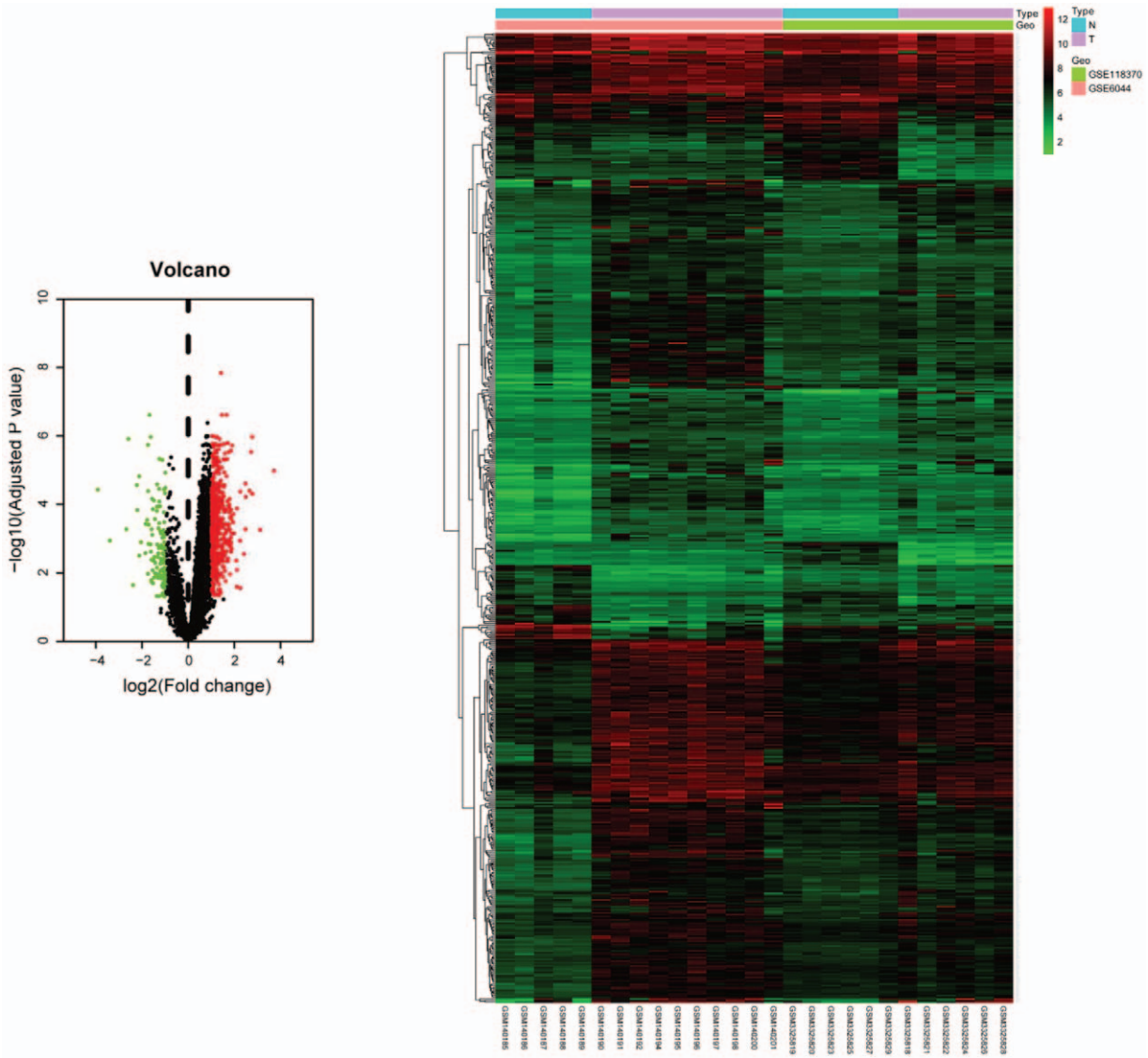

Figure 1. Volcano plot and heatmap of DEGs. (A) Volcano plot and (B) heatmap of DEGs of normal lung tissues and LUAD specimens in GSE6044 and GSE118370 integrated dataset. The black points represent genes without significant difference. The red points are the upregulated genes. The green points are the downregulated genes. DEG = differentially expressed genes; LUAD = lung adenocarcinoma.

3. Results

3.1. DEGs identification

To identify the DEGs between LUAD specimens and normal lung tissues, GSE6044 and GSE118370 were screened using the limma package after data preprocessing and eliminating batch effects. A total of 589 DEGs were identified in LUAD, including 476 upregulated genes and 113 downregulated genes (Fig. 1A and B).

3.2. GO and KEGG enrichment analysis of DEGs

The biological annotation of DEGs in LUAD specimens identified from the integrated analysis of microarray data was performed using clusterprofiler package for R. The upregulated DEGs were significantly enriched in "protein serine/threonine kinase activity," "RNA polymerase II proximal promoter sequence-specific DNA binding," and "catalytic activity, acting on DNA" (Fig. 2A). The downregulated DEGs were significantly enriched in "serine-type endopeptidase inhibitor activity," "peptide binding," and "enzyme inhibitor activity" (Fig. 2B). The KEGG analysis showed the upregulated DEGs were significantly enriched in "cell cycle," "oocyte meiosis," and "DNA replication" signaling pathways (Fig. 3A). Meanwhile, the downregulated DEGs were significantly enriched in "drug metabolism-cytochrome P450," "metabolism of xenobiotics by cytochrome P450," and "glutathione metabolism" signaling pathways

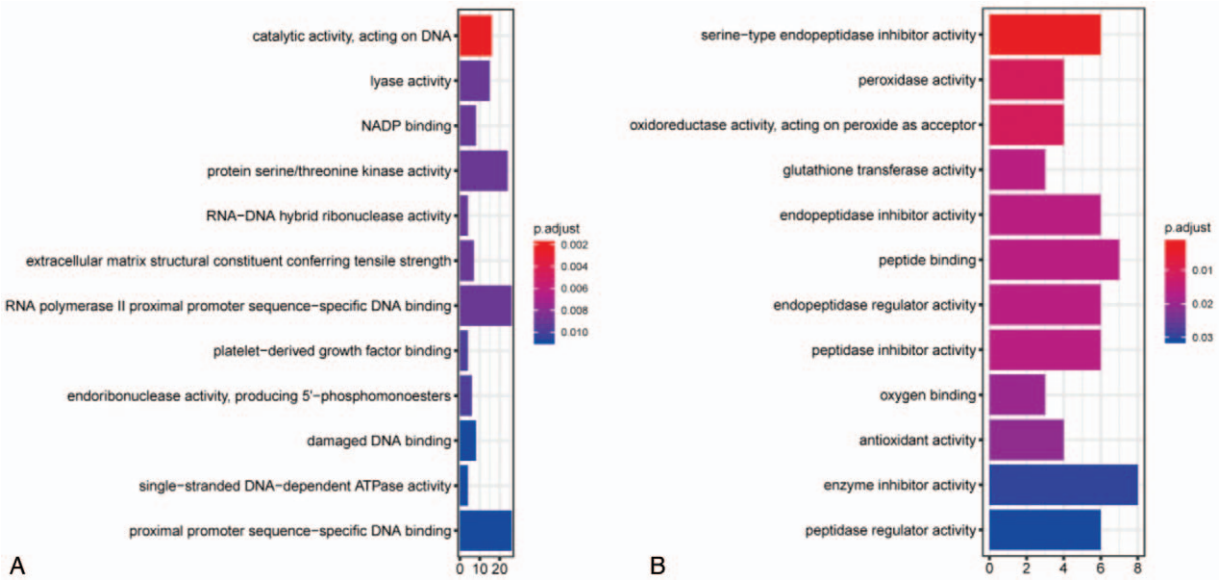

Figure 2. The top 12 GO terms with $P < .05$ and adjusted $P < .05$ in the enrichment analysis of (A) upregulated and (B) downregulated DEGs. GO = Gene Ontology.

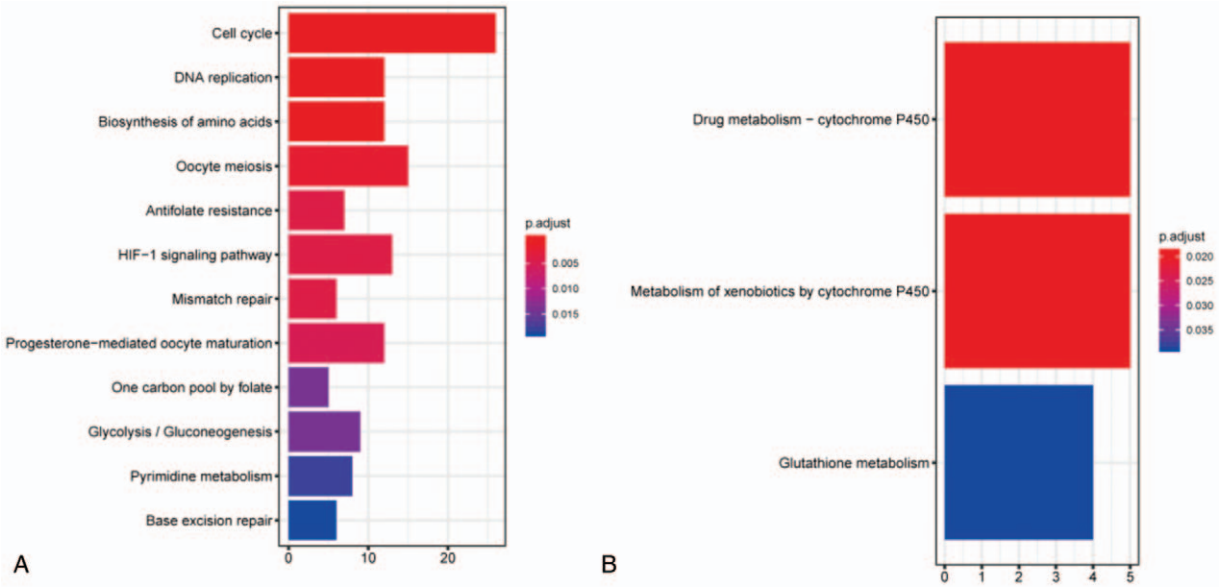

Figure 3. The top 12 KEGG pathways with $P < .05$ and adjusted $P < .05$ in the enrichment analysis of (A) upregulated and (B) downregulated DEGs. KEGG = Kyoto Encyclopedia of Genes and Genomes.

Table 1

The hub gene in the GSE6044 and GSE118370 integrated dataset.

Gene	Log fold change	Adjust *P* value
AURKA	1.58	1.66e–04
CCNA2	1.10	2.01e–03
CDK1	1.85	1.47e–04
FEN1	1.17	3.21e–03
GAPDH	1.10	1.11e–03
PCNA	1.53	3.13e–05
PLK1	1.05	2.16e–04
TOP2A	2.64	4.06e–05
MYC	1.54	3.28e–03
AKT1	1.10	1.77e–03

(Fig. 3B). These results suggest that most of DEGs are related to cell proliferation and drug metabolism, which could promote further understanding of the key roles of these DEGs in the occurrence and development of LUAD.

3.3. Identification and alteration analysis of the hub genes

To predict the interactions of DEGs, we analyzed the all candidate proteins in STRING database. Eventually, AURKA, CCNA2, CDK1, FEN1, GAPDH, PCNA, PLK1, TOP2A, AKT1, and MYC were considered as hub genes (Table 1). The results showed that all these genes were upregulated in the integrated dataset. We showed the node numbers (Fig. 4A), protein–protein interaction network (Fig. 4B), and heatmap (Fig. 4C) of the 10

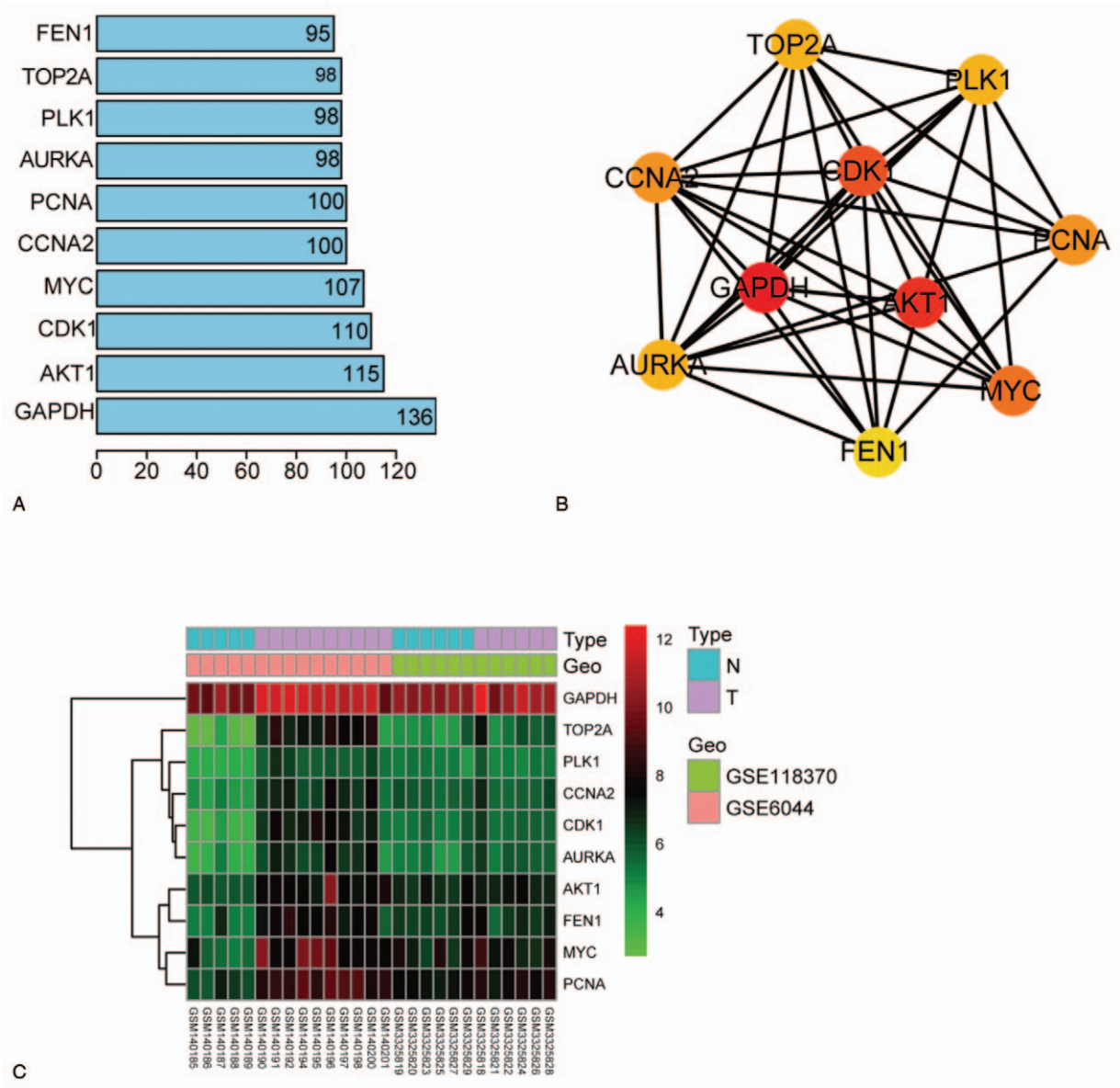

Figure 4. Interaction and expression of hub genes. (A) Vertical and horizontal coordinates represent hub gene name and degree of hub genes, respectively. (B) Protein–protein interaction network of the 10 hub genes visualized by plugin cytohubba of Cytoscape software (version 3.7.1). (C) Heatmap of the hub genes in GSE6044 and GSE118370 integrated dataset.

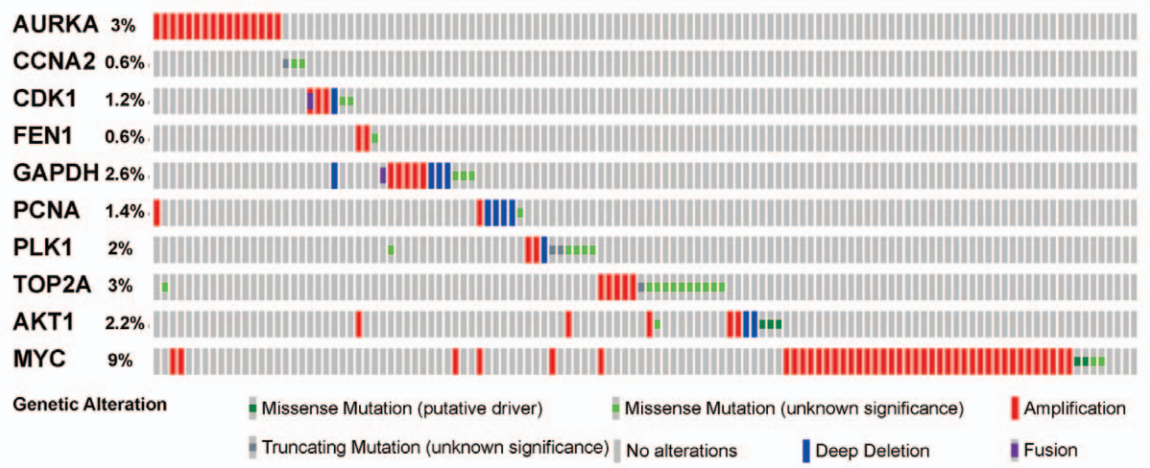

Figure 5. Mutation analysis in LUAD using cBioportal. LUAD = lung adenocarcinoma.

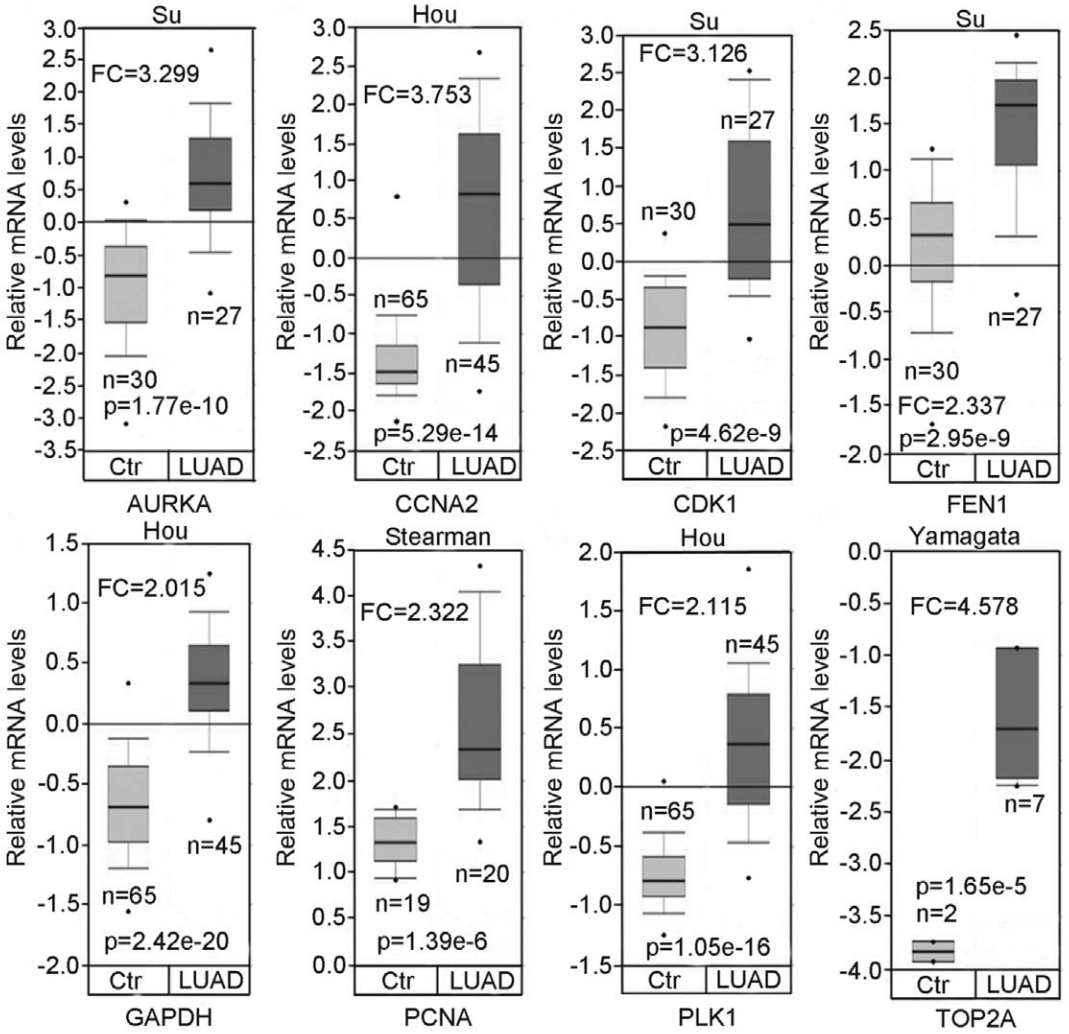

Figure 6. Expression levels of hub genes in the Oncomine database. We acquired the gene expression data from Su, Hou, Stearman, and Yamagata lung datasets, which were analyzed with Oncomine. The mRNA expression levels of (A) AURKA, (B) CCNA2, (C) CDK1, (D) FEN1, (E) GAPDH, (F) PCNA, (G) PLK1, and (H) TOP2A in normal lung tissues and LUAD specimens were compared. Preprocessed expression levels are Log2 normalized and median centered. Ctr = control, FC = fold change, LUAD = lung adenocarcinoma.

Table 2

The hub genes in the TCGA.

Gene	Log fold change	Log CPM	Adjust *P* value
AURKA	2.67	4.50	2.12e–20
CCNA2	2.93	4.30	7.32e–45
CDK1	2.48	5.04	5.83e–37
FEN1	1.62	4.99	1.90e–34
GAPDH	1.75	11.05	2.05e–31
PCNA	1.14	6.43	2.50e–20
PLK1	3.25	4.63	1.08e–55
TOP2A	3.90	6.85	9.56e–65
MYC	–0.28	6.22	6.10e–02
AKT1	–0.08	7.28	0.41

CPM = count per million, TCGA = The Cancer Genome Atlas.

hub genes. In addition, alterations of the hub genes were investigated using cBioportal for LUAD (The Cancer Genome Atlas, provisional) (Fig. 5). The hub genes were altered in 118 cases of 507 patients with LUAD (23%). These results suggest that there are missense mutations, amplification, and deep deletion in the hub genes.

3.4. Validation of hub gene expression in TCGA and Oncomine database

To validate the expression of hub genes, TCGA and Oncomine databases were performed. According to the Oncomine database, the mRNA expression levels of AURKA, CCNA2, CDK1, FEN1, GAPDH, PCNA, PLK1, and TOP2A significantly increased in LUAD specimens compared with normal lung tissues (Fig. 6). In the TCGA database, the expression of 8 hub genes mentioned above was consistent with their expression tendency in GSE6044 and GSE118370 integrated dataset (Table 2). However, according to the LUAD data in TCGA, the 2 hub genes, MYC and AKT1, were downregulated and not significantly differentially expressed at the same filtering criteria, which are the adjusted $P < .05$ and $|\log FC| > 1$. In the Oncomine database, the MYC

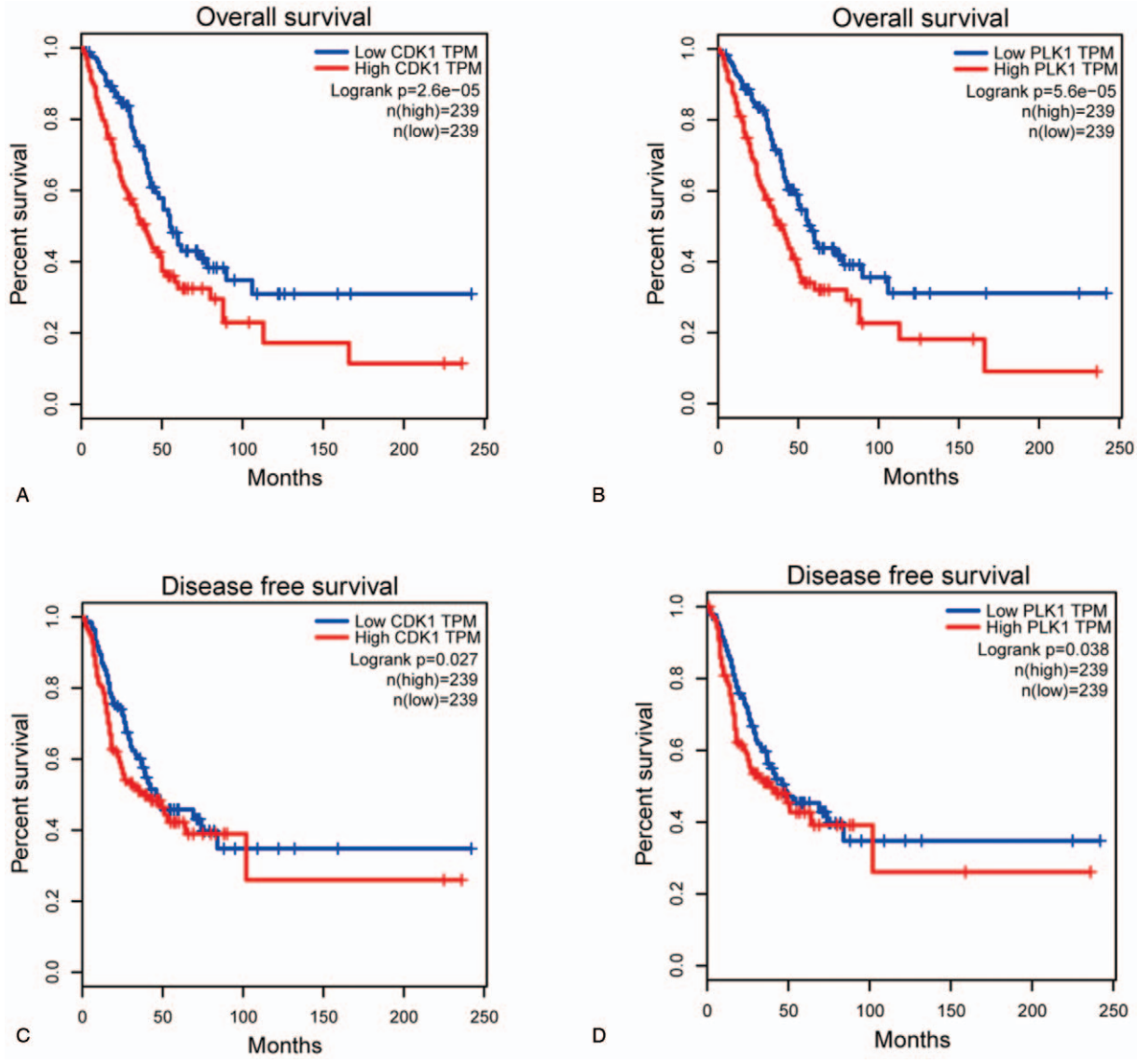

Figure 7. Overall survival (OS) (A and C) and disease-free survival (DFS) (B and D) of CDK1 and PLK1 in LUAD patients. LUAD = lung adenocarcinoma.

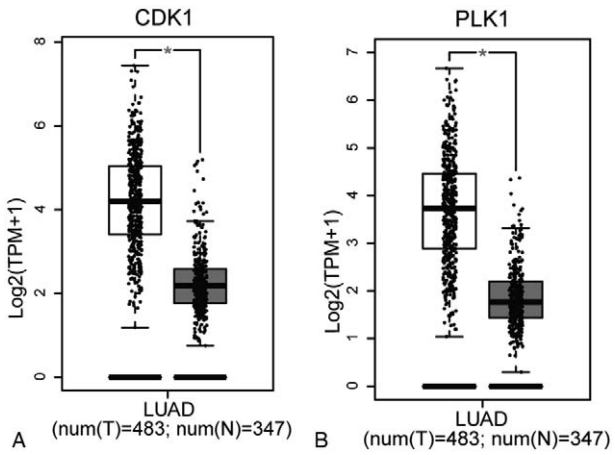

Figure 8. The expression of (A) CDK1 and (B) PLK1 in LUAD using GEPIA. *$P < .01$. LUAD = lung adenocarcinoma, TPM = transcripts per million.

and AKT1 mRNA expression in normal lung tissues had no corresponding data.

3.5. Association of CDK1 and PLK1, 2 cell cycle-related hub genes, in GEPIA with survival and clinicopathological stage of LUAD

To evaluate the prognosis of patients with LUAD, 2 cell cycle-related genes, CDK1 and PLK1, were selected from the 10 hub genes with $P < .05$ of both OS and DFS using the GEPIA database. Survival analysis suggested that higher expression levels of CDK1 and PLK1 predicted poor OS (Fig. 7A and C), and their expression levels were associated with DFS (Fig. 7B and D). We compared the mRNA expression of CDK1 and PLK1 between LUAD specimens and normal lung tissues (Fig. 8), which included a lot of sequenced data of normal tissues from GTEx database. The results further confirmed that their expression levels were higher in LUAD specimens than those in normal lung tissues. We also analyzed the correlation between mRNA expression and clinicopathological stage for the patients with LUAD (Fig. 9). The mRNA expressions of CDK1 and PLK1 were significantly different in varied tumor stages. These results suggest that their mRNA expression levels are associated with advanced tumor stage.

3.6. Analysis of the methylation of CDK1 and PLK1

DiseaseMeth version 2.0 was used to analyze the methylation levels of CDK1 and PLK1. The results showed that the mean methylation level in the promoter region of CDK1 ($P = 5.621e{-}03$) and PLK1 ($P = 7.344e{-}08$) were significantly lower in LUAD specimens, compared with those in normal tissues (Fig. 10). This indicates that decreased methylation level in the promoter region may contribute to higher expression of these 2 genes.

3.7. Association of CDK1 and PLK1 with tumor purity and immune cells

Immune cells are important components of tumor microenvironment. TIMER database was used to explore the correlation of CDK1 and PLK1 with tumor purity and infiltration of immune cells. The results demonstrated that CDK1 and PLK1 were not related to tumor purity. However, negatively correlations were observed between these 2 genes and infiltration of B cells, $CD4^+$ T cells, and macrophages (Fig. 11). These results suggest that the decrease of immune cell content may be one of the reasons for the poor prognosis of patients with high expression of CDK1 and PLK1.

3.8. Analysis of the cancer-related pathways of CDK1 and PLK1

To analyze the associated pathways of CDK1 and PLK1, GSEA analysis was performed using the TCGA-LUAD (FPKM) dataset. The results demonstrated that cell cycle, DNA replication, homologous recombination, mismatch repair, P53 signaling pathway, and small cell lung cancer signaling were significantly

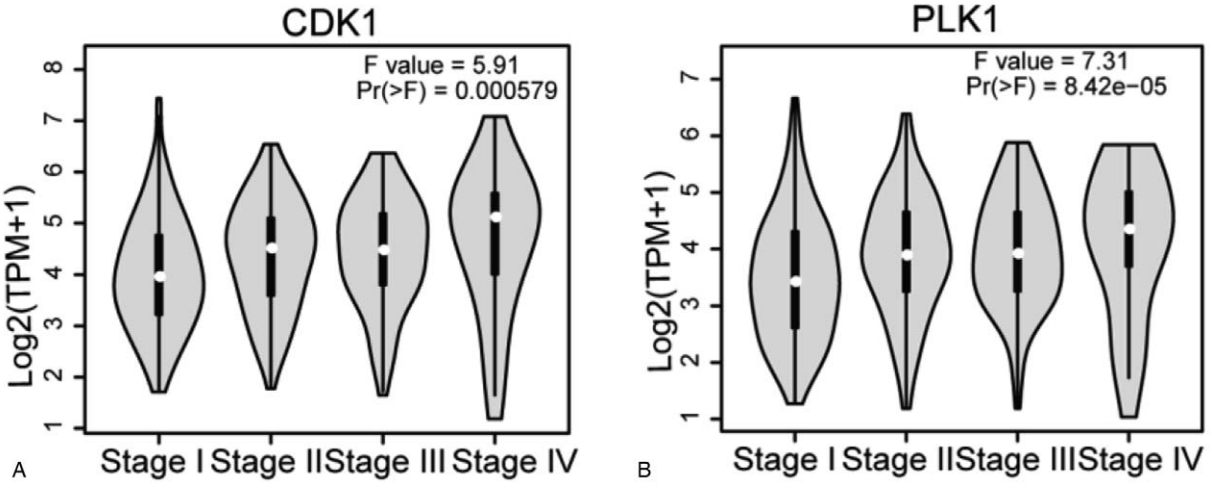

Figure 9. Correlation between the expression of (A) CDK1 and (B) PLK1 and tumor stage in LUAD using GEPIA. LUAD = lung adenocarcinoma, TPM = transcripts per million.

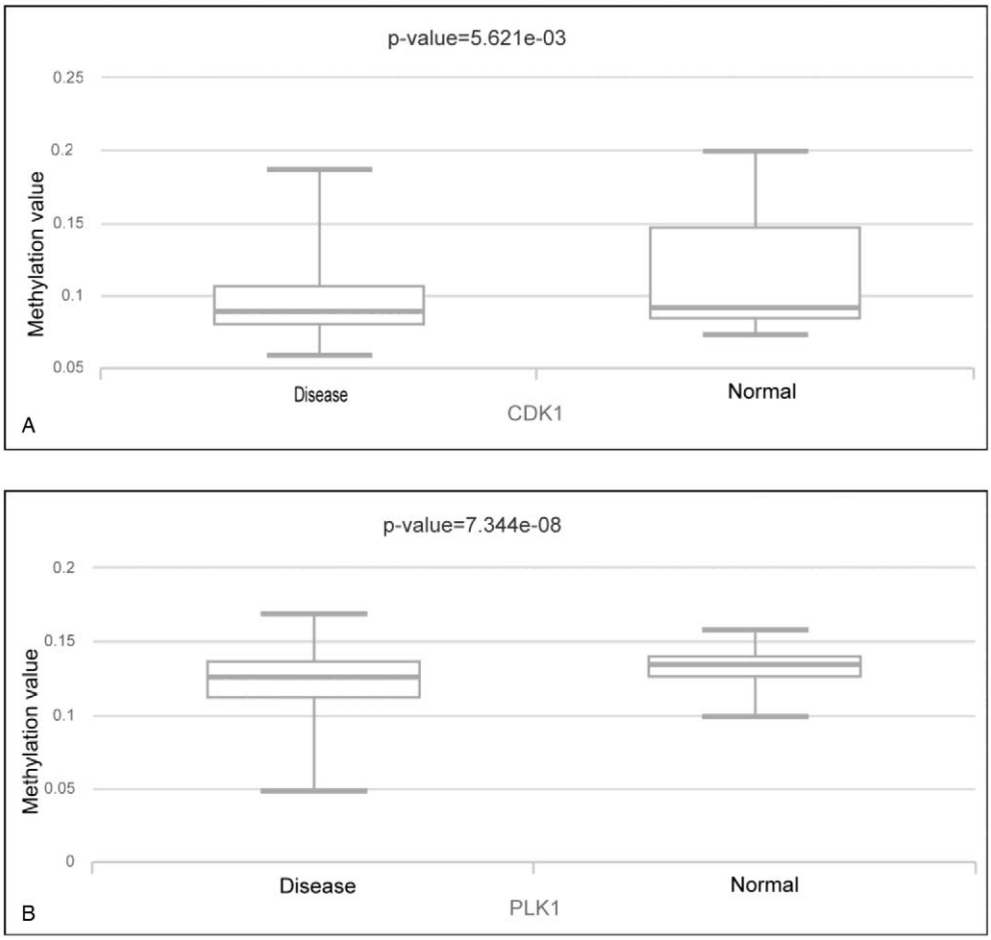

Figure 10. Methylation analysis of the two cell cycle-related hub genes in LUAD. The methylation levels in the promoter region of (A) CDK1 and (B) PLK1 in LUAD and normal lung tissues were analyzed using DiseaseMeth 2.0. LUAD = lung adenocarcinoma.

enriched in CDK1 and PLK1 high-expression phenotype (Fig. 12 and Table 3). These results suggest that these 2 genes may participate in LUAD tumorigenesis through above cancer-related pathways.

4. Discussion

Lung cancer is the most frequently diagnosed cancer and also the leading cause of cancer-related deaths.[30] LUAD, the most common histologic subtype of NSCLC, is one of the most malignant tumors and has a high mortality due to the absence of effective early diagnosis methods. To find therapeutic targets and diagnostic biomarkers, in the present study, we identified 10 hub genes with the highest connective degree, including AURKA, CCNA2, CDK1, FEN1, GAPDH, PCNA, PLK1, TOP2A, MYC, and AKT1, which were connected with cell cycle, DNA replication, base excision, and oocyte meiosis signaling pathways.

Over the past few decades, researchers.[31–35] have demonstrated that these hub genes are highly associated with prognosis of different malignancies. Studies have shown that overexpression of AURKA, a key kinase that regulates G2/M transition, has been identified in various malignancies, including bladder cancer, breast cancer, and prostate cancer.[32,36,37] Knockdown of AURKA and FEN1 inhibit the migration, invasion, and proliferation of LUAD cells and induce apoptosis.[38,39] More importantly, AURKA is a driving force for the evolution of resistance to third-generation EGFR inhibitors in LUAD and associated with epithelial mesenchymal transition.[40] Recently, it is found that the deregulation of transcriptional levels and posttranscriptional modification of GAPDH, which is traditionally considered as a critical enzyme of glycolytic process, is an important determinant for tumor cell survival.[41]

Cell cycle is the essential life activity for cellular growth. In this study, we chose 2 cell cycle-related hub genes, CDK1 and PLK1, for further investigation. There is a interplay between AURKA, CDK1, and PLK1, forming an axis that regulates mitotic entry and mediates G2/M transition.[42] Some inhibitors targeting CDK1 and PLK1 are being tested in clinical trials, showing effects of mitotic block and cell apoptosis for certain cancers.[43,44]

In this work, the cBioportal database analysis indicated that there were missense mutation, gene fusion, amplification, and deep deletion for certain hub genes in LUAD patients. To assess the therapeutic and diagnostic value of CDK1 and PLK1, we explored the GEPIA database. It was demonstrated that CDK1 and PLK1 were not only significantly upregulated, but also positively correlated with advanced tumor stage in LUAD patients. The patients with high expression of CDK1 and PLK1

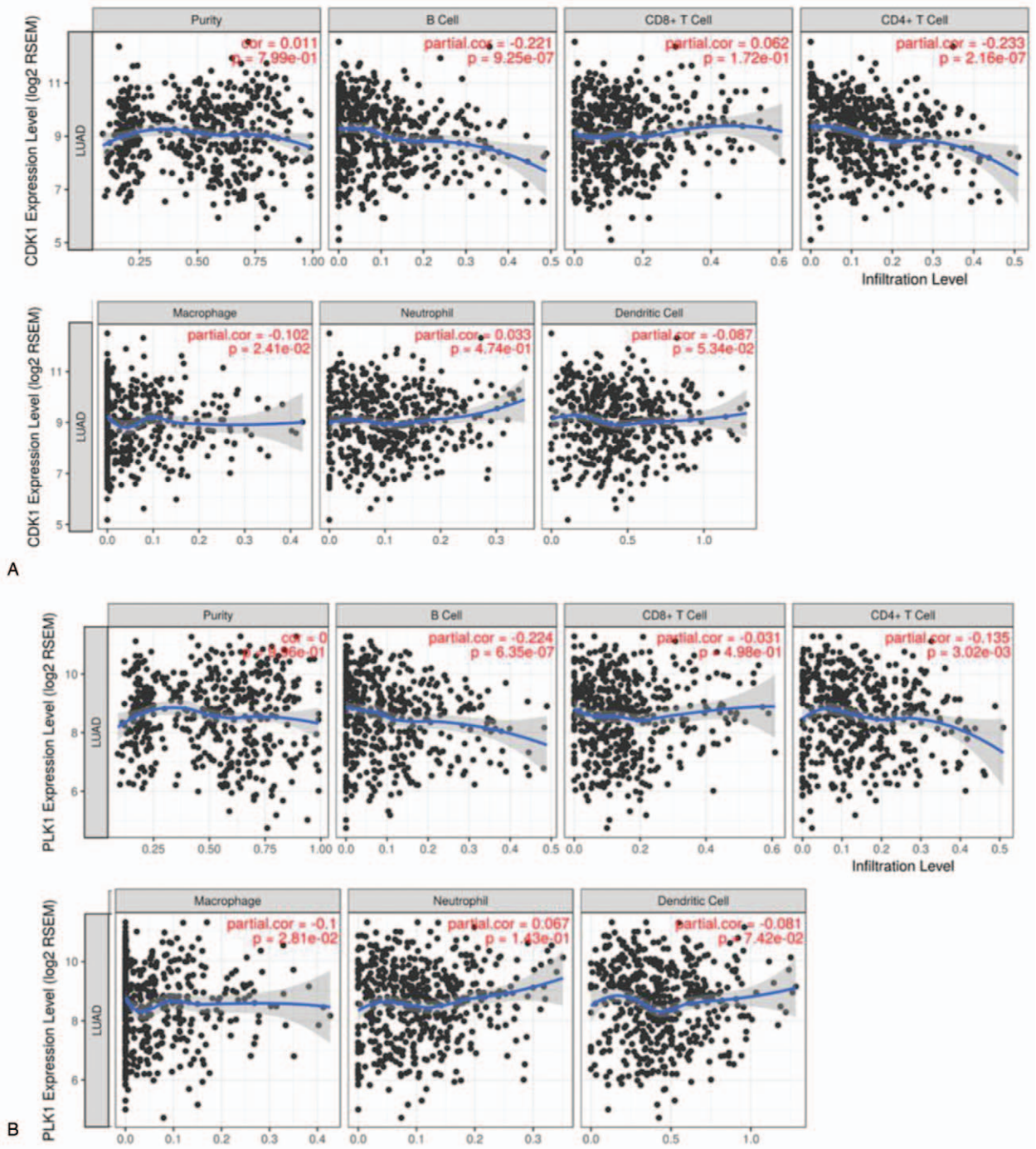

Figure 11. Association of the expression of the 2 cell cycle-related hub genes with immune infiltration in LUAD. (A) CDK1, (B) PLK1. $P < .05$. LUAD = lung adenocarcinoma.

had poorer OS than those with low expression, and their expression was also significantly associated with DFS. Epigenetic alteration, especially DNA methylation pattern and level, is thought to play a critical role in cancer-related gene expression.[45] We found that hypomethylation level in the promoter region of CDK1 and PLK1 may give a rise to their abnormal upregulation in LUAD tissues. Tumor microenvironment including tumor cells, stromal cells, and immune cells has a great impact on immunotherapy and clinical outcome.[46,47] Herein, we explored the relevance between immune infiltration and the expression of CDK1 and PLK1. Results showed that there was negative correlation between the expression of CDK1 and PLK1 and the infiltration abundance of B cells, $CD4^+$ T cells, and macrophages, which implies poor prognosis of LUAD patients with high expression of CDK1 and PLK1. To further explore the roles of CDK1 and PLK1, we performed GSEA analysis. Some important

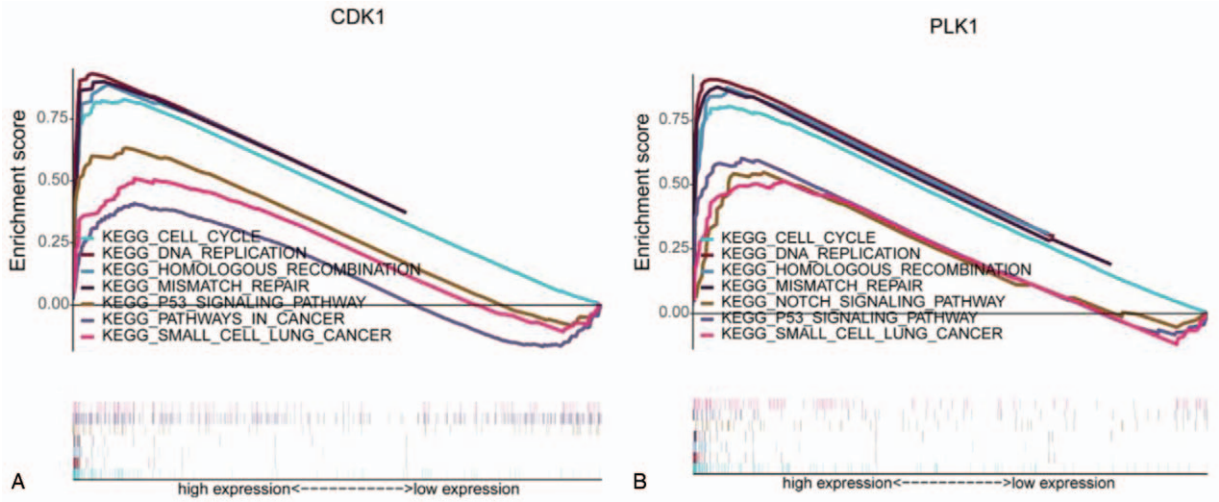

Figure 12. Multiple GSEA plots enriched in (A) CDK1 and (B) PLK1-related LUAD. GSEA = Gene Set Enrichment Analysis, LUAD = lung adenocarcinoma.

Table 3

Gene sets enriched in phenotype high.

Gene set name	NES	NOM *P* value	FDR *q* value
CDK1			
KEGG_CELL_CYCLE	2.603	.000	0.000
EGG_P53_SIGNALING_PATHWAY	2.329	.000	0.000
KEGG_HOMOLOGOUS_RECOMBINATION	2.319	.000	0.000
KEGG_MISMATCH_REPAIR	2.178	.000	4.42e–04
KEGG_DNA_REPLICATION	2.117	.000	0.001
KEGG_SMALL_CELL_LUNG_CANCER	1.908	.009	0.010
KEGG _PATHWAY_IN_CANCER	1.738	.011	0.037
PLK1			
KEGG_ CELL_CYCLE	2.512	.000	0.000
KEGG_P53_SIGNALING_PATHWAY	2.229	.000	6.36e–04
KEGG__HOMOLOGOUS_RECOMBINATION	2.264	.000	8.88e–04
KEGG_MISMATCH_REPAIR	2.104	.000	0.002
KEGG_DNA_REPLICATION	2.058	.000	0.003
KEGG_SMALL_CELL_LUNG_CANCER	1.930	.006	0.011
KEGG_NOTCH_ SIGNALING_PATHWAY	1.780	.010	0.033

FDR = false discovery rate, NES = normalized enrichment score, NOM = nominal.
Gene sets with NOM *P* value < .05 and FDR *q* value < 0.05 are considered as significant.

pathways including cell cycle, DNA replication, homologous recombination, mismatch repair, P53 signaling pathway, and small cell lung cancer signaling were significantly enriched in CDK1 and PLK1 high expression phenotype, suggesting their vital contribution to the growth and proliferation of LUAD cells.

5. Conclusions

In conclusion, CDK1 and PLK1 may be potential biomarkers and therapeutic targets for patients with LUAD. Low methylation levels and gene alterations may lead to expression deregulation of these 2 genes. Moreover, they are both enriched in some cancer-related pathways. These findings may provide a novel insight into the investigation of biomarkers and underlying molecular mechanisms of LUAD, thereby contributing to the development of potential diagnosis and therapeutic methods for this disease. Further studies, however, are needed to expound the biological functions of DEGs and hub genes in LUAD.

Author contributions

Conceptualization: Chunmei Dai, Ning Li.
Data curation: Shuzhen Li, Hua Li, Yajie Cao, Haiying Geng, Fu Ren, Keyan Li.
Formal analysis: Shuzhen Li, Hua Li, Yajie Cao, Haiying Geng, Fu Ren, Keyan Li.
Funding acquisition: Chunmei Dai, Ning Li.
Methodology: Shuzhen Li, Hua Li, Yajie Cao.
Project administration: Chunmei Dai, Ning Li.
Software: Shuzhen Li, Haiying Geng, Fu Ren, Keyan Li.
Supervision: Chunmei Dai, Ning Li.
Writing – original draft: Shuzhen Li.
Writing – review & editing: Chunmei Dai, Ning Li.

References

[1] Boloker G, Wang C, Zhang J. Updated statistics of lung and bronchus cancer in United States (2018). J Thorac Dis 2018;10:1158–61.

[2] Kuribayashi K, Funaguchi N, Nakano T. Chemotherapy for advanced non-small cell lung cancer with a focus on squamous cell carcinoma. J Cancer Res Ther 2016;12:528–34.
[3] Liu WT, Wang Y, Zhang J, et al. A novel strategy of integrated microarray analysis identifies CENPA, CDK1 and CDC20 as a cluster of diagnostic biomarkers in lung adenocarcinoma. Cancer Lett 2018; 425:43–53.
[4] Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 2008;83:584–94.
[5] Chen D, Song Z, Cheng G. Clinical efficacy of first-generation EGFR-TKIs in patients with advanced non-small-cell lung cancer harboring EGFR exon 20 mutations. Onco Targets Ther 2016;9:4181–6.
[6] Nie K, Jiang H, Zhang C, et al. Mutational profiling of non-small-cell lung cancer resistant to osimertinib using next-generation sequencing in chinese patients. Biomed Res Int 2018;2018:9010353.
[7] Shao C, Yang B, Zhao L, Wang S, Zhang J, Wang K. Tumor suppressor gene RBM5 delivered by attenuated Salmonella inhibits lung adenocarcinoma through diverse apoptotic signaling pathways. World J Surg Oncol 2013;11:123.
[8] Kim Y, Ko J, Cui Z, et al. The EGFR T790 M mutation in acquired resistance to an irreversible second-generation EGFR inhibitor. Mol Cancer Ther 2012;11:784–91.
[9] Xu M, Xie Y, Ni S, Liu H. The latest therapeutic strategies after resistance to first generation epidermal growth factor receptor tyrosine kinase inhibitors (EGFR TKIs) in patients with non-small cell lung cancer (NSCLC). Ann Transl Med 2015;3:96.
[10] Karachaliou N, Sosa A, Barron F, Gonzalez Cao M, Santarpia M, Rosell R. Pharmacological management of relapsed/refractory NSCLC with chemical drugs. Expert Opin Pharmacother 2017;18:295–304.
[11] Al Zeyadi M, Dimova I, Ranchich V, et al. Whole genome microarray analysis in non-small cell lung cancer. Biotechnol Biotechnol Equip 2015;29:111–8.
[12] Barrett T, Edgar R. Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. Methods Enzymol 2006;411: 352–69.
[13] Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015;19:A68–77.
[14] Liu Y, Xie D, He Z, Zheng L. Integrated analysis reveals five potential ceRNA biomarkers in human lung adenocarcinoma. PeerJ 2019;7: e6694.
[15] Xie L, Dang Y, Guo J, et al. High KRT8 expression independently predicts poor prognosis for lung adenocarcinoma patients. Genes (Basel) 2019;10:36.
[16] Lu CJ, Chen HZ, Shan ZX, Yang LX. Identification of differentially expressed genes between lung adenocarcinoma and lung squamous cell carcinoma by gene expression profiling. Mol Med Rep 2016;14:1483–90.
[17] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.
[18] Xia L, Su X, Shen J, et al. ANLN functions as a key candidate gene in cervical cancer as determined by integrated bioinformatic analysis. Cancer Manag Res 2018;10:663–70.
[19] Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16:284–7.
[20] Klonowska K, Czubak K, Wojciechowska M, et al. Oncogenomic portals for the visualization and analysis of genome-wide cancer data. Oncotarget 2016;7:176–92.
[21] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.
[22] Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 2004;6:1–6.
[23] Su LJ, Chang CW, Wu YC, et al. Selection of DDX5 as a novel internal control for Q-RT-PCR from microarray data using a block bootstrap re-sampling scheme. BMC Genomics 2007;8:140.
[24] Hou J, Aerts J, den Hamer B, et al. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. PLoS One 2010;5:e10312.
[25] Stearman RS, Dwyer-Nield L, Zerbe L, et al. Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. Am J Pathol 2005;167:1763–75.
[26] Yamagata N, Shyr Y, Yanagisawa K, et al. A training-testing approach to the molecular classification of resected non-small cell lung cancer. Clin Cancer Res 2003;9:4695–704.
[27] Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98–102.
[28] Xiong Y, Wei Y, Gu Y, et al. DiseaseMeth version 2.0: a major expansion and update of the human disease methylation database. Nucleic Acids Res 2017;45:D888–95.
[29] Li T, Fan J, Wang B, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017;77: e108–10.
[30] 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.
[31] Hwang J, Sung W, Tu H, et al. The overexpression of FEN1 and RAD54B may act as independent prognostic factors of lung adenocarcinoma. PLoS One 2015;10:e0139435.
[32] Guo M, Lu S, Huang H, et al. Increased AURKA promotes cell proliferation and predicts poor prognosis in bladder cancer. BMC Syst Biol 2018;12:118.
[33] Sarin SK, Kedarisetty CK, Abbas Z, et al. Acute-on-chronic liver failure: consensus recommendations of the Asian Pacific Association for the Study of the Liver (APASL) 2014. Hepatol Int 2014;8:453–71.
[34] Abdel-Fatah T, Russell R, Albarakati N, et al. Genomic and protein expression analysis reveals flap endonuclease 1 (FEN1) as a key biomarker in breast and ovarian cancer. Mol Oncol 2014;8:1326–38.
[35] Piao J, Zhu L, Sun J, et al. High expression of CDK1 and BUB1 predicts poor prognosis of pancreatic ductal adenocarcinoma. Gene 2019; 701:15–22.
[36] Cirak Y, Furuncuoglu Y, Yapicier O, Aksu A, Cubukcu E. Aurora A overexpression in breast cancer patients induces taxane resistance and results in worse prognosis. J BUON 2015;20:1414–9.
[37] Li S, Xuan Y, Gao B, et al. Identification of an eight-gene prognostic signature for lung adenocarcinoma. Cancer Manag Res 2018;10: 3383–92.
[38] He L, Luo L, Zhu H, et al. FEN1 promotes tumor progression and confers cisplatin resistance in non-small-cell lung cancer. Mol Oncol 2017;11:640–54.
[39] Zhong N, Shi S, Wang H, et al. Silencing Aurora-A with siRNA inhibits cell proliferation in human lung adenocarcinoma cells. Int J Oncol 2016;49:1028–38.
[40] Shah KN, Bhatt R, Rotow J, et al. Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. Nat Med 2019;25:111–8.
[41] Zhang JY, Zhang F, Hong CQ, et al. Critical protein GAPDH and its regulatory mechanisms in cancer cells. Cancer Biol Med 2015;12: 10–22.
[42] Asteriti IA, De Mattia F, Guarguaglini G. Cross-talk between AURKA and Plk1 in mitotic entry and spindle assembly. Front Oncol 2015;5:283.
[43] Gutteridge RE, Ndiaye MA, Liu X, Ahmad N. Plk1 inhibitors in cancer therapy: from laboratory to clinics. Mol Cancer Ther 2016;15:1427–35.
[44] Heptinstall AB, Adiyasa I, Cano C, Hardcastle IR. Recent advances in CDK inhibitors for cancer therapy. Future Med Chem 2018;10: 1369–88.
[45] Akar RO, Selvi S, Ulukaya E, Aztopal N. Key actors in cancer therapy: epigenetic modifiers. Turk J Biol 2019;43:155–70.
[46] Reina-Campos M, Moscat J, Diaz-Meco M. Metabolism shapes the tumor microenvironment. Curr Opin Cell Biol 2017;48:47–53.
[47] Tang H, Qiao J, Fu YX. Immunotherapy and tumor microenvironment. Cancer Lett 2016;370:85–90.