



Research article

Identifying M1-like macrophage related genes for prognosis prediction in lung adenocarcinoma based on a gene co-expression network

Zhiyuan Wang^{a,b,1}, Shan Yan^{c,1}, Ying Yang^b, Xuan Luo^a, Xiaofang Wang^d, Kun Tang^e, Juan Zhao^a, Yongwen He^{f,g,**}, Li Bian^{b,*}

^a School of Basic Medicine, Kunming Medical University, Kunming, 650500, China

^b Department of Pathology, The First Affiliated Hospital of Kunming Medical University, Kunming, 650031, China

^c Institute of Biomedical Engineering, Kunming Medical University, Kunming, 650031, China

^d Department of Pathology, The Second Affiliated Hospital of Kunming Medical University, Kunming, 650031, China

^e Intensive Care Unit, The First Affiliated Hospital of Kunming Medical University, Kunming, 650031, China

^f School of Stomatology, Kunming Medical University, Kunming, 650021, China

^g Qujing Medical College, Qujing, 655099, China



ARTICLE INFO

Keywords:

Lung adenocarcinoma
Macrophages
Immune landscape
WGCNA
Prognostic hub genes

ABSTRACT

Macrophages are one of the most important players in the tumor microenvironment. But the contribution of macrophages to lung adenocarcinoma (LUAD) is still controversial. The current study aimed to display an immune landscape to clarify the function of macrophages and detect prognostic hub genes in LUAD. The transcriptome data were adopted to screen differently expressed genes (DEGs) in The Cancer Genome Atlas database (TCGA). The cell type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm was used to reveal the immune landscape. Weighted gene co-expression network analysis (WGCNA) analysis was performed to identify the hub module associated with macrophages. Function Enrichment analysis was conducted on hub module genes. Moreover, univariate and multivariate Cox regression analyses were performed to identify prognostic hub genes. Kaplan-Meier (KM) and Time-dependent receiver operating characteristic (ROC) curves were plotted to assess the prognostic capacity of the four prognostic hub genes. The GSE1196959 dataset from the Gene Expression Omnibus (GEO) database was downloaded to verify the differential expression of the 4 prognostic hub genes.

1. Introduction

Lung cancer is a common malignant tumor and the leading cause of cancer-related mortality worldwide [1]. Lung adenocarcinoma (LUAD) is the dominant histologic type of lung cancer, which accounts for approximately 40% of lung cancer cases [2]. Despite major progress has been made in LUAD diagnostics and therapy, the five-year survival rate for advanced-stage LUAD remains dismal. Early

* Corresponding author.

** Corresponding author. School of Stomatology, Kunming Medical University, Kunming, 650021, China.

E-mail addresses: heyongwen2@sina.com (Y. He), bianli1976@sina.com (L. Bian).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.heliyon.2023.e12798>

Received 23 July 2022; Received in revised form 25 December 2022; Accepted 2 January 2023

Available online 4 January 2023

2405-8440/© 2023 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

detection and timely treatment are essential to improve patient survival. Therefore, the discovery of novel early prognostic molecular markers is extremely pressing.

The tumor microenvironment (TME) is a complex ecosystem that includes a variety of components, including tumor cells, vascular vessels, fibroblasts, immune cells, and extracellular matrix components. Emerging evidence suggests that the degree of immune cell infiltration correlates with prognosis in many types of cancer [3]. Immune-related genes (IRGs) have been reported to correlate with the progression, recurrence, and metastasis of lung cancer, suggesting that IRGs have acceptable prognostic values for disease outcomes [4,5,6]. Macrophages resident in the TME referred to as tumor-associated macrophages (TAMs), are the main immune cells found in the tumor microenvironment and can affect tumor progression and metastasis. Based on their distinct abilities, TAMs can be categorized into two different states, M1 and M2 phenotypes [7,8]. Recently, increasing evidence has suggested that TAMs are not composed of a homogeneous population but rather a mixed population of macrophages. Both M1 and M2 phenotypes that have been detected in several malignant solid tumours [9,10,11,12]. In addition to promoting inflammation, M1 macrophage-associated genes are also involved in tumor progression and metastasis. Meng Xiao reported that conditioned media from exosomes induced M1-like TAMs and significantly promoted migration of oral squamous cell carcinoma cells [13]. Another study found that TNF- α released from M1 macrophages increased the metastatic potential of ovarian cancer cells through the activation of the NF- κ B signaling pathway [14]. However, the biological impact of M1 macrophage-associated gene activities should be further investigated. The present study focuses on finding prognosis-relevant M1 related genes in LUAD.

2. Material and methods

2.1. Data acquisitions and processing

The transcriptome data and clinical information of 572 samples from 513 LUAD patients, were downloaded from The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>) database. GSE116959 dataset included 57 LUAD and 11 normal samples were downloaded from the gene expression omnibus (GEO) database and served as validation set. The clinical informations of TCGA cohort and GSE116959 dataset was shown in Supplementary Table S1.

2.2. Evaluation of tumor-infiltrating immune cells

In this study, we used the R package ‘CIBERSORT’ to estimate the fraction of immune cells in TCGA cohort (Supplementary Table S1). Specifically, the cell type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm was used to calculate the fractions of the 22 types of TIICs [15]. Naive CD4⁺ T cells was not detected in all samples (that was, the proportion of immune cells in each sample was 0), and the remaining 21 immune cells were used for subsequent analysis. In this study, 572 samples were analyzed for immune cell infiltration, and 530 samples (58 Normal and 472 Tumor) were obtained after filtering with a $P < 0.05$ were used for subsequent analysis.

2.3. Differential expression analysis

Differential expression analysis was performed on the counter matrix of LUAD and normal samples from the TCGA cohort using limma package. The screening conditions for the differential genes were: Fold Change $> |\pm 2|$, $P < 0.05$. Volcano Plot was drawn using ggplot2 package to visualize the differentially expressed genes (DEGs) between normal and LUAD samples (Supplementary Figure S1).

2.4. Co-expression network constructions

The aforementioned DEGs were utilized to establish a weight co-expression network using the R package ‘WGCNA’ [16]. First, based on Pearson’s correlation value between paired genes, the expression levels of individual transcripts were converted into a similarity matrix. Parameter β can improve strong correlations and decrease weak correlations between genes. The adjacency matrix was then converted into a topological overlap matrix based on the optimal soft-thresholding parameter β which could enhance strong correlations between genes and penalize weak correlations. Then, turned adjacency matrix into a topological overlap matrix (TOM). To categorize genes with similar expression patterns into different modules, we applied a dynamic hybrid cutting method.

2.5. Identification of the hub module

Module eigengenes were used to perform a correlation analysis of each module. We calculated the correlation between module eigengenes and the infiltration level of macrophage to determine the significance of the modules by the Pearson test. An individual module was considered significantly correlated with Macrophage when $P < 0.05$. We selected the module with the highest correlation coefficient and defined that as a hub module.

2.6. Enrichment analysis

The function of the genes in the aforementioned hub module was annotated using the R package GOplot for pathway and process enrichment analysis [17], the corresponding Biological Processes (BP), Cell Components (CC), and the function of Gene Ontology (GO)

and the signaling pathways of the Kyoto Encyclopedia of Genes and Genomes (KEGG) were conducted. The enrichplot package was utilized to visualize the results of the enrichment analysis.

2.7. Construction of PPI network and identification of hub genes

The protein-protein interaction (PPI) network was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database based on the hub module genes [18] with minimum combined score >0.4 as a threshold. Then visualized the PPI network using the Cytoscape software [19]. According to the degree value, the top 20 genes were designated as hub genes.

2.8. Risk model development and validation

Univariate Cox proportional hazards model was performed to screen genes significantly correlated with the survival probability of LUAD patients from the 20 hub genes, and $P < 0.05$ was the threshold. The genes detected by univariate analysis were subjected to a multivariate Cox regression analysis for identifying optimal hub genes. The risk score of each sample was calculated by the following formula: Risk score = $(\beta_{\text{gene1}} \times \text{expression level of gene1}) + (\beta_{\text{gene2}} \times \text{expression level of mRNA2}) + (\beta_{\text{gene3}} \times \text{expression level of mRNA3}) + \dots + (\beta_{\text{geneN}} \times \text{expression level of geneN})$. (β : coefficient value).

Subsequently, LUAD patients in the TCGA cohort were divided into high- and low-risk groups by the median risk score as the cutoff value. Kaplan-Meier (K-M) curves were plotted using the R packages survival, and survival differences between high-risk and low-risk groups were compared by the log-rank test. Receiver operating characteristic (ROC) curve was plotted to assess the accuracy of the risk score model using R package pROC. The area under the curve (AUC) of the ROC curve was estimated.

2.9. Validation of the expression of the optimal hub genes

The GSE116959 dataset was selected to validate the expression difference of the optimal hub genes between normal samples and LUAD samples.

2.10. Statistical analyses

Statistical analysis in this study was performed using R software (version 3.6.0). The differences between two group of data were compare by Wilcoxon test, and $P < 0.05$ was considered as significant.

3. Result

3.1. Evaluation of tumor-infiltrating immune cells (TIICs)

CIBERSORT algorithm was conducted to assess the abundance of 22 TIICs in 530 samples (58 normal and 472 tumor samples) using the R package 'CIBERSORT'. LUAD patients were infiltrated with affluent Plasma cells, as statistics have shown. Increased numbers of naïve B cells, resting dendritic cells, activated memory CD4⁺ T cells, macrophages M1, and especially Plasma cells, were observed in LUAD patients, while the other immune cell types (including macrophages M2, Mast cells resting, Monocytes) were depleted in

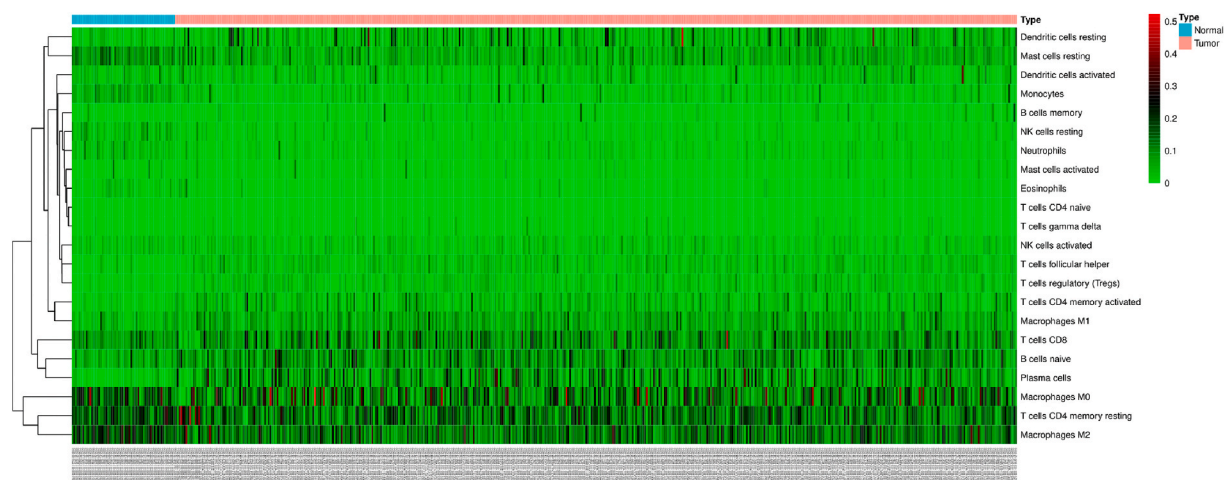


Fig. 1. Immune landscape of TCGA-LUAD. Bar plot showing the proportion of 22 kinds of TIICs in the normal and LUAD tumor samples. Column names of plot were sample ID.

number. Gamma delta T cells, activated NK cells, activated mast cells, and activated dendritic cells were kept a balance in their number. No naive CD4⁺ T cells was detected from the CIBERSORT results (Fig. 1, Supplementary Table S2).

3.2. Gene co-expression network in LUAD

A total of 12,967 DEGs were obtained, of which 7525 were up-regulated genes and 5442 were down-regulated genes (Supplementary Table S3). Expression values of the 12,967 DEGs were used to construct a co-expression network. Using the R package 'WGCNA', we calculated average linkage and Pearson's correlation values to cluster the samples of the TCGA database. To build a scale-free network, we picked $\beta = 5$ (scale-free $R^2 = 0.9$) as the soft-thresholding power (Fig. 2A). A total of nine modules were generated through hierarchical clustering (Supplementary Table S4) A hierarchical clustering tree was constructed using dynamic hybrid cutting. Each leaf on the tree represented a single gene, and genes with similar expression data were close together and formed a branch of the tree, representing a gene module (Fig. 2B).

3.3. Identification of hub modules and enrichment analysis

To identify the modules significantly associated with macrophages, we conducted a Pearson correlation analysis. Among the

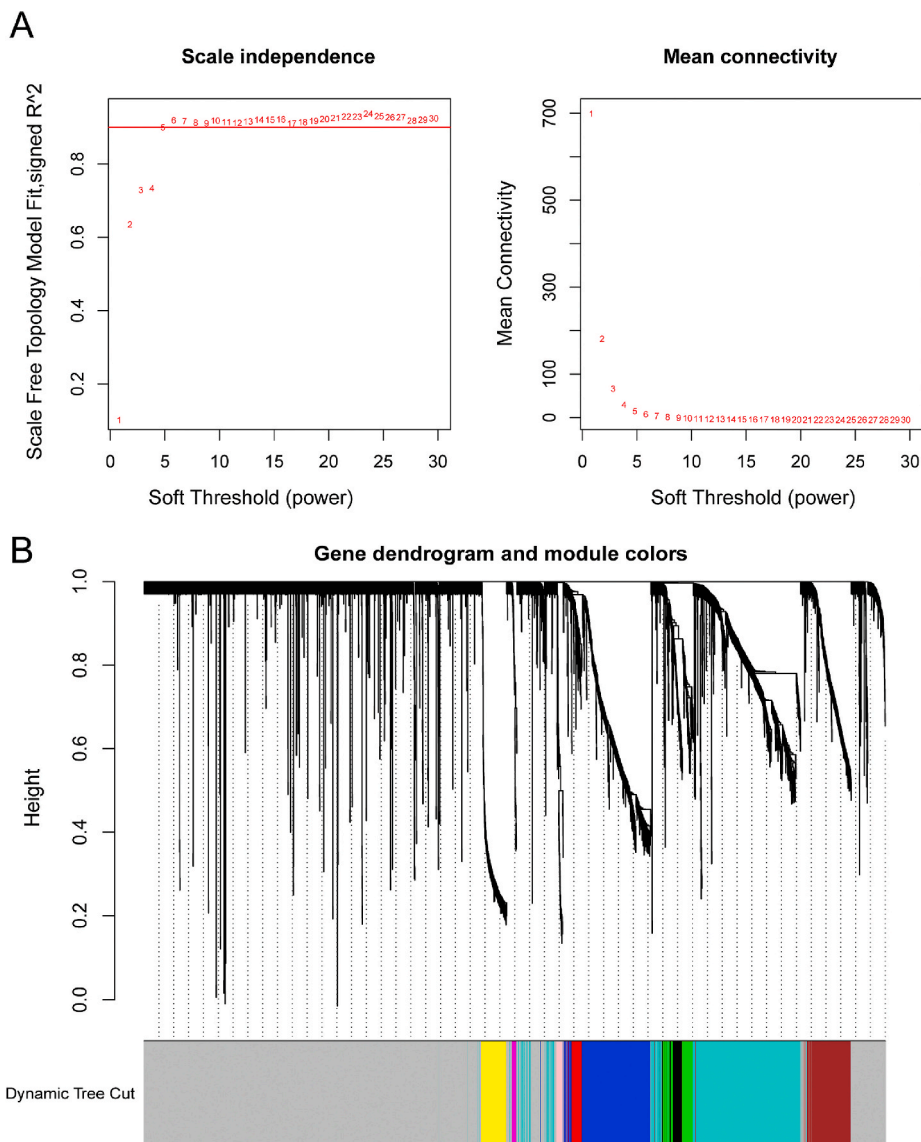
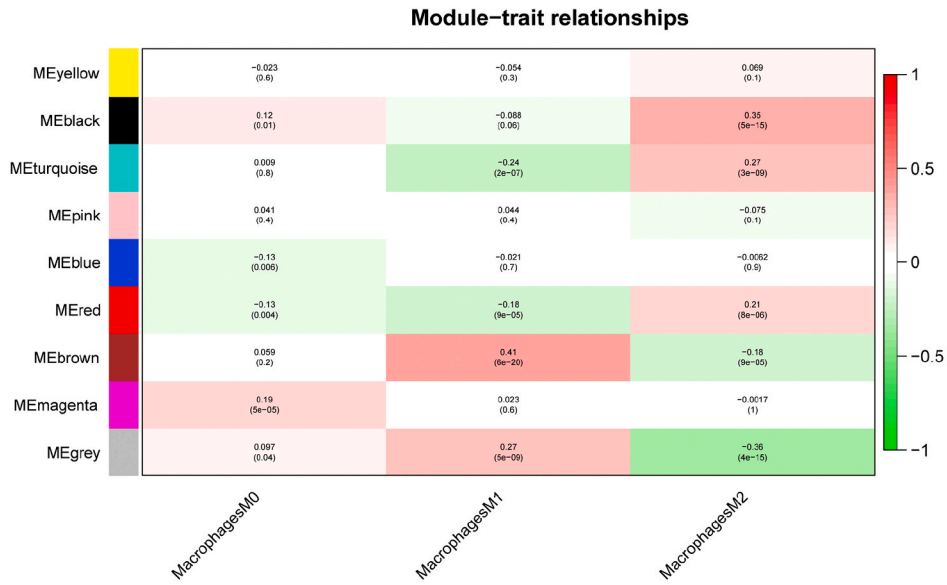
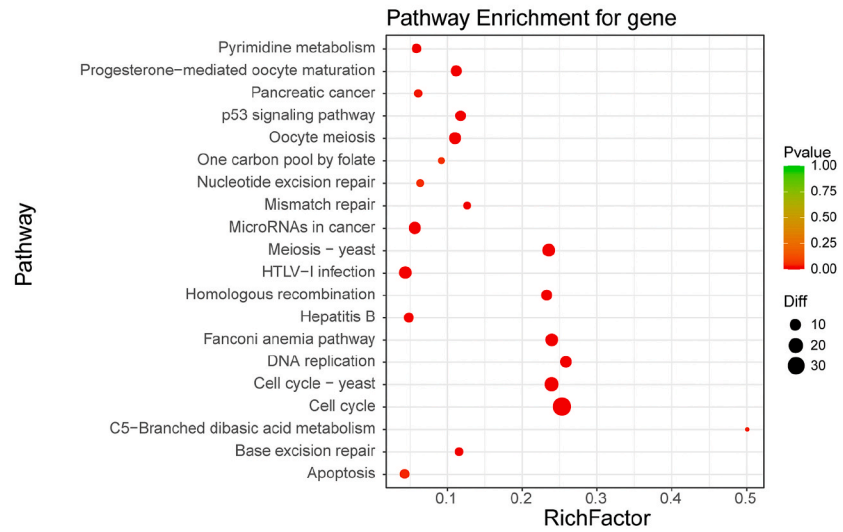


Fig. 2. Hub module selection. (A). Soft-thresholding for the adjacency matrix. (B). The cluster dendrogram of all the genes in TCGA- LUAD database. Each leaf represents a separate gene, and each branch represents a co-expression gene module.

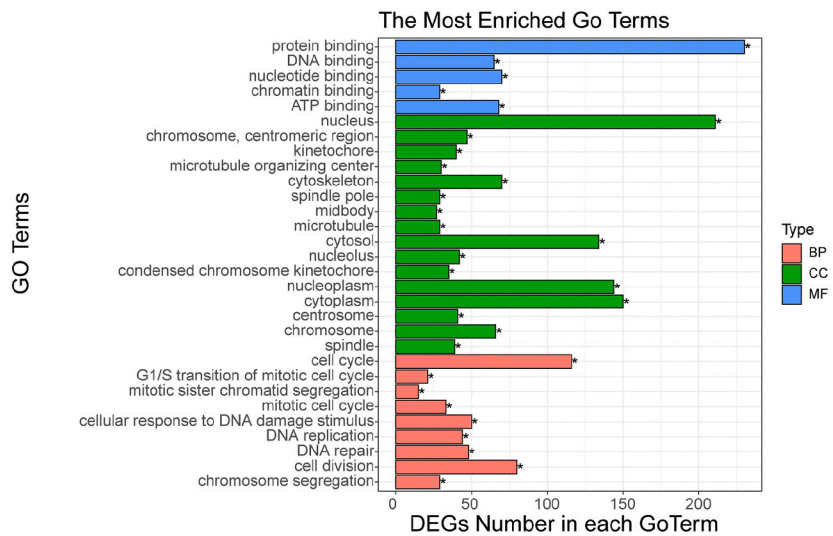
A



B



C



(caption on next page)

Fig. 3. Functional enrichment analysis of 307 hub module genes. (A). Correlation between modules and traits. The upper number in each cell refers to the correlation coefficient of each module in the trait, and the lower number is the corresponding p-value. (B). GO enrichment analysis. The y-axis shows significantly enriched GO terms, and x-axis represents the number of gene enrichment in this term. (C). KEGG pathway enrichment analysis. In this sub-figure, the x-axis represents the rich factor, and the y-axis represents the different KEGG pathways. The size of the bubbles grows as the number of involved genes increases.

above-mentioned nine modules, the brown module was highly correlated to Macrophages M1 ($R^2 = 0.4$, $P = 6e-20$; Fig. 3A). The correlation between other modules and macrophages was less than 0.4. So, we focused on the brown module that was identified as a hub module. 308 genes in this module were analyzed using the web tool “Matascape” for the next pathway and process enrichment analysis. 30 highest enrichment terms were all cell cycle- and cellular structure-related terms (Fig. 3B), and the three most highly enriched terms were protein binding (molecular composition (MF)), nucleus (cellular component (CC)), and cell cycle (biological process (BP)) (Fig. 3B, Supplementary Table S5).

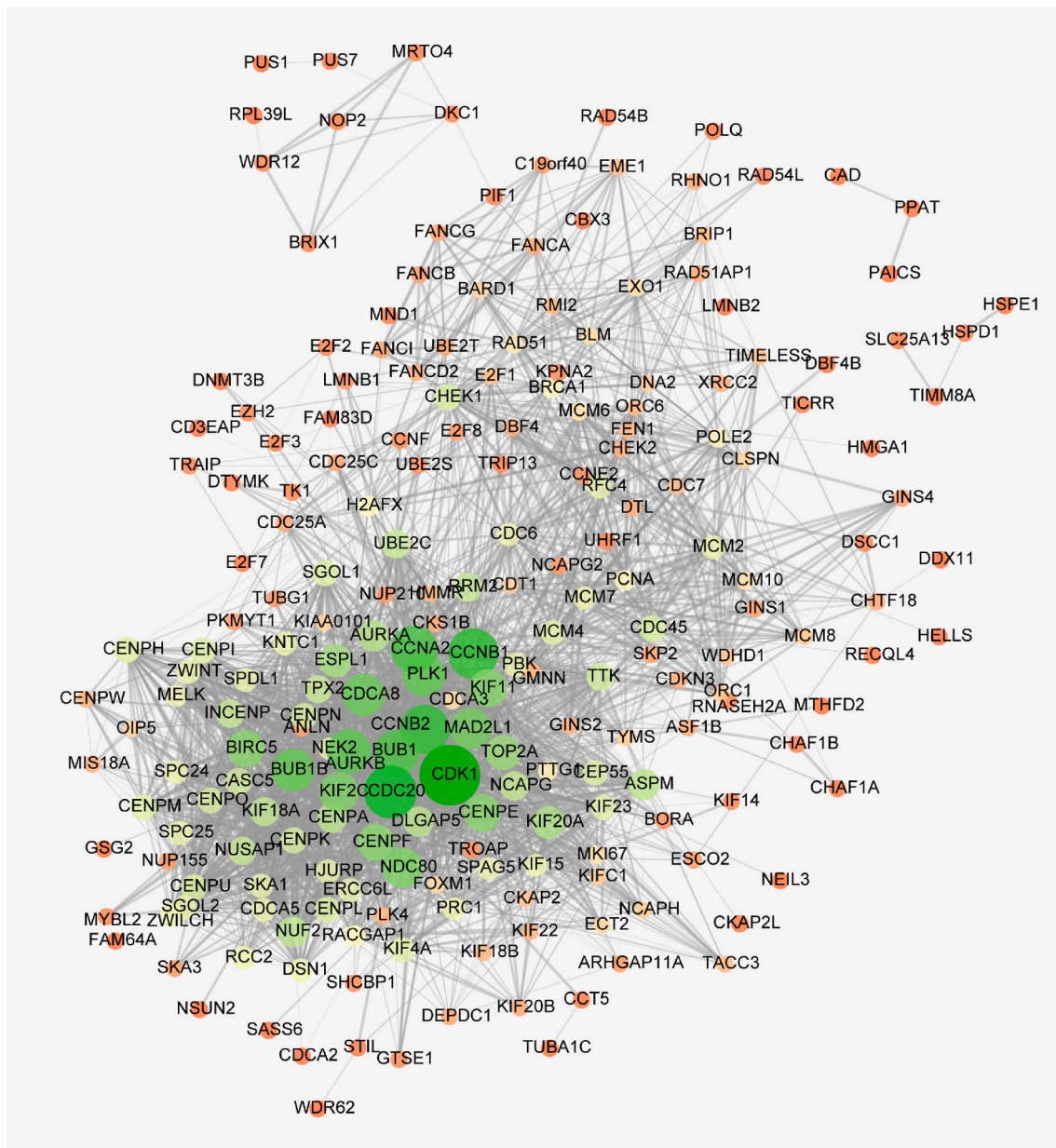


Fig. 4. PPI network of WGCNA hub module genes. A total of 307 genes were filtered into the hub module genes PPI network complex that contained 224 nodes and 2836 edges.

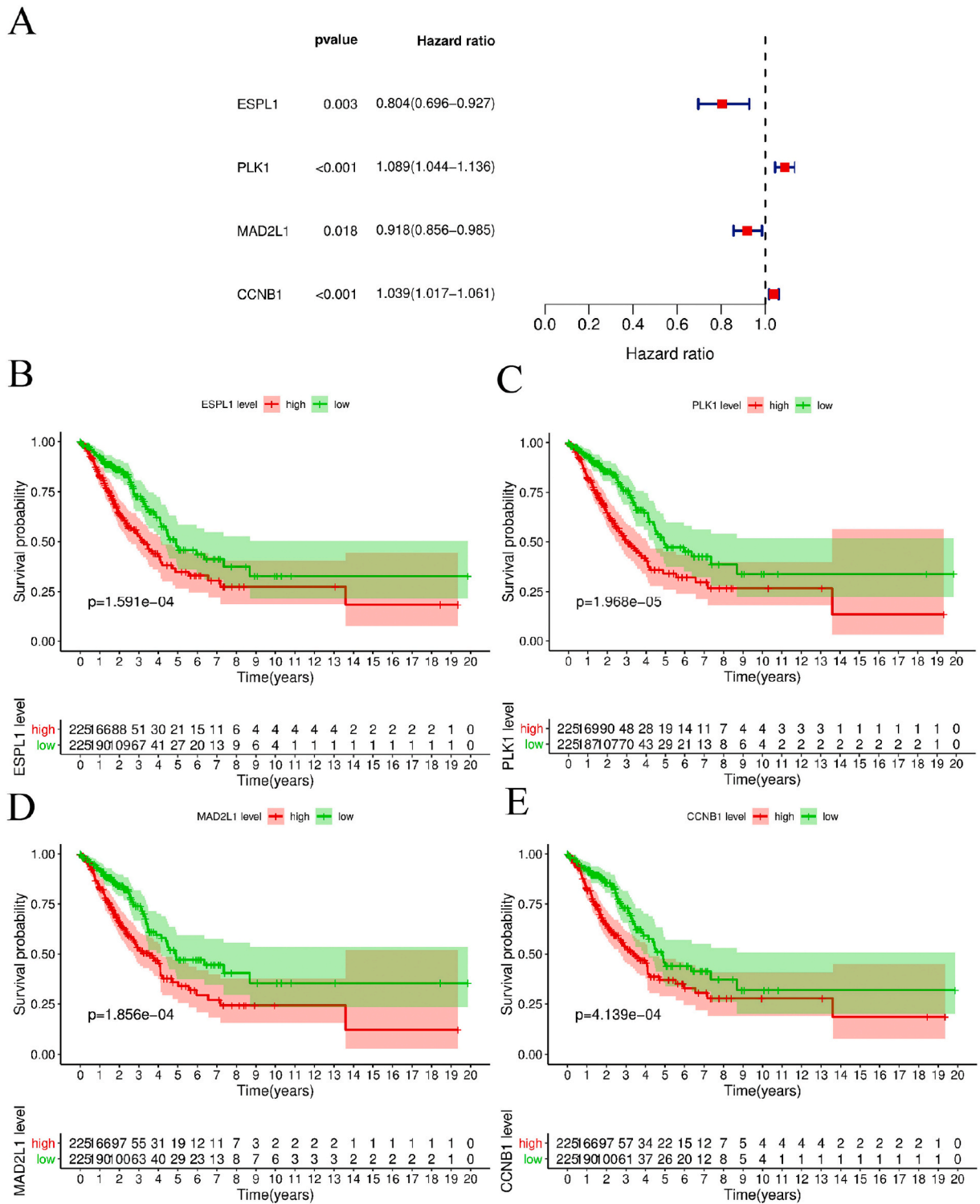


Fig. 5. Association of prognostic hub genes with OS in LUAD patients (A). Forest plot of multivariate Cox regression analysis to assess the association between hub genes and LUAD patients' OS. The overall survival of LUAD patients with high or low expression of ESPL1 (B), PLK1 (C), MAD2L1 (D), and CCNB1 (E).

Pathway analysis based on KEGG showed that these genes in the hub module considerably were enriched in 107 pathways (Fig. 3C, Supplementary Table S6). Among these terms, the cell cycle pathway was the most significant term, which also enriched more genes than other terms. Several enriched KEGG pathways were directly associated with cancer, including the ‘MicroRNAs in cancer’ and ‘Pancreatic cancer’ pathways. Furthermore, there was the enrichment of certain other terms that were potentially involved in the development and progression of LUAD via various biological processes, including the ‘DNA replication’, ‘p53 signaling pathway’, ‘Apoptosis’, and ‘Nucleotide excision repair’.

3.4. Identification of prognostic hub genes

The highly connected genes of the module were investigated as potential key factors related to M1-like macrophages’ infiltration level. According to the cutoff standard (combined score >0.9), we selected the eligible genes to construct the PPI network and designated the top 20 genes as hub genes based on their degree values (Fig. 4; Supplementary Table S7). We visualized these results using Cytoscape.

We analyzed 20 hub genes by univariate Cox proportional hazard regression analysis. The results indicated that 18 genes except *AURKB* and *TOP2A* were statistically significant ($P < 0.05$; Supplementary Table S8). Through the subsequent multivariate Cox analysis, 4 genes (*ESPL1*, *PLK1*, *MAD2L1*, and *CCNB1*) were regarded as prognostic hub genes (Fig. 5A). Besides, we further analyzed the relationship between the expression of the above four prognostic hub genes and the survival probability of patients (Fig. 5B–E). The results showed that the survival probability of patients with low expression of the hub gene was better than that of patients with high expression, indicating that the above-mentioned prognostic hub genes were significantly related to survival probability. Then we selected the above genes to construct a risk model and further evaluate its prognostic significance.

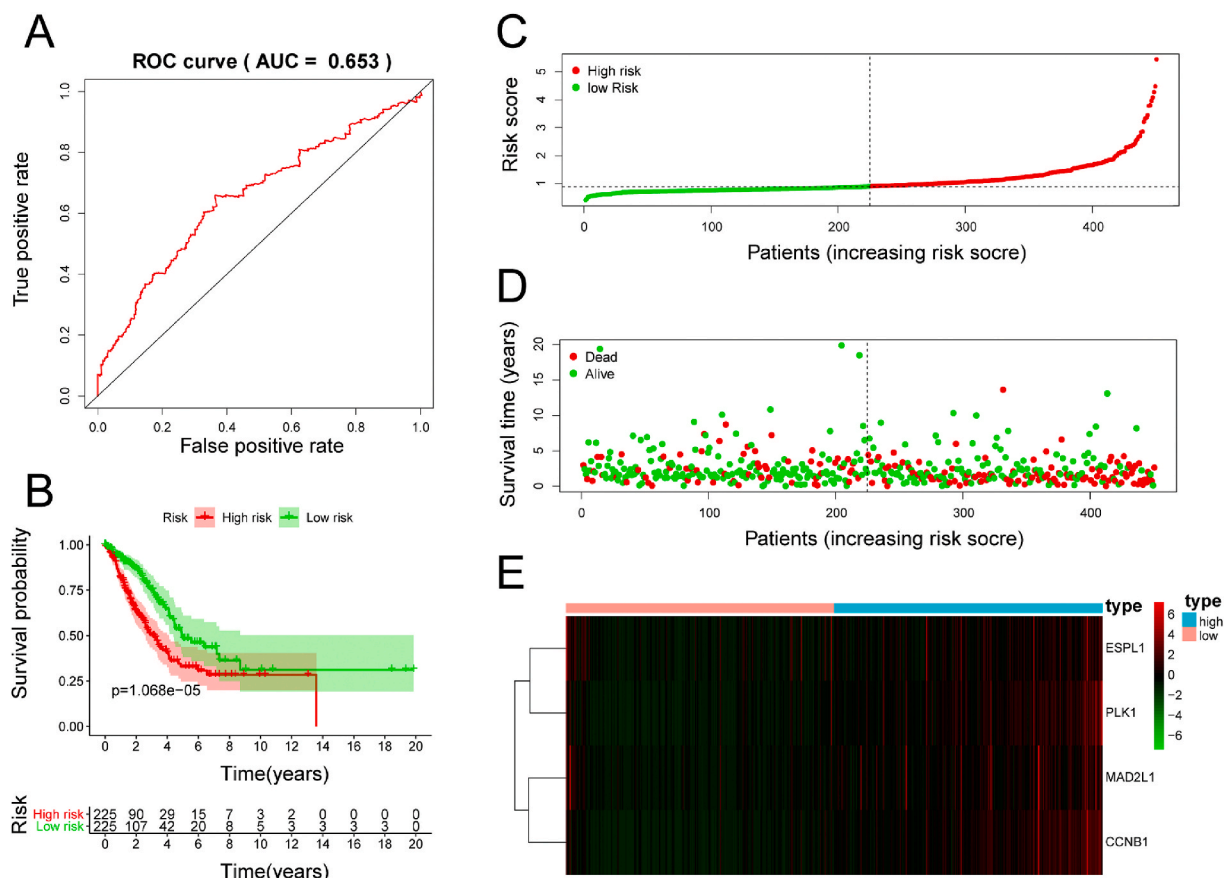


Fig. 6. Construction of the prognostic model. (A). ROC curve for the TCGA- LUAD cohort. (B). Kaplan-Meier survival curves of the TCGA- LUAD cohort. (C). The risk score distribution for 450 LUAD patients (TCGA dataset). (D). Distribution of duration of survival in the TCGA- LUAD data. The x-axis is arranged in order of patient risk score and y-axis represents patient survival time. (E). The expression pattern of the prognostic signature genes in the classifiers of the high- and low-risk groups.

3.5. Prognostic value of risk model

The ROC curve was used to prove the validity of the 4-prognosis hub gene risk model (AUC = 0.653; Fig. 6A). To evaluate the prognostic value of the risk model, the risk value of each patient was calculated by the formula below: Risk score = $ESPL1 \times (-0.218668867) + PLK1 \times 0.085289337 + MAD2L1 \times (-0.085415913) + CCNB1 \times 0.037918018$, and the patients were divided into high- and low-risk groups according to the median risk value. The K-M curve showed a significant difference in survival between the

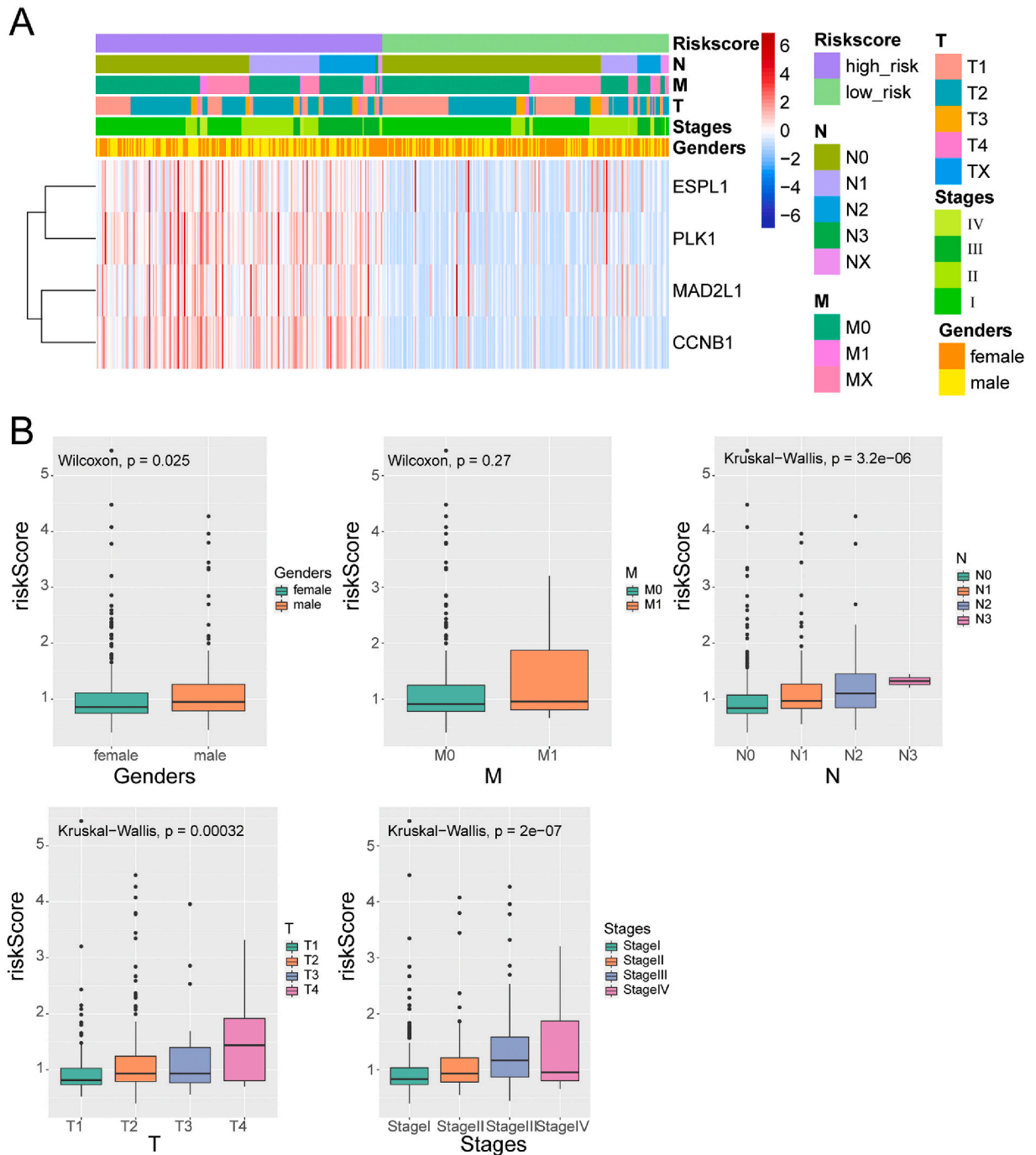


Fig. 7. Association between the prognostic hub genes and clinic pathological features. (A) Heatmap shows the association of risk scores and clinic pathological features based on the prognostic hub genes. (B) Distribution of the risk score in patients stratified by gender, M stage, N stage, T stage, and stages.

two groups (Fig. 6B). We then ranked the risk score of patients for survival probability and analyzed their distribution (Fig. 6C). Dot plots showed the survival status of individual LUAD patients in the training groups (Fig. 6D). The heatmap showed the expression patterns of the risk genes in the high- and low-risk patient groups (Fig. 6E). As shown, patients with high-risk scores showed upregulation of *ESPL1*, *PLK1*, *MAD2L1*, and *CCNB1* (high-risk genes), whereas, patients with low-risk scores showed downregulation of the above genes. Combining the previous results, we concluded that the higher the expression of the four prognostic hub genes, the greater the risk of the patient, and the higher the mortality rate. Coincidentally, this also confirmed the result of Fig. 5.

3.6. Distribution of risk scores for clinicopathological characteristics

Previous results indicated that macrophages may have important significance in the classification of cancer clinicopathology [20]. Furthermore, prognostic hub genes used to construct the risk model were screened from the macrophages relate-brown module. Therefore, we further investigated the correlation between the risk score and the clinicopathological characteristics of LUAD. The heatmap showed the expression of the prognostic hub genes in the high- and low-risk groups and the correlation between the risk model and clinicopathological characteristics (Fig. 7A). The relationships between risk score and clinicopathological stages were shown by boxplot (Fig. 7B–D). Except for the M stage, all risk score levels showed significant differences in clinicopathological stages ($P < 0.05$) and showed an upward trend with increased stages. Although no significant difference was detected for M0 and M1, the risk

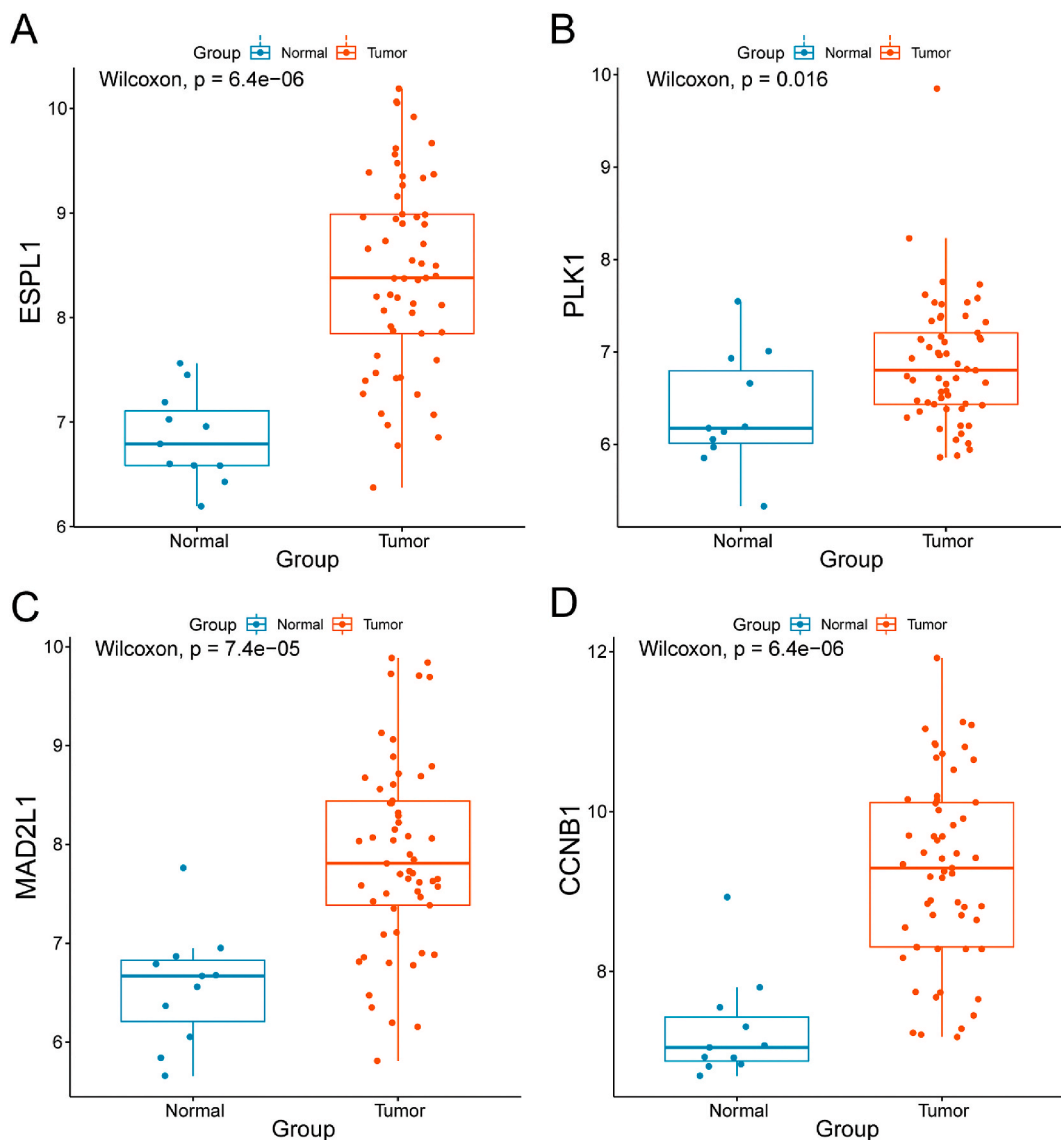


Fig. 8. Expression pattern validation for the prognostic genes. *ESPL1* (A), *PLK1* (B), *MAD2L1* (C), and *CCNB1* (D) in the GSE116959 database for expressing validation. Statistical differences in these data were calculated using Wilcoxon rank-sum test.

score level showed an upward trend with increased tumor stage.

3.7. Validation of risk model gene expression

The GEO dataset, GSE116959 (68 samples in total, including 11 Normal samples and 57 Tumor samples) was used to verify the expression of the prognostic hub gene. The results were consistent with the trend in TCGA. *ESPL1*, *PLK1*, *MAD2L1*, and *CCNB1* were all significantly expressed in the Tumor group (all $P < 0.05$; Fig. 8A–D).

4. Discussion

Tumor-associated macrophages is heterogeneous in the tumor microenvironment, different types of TAMs play different roles within the tumor ecosystem. M1-like macrophage-related genes promote inflammatory processes and also play a role in the progression of a variety of malignancies. However, although it has been confirmed that some M1 macrophage-related genes play carcinogenic roles in LUAD, the functions and mechanisms of M1 macrophage-related genes in LUAD remain to be further studied. We believe that, the analysis of LUAD expression data demonstrates that M1 macrophage-related genes have prognostic value.

CIBERSORT, a gene expression-based deconvolution algorithm, was recently developed to accurately estimate the fractions of the infiltration of immune cells subsets in tumor samples [15]. In the present study, weighted gene coexpression network analysis (WGCNA) and CIBERSORT were used to analyse LUAD expression data and identify cellular biomarkers related to M1-like macrophages. We identified 4 hub genes (*ESPL1*, *PLK1*, *MAD2L1*, and *CCNB1*) and established a prognostic risk model composed of the 4 hub genes to evaluate the prognosis of LUAD.

ESPL1 gene, encoding the Separase protein, is a protease that cleaves chromosomal cohesin during mitosis. It is overexpressed in a wide range of cancers [21,22,23]. Moreover, the overexpression of *ESPL1* leads to aneuploidy and tumorigenesis in animal models. Mukherjee, et al. showed that *MMTV-Espl1* mice in a C57BL/6 genetic background develop aggressive, highly aneuploid, and estrogen receptor alpha positive (ER α +) mammary adenocarcinomas [24]. More recently, the Liu Jie group reported that abnormal expression of *ESPL1* in endometrial cancer cells promoted metastasis and invasion, resulting in poor prognosis of endometrial cancer [25]. In a clinical study with long-term follow-up data, abnormal separase expression could predicted impaired survival for luminal breast carcinoma. In multivariate analyses, abnormal separase expression showed independent prognostic value [26]. There are few studies have investigated the relationship between *ESPL1* and lung cancer. Xiaona He et al. identified a total of 2932 DEGs, of which five genes, including *ESPL1*, may link lung adenocarcinoma to smoking history [27]. However, the exact mechanism of *ESPL1* in cancer-promoting processes remain unknown. Our results suggest that *ESPL1* was significantly overexpressed in lung cancer, and it was a risk factor in LUAD patients.

PLK1 (polo like kinase 1) is a highly conserved serine/threonine protein kinase that is broadly expressed in eukaryotic cells and has crucial functions in the cell cycle and proliferation [28,29]. Overexpression of *PLK1* has been validated as a marker for poor prognosis in many cancers. *PLK1* knockdown decreases the survival of cancer cells [30]. A recent study revealed high *PLK1* expression was associated with shorter cancer specific overall survival and disease-free survival in patients with breast cancer (follow-up 15 years). In this study, multivariate analysis confirmed the negative prognostic significance of *PLK1* overexpression for cancer-specific overall survival [31]. *PLK1* expression was found to be higher in the lung adenocarcinoma tissues of patients, which was consistent with survival prediction [32]. Similarly, *PLK1* is associated with the aggressive progression and poor prognosis in lung squamous cell carcinoma patients [33].

Human *MAD2* mitotic arrest deficient-like1 (*MAD2L1*) is a member of the spindle checkpoint proteins. The *MAD2L1* gene is located on human chromosome 4 [34,35]. To date, studies have shown that the disruption of *MAD2L1* function in mammalian cells can affect the function of the spindle checkpoint, leading to aneuploidy or tumor development. The overexpression of *MAD2L1* in mice lead to chromosome instability [36]. A large number of studies have shown that deletion of *MAD2L1* gene can cause chromosomal instability (CIN), which then drives the occurrence of tumours [35,37]. A clinical study with 203 female patients diagnosed as primary breast cancer uncovered that high expression of *MAD2L1* was significantly associated with increased risk of disease recurrence and death [38].

CCNB1 is a member of the cyclin family [39]. *CCNB1* interacts with *CDK1* to form a complex that regulates the transition from G2 phase to M phase during mitosis. Recent studies have shown that the overexpression of *CCNB1* is closely related to the low survival rate of most solid tumours, indicating that the expression status of *CCNB1* is an important indicator of the prognosis of solid tumours [40]. A previous study noted that *Islet-1* (*ISL1*) promotes gastric cancer cell cycle progression and cell proliferation by upregulating *CCNB1*, *CCNB2* and *C-MYC* gene expression [41]. Chai [42] reported that *FOXM1* and *CCNB1* are co-overexpressed in liver tumor tissues, and that high levels of *FOXM1* and *CCNB1* are associated with poor prognosis in HCC patients. Zilong Li and colleagues demonstrated that *BRG1* regulated *CCNB1* and *LTBP2* transcription by altering histone modifications on target promoters in lung cancer cells [43]. As a G2/M checkpoint member, *PLK1* can be activated by the *CDK1/CCNB1* complex, forming a positive feedback loop [42].

Our data showed a significant difference in survival between high- and low-risk groups according to the median risk value. The higher the expression of the four prognostic hub genes, the greater the risk of the patient, and the higher the mortality rate. K-M analysis revealed significant differences in survival rates between high- and low-risk groups. The area under curve (AUC) of the ROC curves of these four hub genes was more than 65%, suggesting that they have the potential as biomarkers for evaluation of the prognosis of LUAD. A previous study has demonstrated that immune-related gene expression models can predict prognosis and microenvironment immune infiltration in lung cancer based on GEO [44]. However, the data used in that study include lung cancers of all histological types, including 17 small cell carcinomas. Moreover, they used TCGA data from 182 LUAD patients as a validation set,

but these data do not contain small cell carcinoma samples. Therefore, there is a possibility of presence of selection bias. A prognosis analyse have identified 6 hub genes which correlated with poor OS and poor PFS in LUAD. However, all the area under receiver operating characteristics (AUCs) were lower than 55, which indicated that the predictive role of 6 hub genes were poor [45]. Another Multi-omics Data Analyses Identify the Immune-Related Prognosis markers in Human LUAD which achieved a high degree of accuracy on the 1-year (AUC = 0.861), 3-year (AUC = 0.850) survival predictions. But the prognosis model comprised 27 variables. Too many variables are required to construct a model, reducing its clinical practicality [46].

Our prognostic model combines four hub genes to guide clinical treatment and the judgement of prognosis. Patients with high risk should be treated with increased precaution and aggressive therapy. At present, Polo-like kinases (PLKs) have emerged as attractive targets of cancer therapeutics. Volasertib, the most advanced and most potent PLK1 inhibitor is currently being tested in Phase I-III trials [47]. Beyond that, inhibition of Plk1 sensitized pancreatic ductal adenocarcinoma to immune checkpoint blockade therapy through activation of an antitumor immune response [48]. The model associated patient prognosis with these four genes; to the best of our knowledge, this has not been reported previously. It is advanced, accurate, and has high clinical practicability. Therefore, we have reason to believe that it may be a useful tool for predicting LUAD prognosis. Nevertheless, the current study also had some shortcomings. First, despite the large number of samples downloaded from TCGA, some subgroup analyses could not be performed due to insufficient numbers of samples. Second, this study was a preliminary bioinformatic analysis, further experimental, clinical research and mechanistic studies of these four genes are needed.

5. Conclusions

In the present study, Weighted gene Co-expression Network Analysis (WGCNA) and the CIBERSORT algorithm were used to analyse LUAD expression data and identify modules related to M1 macrophages. A risk scoring model based on the ESPL1, PLK1, MAD2L1, and CCNB1 genes was established to predict the prognosis of LUAD patients. However, the biological roles and functions of these four genes in LUAD still require additional study, and more detailed clinical data are needed to construct a more widely applicable prognostic model. We hope that this study will provide important perspectives for the prognosis of LUAD and eventually provide aid in clinical decision-making.

Declarations

Author contributions

Zhiyuan Wang conceived and designed the experiments, performed the experiments, analyzed and interpreted the data, wrote the paper.

Shan Yan conceived and designed the experiments, performed the experiments, analyzed and interpreted the data, wrote the paper.

Ying Yang performed the experiments, wrote the paper.

Xuan Luo performed the experiments, analyzed and interpreted the data.

Xiaofang Wang performed the experiments.

Kun Tang performed the experiments.

Juan Zhao performed the experiments.

Yongwen He conceived and designed the experiments, contributed reagents, materials, analysis tools or data.

Li Bian conceived and designed the experiments, contributed reagents, materials, analysis tools or data.

All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the National Natural Science Foundation of China [grant number 82060423] and Yunnan Provincial Department of science and Technology-Kunming Medical University applied basic research joint special major project [grant number 202001AY0700 1-007] to L. Bian. Z.Y. Wang was supported by a grant from Yunnan Provincial Department of science and Technology-Kunming Medical University applied basic research joint special general project [grant number 2019FE001(-219)].

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank Zhongfa Liu for expert assistance with the preparation of this study.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.heliyon.2023.e12798>.

References

- [1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA A Cancer J. Clin.* 71 (2021) 209–249, <https://doi.org/10.3322/caac.21660>.
- [2] W.D. Travis, E. Brambilla, A.P. Burke, A. Marx, A.G. Nicholson, Introduction to the 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart, *J. Thorac. Oncol.* 10 (2015) 1240–1242, <https://doi.org/10.1097/JTO.0000000000000663>.
- [3] O. Pearce, R.M. Delaine-Smith, E. Maniati, S. Nichols, J. Wang, S. Böhm, V. Rajeeve, D. Ullah, P. Chakravarty, R.R. Jones, et al., Deconstruction of a metastatic tumor microenvironment reveals a common matrix response in human cancers, *Cancer Discov.* 8 (2018) 304–319, <https://doi.org/10.1158/2159-8290.CD-17-0284>.
- [4] P. Muppa, S.B.S.P. Terra, A. Sharma, A.S. Mansfield, M.C. Aubry, K. Bhingre, M.K. Asiedu, M. Andrade, N. Janaki, S.J. Murphy, A. Nasir, V.V. Keulen, G. Vasmatazis, D.A. Wigle, P. Yang, E.S. Yi, T. Peikert, F. Kosari, Immune cell infiltration may be a key determinant of long-term survival in small cell lung cancer, *J. Thorac. Oncol.* 14 (2019) 1286–1295, <https://doi.org/10.1016/j.jtho.2019.03.028>.
- [5] R. Jia, Z. Sui, H.D. Zhang, Z.T. Yu, Identification and validation of immune-related gene signature for predicting lymph node metastasis and prognosis in lung adenocarcinoma, *Front. Mol. Biosci.* 8 (2021), 679031, <https://doi.org/10.3389/fmolb.2021.679031>.
- [6] Q. Wang, D.T. Zhou, F. Wu, Q.C. Liang, Q.Z. He, M.Y. Peng, T.Y. Yao, Y. Hu, B.L. Qian, J.Q. Tang, X. Wang, W.L. Liu, F.L. Yu, C. Chen, Immune microenvironment signatures as biomarkers to predict early recurrence of stage ia-b lung cancer, *Front. Oncol.* 11 (2021), 680287, <https://doi.org/10.3389/fonc.2021.680287>.
- [7] F.O. Martinez, S. Gordon, The M1 and M2 paradigm of macrophage activation: time for reassessment, *F1000prime Rep.* 6 (2014) 13, <https://doi.org/10.12703/P6-13>.
- [8] N. Wang, H. Liang, K. Zen, Molecular mechanisms that influence the macrophage m1-m2 polarization balance, *Front. Immunol.* 5 (2014) 614, <https://doi.org/10.3389/fimmu.2014.00614>.
- [9] H. Wang, X. Wang, X. Li, Y. Fan, G. Li, C. Guo, F. Zhu, L. Zhang, Y. Shi, CD68(+)/HLA-DR(+) M1-like macrophages promote motility of HCC cells via NF- κ B/FAK pathway, *Cancer Lett.* 345 (2014) 91–99, <https://doi.org/10.1016/j.canlet.2013.11.013>.
- [10] M. Estko, S. Baumgartner, K. Urech, M. Kunz, U. Regueiro, P. Heusser, U. Weissenstein, Tumour cell derived effects on monocyte/macrophage polarization and function and modulatory potential of viscum album lipophilic extract in vitro, *BMC Compl. Alternative Med.* 15 (2015) 130, <https://doi.org/10.1186/s12906-015-0650-3>.
- [11] T.H. Wu, Y.Y. Li, T.L. Wu, J.W. Chang, W.C. Chou, L.L. Hsieh, J.R. Chen, K.Y. Yeh, Culture supernatants of different colon cancer cell lines induce specific phenotype switching and functional alteration of THP-1 cells, *Cell. Immunol.* 290 (2014) 107–115, <https://doi.org/10.1016/j.cellimm.2014.05.015>.
- [12] E. Karnevi, R. Andersson, A.H. Rosendahl, Tumour-educated macrophages display a mixed polarisation and enhance pancreatic cancer cell invasion, *Immunol. Cell Biol.* 92 (2014) 543–552, <https://doi.org/10.1038/icb.2014.22>.
- [13] M. Xiao, J. Zhang, W. Chen, W. Chen, M1-like tumor-associated macrophages activated by exosome-transferred THBS1 promote malignant migration in oral squamous cell carcinoma, *J. Exp. Clin. Cancer Res.* 37 (2018) 143, <https://doi.org/10.1186/s13046-018-0815-2>.
- [14] U. Cho, B. Kim, S. Kim, Y. Han, Y.S. Song, Pro-inflammatory M1 macrophage enhances metastatic potential of ovarian cancer cells through NF- κ B activation, *Mol. Carcinog.* 57 (2018) 235–242, <https://doi.org/10.1002/mc.22750>.
- [15] A.M. Newman, C.B. Steen, C.L. Liu, A.J. Gentles, A.A. Chaudhuri, F. Scherer, M.S. Khodadoust, M.S. Eshfahani, B.A. Luca, D. Steiner, M. Diehn, A.A. Alizadeh, Determining cell type abundance and expression from bulk tissues with digital cytometry, *Nat. Biotechnol.* 37 (2019) 773–782, <https://doi.org/10.1038/s41587-019-0114-2>.
- [16] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, *BMC Bioinf.* 9 (2008) 559, <https://doi.org/10.1186/1471-2105-9-559>.
- [17] W. Walter, F. Sánchez-Cabo, M. Ricote, GPlot: an R package for visually combining expression data with functional analysis, *Bioinformatics (Oxford, England)* 31 (2015) 2912–2914, <https://doi.org/10.1093/bioinformatics/btv300>.
- [18] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K.P. Tsafou, M. Kuhn, P. Bork, L.J. Jensen, C. von Mering, STRING v10: protein-protein interaction networks, integrated over the tree of life, *Nucleic Acids Res.* 43 (2015) D447–D452, <https://doi.org/10.1093/nar/gku1003>.
- [19] P. Shannon, A. Markiel, O. Zier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res.* 13 (2003) 2498–2504, <https://doi.org/10.1101/gr.1239303>.
- [20] N. Harimoto, K. Hoshino, R. Muranushi, K. Hagiwara, T. Yamanaka, N. Ishii, M. Tsukagoshi, T. Igarashi, H. Tanaka, A. Watanabe, N. Kubo, K. Araki, Y. Hosouchi, H. Suzuki, K. Arakawa, K. Hirai, T. Fukazawa, H. Ikota, K. Shirabe, Prognostic significance of neutrophil-lymphocyte ratio in resectable pancreatic neuroendocrine tumors with special reference to tumor-associated macrophages, *Pancreatology* 19 (2019) 897–902, <https://doi.org/10.1016/j.pan.2019.08.003>.
- [21] D.Y. Wen, P. Lin, Y.Y. Pang, G. Chen, Y. He, Y.W. Dang, H. Yang, Expression of the long intergenic non-protein coding RNA 665 (LINC00665) gene and the cell cycle in hepatocellular carcinoma using the cancer Genome Atlas, the gene expression omnibus, and quantitative real-time polymerase chain reaction, *Med. Sci. Mon. Int. Med. J. Exp. Clin. Res.* 24 (2018) 2786–2808, <https://doi.org/10.12659/MSM.907389>.
- [22] P. Finetti, A. Guille, J. Adelaide, D. Birnbaum, M. Chaffanet, F. Bertucci, ESPL1 is a candidate oncogene of luminal B breast cancers, *Breast Cancer Res. Treat.* 147 (2014) 51–59, <https://doi.org/10.1007/s10549-014-3070-z>.
- [23] R. Meyer, V. Fofanov, A. Panigrahi, F. Merchant, N. Zhang, D. Pati, Overexpression and mislocalization of the chromosomal segregation protein separase in multiple human cancers, *Clin. Cancer Res.* 15 (2009) 2703–2710, <https://doi.org/10.1158/1078-0432.CCR-08-2454>.
- [24] M. Mukherjee, G. Ge, N. Zhang, D.G. Edwards, P. Sumazin, S.K. Sharan, P.H. Rao, D. Medina, D. Pati, MMTV-Espl1 transgenic mice develop aneuploid, estrogen receptor alpha (ER α)-positive mammary adenocarcinomas, *Oncogene* 33 (2014) 5511–5522, <https://doi.org/10.1038/onc.2013.493>.
- [25] J. Liu, Y. Wan, S. Li, H. Qiu, Y. Jiang, X. Ma, S. Zhou, W. Cheng, Identification of aberrantly methylated differentially expressed genes and associated pathways in endometrial cancer using integrated bioinformatic analysis, *Cancer Med.* 9 (2020) 3522–3536, <https://doi.org/10.1002/cam4.2956>.
- [26] N. Gurvits, E. Löyttyniemi, M. Nykänen, T. Kuopio, P. Kronqvist, K. Talvinen, Separase is a marker for prognosis and mitotic activity in breast cancer, *Br. J. Cancer* 117 (2017) 1383–1391, <https://doi.org/10.1038/bjc.2017.301>.
- [27] X. He, C. Zhang, C. Shi, Q. Lu, Meta-analysis of mRNA expression profiles to identify differentially expressed genes in lung adenocarcinoma tissue from smokers and non-smokers, *Oncol. Rep.* 39 (2018) 929–938, <https://doi.org/10.3892/or.2018.6197>.
- [28] F.A. Barr, H.H. Silljé, E.A. Nigg, Polo-like kinases and the orchestration of cell division, *Nat. Rev. Mol. Cell Biol.* 5 (2004) 429–440, <https://doi.org/10.1038/nrm1401>.
- [29] V. Archambault, D.M. Glover, Polo-like kinases: conservation and divergence in their functions and regulation, *Nat. Rev. Mol. Cell Biol.* 10 (2009) 265–275, <https://doi.org/10.1038/nrm2653>.
- [30] W.T. Weng Ng, J.S. Shin, T.L. Roberts, B. Wang, C.S. Lee, Molecular interactions of polo-like kinase 1 in human cancers, *J. Clin. Pathol.* 69 (2016) 557–562, <https://doi.org/10.1136/jclinpath-2016-203656>.
- [31] P. Donizy, A. Halon, P. Surowiak, M. Kaczorowski, C. Kozyra, R. Matkowski, Augmented expression of Polo-like kinase 1 is a strong predictor of shorter cancer-specific overall survival in early stage breast cancer at 15-year follow-up, *Oncol. Lett.* 12 (2016) 1667–1674, <https://doi.org/10.3892/ol.2016.4890>.
- [32] Y. Bian, Q. Sui, G. Bi, Y. Zheng, M. Zhao, G. Yao, L. Xue, Y. Zhang, H. Fan, Identification and validation of a proliferation-associated score model predicting survival in lung adenocarcinomas, *Dis. Markers* 2021 (2021), 3219594, <https://doi.org/10.1155/2021/3219594>.
- [33] H. Li, H. Wang, Z. Sun, Q. Guo, H. Shi, Y. Jia, The clinical and prognostic value of polo-like kinase 1 in lung squamous cell carcinoma patients: immunohistochemical analysis, *Biosci. Rep.* 37 (2017), <https://doi.org/10.1042/BSR20170852>.

- [34] Z. Wang, D. Katsaros, Y. Shen, Y. Fu, E.M. Canuto, C. Benedetto, L. Lu, W.M. Chu, H.A. Risch, H. Yu, Biological and clinical significance of MAD2L1 and BUB1, genes frequently appearing in expression signatures for breast cancer prognosis, *PLoS One* 10 (2015), e0136246, <https://doi.org/10.1371/journal.pone.0136246>.
- [35] Y. Guo, X. Zhang, M. Yang, X. Miao, Y. Shi, J. Yao, W. Tan, T. Sun, D. Zhao, D. Yu, J. Liu, D. Lin, Functional evaluation of missense variations in the human MAD1L1 and MAD2L1 genes and their impact on susceptibility to lung cancer, *J. Med. Genet.* 47 (2010) 616–622, <https://doi.org/10.1136/jmg.2009.074252>.
- [36] R. Sotillo, E. Hernando, E. Díaz-Rodríguez, J. Teruya-Feldstein, C. Cordón-Cardo, S.W. Lowe, R. Benezra, Mad2 overexpression promotes aneuploidy and tumorigenesis in mice, *Cancer Cell* 11 (2007) 9–23, <https://doi.org/10.1016/j.ccr.2006.10.019>.
- [37] M.B. Kastan, J. Bartek, Cell-cycle checkpoints and cancer, *Nature* 432 (2004) 316–323, <https://doi.org/10.1038/nature03097>.
- [38] Z. Wang, D. Katsaros, Y. Shen, Y. Fu, E.M. Canuto, C. Benedetto, L. Lu, W.M. Chu, H.A. Risch, H. Yu, Biological and clinical significance of MAD2L1 and BUB1, genes frequently appearing in expression signatures for breast cancer prognosis, *PLoS One* 10 (2015), e0136246, <https://doi.org/10.1371/journal.pone.0136246>.
- [39] P. Zhao, P. Zhang, W. Hu, H. Wang, G. Yu, Z. Wang, C. Li, J. Bai, Y. Zhang, Upregulation of cyclin B1 plays potential roles in the invasiveness of pituitary adenomas, *J. Clin. Neurosci.* 43 (2017) 267–273, <https://doi.org/10.1016/j.jocn.2017.05.005>.
- [41] E. Ersvar, W. Kildal, L. Vlatkovic, K. Cyll, M. Pradhan, A. Kleppe, T.S. Hveem, H.A. Askautrud, M. Novelli, H. Wæhre, K. Liestøl, H.E. Danielsen, Prognostic value of mitotic checkpoint protein BUB3, cyclin B1, and pituitary tumor-transforming 1 expression in prostate cancer, *Mod. Pathol.* 33 (2020) 905–915, <https://doi.org/10.1038/s41379-019-0418-2>.
- [42] Q. Shi, W. Wang, Z. Jia, P. Chen, K. Ma, C. Zhou, ISL1, a novel regulator of CCNB1, CCNB2 and c-MYC genes, promotes gastric cancer cell proliferation and tumor growth, *Oncotarget* 7 (2016) 36489–36500, <https://doi.org/10.18632/oncotarget.9269>.
- [43] N. Chai, H.H. Xie, J.P. Yin, K.D. Sa, Y. Guo, M. Wang, J. Liu, X.F. Zhang, X. Zhang, H. Yin, Y.Z. Nie, K.C. Wu, A.G. Yang, R. Zhang, FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1, *Biochem. Biophys. Res. Commun.* 500 (2018) 924–929, <https://doi.org/10.1016/j.bbrc.2018.04.201>.
- [44] Z. Li, J. Xia, M. Fang, Y. Xu, Epigenetic regulation of lung cancer cell proliferation and migration by the chromatin remodeling protein BRG1, *Oncogenesis* 8 (2019) 66, <https://doi.org/10.1038/s41389-019-0174-7>.
- [45] S. Wen, W. Peng, Y. Chen, X. Du, J. Xia, B. Shen, G. Zhou, Four differentially expressed genes can predict prognosis and microenvironment immune infiltration in lung cancer: a study based on data from the GEO, *BMC Cancer* 22 (2022) 193, <https://doi.org/10.1186/s12885-022-09296-8>.
- [46] Q. Zheng, S. Min, Q. Zhou, Identification of potential diagnostic and prognostic biomarkers for LUAD based on TCGA and GEO databases, *Biosci. Rep.* 41 (2021), BSR20204370, <https://doi.org/10.1042/BSR20204370>.
- [47] Y. Zhang, M. Yang, D.M. Ng, M. Haleem, T. Yi, S. Hu, H. Zhu, G. Zhao, Q. Liao, Multi-omics data analyses construct TME and identify the immune-related prognosis signatures in human LUAD, *Mol. Ther. Nucleic Acids* 21 (2020) 860–873, <https://doi.org/10.1016/j.omtn.2020.07.024>.
- [48] [doi.org/10.1016/j.bbcan.2020.188467]. From the American Association of Neurological Surgeons (AANS), American Society of Neuroradiology (ASNR), Cardiovascular and Interventional Radiology Society of Europe (CIRSE), Canadian Interventional Radiology Association (CIRA), Congress of Neurological Surgeons (CNS), European Society of Minimally Invasive Neurological Therapy (ESMINT), European Society of Neuroradiology (ESNR), European Stroke Organization (ESO), Society for Cardiovascular Angiography and Interventions (SCAI), Society of Interventional Radiology (SIR), Society of NeuroInterventional Surgery (SNIS), and World Stroke Organization (WSO) D. Sacks, B. Baxter, B.C.V. Campbell, J.S. Carpenter, C. Cognard, D. Dippel, M. Eesa, U. Fischer, et al., Multisociety consensus quality improvement revised consensus statement for endovascular therapy of acute ischemic stroke, *Int. J. Stroke* 13 (2018) 612–632, <https://doi.org/10.1177/1747493018778713>.
- [49] Z. Zhang, L. Cheng, J. Li, Q. Qiao, A. Karki, D.B. Allison, N. Shaker, K. Li, S.M. Utturkar, N.M. Atallah Lanman, X. Rao, P. Rychahou, D. He, S.F. Konieczny, C. Wang, Q. Shao, B.M. Evers, X. Liu, Targeting Plk1 sensitizes pancreatic cancer to immune checkpoint therapy, *Cancer Res.* 82 (2022) 3532–3548, <https://doi.org/10.1158/0008-5472.CAN-22-0018>.