

WITHDRAWN: PLK1 is a Potential Prognostic Factor Associated With the Tumor Microenvironment During Lung Adenocarcinoma Progression

Zhehao Huang

China-Japan Union Hospital of Jilin University

Xianglan Li

fatcell@jlu.edu.cn

Department of Dermatology, China-Japan Union Hospital of Jilin University, Changchun 130033, Jilin, China. <https://orcid.org/0000-0001-9319-8156>

Primary research

Keywords: lung adenocarcinoma, tumor immune microenvironment, prognosis, PLK1, tumor-infiltrating immune cells

Posted Date: November 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-105349/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

EDITORIAL NOTE:

The full text of this preprint has been withdrawn by the authors while they make corrections to the work. Therefore, the authors do not wish this work to be cited as a reference. Questions should be directed to the corresponding author.

Abstract

Background: Lung adenocarcinoma (LUAD) accounts for more than 40% of lung cancer cases worldwide, and the 5-year survival rate of LUAD patients is less than 10% due to a lack of reliable therapeutics. Here, we sought to identify a new therapeutic target for LUAD via bioinformatics analysis.

Methods: Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was conducted for the differentially expressed genes (DEGs) identified in 551 samples from the Cancer Genome Atlas (TCGA) database. Gene set variation analysis (GSVA) and the CIBERSORT method were performed to estimate the expression profile of biological pathways and the population of tumor-infiltrating immune cells (TICs) in the TCGA dataset. DEGs were further analyzed by protein-protein interaction (PPI) network and Cox regression analyses, followed by RT-PCR and western blotting for confirmation.

Results: Cell cycle was the only shared pathway identified by the KEGG and GSVA interaction analyses. Cell cycle score was positively associated with clinical characteristics (age, clinical stage, and metastasis) and negatively associated with overall survival in LUAD patients. PPI and Cox analyses identified *PLK1* as a prognostic factor, which was positively correlated with clinical stage and negatively correlated with overall survival in LUAD patients. CIBERSORT analysis indicated that *PLK1* expression was significantly positively correlated with CD8+ and activated memory CD4+ T cells, and negatively correlated with activated natural killer cells. Additionally, *PLK1* overexpression resulted in increased immune cytotoxic metrics, such as cytolytic activity score, IFN- γ score, and IFN- γ level.

Conclusions: *PLK1* may be useful for survival estimation in LUAD patients due to its strong correlation with features of TICs in the tumor immune microenvironment.

Background

Lung cancer is the most common solid tumor in the world, and mainly consists of small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC) [1]. The incidence of NSCLC accounts for about 85% of lung cancers, and it can be further categorized into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma [2, 3]. Lung adenocarcinoma (LUAD) accounts for approximately 60% of NSCLC [4]. Although there have been significant advances in therapeutic strategies, such as multiple targeted therapies and novel immunotherapy approaches, LUAD patients have a 5-year survival rate of less than 10% [5]. This is likely due to several LUAD characteristics, such as early dissemination and metastasis. In addition, the molecular mechanisms behind the initiation and progression of LUAD are poorly understood. Further exacerbating this problem, the underlying interactions between LUAD and tumor-infiltrating immune cells (TICs) have not been clearly elucidated.

Many genes have been reported to be correlated with LUAD progression in various different studies. *PIG3*, which is induced by p53, has been shown to have higher expression in LUAD patients with lymph node metastasis than those without, and it is thought to play essential roles in LUAD progression via activation of the FAK/Src/paxillin pathway [6]. ENO1, a glycolysis enzyme, is responsible for the glucose

metabolism and has also been shown to promote tumor progression in LUAD. In a recent study, it was suggested that silencing circ-ENO1 reduced glycolysis, repressed cell growth, and subsequently induced apoptosis via suppression of the *ENO1* gene expression in LUAD patients [7]. Foster *et al.* suggested that *ATMIN* typically had lower expression in LUAD tissues, and heterozygous *ATMIN* deletion contributed to cancer cell proliferation and cancer grade in a mouse model of LUAD, indicating that *ATMIN* serves as a tumor suppressor for LUAD [8]. A previous study has identified *PLK1* as a key gene correlated with poor survival for LUAD [9]. However, the correlation of *PLK1* expression with TICs has not been demonstrated.

In recent years, significant evidence has been found indicating that TICs in the tumor microenvironment (TME) play essential roles in the regulation of malignance and improvement of clinical outcomes in many tumors [10-12]. In particular, CD4+ and CD8+ T cells, regulatory T-cells (Tregs), dendritic cells (DCs), and macrophages appear to be highly predictive of patient outcome [13-15]. Zhang *et al.* found that 55% of ovarian cancer patients had tumor-infiltrating T cells, and the 5-year overall survival rate for those patients with tumor-infiltrating T cells was significantly higher than those without [16]. Based on 179 triple-negative breast cancer (TNBC) patients, Vihervuori *et al.* found that the survival rate of patients with at least a 14% fraction of CD8+ T cells in an 18-year follow-up was significantly higher than those with less than 14%, indicating that CD8+ T cells are correlated with better clinical outcomes [17]. Kim *et al.* also found that a high level of tumor-infiltrating lymphocytes (TILs) was significantly negatively correlated with histologic grade of tumors [18]. Meanwhile, LUAD patients with high TILs had better progression-free survival and overall survival compared to those with low TILs [19]. The focus of this paper was to identify different TIC compositions among LUAD and normal lung tissues, verify the association of hub genes with TICs, and further elucidate the underlying mechanisms behind the LUAD. The detailed workflow of this work is provided in **Fig. 1**.

Methods

Data collection and identification of differentially expressed genes (DEGs)

Transcriptome RNA-sequencing data (HTSeq-FPKM) of LUAD patients were retrieved from the Cancer Genome Atlas (TCGA) database (<https://cancergenome.nih.gov/>), which contained 497 samples from the lungs of LUAD patients and 54 samples from normal lung tissue. The GSE31210 gene expression profile was downloaded from Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>), including 226 samples from LUAD patients and 20 normal samples. The R packages “edgeR” and “limma” were applied to identify DEGs in TCGA dataset and normalize the GSE31210 matrix data, respectively. An adjusted p -value <0.05 and $|\log\text{FC}|>2$ were used for the cutoff criteria for DEGs.

Gene set variation analysis (GSVA) and Kyoto Encyclopedia of genes and Genomes (KEGG) enrichment analysis

The R packages “GSVA” and “clusterProfiler” were used for GSVA ($|\log\text{FC}|>0.3$ and adjusted p -value <0.05 set as cutoff criteria) and KEGG enrichment analysis (adjusted p -value <0.05 set as cutoff criteria), respectively. The “c2.cp.kegg.v7.1.symbols.gmt” dataset was downloaded from Molecular Signature

Database (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>) and served as the reference gene set.

Protein-protein interaction (PPI) network construction and Cox regression analysis

Data containing known interactions among different proteins was retrieved from the Search Tool for the Retrieval of Interacting Genes database (STRING, <https://string-db.org/>), and a combined score of ≥ 0.4 was used as a cutoff. Next, a PPI network was visualized by Cytoscape software (version 3.6.0, <https://cytoscape.org/>). The R package “survival” was used for Cox regression analysis, and genes with p -value <0.001 were shown in the resulting forest plot.

Survival analysis and correlation analysis

The R packages “survival” and “survminer” were used for estimation of overall survival in LUAD patients. The Kaplan-Meier method was used to plot the survival curves and p -value <0.05 was considered significant. The Pearson coefficient was used for correlation analysis via Graphpad prism 7.0.

Validation of gene expression

The TCGA dataset and GSE31210 profile were used to perform statistical analyses for the hub genes. A meta-analysis was developed on the Oncomine database website (<https://www.oncomine.org>) to identify the expression pattern of hub genes among LUAD and normal lung samples. The immunohistochemical results of corresponding genes from the Human Protein Atlas (HPA, <https://www.proteinatlas.org>) were adopted to illustrate the protein expression levels. Meanwhile, the mRNA and protein expression levels of certain genes were measured by quantitative real-time polymerase chain reaction (qRT-PCR) and western blotting.

Assessment of immune activity of the TME

The CIBERSORT computational method was adopted to estimate the relative fractions of TICs in all LUAD samples. The R package “corrplot” was applied for plotting the heatmap of correlations among TICs. The FPKM values of TCGA dataset were converted to Transcript Per Million (TPM) values, and the immune cytolytic activity score (CAS) was then calculated as the geometric mean of the GZMA and PRF1 expression in TPM. IFN- γ score was defined as the mean of CD8A, GZMA, GZMB, IFN- γ , EOMES, CXCL9, CXCL10, and TBX21 expression in FPKM.

Cell culture

HBE, A549, and PC-9 cells were cultured in high glucose Dulbecco's Modified Eagle's media (DMEM; Hyclone, Logan, Utah, USA) containing 10% (v/v) fetal bovine serum (FBS; Gibco, Grand Island, NY, USA), and 1% penicillin/streptomycin (MRC, Jintan, China) at 37 °C and 5% CO₂. All cell lines were obtained from the Shanghai Cell Bank of Chinese Academy of Medical Sciences (Shanghai, China).

qRT-PCR

Total RNA was extracted using Total RNA Extraction Kit (Solarbo, Beijing, China), and reverse transcription was performed using the first-strand cDNA synthesis kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturers' instructions. RT-PCR was conducted using Premix Ex Taq SYBR Green PCR (TaKaRa, Dalian, China) on an ABI PRISM 7300 plus (Applied Biosystems, Foster City, CA, USA) following the manufacturer's protocols. The primer sequences used in the study were as followed: PLK1 forward primer AAAGAGATCCCGGAGGTCCTA, PLK1 reverse primer GGCTGCGGTGAATGGATATTTTC; GAPDH forward primer GGAGCGAGATCCCTCCAAAAT, GAPDH reverse primer GGCTGTTGTCATACTTCTCATGG.

Western blotting assay

After extraction of cell lysate and quantification of protein concentrations, proteins were separated on 10% SDS-PAGE gels and then transferred to 0.45 μ m PVDF membranes (ThermoFisher, Waltham, MA, USA). The membranes were blocked with 5% non-fat milk, incubated with primary antibody of PLK1 (Abcam, Cambridge, UK, dilute 1:1000, ab189139) and β -actin (Abcam, dilute 1 1000, ab8226) at 4 °C overnight, treated with horseradish peroxidase-conjugated secondary antibody (Bioss, Beijing, China) at room temperature for 1 hour, and visualized on a Tanon 5200 (Tanon, Shanghai, China).

Statistical analysis

Statistical analysis was performed using R software (version 3.6.0) and Graphpad prism 7.0. The correlation analysis was carried out via Pearson's correlation tests. A log-rank test was used to compare statistical differences in the Kaplan-Meier analysis. The results of qRT-PCR and western blotting were analyzed with Student's *t*-test. All experiments were performed at least three independent times and *p*-value<0.05 was considered statistically significant.

Results

Upregulation of the cell cycle pathway in LUAD

GSVA was used to explore different KEGG pathways in samples based on the TCGA dataset. Compared with normal lung samples, a total of 48 KEGG pathways (19 downregulated and 29 upregulated) were significantly different in LUAD samples (**Fig. 2a** and **Supplementary Table S1**), the patterns of which were visualized in a heatmap (**Fig. 2b**). The 1,419 DEGs (310 downregulated and 1,109 upregulated) identified from TCGA dataset (**Fig. 2c**) were mainly enriched in cell cycle, protein digestion and absorption, ECM-receptor interaction, complement, coagulation cascades, and p53 signaling pathway (**Fig. 2d** and **supplementary Table S2**). Cell cycle was the only common pathway identified by GSVA and KEGG enrichment analyses (**Fig. 2e**).

Correlation of cell cycle pathway with overall survival (OS) rate and clinical characteristics in LUAD patients

We further analyzed the TCGA dataset to evaluate the possibility of utilizing cell cycle pathway genes as prognostic markers for LUAD patients. The results of this analysis indicated that there were higher cell cycle scores in LUAD samples than normal samples (**Fig. 3a**). The cell cycle score was significantly higher in male ($p=0.0036$), M1 ($p=0.0186$), N1-3 ($p=0.0007$), and stage III&IV ($p=0.0007$) subgroups compared with other subgroups (**Fig. 3b-e**). Additionally, the LUAD patients were stratified into high and low cell cycle score subgroups according to their median cell cycle scores. The high cell cycle score subgroup was distinctively associated with worse OS in LUAD patients as shown in **Fig. 3f**.

Interaction analysis of PPI network and univariate Cox regression analysis

The 23 DEGs involved in the cell cycle pathway were provided in **Table 1** and were used for construction of a PPI network. This analysis revealed that there were 23 nodes and 240 edges in the PPI network (**Fig. 4a**), and the connectivity degree of each gene in the network was calculated and displayed in the histogram (**Fig. 4b**). *BUB1B*, *CDC6*, *CDK1*, *MCM2*, *CDC20*, *CCNA2*, *CCNB1*, *CCNB2*, *CDC45*, *PLK1*, *CHEK1*, and *CCNE1* were the DEGs with the highest degrees of connectivity. Additionally, *CCNA2*, *PTTG1*, *CDC25C*, *CCNB1*, and *PLK1* were significantly associated with OS in LUAD patients, as determined by univariate Cox analysis for the 23 DEGs (**Fig. 4c**). We found that *CCNA2*, *CCNB1*, and *PLK1* were shared between the twelve leading genes in the PPI network and the five genes determined by univariate Cox analysis (**Fig. 4d**).

Identification of *PLK1* as a prognostic factor for the LUAD patients

To further evaluate the prognostic value of *PLK1*, we conducted univariate and multivariate Cox regression analyses based on the TCGA dataset and GSE31210 profile. As indicated by **Table 2**, *PLK1* expression and clinical stage had a significant association with OS in LUAD patients determined by Cox regression analysis in both the TCGA dataset and GSE31210 profile ($p<0.05$). Compared with stage I, patients with stage II, III, and IV had higher *PLK1* expression according to the TCGA dataset ($p<0.05$, **Fig. 5a**), as well patients with stage II in the GSE31210 profile ($p<0.05$, **Fig. 5b**). *PLK1* expression had a significantly positive correlation with stage classification in both the TCGA dataset ($r=0.2851$, $p<0.0001$, **Fig. 5c**) and GSE31210 profile ($r=0.2996$, $p<0.0001$, **Fig. 5d**). In addition, the LUAD patients were grouped into subgroups with high and low median *PLK1* expression. The *PLK1* high expression subgroup was significantly associated with worse OS of LUAD patients in both the TCGA dataset (**Fig. 5e**) and GSE31210 profile (**Fig. 5f**).

High expression of *PLK1* in LUAD samples

Both paired and unpaired difference analyses of the TCGA dataset revealed that *PLK1* was expressed at a significantly higher level in LUAD samples compared to normal samples (**Fig. 6a and b**). Meta-analysis indicated that *PLK1* expression was significantly higher in LUAD tissues in five different datasets (Beer Lung [20], Garber Lung [21], Hou Lung [22], Stearman Lung [23], and Su Lung [24], **Fig. 6c**). Immunohistochemical images from HPA indicated that *PLK1* was upregulated in LUAD tissues compared to normal samples (**Fig. 6d**). We also found that *PLK1* expression was significantly upregulated in LUAD

versus normal lung tissues in the GSE31210 profile (**Fig. 6e**). Additionally, both qRT-PCR and western blotting suggested that *PLK1* was distinctively elevated in A549 and PC-9 cell lines compared to HBE cells (**Fig. 6f and g**).

Association of *PLK1* with immune activity of TME

To investigate the correlation between *PLK1* and LUAD in the immune microenvironment, we further analyzed the TCGA dataset with a focus on TICs, CAS, and IFN- γ score. As shown in **Fig. 7**, the proportion of 21 kinds of infiltrating immune cells in LUAD samples was calculated with the CIBERSORT algorithm. The TIC profiles were used to generate a histogram and the correlation among TICs was visualized in a heatmap. The proportion scores of 12 kinds of TICs were distinctively different between high and low *PLK1* expression subgroups (**Fig. 8a**). Additionally, 13 kinds of TICs were significantly associated with *PLK1* expression via Pearson correlation analysis (**Fig. 8b**). The 12 kinds of common TICs were determined by both difference analyses and correlation tests (**Fig. 8c and Supplementary Table S3**), and the resulting altered TICs were termed *PLK1*-related TICs. Compared with normal tissues, CAS, IFN- γ expression, and IFN- γ score were significantly elevated in LUAD samples (**Fig. 9a-c**). Compared to the low *PLK1* expression subgroup, CAS, IFN- γ expression, and IFN- γ score were significantly increased in high *PLK1* expression subgroup (**Fig. 9d-f**). Additionally, CAS, IFN- γ expression, and IFN- γ score had significant associations with *PLK1* expression ($p < 0.0001$, $p < 0.0001$, and $p = 0.0013$, respectively) (**Fig. 9g-i**).

Discussion

LUAD accounts for more than 40% of lung cancer worldwide and is one of the deadliest cancers [25]. Currently, the most common strategy for treating LUAD involves surgery, which benefits greatly from early detection. The survival rate of patients who receive surgical intervention in combination with chemotherapy further increases by 5 to 10% [26]. Despite these interventions, the 5-year survival rate for LUAD is still unacceptably low. Additionally, limited knowledge about molecular and cellular mechanisms involved in LUAD have made it difficult to identify new therapeutics and strategies for LUAD patients. In this work, *PLK1* was identified as a hub gene involved in cell cycle progression which had positive correlations with clinical stage and cytotoxic activity in the TME.

PLK1 belongs to the polo-like kinases (PLKs) family, which is a subgroup of serine/threonine protein kinases [27]. It impacts the cell cycle in multiple ways, including acting as a G2/M checkpoint, regulating the centrosome during cell cycle, and coordinating the spindle assembly and chromosome segregation [28]. Several studies have demonstrated that high expression of *PLK1* commonly occurs in malignant tumors, and its overexpression is negatively correlated with prognosis in tumor patients [29-31]. Based on 25 datasets and 12 publications, Lin et al. found that the mRNA and protein expression levels of *PLK1* were significantly increased in gastric cancer compared with normal gastric tissues, and high *PLK1* expression was correlated with unsatisfactory OS ($p < 0.001$), lymph node metastasis ($p = 0.013$), and advanced TNM stage ($p = 0.038$) [32]. Meanwhile, high expression of *PLK1* protein had a prominent association with patients of ovarian clear cell carcinoma (OCCC), and silencing *PLK1* sensitized OCCC

cell lines (OVTOKO and KK) to the cisplatin treatment via activation of autophagy and apoptosis [33]. Based on 3,173 samples, Takeshita et al. demonstrated that *PLK1* was highly expressed in TNBC and its overexpression was dominantly associated with TICs (CD8+ T cells and macrophage) and unfavorable prognosis in the whole cohort [34]. They also found that both cell cycle-related gene sets (G2/M check point and E2F targets) and MYC target gene sets were enriched in samples with high *PLK1* expression via Gene Set Enrichment Analysis. It also has been reported that *PLK1* up-regulated TGF- β signaling to drive cancer cell invasiveness in vitro and served as a strong predictor for diminished survival rates in metastatic NSCLC patients [35]. Many previous studies have suggested that downregulating *PLK1* expression could strongly repress cell growth and lead to apoptosis in many types of tumors, indicating that *PLK1* could be a promising treatment target for cancers [36-38].

Previous studies have demonstrated that TICs play a significant role in the tumor immune microenvironment, and directly or indirectly affect tumor progression and survival of cancer patients [39-41]. Based on 212 pancreatic cancer cases, Ino et al. found that patients with high levels of tumor-infiltrating CD4+ and CD8+ T cells survived longer than those with low levels [42]. In NSCLC, most tumor-infiltrating CD4+ and CD8+ T cells expressed CD69 and a subset also produced CD103, both of which are biomarkers of resident memory T cells [43]. It has been reported that CD103+ T cells not only served as a guard against secondary infections, but also attacked tumor cells [44, 45]. Both CD8+ and CD4+ T cells produced inhibitory receptors PD-1, 2B4, and CTLA-4. Particularly, CD103+CD4+ T cells were the most dominant producer of TNF- α and IFN- γ [46-49]. Additionally, cytotoxic T cells could release two key cytolytic effectors, perforin (PRF1) and pro-apoptotic granzymes (GZMA), to destroy cancer cells via the granzyme-perforin pathway [50]. Based on 280 tumor samples, Mandal et al. found that CD56dim natural killer (NK) cells were the most predominant infiltrating cells in advanced head and neck squamous cell carcinoma (HNSCC), and their presence was positively correlated with survival rate in HNSCC patients [51]. Another study found that the population of both infiltrating NK cells and macrophages served as strong predictors for the survival of stage II&III esophageal cancer patients [52]. Previous studies have also suggested that despite the powerful cytotoxic activity of NK cells in response to tumor cells, this activity might be limited by the TME due to interference with NK cell activation pathways and NK cell infiltration [53, 54]. In addition, the TME-mediated interference has also been shown to suppress the functions of infiltrating DCs in maintaining balance between CD8+ T cell immunity and tolerance to tumor antigens, which can further promote tumor progression [55]. This TME-mediated suppression has prevented the development of many tumor immunotherapies, making a better understanding of the mechanisms behind it key to developing new treatments.

Although *PLK1* was shown to be associated with clinical stage and the tumor immune microenvironment, this study still contains some limitations. First, although qRT-PCR and western blotting were used to verify *PLK1* expression in LUAD cell lines, the proposed mechanism of *PLK1* regulation of tumor progression via the cell cycle needs further validation. Second, clinical samples were not applied in the work because of limitations on experimental conditions. Third, the correlation analyses of clinical characteristics were not satisfactory due to the unbalanced clinical data in datasets. Finally, the association of *PLK1* with TICs and molecular mechanisms were not fully elucidated.

Conclusions

This work represents a significant step forward in elucidating the molecular mechanisms behind LUAD progression. *PLK1* was revealed as both a potential therapeutic target for LUAD, as well as a novel prognostic factor for LUAD patient survival. *PLK1* is strongly correlated with several TICs in the tumor immune microenvironment. Further investigation should be conducted to improve the accuracy of the combined analysis and further illuminate the roles of *PLK1* in the TIC community during LUAD progression.

Abbreviations

LUAD: Lung adenocarcinoma; KEGG: Kyoto Encyclopedia of Genes and Genomes; DEGs: Differentially expressed genes; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; GSEA: Gene set variation analysis; TICs: Tumor-infiltrating immune cells; TME: Tumor microenvironment; PPI: Protein-protein interaction; STRING: Search tool for the retrieval of Interacting genes database; HPA: Human Protein Atlas; qRT-PCR: Quantitative real-time polymerase chain reaction; OS: Overall survival; DCs: Dendritic cells; NK: Natural killer; SCLC: Small cell lung carcinoma; NSCLC: Non-small cell lung carcinoma; FC: Fold change; HR: Hazard ratio; CI: Confidence interval; CAS: Cytolytic activity score.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

ZH performed the data analysis, conducted all experimental procedures, and drafted this manuscript. XL edited and revised the manuscript. ZH and XL were in charge of analysis tools. All authors read and approved the final manuscript.

Acknowledgments

We are extremely grateful for reviewers' input in helping this manuscript.

References

1. Chen M, Liu X, Du J, Wang XJ, Xia L. Differentiated regulation of immune-response related genes between LUAD and LUSC subtypes of lung cancers. *Oncotarget*. 2017;8(1):133-44.
2. Wang C, Tan S, Liu WR, Lei Q, Qiao W, Wu Y, et al. RNA-Seq profiling of circular RNA in human lung adenocarcinoma and squamous cell carcinoma. *Mol Cancer*. 2019;18(1):134.
3. Selvaraj G, Kaliampurthi S, Kaushik AC, Khan A, Wei YK, Cho WC, et al. Identification of target gene and prognostic evaluation for lung adenocarcinoma using gene expression meta-analysis, network analysis and neural network algorithms. *J Biomed Inform*. 2018;86:120-34.
4. Guo T, Ma H, Zhou Y. Bioinformatics analysis of microarray data to identify the candidate biomarkers of lung adenocarcinoma. *PeerJ*. 2019;7:e7313.
5. Zhao X, Li X, Zhou L, Ni J, Yan W, Ma R, et al. LncRNA HOXA11-AS drives cisplatin resistance of human LUAD cells via modulating miR-454-3p/Stat3. *Cancer Sci*. 2018;109(10):3068-79.
6. Gu MM, Gao D, Yao PA, Yu L, Yang XD, Xing CG, et al. p53-inducible gene 3 promotes cell migration and invasion by activating the FAK/Src pathway in lung adenocarcinoma. *Cancer Sci*. 2018;109(12):3783-93.
7. Zhou J, Zhang S, Chen Z, He Z, Xu Y, Li Z. CircRNA-ENO1 promoted glycolysis and tumor progression in lung adenocarcinoma through upregulating its host gene ENO1. *Cell Death Dis*. 2019;10(12):885.
8. Foster H, Ruiz EJ, Moore C, Stamp GWH, Nye EL, Li N, et al. ATMIN Is a Tumor Suppressor Gene in Lung Adenocarcinoma. *Cancer Res*. 2019;79(20):5159-66.
9. Zhang Y, Tseng JT, Lien IC, Li F, Wu W, Li H. mRNAsi Index: Machine Learning in Mining Lung Adenocarcinoma Stem Cell Biomarkers. *Genes (Basel)*. 2020;11(3):257.
10. Andersen R, Donia M, Westergaard MC, Pedersen M, Hansen M, Svane IM. Tumor infiltrating lymphocyte therapy for ovarian cancer and renal cell carcinoma. *Hum Vaccin Immunother*. 2015;11(12):2790-5.
11. Garaud S, Buisseret L, Solinas C, Gu-Trantien C, de Wind A, Van den Eynden G, et al. Tumor infiltrating B-cells signal functional humoral immune responses in breast cancer. *JCI Insight*. 2019;5(18):e129641.
12. Tanaka A, Sakaguchi S. Targeting Treg cells in cancer immunotherapy. *Eur J Immunol*. 2019;49(8):1140-6.

13. Li F, Sun Y, Huang J, Xu W, Liu J, Yuan Z. CD4/CD8 + T cells, DC subsets, Foxp3, and IDO expression are predictive indicators of gastric cancer prognosis. *Cancer Med.* 2019;8(17):7330-44.
14. Ye L, Zhang T, Kang Z, Guo G, Sun Y, Lin K, et al. Tumor-Infiltrating Immune Cells Act as a Marker for Prognosis in Colorectal Cancer. *Front Immunol.* 2019;10:2368.
15. Yang S, Liu T, Cheng Y, Bai Y, Liang G. Immune cell infiltration as a biomarker for the diagnosis and prognosis of digestive system cancer. *Cancer Sci.* 2019;110(12):3639-49.
16. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med.* 2003;348(3):203-13.
17. Vihervuori H, Autere TA, Repo H, Kurki S, Kallio L, Lintunen MM, et al. Tumor-infiltrating lymphocytes and CD8+ T cells predict survival of triple-negative breast cancer. *J Cancer Res Clin Oncol.* 2019;145(12):3105-14.
18. Kim JH, Chon SK, Im KS, Kim NH, Sur JH. Correlation of tumor-infiltrating lymphocytes to histopathological features and molecular phenotypes in canine mammary carcinoma: A morphologic and immunohistochemical morphometric study. *Can J Vet Res.* 2013;77(2):142-9.
19. Kim A, Lee SJ, Ahn J, Park WY, Shin DH, Lee CH, et al. The prognostic significance of tumor-infiltrating lymphocytes assessment with hematoxylin and eosin sections in resected primary lung adenocarcinoma. *PLoS One.* 2019;14(11):e0224430.
20. Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med.* 2002;8(8):816-24.
21. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A.* 2001;98(24):13784-9.
22. Hou J, Aerts J, den Hamer B, van Ijcken W, den Bakker M, Riegman P, et al. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PLoS One.* 2010;5(4):e10312.
23. Stearman RS, Dwyer-Nield L, Zerbe L, Blaine SA, Chan Z, Bunn PA Jr, et al. Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. *Am J Pathol.* 2005;167(6):1763-75.
24. Su LJ, Chang CW, Wu YC, Chen KC, Lin CJ, Liang SC, 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.
25. Liu W, Ouyang S, Zhou Z, Wang M, Wang T, Qi Y, et al. Identification of genes associated with cancer progression and prognosis in lung adenocarcinoma: Analyses based on microarray from Oncomine and The Cancer Genome Atlas databases. *Mol Genet Genomic Med.* 2019;7(2):e00528.
26. Burdett S, Pignon JP, Tierney J, Tribodet H, Stewart L, Le Pechoux C, et al. Adjuvant chemotherapy for resected early-stage non-small cell lung cancer. *Cochrane Database Syst Rev.* 2015;(3):CD011430.
27. Golsteyn RM, Lane HA, Mundt KE, Arnaud L, Nigg EA. The family of polo-like kinases. *Prog Cell Cycle Res.* 1996;2:107-14.

28. van de Weerd BC, Medema RH. Polo-like kinases: a team in control of the division. *Cell Cycle*. 2006;5(8):853-64.
29. Ramani P, Nash R, Sowa-Avugrah E, Rogers C. High levels of polo-like kinase 1 and phosphorylated translationally controlled tumor protein indicate poor prognosis in neuroblastomas. *J Neurooncol*. 2015;125(1):103-11.
30. Takeshita T, Asaoka M, Katsuta E, Photiadis SJ, Narayanan S, Yan L, et al. High expression of polo-like kinase 1 is associated with TP53 inactivation, DNA repair deficiency, and worse prognosis in ER positive Her2 negative breast cancer. *Am J Transl Res*. 2019;11(10):6507-21.
31. Wu J, Lv Q, Huang H, Zhu M, Meng D. Screening and Identification of Key Biomarkers in Inflammatory Breast Cancer Through Integrated Bioinformatic Analyses. *Genet Test Mol Biomarkers*. 2020;24(8):484-91.
32. Lin P, Xiong DD, Dang YW, Yang H, He Y, Wen DY, et al. The anticipating value of PLK1 for diagnosis, progress and prognosis and its prospective mechanism in gastric cancer: a comprehensive investigation based on high-throughput data and immunohistochemical validation. *Oncotarget*. 2017;8(54):92497-521.
33. Chan KK, Wong OG, Wong ES, Chan KK, Ip PP, Tse KY, et al. Impact of iASPP on chemoresistance through PLK1 and autophagy in ovarian clear cell carcinoma. *Int J Cancer*. 2018;143(6):1456-69.
34. Takeshita T, Asaoka M, Katsuta E, Photiadis SJ, Narayanan S, Yan L, et al. High expression of polo-like kinase 1 is associated with TP53 inactivation, DNA repair deficiency, and worse prognosis in ER positive Her2 negative breast cancer. *Am J Transl Res*. 2019;11(10):6507-21.
35. Shin SB, Jang HR, Xu R, Won JY, Yim H. Active PLK1-driven metastasis is amplified by TGF- β signaling that forms a positive feedback loop in non-small cell lung cancer. *Oncogene*. 2020;39(4):767-85.
36. Gao Z, Man X, Li Z, Bi J, Liu X, Li Z, et al. PLK1 promotes proliferation and suppresses apoptosis of renal cell carcinoma cells by phosphorylating MCM3. *Cancer Gene Ther*. 2020;27(6):412-23.
37. Mao Y, Xi L, Li Q, Cai Z, Lai Y, Zhang X, et al. Regulation of cell apoptosis and proliferation in pancreatic cancer through PI3K/Akt pathway via Polo-like kinase 1. *Oncol Rep*. 2016;36(1):49-56.
38. Wu ZY, Wei N. Knockdown of PLK1 inhibits invasion and promotes apoptosis in glioma cells through regulating autophagy. *Eur Rev Med Pharmacol Sci*. 2018;22(9):2723-33.
39. Ge P, Wang W, Li L, Zhang G, Gao Z, Tang Z, et al. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of colorectal cancer. *Biomed Pharmacother*. 2019;118:109228.
40. Garnelo M, Tan A, Her Z, Yeong J, Lim CJ, Chen J, et al. Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma. *Gut*. 2017;66(2):342-51.
41. Hladíková K, Koucký V, Bouček J, Laco J, Grega M, Hodek M, et al. Tumor-infiltrating B cells affect the progression of oropharyngeal squamous cell carcinoma via cell-to-cell interactions with CD8+ T cells. *J Immunother Cancer*. 2019;7(1):261.

42. Ino Y, Yamazaki-Itoh R, Shimada K, Iwasaki M, Kosuge T, Kanai Y, et al. Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. *Br J Cancer*. 2013;108(4):914-23.
43. Oja AE, Piet B, van der Zwan D, Blaauwgeers H, Mensink M, de Kivit S, et al. Functional Heterogeneity of CD4+ Tumor-Infiltrating Lymphocytes With a Resident Memory Phenotype in NSCLC. *Front Immunol*. 2018;9:2654.
44. Hofmann M, Oschowitz A, Kurzhals SR, Krüger CC, Pircher H. Thymus-resident memory CD8+ T cells mediate local immunity. *Eur J Immunol*. 2013;43(9):2295-304.
45. Komdeur FL, Prins TM, van de Wall S, Plat A, Wisman GBA, Hollema H, et al. CD103+ tumor-infiltrating lymphocytes are tumor-reactive intraepithelial CD8+ T cells associated with prognostic benefit and therapy response in cervical cancer. *Oncoimmunology*. 2017;6(9):e1338230.
46. Ahrends T, Bąbała N, Xiao Y, Yagita H, van Eenennaam H, Borst J. CD27 Agonism Plus PD-1 Blockade Recapitulates CD4+ T-cell Help in Therapeutic Anticancer Vaccination. *Cancer Res*. 2016;76(10):2921-31.
47. Ahrends T, Spanjaard A, Pilzecker B, Bąbała N, Bovens A, Xiao Y, et al. CD4+ T Cell Help Confers a Cytotoxic T Cell Effector Program Including Coinhibitory Receptor Downregulation and Increased Tissue Invasiveness. *Immunity*. 2017;47(5):848-861.e5.
48. Quezada SA, Simpson TR, Peggs KS, Merghoub T, Vider J, Fan X, et al. Tumor-reactive CD4(+) T cells develop cytotoxic activity and eradicate large established melanoma after transfer into lymphopenic hosts. *J Exp Med*. 2010;207(3):637-50.
49. Friedman KM, Prieto PA, Devillier LE, Gross CA, Yang JC, Wunderlich JR, et al. Tumor-specific CD4+ melanoma tumor-infiltrating lymphocytes. *J Immunother*. 2012;35(5):400-8.
50. Rooney MS, Shukla SA, Wu CJ, Getz G, Hachohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1-2):48-61.
51. Mandal R, Şenbabaoğlu Y, Desrichard A, Havel JJ, Dalin MG, Riaz N, et al. The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight*. 2016;1(17):e89829.
52. Xu B, Chen L, Li J, Zheng X, Shi L, Wu C, et al. Prognostic value of tumor infiltrating NK cells and macrophages in stage II+III esophageal cancer patients. *Oncotarget*. 2016;7(46):74904-16.
53. Vitale M, Cantoni C, Pietra G, Mingari MC, Moretta L. Effect of tumor cells and tumor microenvironment on NK-cell function. *Eur J Immunol*. 2014;44(6):1582-92.
54. Frazao A, Rethacker L, Messaoudene M, Avril MF, Toubert A, Dulphy N, et al. NKG2D/NKG2-Ligand Pathway Offers New Opportunities in Cancer Treatment. *Front Immunol*. 2019;10:661.
55. Fu C, Jiang A. Dendritic Cells and CD8 T Cell Immunity in Tumor Microenvironment. *Front Immunol*. 2018;9:3059.

Tables

Table 1 Differentially expressed genes in the cell cycle pathway

Gene	logFC	p-value	adjust p-value	Gene	logFC	p-value	adjust p-value
<i>CDC20</i>	3.681	8.76E-47	3.21E-45	<i>PKMYT1</i>	2.599	9.12E-26	1.29E-24
<i>CDKN2A</i>	3.299	9.88E-19	8.65E-18	<i>ESPL1</i>	2.566	7.64E-21	7.91E-20
<i>CDC6</i>	3.140	4.56E-35	1.05E-33	<i>ORC6</i>	2.528	1.12E-23	1.40E-22
<i>BUB1B</i>	3.042	2.42E-32	4.84E-31	<i>BUB1</i>	2.508	5.02E-28	8.14E-27
<i>CCNE1</i>	3.000	1.78E-24	2.34E-23	<i>ORC1</i>	2.431	6.50E-23	7.74E-22
<i>PLK1</i>	2.998	5.85E-37	1.45E-35	<i>CDK1</i>	2.424	1.63E-30	3.01E-29
<i>CCNB1</i>	2.946	2.33E-44	7.85E-43	<i>MAD2L1</i>	2.384	2.76E-24	3.61E-23
<i>CCNB2</i>	2.921	2.60E-37	6.54E-36	<i>PTTG1</i>	2.305	3.80E-32	7.53E-31
<i>CDC45</i>	2.894	2.51E-30	4.56E-29	<i>MCM4</i>	2.219	1.60E-37	4.05E-36
<i>TTK</i>	2.849	1.96E-25	2.71E-24	<i>CHEK1</i>	2.069	2.99E-22	3.41E-21
<i>CDC25C</i>	2.811	9.20E-23	1.09E-21	<i>MCM2</i>	2.014	2.82E-29	4.81E-28
<i>CCNA2</i>	2.791	4.82E-34	1.04E-32				

Abbreviation: FC, fold change.

Table 2 Cox regression analysis for *PLK1* expression versus overall survival rate of LUAD patients

Parameter	Univariate analysis			Multivariate analysis		
	HR	95%CI	<i>p</i> -value	HR	95%CI	<i>p</i> -value
TCGA dataset						
<i>PLK1</i>	1.045	1.020-1.070	0.001	1.053	1.023-1.084	0.001
age	1.002	0.983-1.048	0.843	1.013	0.993-1.033	0.192
gender	1.035	0.717-1.495	0.852	0.927	0.638-1.347	0.691
stage	1.654	1.401-1.951	0.000	2.037	1.275-3.254	0.003
T	1.632	1.315-2.024	0.000	1.189	0.940-1.503	0.149
M	1.757	0.964-2.024	0.066	0.353	0.103-1.210	0.098
N	1.790	1.459-2.196	0.000	0.970	0.651-1.444	0.881
GSE31210 profile						
<i>PLK1</i>	1.658	1.192-2.306	0.003	1.538064	1.068-2.215	0.021
age	1.025	0.977-1.075	0.306	1.035388	0.985-1.088	0.169
gender	1.519	0.780-2.955	0.219	1.100991	0.547-2.215	0.787
stage	4.232	2.175-8.236	0.000	3.720136	1.886-7.337	0.000

Abbreviation: LUAD, lung adenocarcinoma; HR, hazard ratio; CI, confidence interval.

Figures

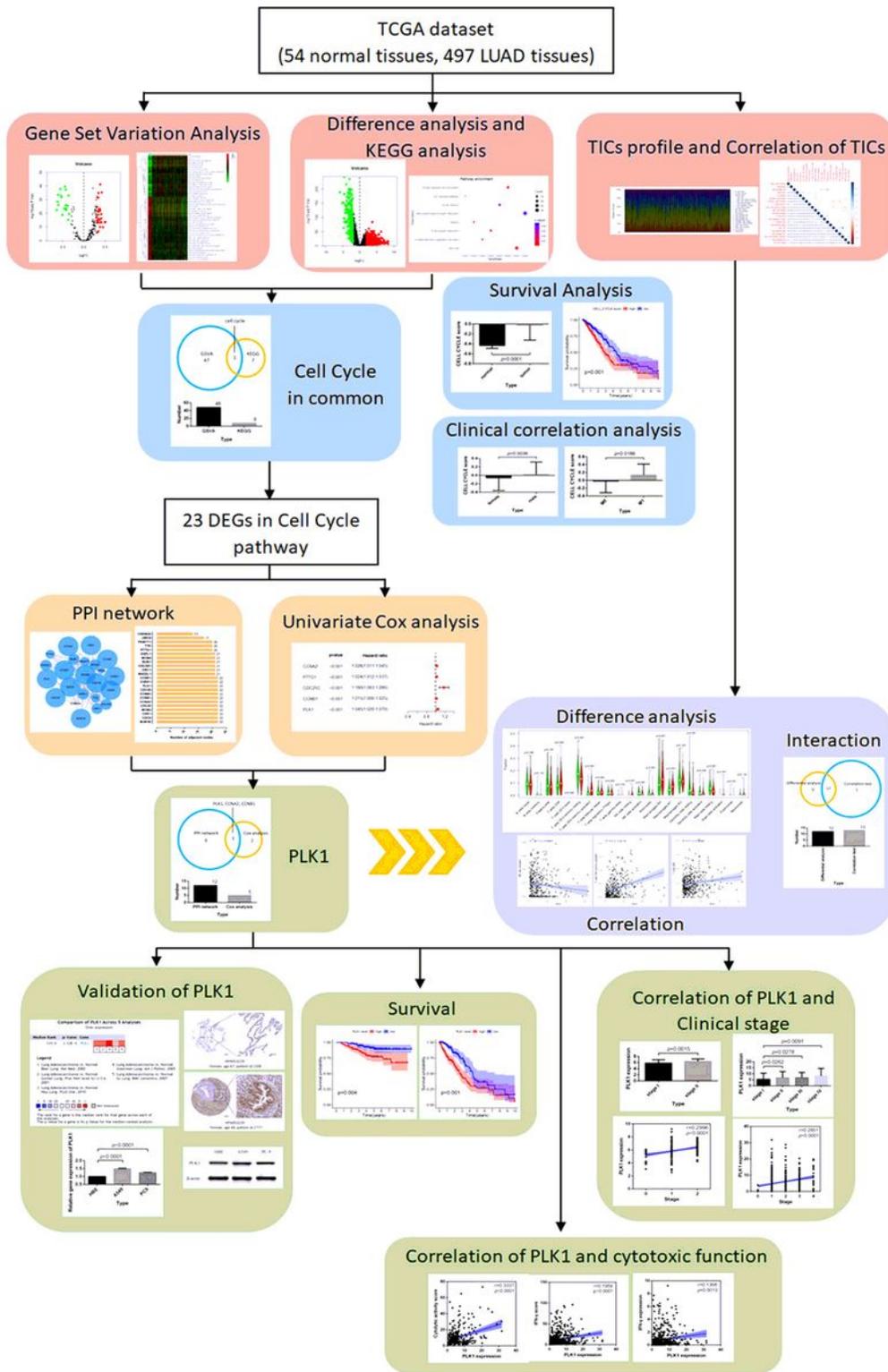


Figure 1

Flow chart for the analysis conducted in this work.

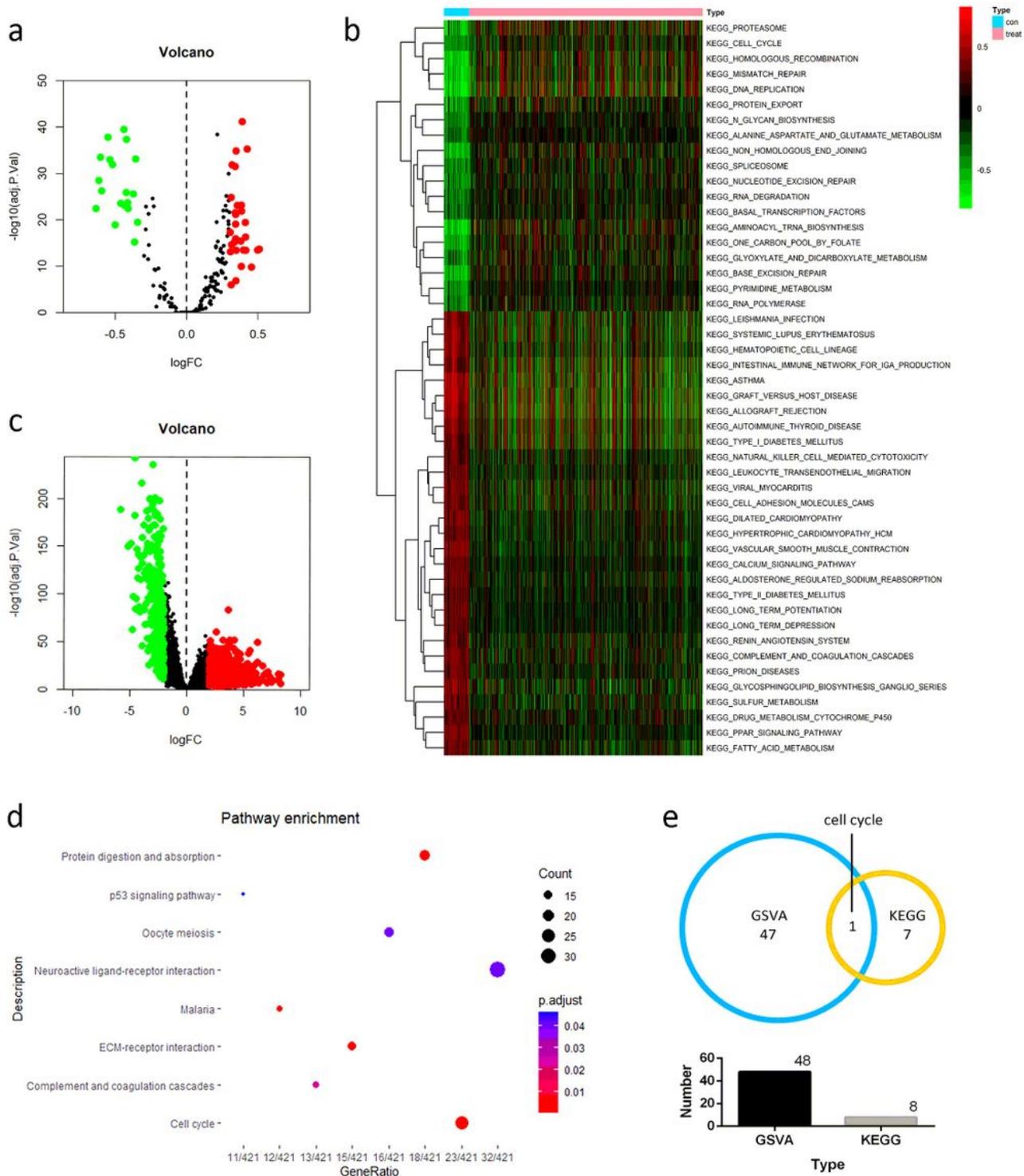


Figure 2

GSVA for samples and KEGG enrichment analysis for DEGs. a 29 upregulated (red spots) and 19 downregulated (green spots) metabolic pathways were identified via GSVAs in LUAD samples versus normal lung samples. Adjusted p -value < 0.05 and $|\log_{2}FC| > 0.3$ were utilized as the cutoff criteria for determining significant differences. b Heatmap for the significantly altered metabolic pathways. c 1,109 upregulated (red spots) and 310 downregulated (green spots) DEGs were identified in LUAD samples

versus normal lung samples. Adjusted p -value <0.05 and $|\log FC|>2$ were utilized as the cutoff criteria for determining significant differences. d KEGG enrichment analysis for DEGs. e Cell cycle was the common pathway identified in both GSVA and KEGG analyses.

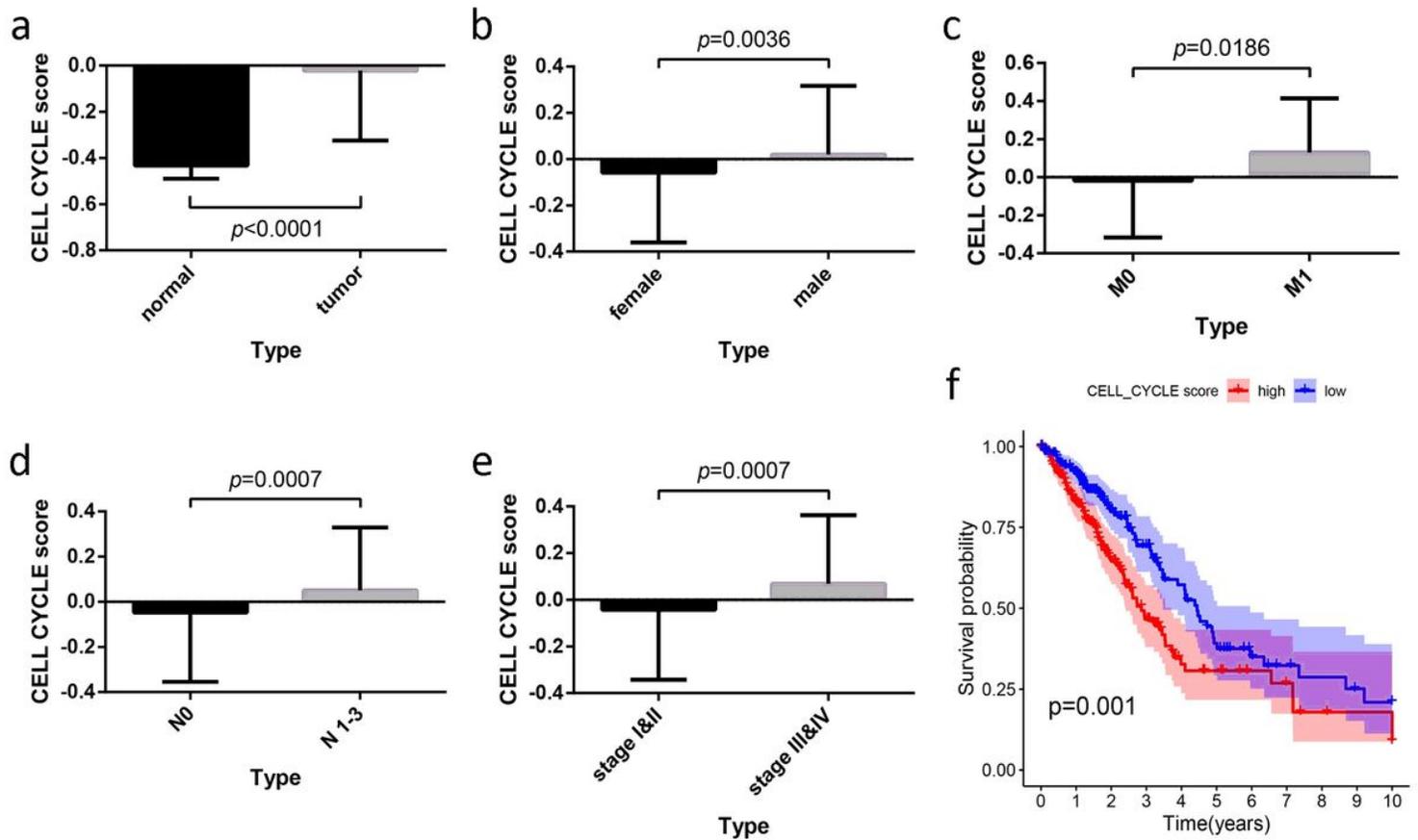


Figure 3

Correlation of cell cycle score with clinical characteristics and survival of LUAD patients. a-e Cell cycle score was significantly higher in LUAD samples versus normal samples (a), male versus female (b), M1 versus M0 classification (c), N1-3 versus N0 classification (d), and stage III&IV versus stage I&II (e). f Kaplan-Meier curves for OS in high and low score patients grouped by the median of Cell Cycle score.

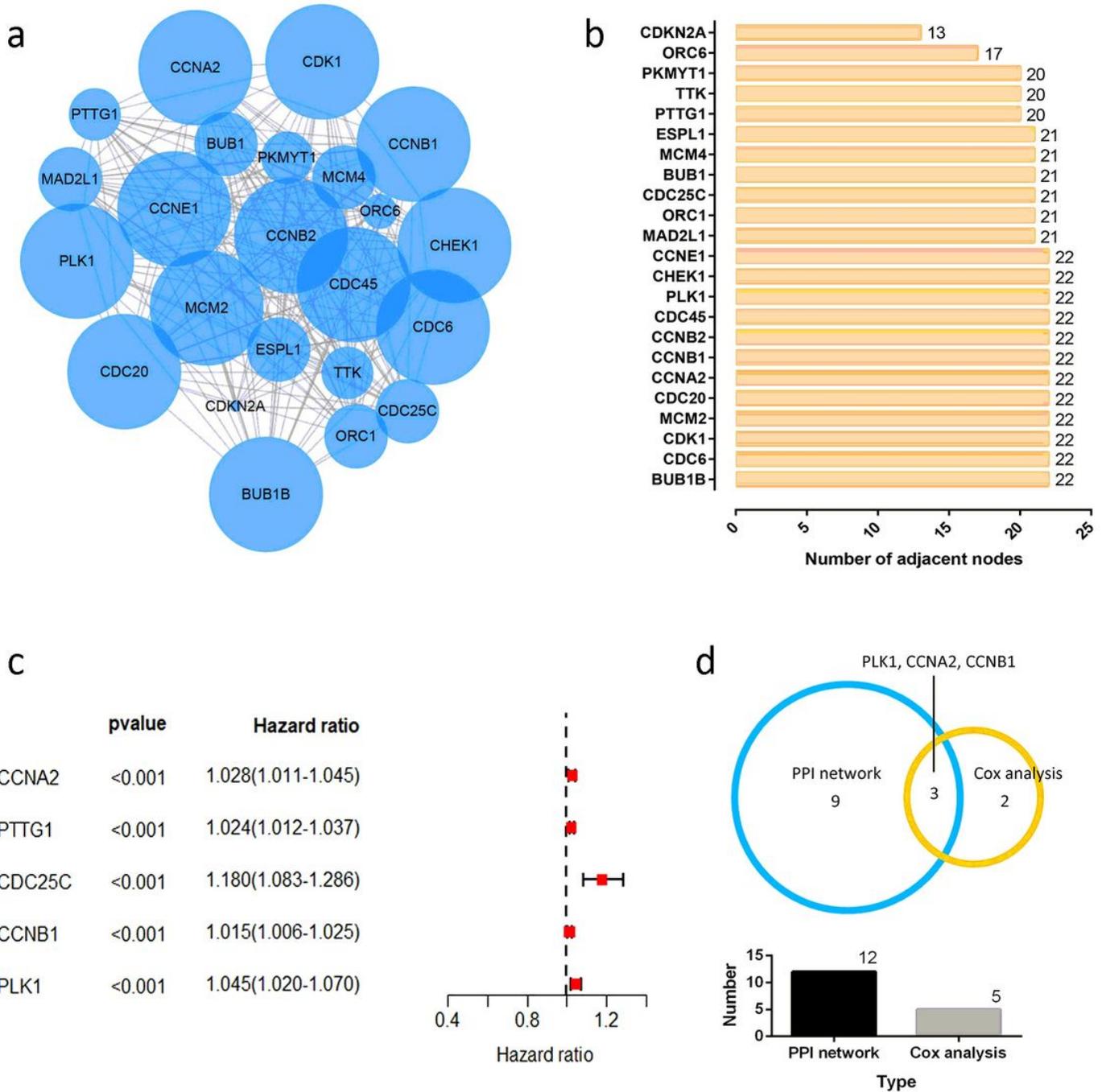


Figure 4

PPI network and univariate Cox analysis for the 23 DEGs in the cell cycle pathway. a The PPI network consisted of 23 nodes (genes) and 240 edges (interactions between genes). b Histogram of the connectivity degrees of genes in the PPI network. c Univariate Cox analysis for the 23 DEGs. A p-value<0.001 was utilized as the cutoff criterion. d Venn diagram indicating 3 common genes which overlapped the 12 leading genes in the PPI network and the 5 prognosis-related genes in the Cox analysis.

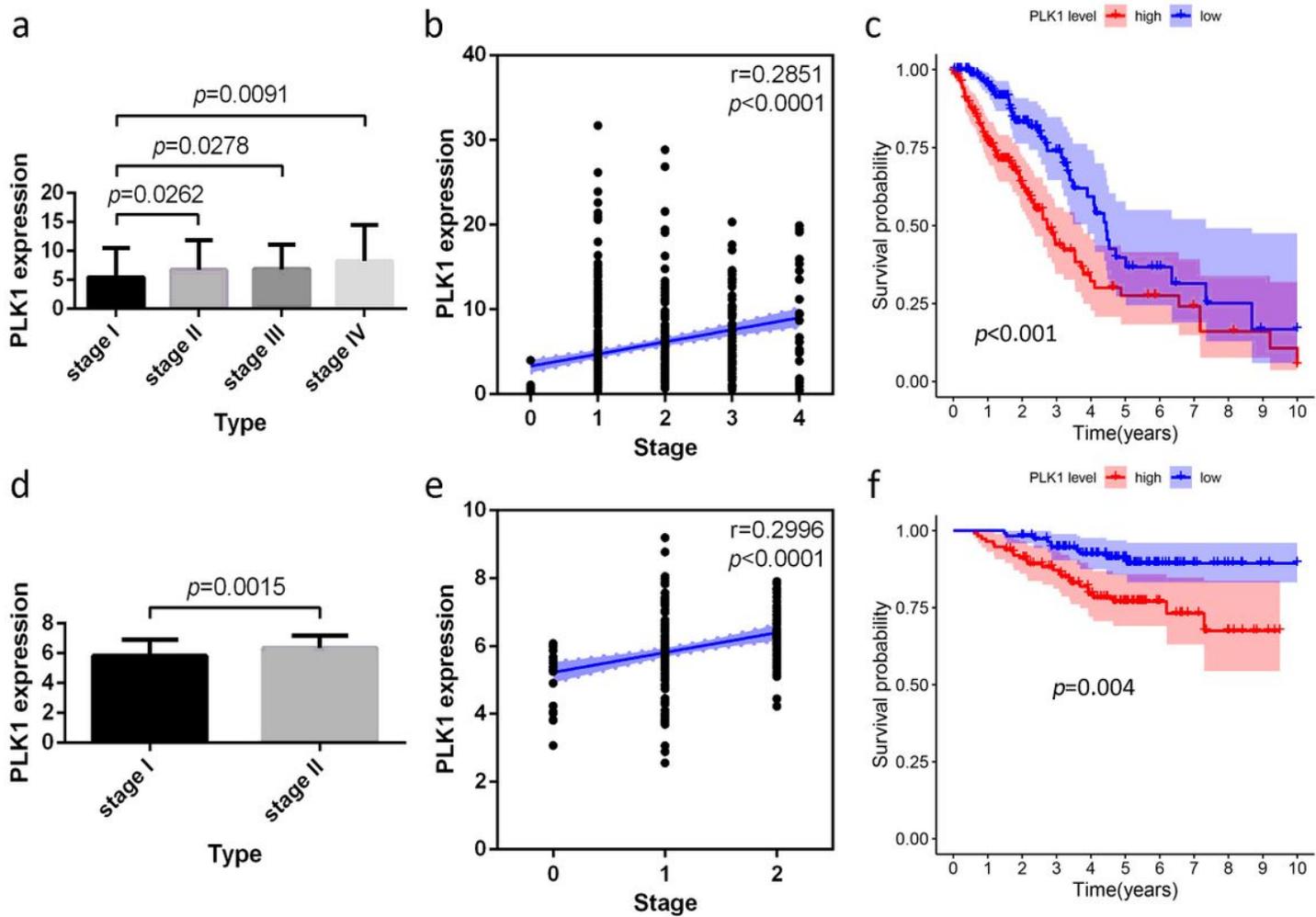


Figure 5

Correlation of PLK1 expression with clinical stage and survival of LUAD patients. a, b Histogram of PLK1 expression in different stages in the TCGA dataset (a) and GSE31210 profile (b). c, d PLK1 expression is significantly correlated with clinical stage in the TCGA dataset (c) and GSE31210 profile (d). e, f Kaplan-Meier curves for the overall survival in high and low PLK1 expression subgroups divided by the median PLK1 level in the TCGA dataset (e) and GSE31210 profile (f).

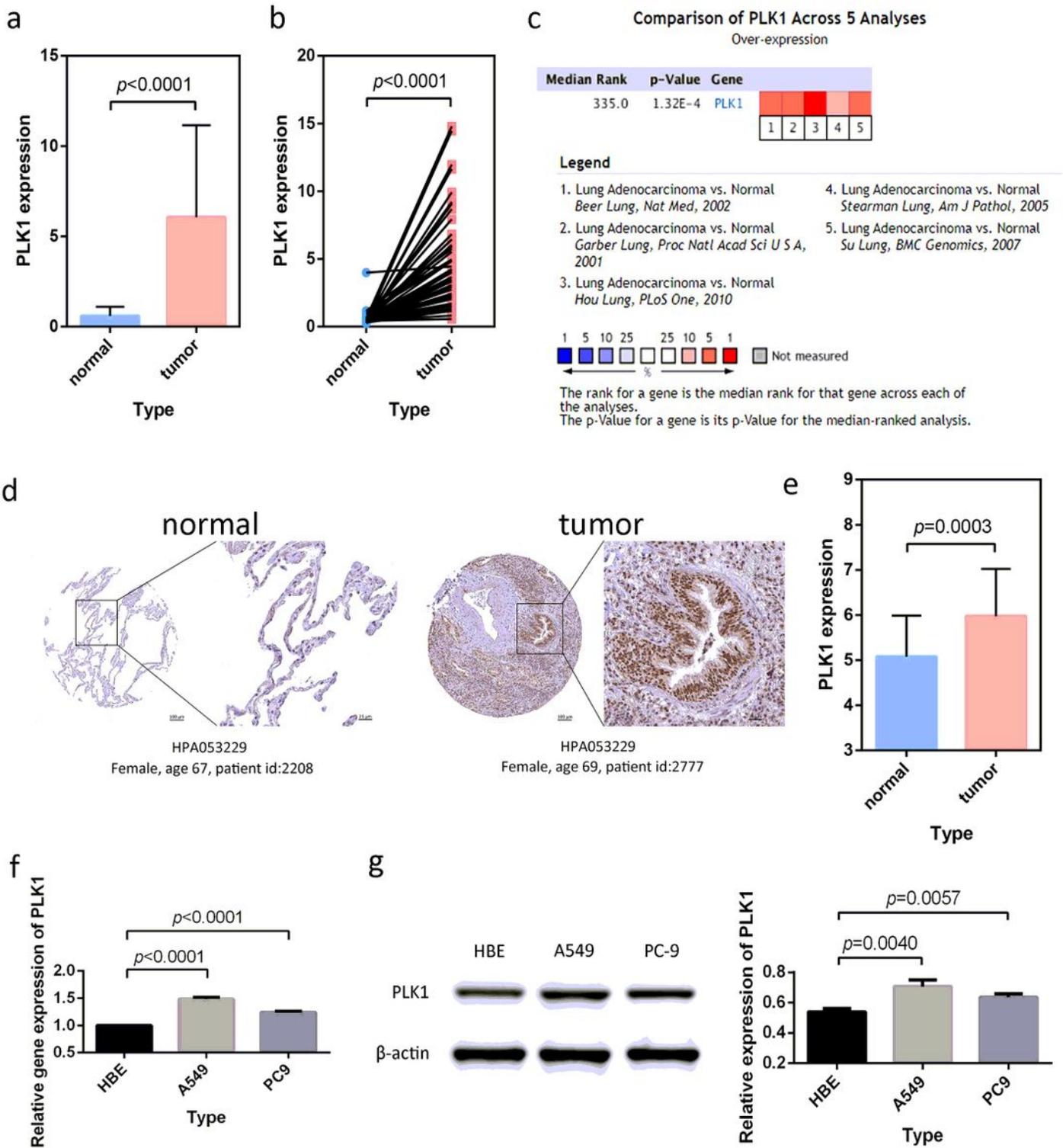


Figure 6

Validation of PLK1 expression. a, b PLK1 expression is significantly upregulated in LUAD samples compared to normal lung samples in both unpaired (a) and paired (b) difference analyses based on the TCGA dataset. c-e PLK1 expression is significantly increased in LUAD samples compared with normal samples in a meta-analysis (c), immunohistochemical analysis (d) and GSE31210 profile (e). f, g Gene

expression level (f) and protein expression level (g) of PLK1 significantly increased in A549 and PC-9 cell lines compared with HBE cells.

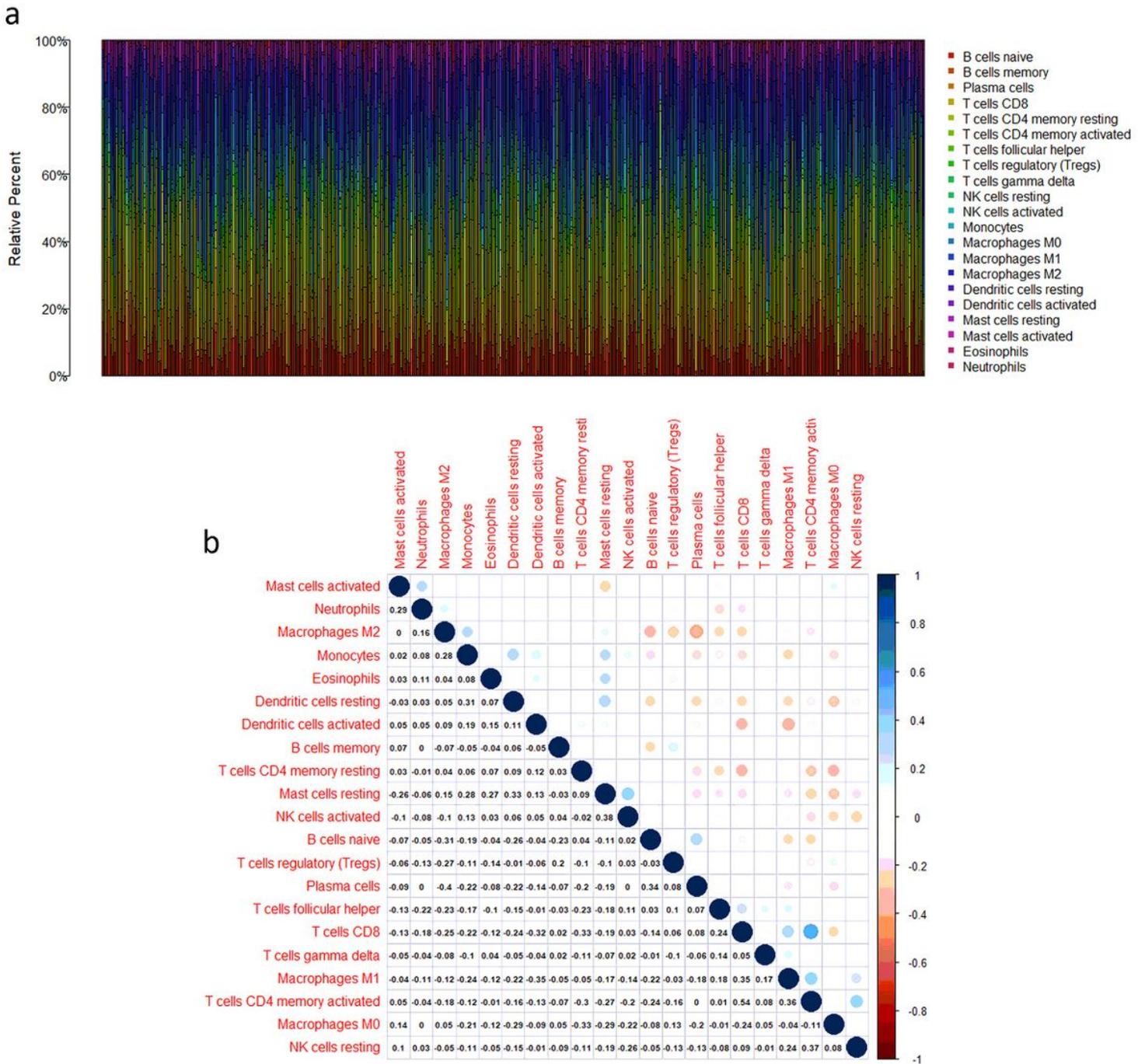


Figure 7

Profile of TICs in tumor samples and their correlation with other factors. a Histogram for proportion of TICs in LUAD samples. The columns represent the sample IDs. b Heatmap for correlation among TICs. The number in each box represents the Pearson coefficient between two kinds of TICs. The blue- and red-colored spots represent the corresponding correlation value on the diagonal of the heatmap.

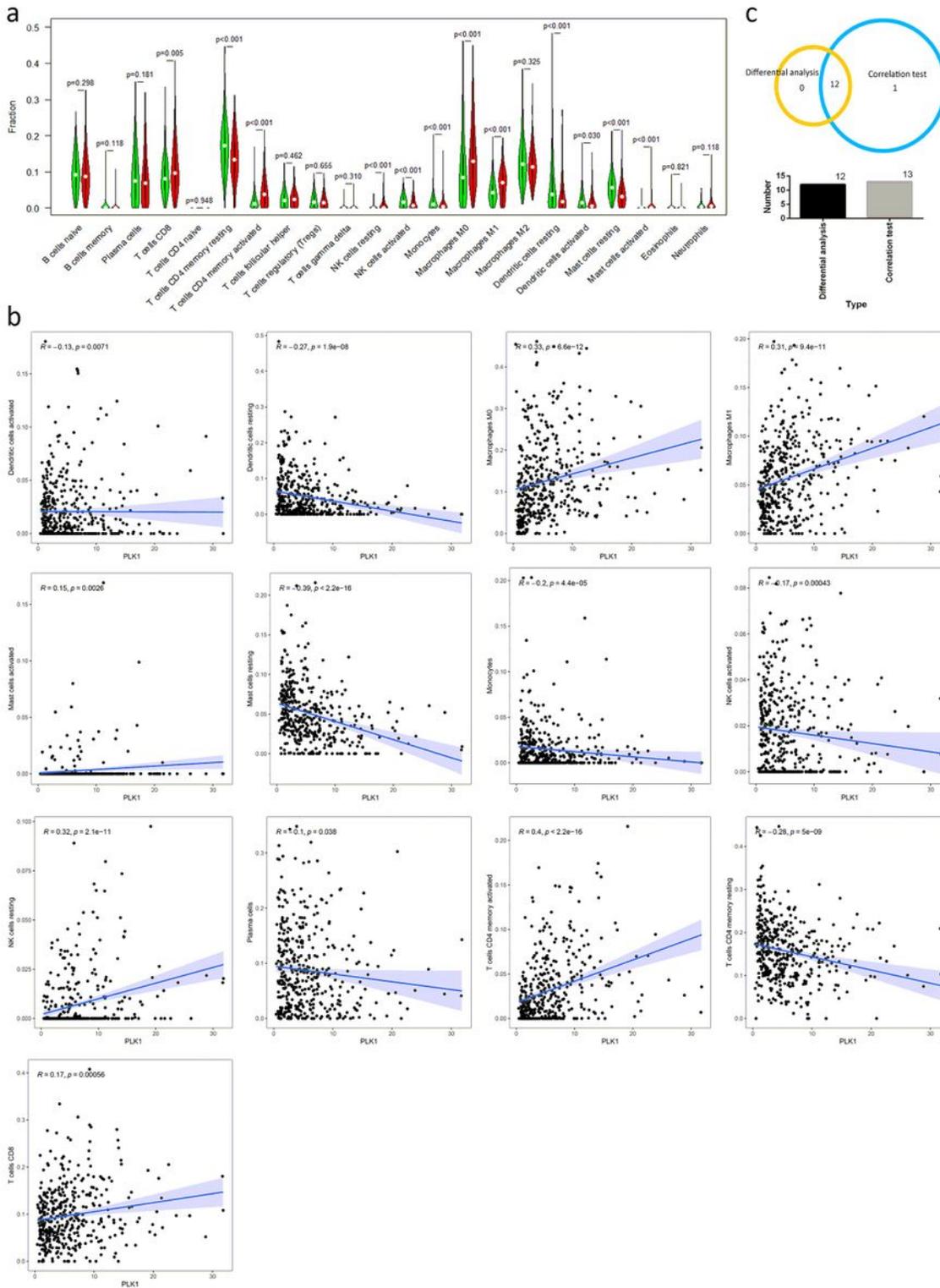


Figure 8

Correlation of PLK1 expression with TICs proportion in LUAD patients. a Violin plot of the relative proportion of TICs in the high and low PLK1 expression subgroups grouped by median PLK1 expression. Green-and red-colored columns represent the low and high PLK1 expression subgroups, respectively. b Scatter plots for the significant correlation of the 13 TICs with PLK1 expression. c Venn diagram indicating 12 kinds of common TICs overlapped by difference analyses and correlation tests.

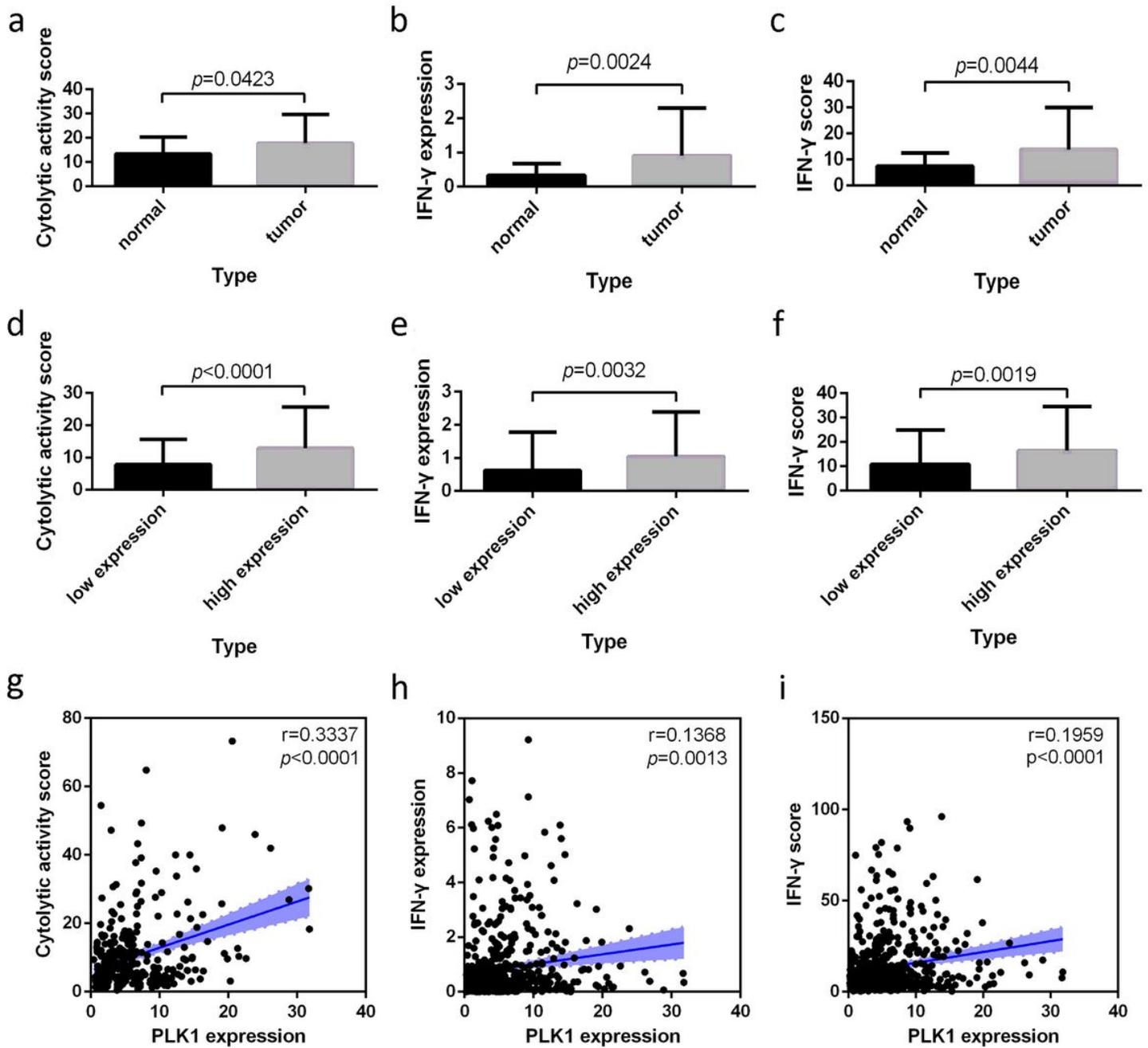


Figure 9

Correlation of PLK1 expression with cytotoxic function. a-c Histogram for cytolytic activity score (a), IFN-γ expression level (b), and IFN-γ score (c), which markedly increased in LUAD samples versus normal lung samples. d-f Histogram for cytolytic activity score (d), IFN-γ expression level (e), and IFN-γ score (f) were significantly higher in the high PLK1 expression subgroup than the low PLK1 expression subgroup. Values displayed were divided by median PLK1 expression. g-i Scatter plots for cytolytic activity score (g), IFN-γ expression level (h), and IFN-γ score (i), which were significantly correlated with PLK1 expression.

Supplementary Files

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

- [SupplementaryTableS1.xlsx](#)
- [SupplementaryTableS2.xlsx](#)
- [SupplementaryTableS3.xlsx](#)