

Co-expression network analysis identified KIF2C in association with progression and prognosis in lung adenocarcinoma

Yuquan Bai, Lecai Xiong, Minglin Zhu, Zetian Yang, Jinping Zhao* and Hexiao Tang*

Department of Thoracic and Cardiovascular Surgery, Zhongnan Hospital of Wuhan University, Wuhan, Hubei 430071, China

Abstract. Lung cancer is a malignant tumor with high morbidity and mortality, of which 80% is non-small cell lung cancer (NSCLC). And lung adenocarcinoma (LUAD) is the most important and common subtype in the NSCLC. In current study, the microarray data GSE31210 containing LUAD ($n = 226$) and normal lung tissue ($n = 20$) was analyzed to identify 965 differentially expressed genes, on which weighted gene co-expression network analysis was performed. Finally, it was confirmed that there was a significant correlation between brown module and LUAD stage. In the significant module, a total of 54 network hub genes were identified, and six of them were also identified as hub genes of the protein-protein interaction network. In validation, KIF2C showed a higher correlation with disease stage than other hub genes ($p < 0.001$, $R^2 = 0.955$). Functional enrichment suggests that KIF2C is associated with cell mitosis and cell cycle. Combined with clinicopathological parameters, we found that the high expression of KIF2C is closely related to the relapse and tumor stage of LUAD. Survival analysis showed a significant reduction in overall survival in LUAD patients with high expression of KIF2C. Gene set enrichment analysis (GSEA) also showed that the “cell cycle signaling pathway” and “P53 related pathway” were significantly enriched in LUAD samples with high expression of KIF2C ($FDR < 0.05$). In conclusion, based on the co-expression analysis, KIF2C was identified in the association with progression and prognosis of LUAD, which might refer a poor prognosis probably by regulating cell cycle signaling pathway.

Keywords: KIF2C, Lung adenocarcinoma, co-expression analysis, cell cycle

1. Introduction

Lung cancer is the leading cause of cancer-related deaths in the worldwide, with 80% of lung cancer being non-small cell lung cancer (NSCLC) [1]. NSCLC includes three major histological subtypes: squamous cell carcinoma, adenocarcinoma and large cell carcinoma. The most common type of NSCLC is lung adenocarcinoma (LUAD), accounting for 40% of all lung cancer [2]. In recent years, despite the great breakthroughs in the diagnosis and treatment of NSCLC, the 5-year survival rate of advanced NSCLC remains

low [3]. Therefore, identification of sensitive biomarkers which can predict tumor recurrence and prognosis is still an urgent task in guiding the treatment of NSCLC.

Currently, with the development of high-throughput chip technologies, gene expression profiling has been used to identify genes which are associated with LUAD progression [4–7]. However, most studies have focused on the screening of differentially expressed genes (DEGs), ignoring the high correlation between genes, although genes with similar patterns of expression may be functionally related [8].

Weighted gene co-expression network analysis (WGCNA) is used to study the association of gene sets with clinical features by constructing free-scale gene co-expression networks [9]. Recently, many studies have used WGCNA to screen hub genes and modules which are associated with clinical features through mi-

*Corresponding authors: Jinping Zhao/Hexiao Tang, Department of Thoracic and Cardiovascular Surgery, Zhongnan Hospital of Wuhan University, Wuhan, Hubei 430071, China. Tel.: +86 138 086 993 92; E-mail: ZhaoJPthx@126.com; Tel.: +86 151 027 235 63; E-mail: thx1245@sina.com.

croarray or RNA sequencing data [10,11]. Therefore, we attempted to construct a co-expression network for the relationship between genes using WGCNA and identify the relationship between the network hub genes and different stages of LUAD [11,12].

2. Materials and method

2.1. Data collection

The research workflow is shown in Fig. 1. The LUAD microarray data set was downloaded from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). Data set GSE31210 [13] performed on Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA, USA) was used to construct a co-expression network and to identify hub genes in this study. The dataset includes 226 LUAD and 20 normal lung tissues. And the clinical information as following: different histological staging (I-II), smoking status (ever-smoker or never-smoker), recurrence (yes or no) and death (yes or no). Another independent data set, GSE40791 [14], was also downloaded from GEO and used to verify our results. This dataset includes 100 non-tumor lung tissues and 94 LUAD with different tumor stages (I-III).

2.2. Data preprocessing

The original microarray data was subjected to RMA background and log₂ transformations and then normalized by quantile. Median-polish probesets were summarized by using the “Affy” R package [15]. At last, probes were annotated by the Affymetrix annotation files. Microarray quality was assessed by sample clustering according to the distance between different samples in Pearson’s correlation matrices and average linkage.

2.3. Differentially expressed genes (DEGs) screening

Differentially expressed data between normal lung tissues and LUAD tissues were screened using the “limma” R package [16]. The significance analysis of microarrays with false discovery rate (FDR) < 0.05 and | log₂ fold change (FC) | > 1.5 was used to screen for genes that require further consideration in network construction.

2.4. Co-expression network construction

First, we examined the expression data characteristics of DEGs and observed whether it is necessary to exclude outlier samples and genes. The co-expression network of DEGs was then constructed using the “WGCNA” R package [17,18]. The Pearson’s correlation matrices were both performed for all pairwise genes. Then, a weighted adjacency matrix was constructed using a power function $a_{mn} = |c_{mn}|^\beta$ (c_{mn} = Pearson’s correlation between gene m and gene n ; a_{mn} = adjacency between gene m and gene n). β was a soft-thresholding parameter that could emphasize strong correlations between genes and penalize weak correlations. Here, $\beta = 5$ (scale-free $R^2 = 0.87$) was used to ensure scale-free network (Fig. 3). After that, the adjacency matrix was transformed into topology overlay matrix (TOM) [19]. To classify genes with similar expression profiles into gene modules, average linkage hierarchical clustering was conducted according to the TOM-based dissimilarity measure. To further analyze the module, we calculated the dissimilarity of the module information, selected a tangent for the module tree, and merged some of the modules.

2.5. Identification of clinical significant modules

Two approaches were used to identify modules related with the progression of LUAD. First, Gene significance (GS) was defined as the log₁₀ transformation of the P value ($GS = \lg P$) in the linear regression between gene expression and pathological stage. In addition, module significance (MS) was defined as the average GS for all the genes in a module. Module eigengenes (MEs) were considered as the major component in the principal component analysis for each gene module and the expression patterns of all genes could be summarized into a single characteristic expression profile within a given module. In addition, we calculated the correlation between MEs and clinical trait to identify the relevant module.

2.6. Analysis and validation of hub gene

In our study, hub genes were defined by module connectivity, measured by absolute value of the Pearson’s correlation ($\text{cor.geneModuleMembership} > 0.8$) and clinical trait relationship, measured by absolute value of the Pearson’s correlation ($\text{cor.geneTraitSignificance} > 0.2$). To screen for the key candidates in the hub gene, a linear regression analysis was performed to cal-

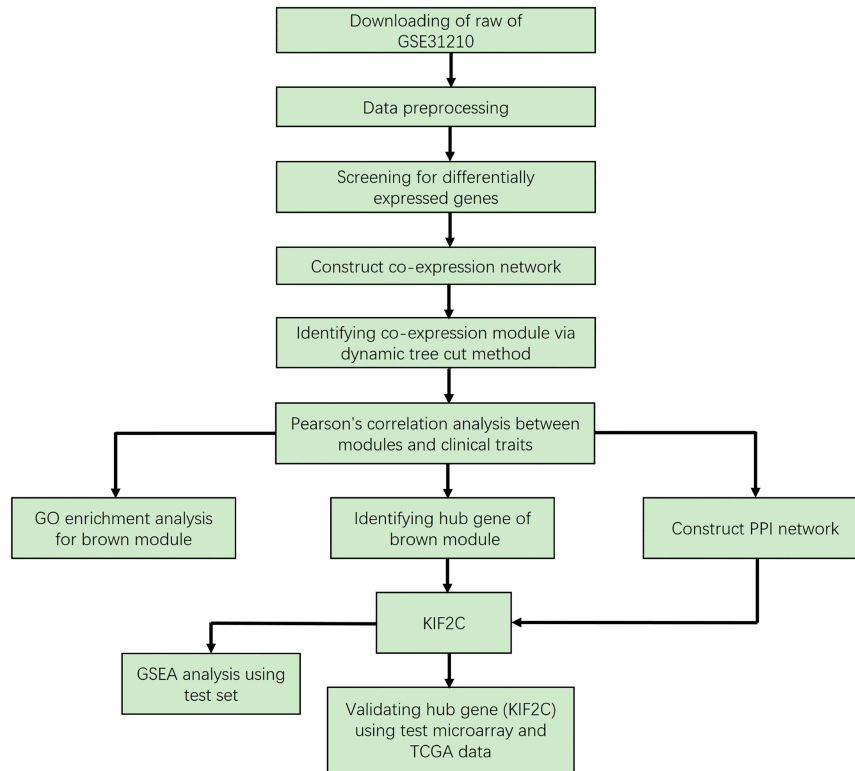


Fig. 1. Flow chart of data preparation, processing, analysis and validation in this study.

culate the relationship between hub gene expression and clinical information. In addition, we uploaded all the genes into the STRING database [20] to construct a protein-protein interaction (PPI) network to screen the central node. To validate our results, we used clinical information in microarray data to explore the correlation between hub gene expression and clinical characteristics of LUAD. The Oncomine database and test set GSE40791 were used to detect the expression of the hub gene in LUAD. In addition, the Gene Expression Profiling Interactive Analysis (GEPIA) database (<http://gepia.cancer-pku.cn/>) was used to verify the progress and prognosis of candidate hub genes [4].

2.7. Functional enrichment analysis

The Database for Annotation, Visualization and Integrate Discovery (DAVID) database [21] (<http://david.abcc.ncifcrf.gov/>) is an online program that provides a comprehensive set of functional annotation tools for investigators to understand biological meaning behind large list of genes. Hub genes in hub module were uploaded to screen enriched Gene Ontology (GO) terms by using DAVID database. $p < 0.05$ and $FDR < 0.01$ was set as the cut-off criteria.

2.8. Gene set enrichment analysis (GSEA)

194 samples from the training set were divided into two groups (high and low) based on the expression level of KIF2C and the median was used as the cut-off point. In order to figure out potential function of KIF2C, GSEA (<http://software.broadinstitute.org/gsea/index.jsp>) [22,23] was performed between the two groups. Annotated gene sets c2.cp.kegg.v6.0.symbols.gmt was chosen as the reference gene sets. $FDR < 0.05$, $|\text{enrichment score (ES)}| > 0.5$ and gene size ≥ 35 were set as the cut-off criteria.

3. Result

3.1. Differentially expressed genes screening

After excluding the abnormalities, the gene profile of 246 samples was analyzed (Fig. 2). Under the screening conditions of $FDR < 0.05$ and $|\log_2FC| > 1.5$, a total of 965 differential genes (572 up-regulated, 393 down-regulated LUAD) were selected for follow-up analysis.

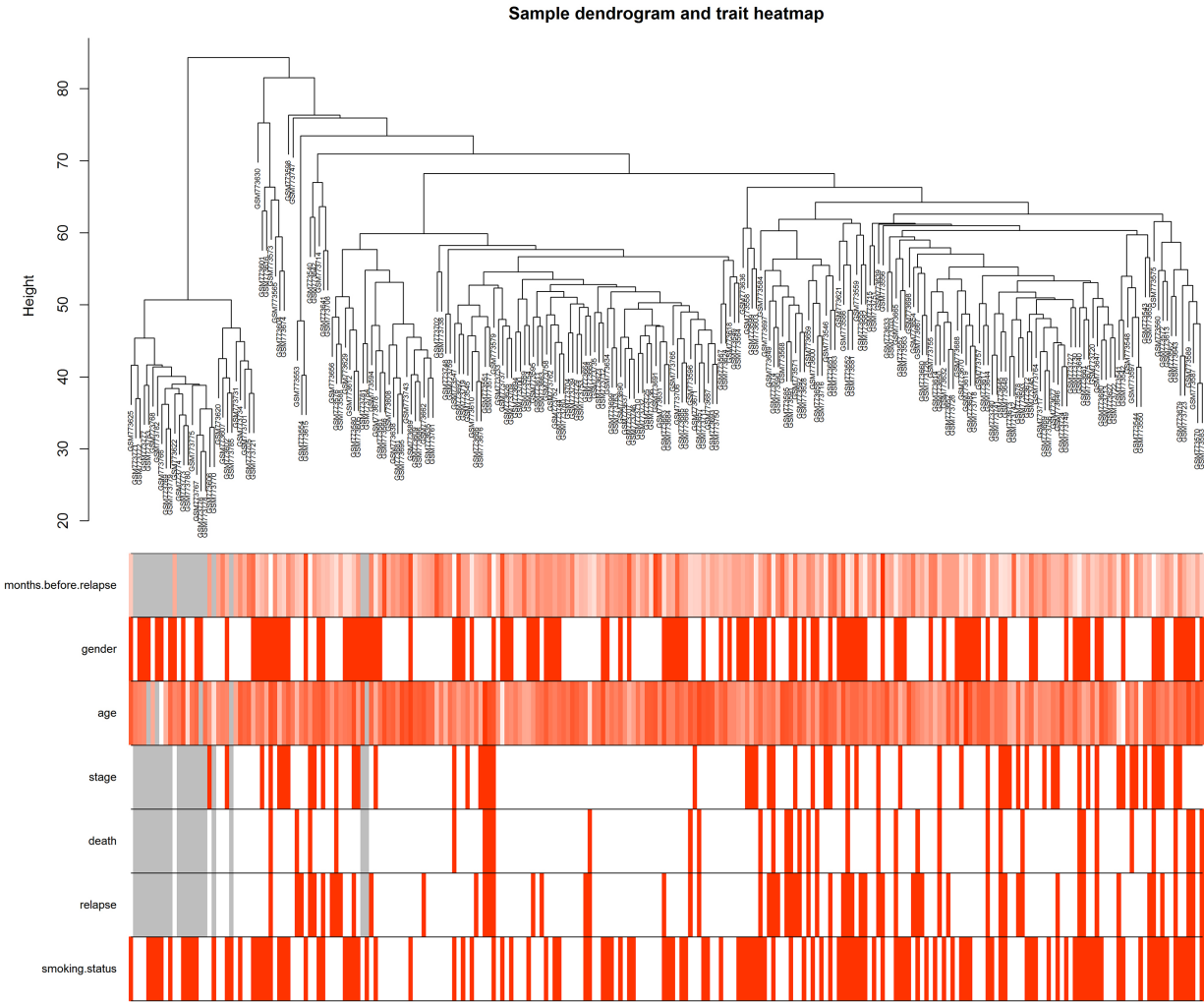


Fig. 2. Clustering dendrogram of GSE31210 and the clinical traits. The clustering was based on the expression data of differentially expressed genes between tumor samples and non-tumor samples in LUAD. The red color represented relapse and female. The color intensity was proportional to older age as well as stage, smoking status.

3.2. Weighted co-expression network construction and key modules identification

We used “WGCNA” R package to put the DEGs with similar expression patterns into modules by average linkage clustering. In current study, the power of $\beta = 5$ (scale free $R^2 = 0.87$) was selected as the soft-thresholding to ensure a scale-free network (Fig. 3), and a total of 3 modules were identified (Fig. 4A). Two methods were used to test the relevance between each module and the clinical information. We found that the ME in the brown module showed a higher correlation with LUAD stage than other modules (Fig. 4B). As modules with greater MS were considered to have more connection with our concerned clinical informa-

tion, we found that the MS of brown module was also higher than those of any other MS (Fig. 4C). Based on the two methods, we identified the brown module was the module most relevant to the disease progression of LUAD.

3.3. Hub gene identification

Defined by module connectivity, measured by absolute value of the Pearson’s correlation ($\text{cor.geneModule Membership} > 0.8$) and clinical trait relationship, measured by absolute value of the Pearson’s correlation ($\text{cor.geneTraitSignificance} > 0.2$), 54 genes with the high connectivity in brown module were taken as hub genes (Fig. 5A–C). In addition, we also constructed

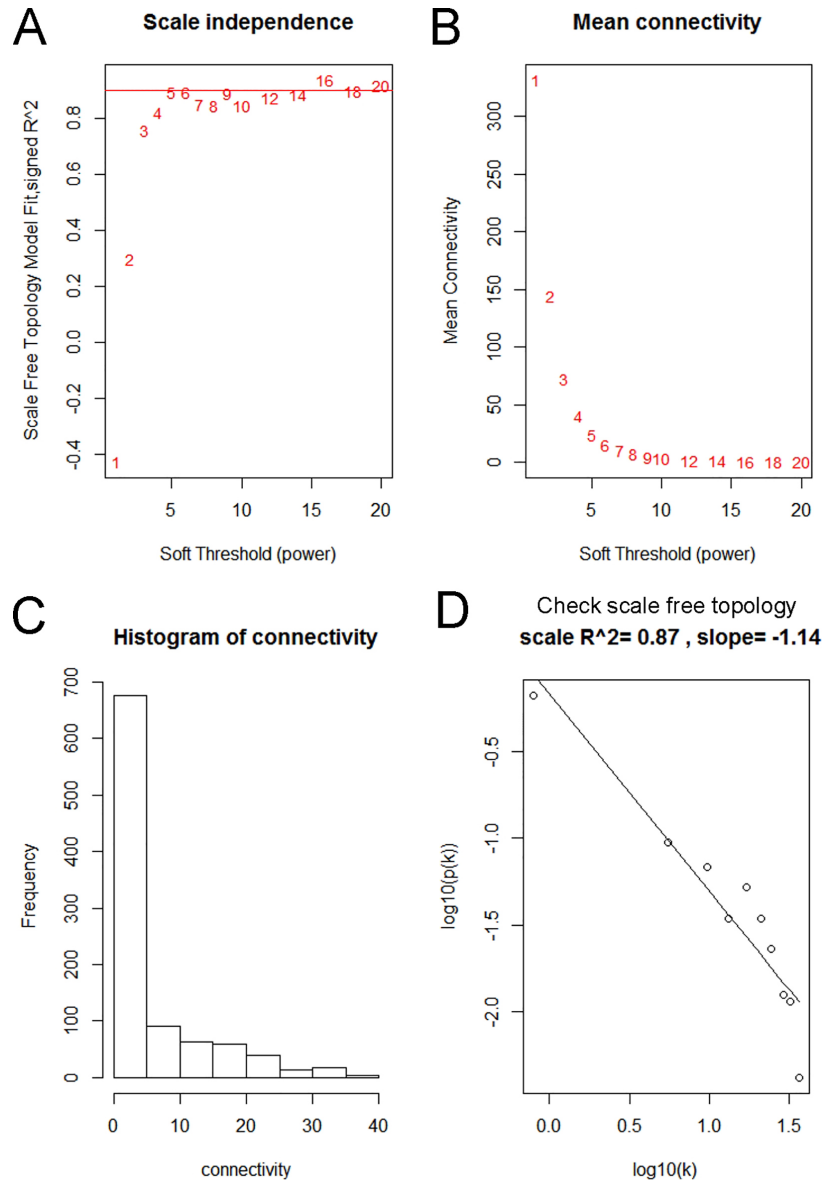


Fig. 3. Determination of soft-thresholding power in the weighted gene co-expression network analysis (WGCNA). (A) Analysis of the scale-free fit index for various soft-thresholding powers (β). (B) Analysis of the mean connectivity for various soft thresholding powers. (C) Histogram of connectivity distribution when $\beta = 5$. (D) Checking the scale free topology when $\beta = 5$.

a PPI network for all genes in the brown module by Cytoscape [24] based on the STRING database (<https://string-db.org/>) (Fig. 5B). And six genes with a degree of centrality over 35 were identified as hub genes in the PPI network. We found that all six genes were contained in 54 genes with high connectivity (Table 1). The six consensus hub genes (KIF2C, CCNB2, CDC20, TOP2A, CCNB1 and CCNA2) in the co-expression and PPI network are considered “real” hub genes for further validation.

3.4. Hub gene validation

The six hub genes (KIF2C, CCNB2, CDC20, TOP2A, CCNB1 and CCNA2) were selected for validation. In training set GSE31210, the association of KIF2C with disease stage ($r = 0.955$) was higher than the other five genes ($r = 0.947, 0.937, 0.931, 0.931$ and 0.885). Therefore, KIF2C was chosen as a candidate gene for further validation. In test set GSE40791, a linear regression analysis was performed

Table 1
Hub genes in the module related with pathological stage

Symbols	Probe	cor.geneModuleMembership	Hub gene in PPI network	DEG analysis	
				logFC	FDR
KIF2C	209408_at	0.955121	Yes	1.771095	1.50E-10
CCNB2	202705_at	0.947118	Yes	1.976688	2.55E-09
CDC20	202870_s_at	0.937495	Yes	1.625766	1.39E-07
TOP2A	201291_s_at	0.931434	Yes	2.621412	1.51E-13
CCNB1	214710_s_at	0.931292	Yes	1.563052	1.54E-08
CCNA2	203418_at	0.885054	Yes	2.120177	5.61E-12

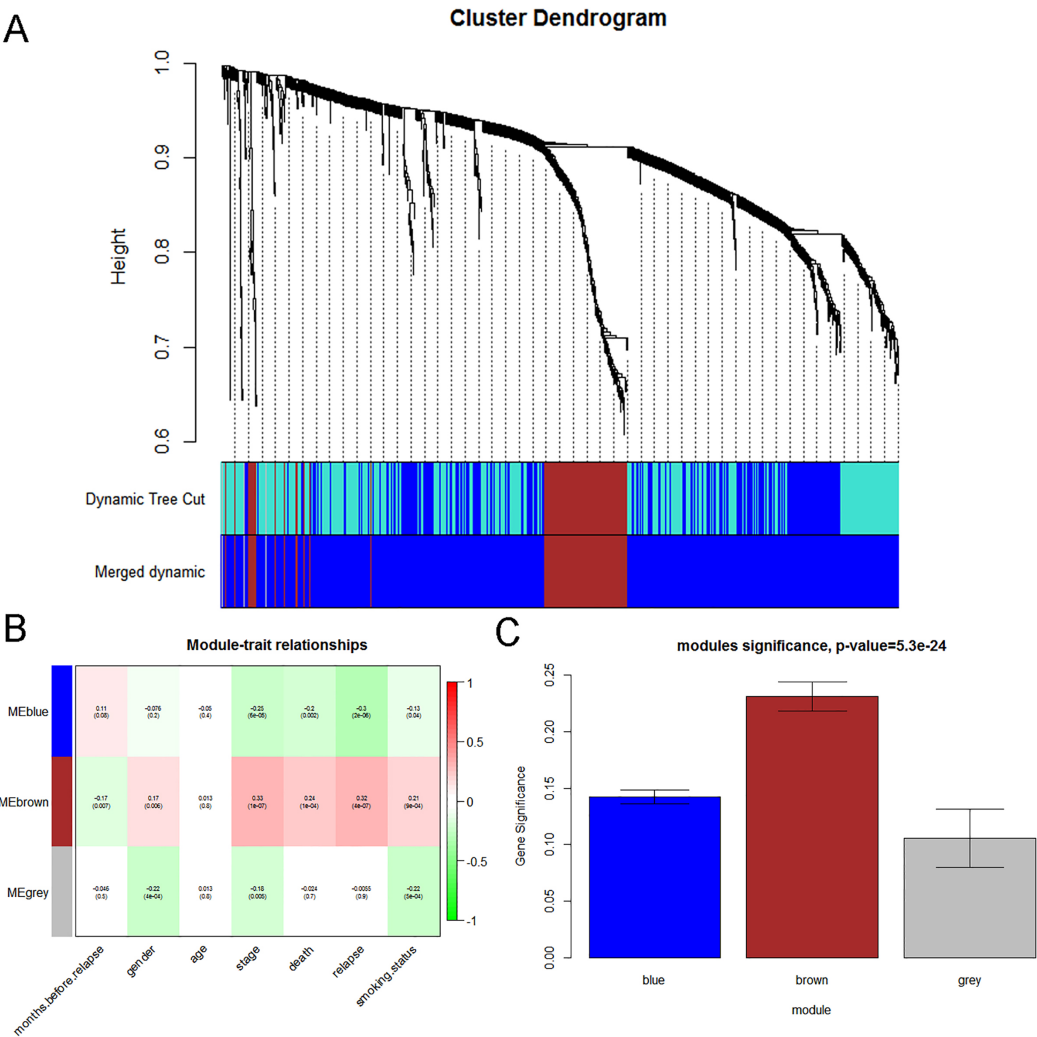


Fig. 4. Identification of modules associated with clinical information. (A) Dendrogram of all differentially expressed genes clustered based on a dissimilarity measure (1-TOM). (B) Heatmap of the correlation between module eigengenes and different clinical information of LUAD (Month before relapse, Gend, Age, Stage, Death, Relapse and Smoking status). (C) Distribution of average gene significance and errors in the modules associated with the progression of LUAD.

to validate KIF2C, showing a strong positive correlation between KIF2C expression and LUAD stage ($p < 0.001$, Fig. 6A). Based on the Oncomine database, we found that KIF2C was highly expressed in LUAD com-

pared to normal lung tissue (Fig. 6D–F). At the same time, KIF2C expression was significantly increased in LUAD compared to normal lung tissue based on the test set GSE40791 and TCGA data (Fig. 6B and C).

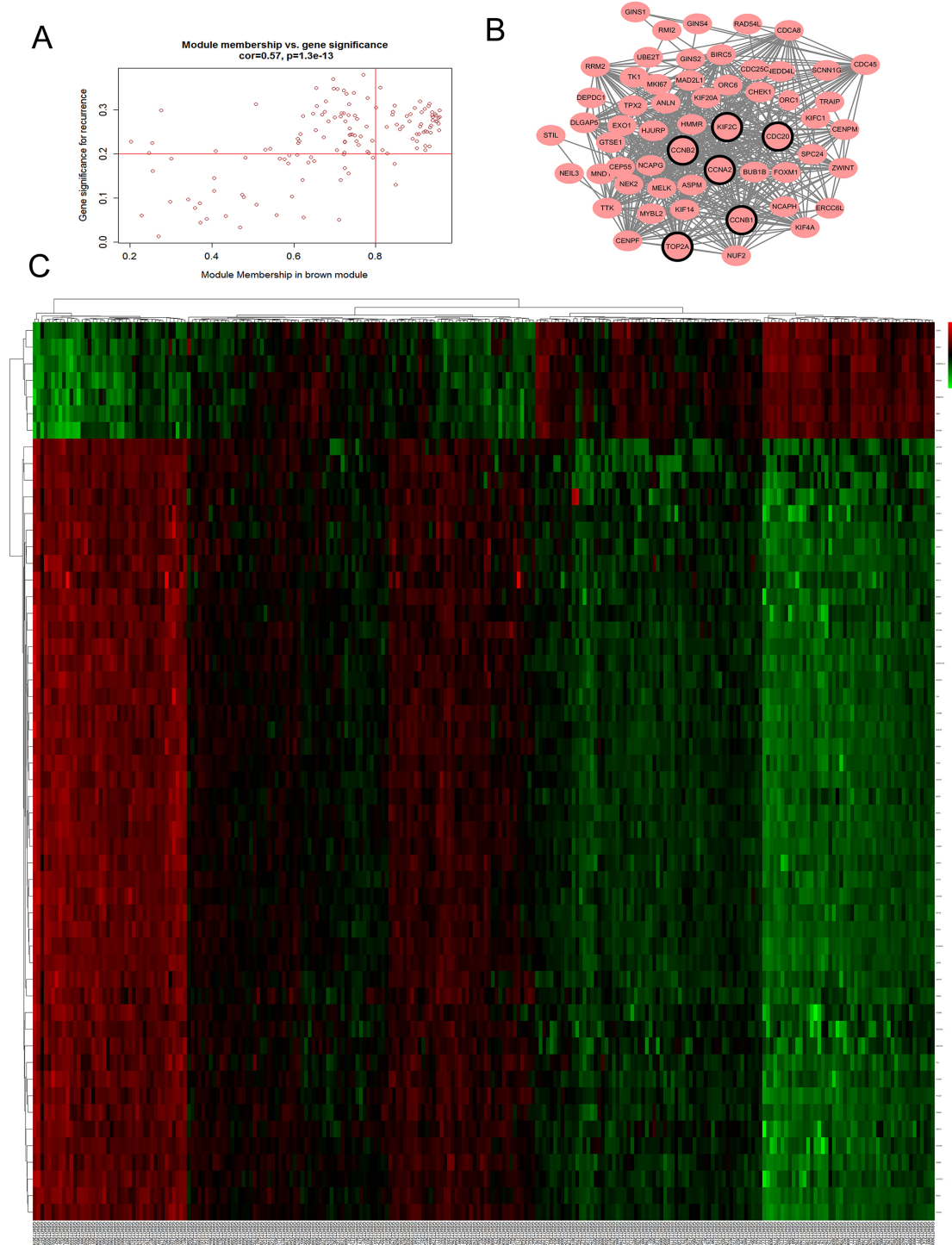


Fig. 5. Hub genes detection and protein-protein network. (A) Scatter plot of module eigengenes in brown module. (B) Protein-protein interaction (PPI) network of genes in the brown module. The positive correlation in red and negative correlation in green, in our PPI network, there's no negative correlation node. The nodes with bold circle represented network hub genes identified by degree>35 in PPI. (C) Heatmap of the expression of hub genes in different stages of LUAD.

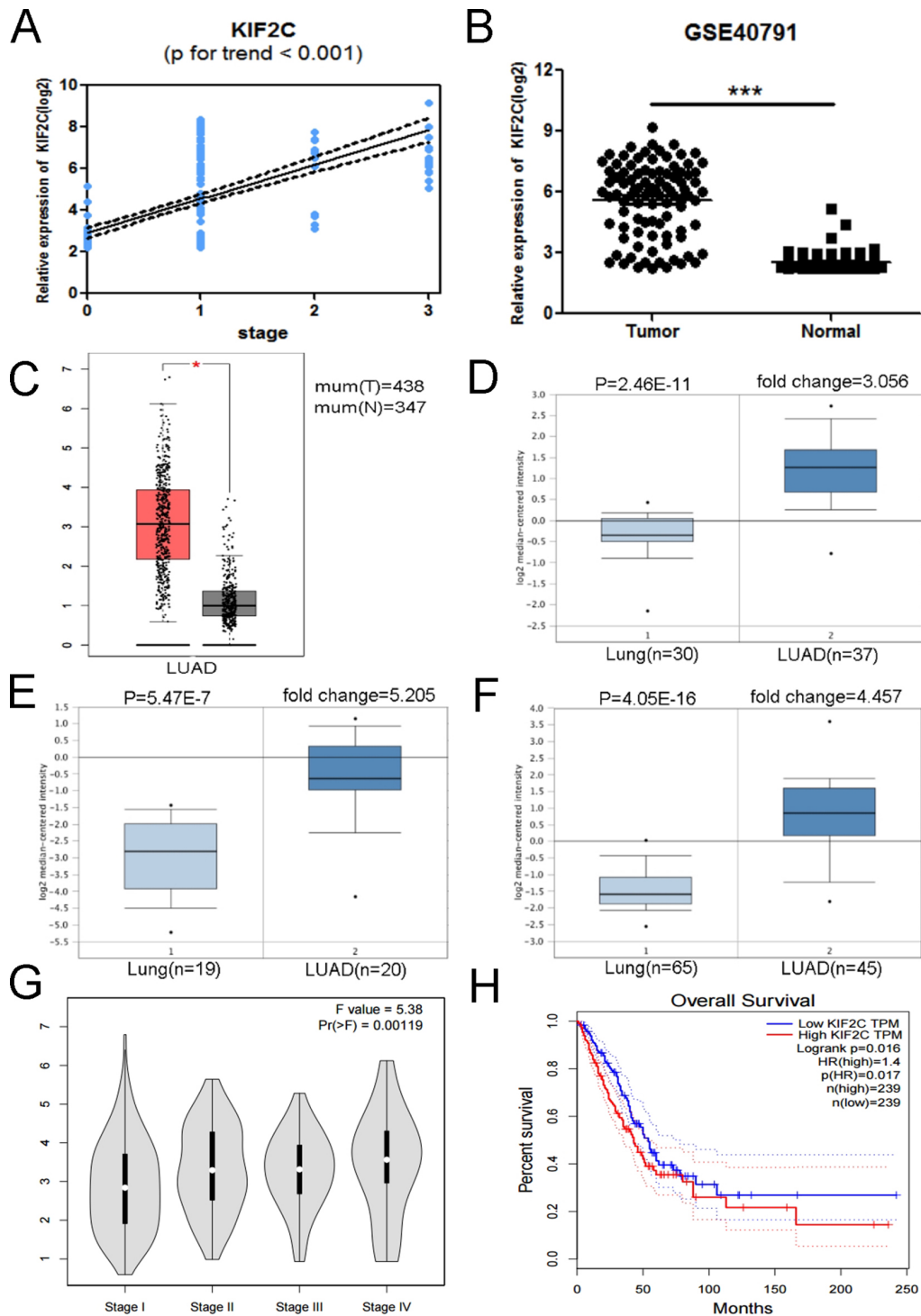


Fig. 6. Validation of KIF2C. (A) KIF2C expression was correlated with the disease progression of LUAD (GSE40791). (B) KIF2C mRNA was validated using GSE40791 ($p < 0.01$). (C) Expression level of KIF2C in cancer and normal tissues. LUAD: Lung Adenocarcinoma; $*p < 0.05$. (D-F) Oncomine database indicated that COL3A1 was up-regulated in LUAD, compared with normal lung tissue. (G) GEPIA database indicated that KIF2C had a strong correlation with the progression of LUAD based on TCGA data. (H) Kaplan-Meier survival curve obtained GEPIA database revealed that LUAD patients with higher expression of KIF2C had a significantly shorter overall survival time.

Table 2
Correlations between KIF2C expression and clinicopathological parameters in lung adenocarcinoma

Parameter	Category	Expression of KIF2C		p value
		Low	High	
Gender	Female	71	59	0.125
	Male	52	64	
Age	< 65	85	92	0.627
	≥ 65	33	31	
Smoking status	Ever-smoker	53	71	0.022
	Never-smoker	70	52	
Stage	I	93	75	< 0.001
	II	10	48	
Months before relapse		7.42 ± 0.56	9.22 ± 0.83	< 0.001
Relapse	Not relapsed	89	73	< 0.001
	Relapsed	14	50	
Death	Alive	95	96	0.003
	Dead	8	27	

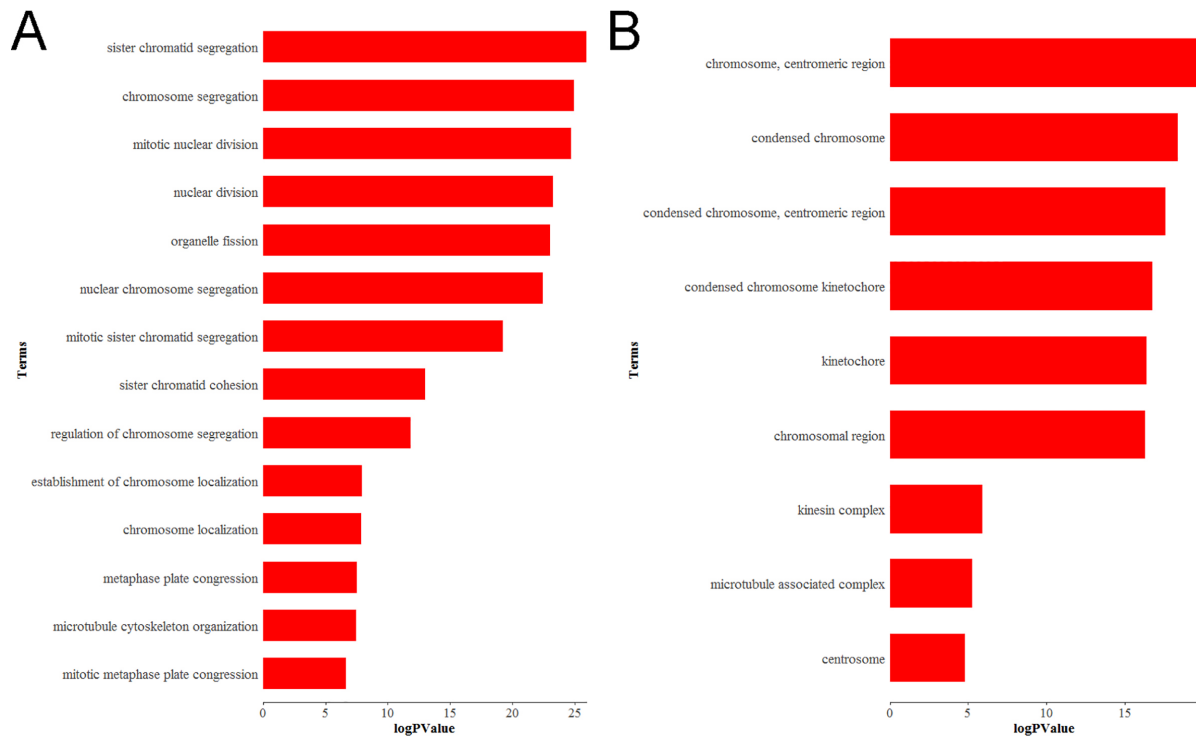


Fig. 7. Bioinformatics analysis of Differentially expressed genes (DEGs). (A) GO biological process (BP) and (B) cellular component (CC) enrichment of KIF2C.

And according to the GEPIA database, we can find significant differences between different stages of LUAD (Fig. 6G). In addition, we found that the overall survival of patients with KIF2C overexpression was significantly shorter (Fig. 6H). Comparing different clinicopathological parameters, we found that KIF2C has significant changes in different stages of LUAD; patients with high KIF2C expression have a history of smoking, more cancer recurrence, and a shorter survival (Table 2).

3.5. Functional enrichment analysis

To learn more about the function of DEGs in the hub module, they were uploaded to the DAVID database (<http://david.abcc.ncifcrf.gov/>). GO analysis revealed that KIF2C is enriched in 14 biological processes and 9 cellular components (Fig. 7, Supplementary Table S1), of which KIF2C is significantly enriched in biological processes such as chromosomal segregation and mitotic mitosis.

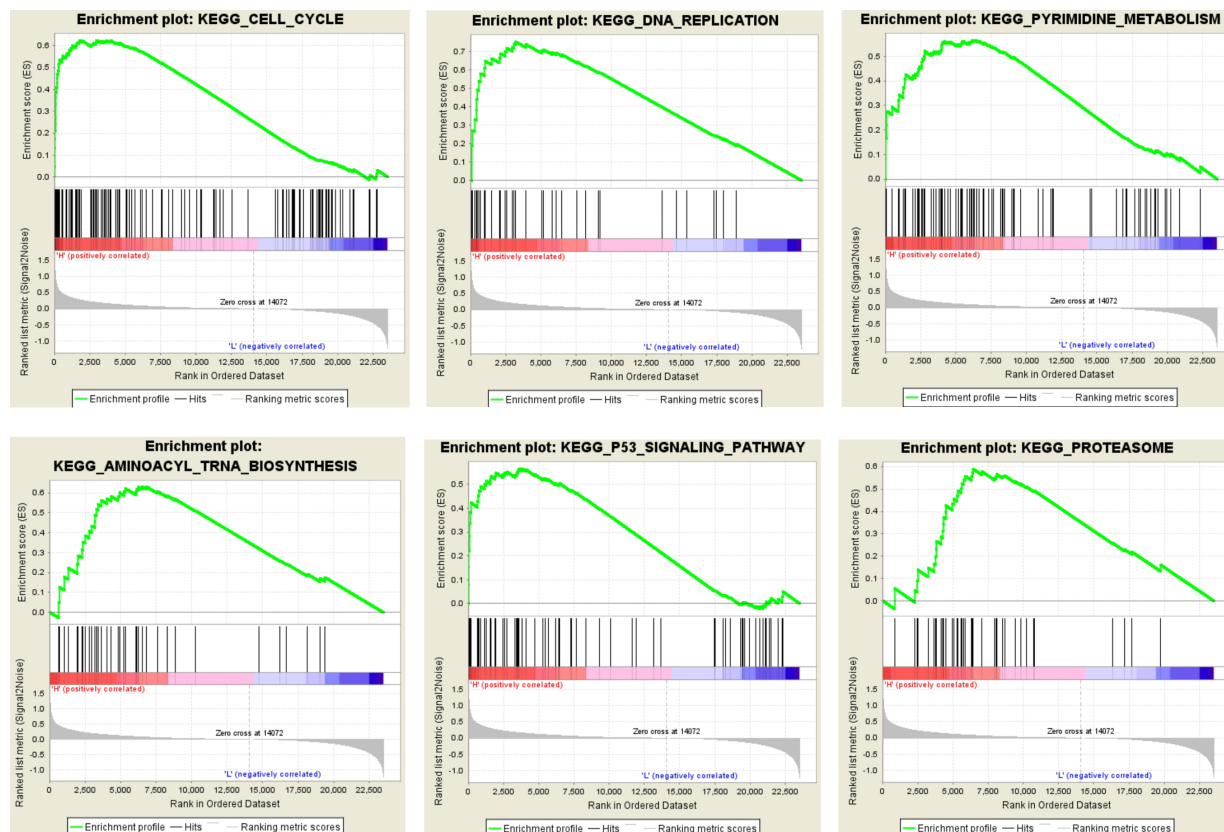


Fig. 8. Gene set enrichment analysis (GSEA) analysis for gene sets related with KIF2C expression. The gene sets of "Cell cycle", "DNA replication", "pyrimidine metabolism", "aminoacyl trna biosynthesis", "p53 signaling pathway" and "proteasome" were enriched in LUAD samples with KIF2C highly expressed (GSE40791).

3.6. Gene set enrichment analysis (GSEA)

In order to identify the potential function of hub genes, GSEA was used to retrieve enrichment of KIF2C-overexpressed samples on the KEGG pathway. Under the cut-off criteria, $FDR < 0.05$, $|\text{enrichment score (ES)}| > 0.5$ and gene size ≥ 35 , a total of 6 functional gene sets were enriched (Fig. 8, Supplementary Table S2). The six representative pathways are "cell cycle", "DNA replication", "pyrimidine metabolism", "aminoacyl trna biosynthesis", "p53 signaling pathway" and "proteasome".

4. Discussions

Lung adenocarcinoma is the most common malignancy in NSCLC. Because the early symptoms of LUAD are mild, the majority of patients are already diagnosed at late stage [25]. Radical resection is the only possible treatment for LUAD patients to obtain

basic cure to restore normal life. The 5-year survival rate of LUAD patients with IA to IIIA stage is approximately 23% to 67% after surgical [26]. However, approximately three-fourths of NSCLC are diagnosed at late stage, and have already lost the opportunity for surgery [27]. Tumor stage as an important clinical indicator is closely related to the malignancy and metastasis of the tumor. There are currently no tumor markers that can accurately predict the progression of LUAD. Therefore, there is an urgent need for effective biomarkers for the prediction of tumor stage and prognosis of LUAD.

In this study, WGCNA was used to identify gene co-expression module (brown module) which was associated with LUAD progression. In addition, we found that KIF2C, CCNB2, CDC20, CCNA2, CCNB1 and TOP2A are the common hub nodes of the co-expression module and PPI network, indicating that these hub genes are closely related to clinical features and biological processes. In validation, KIF2C was

highly correlated with the clinical trait estimated by log rank test (Fig. 6A).

KIF2C, the mitotic centromere associated kinesin, is the most representative member of Kinesin-13. It participates in microtubule disassembly, bipolar spindle formation, and chromosome segregation to regulate mitosis and cell cycle [28]. KIF2C is important for normal spindle assembly to repair microtubules and chromosomal abnormalities. Therefore, in theory, the abnormal expression of KIF2C may play an important role in the development of cancer. The activity and localization of KIF2C is regulated by many kinases and phosphatases [29,30], so that deregulation of mitogenic kinases and phosphatases can lead to mitotic anomalies, chromosomal instability and uncontrolled transcription. Some scholars found that the expression of KIF2C protein in tongue cancer tissues was significantly higher than that in paracancerous tissues, and the high expression of KIF2C was correlated with lymph node metastasis and tumor stage [31]. Our study also found that high expression of KIF2C was correlated with tumor stage in patients with LUAD ($p < 0.001$, Fig. 6G, Table 2), suggesting that the high expression of KIF2C may be related with tumor stage. With compared the expression levels of KIF2C in different types of cancer, we found that KIF2C was overexpressed in gastric cancer, breast cancer and rectal cancer [32]. This study found that KIF2C was highly expressed in LUAD tissues compared to normal lung tissue (Fig. 6B). These findings suggest that KIF2C may play an important role in cell mitosis and is an essential factor in spindle assembly and chromosome segregation. And the overexpression of KIF2C is highly likely to promote tumorigenesis. KIF2C, a cytoskeletal motor protein, is not only involved in the division and proliferation of LUAD cells, but also closely related to the distant metastasis of tumor cells. Therefore, KIF2C can be regarded as a protooncogene.

As oncogenes, based on the Oncomine and GEPIA database, we found that KIF2C expression was significantly increased in LUAD compared to normal lung tissue (Fig. 6C–F). We found that the expression of KIF2C was increased significantly with the progression of LUAD based on the test set GSE40791 and the GEPIA database (Fig. 6A and G), so we speculate that KIF2C plays an important role in the progression of LUAD and can be used as a candidate biomarker to distinguish different LUAD stage. As for the prognosis, we observed that higher expression of KIF2C resulted in shorter overall survival (Fig. 6H) and lower survival ($p = 0.003$, Table 2). Thus, there is

a significant relationship between the overexpression of KIF2C and the poor prognosis in LUAD patients, and KIF2C can be used as an important prognostic factor. Taking into account functional enrichment analysis, KIF2C performed better on chromosome segregation and mitotic mitosis (Fig. 7). Meanwhile, GSEA analysis showed that the overexpression of KIF2C of LUAD samples were significantly enriched in the cell cycle and DNA replication signaling pathways (Fig. 8), suggesting that KIF2C may affect cell mitosis and proliferation through cell cycle signaling pathways. Therefore, we believe that KIF2C plays an important role in the progression of LUAD, which may affect the prognosis by regulating the cell cycle signaling pathway, leading to poor prognosis of LUAD. More interestingly, based on the correlation between KIF2C expression and clinicopathological parameters of LUAD, we found that high expression of KIF2C may lead to the progression of LUAD and higher recurrence rate ($p < 0.001$, Table 2).

In summary, we constructed a co-expression network using weighted co-expression analysis, identify and validate network hub genes associated with the progression and poor prognosis of LUAD. Finally, KIF2C was identified as involved in the progression of LUAD and the development of poor prognosis by regulating cell cycle signaling pathways.

Supplementary data

The supplementary files are available to download from <http://dx.doi.org/10.3233/CBM-181512>.

References

- [1] K.Y. Chuang, Y.C. Chuang and Y.S. Ho, Global influence of Cancer Statistics articles, *Current Science* **109** (2015), 1552–1554.
- [2] C. Zappa and S.A. Mousa, Non-small cell lung cancer: current treatment and future advances, *Transl Lung Cancer Res* **5** (2016), 288–300.
- [3] K. Cetin, D.S. Ettinger, Y.J. Hei and C.D. O'Malley, Survival by histologic subtype in stage IV nonsmall cell lung cancer based on data from the Surveillance, Epidemiology and End Results Program, *Clin Epidemiol* **3** (2011), 139–48.
- [4] Z.F. Tang, C.W. Li, B.X. Kang, G. Gao, C. Li and Z.M. Zhang, GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses, *Nucleic Acids Research* **45** (2017), W98–W102.
- [5] M. Yamauchi, R. Yamaguchi, A. Nakata, T. Kohno, M. Nagasaki, T. Shimamura, S. Imoto, A. Saito, K. Ueno, Y. Hatanaka, R. Yoshida, T. Higuchi, M. Nomura, D.G. Beer, J. Yokota, S. Miyano and N. Gotoh, Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma, *PLoS One* **7** (2012), e43923.

- [6] J. Xiao, X. Lu, X. Chen, Y. Zou, A. Liu, W. Li, B. He, S. He and Q. Chen, Eight potential biomarkers for distinguishing between lung adenocarcinoma and squamous cell carcinoma, *Oncotarget* **8** (2017), 71759–71771.
- [7] J. Botling, K. Edlund, M. Lohr, B. Hellwig, L. Holmberg, M. Lambe, A. Berglund, S. Ekman, M. Bergqvist, F. Ponten, A. König, O. Fernandes, M. Karlsson, G. Helenius, C. Karlsson, J. Rahnenführer, J.G. Hengstler and P. Micke, Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis, and tissue microarray validation, *Clin Cancer Res* **19** (2013), 194–204.
- [8] S. Tavazoie, J.D. Hughes, M.J. Campbell, R.J. Cho and G.M. Church, Systematic determination of genetic network architecture, *Nat Genet* **22** (1999), 281–5.
- [9] P. Langfelder and S. Horvath, WGCNA: an R package for weighted correlation network analysis, *BMC Bioinformatics* **9** (2008), 559.
- [10] Z. He, M. Sun, Y. Ke, R. Lin, Y. Xiao, S. Zhou, H. Zhao, Y. Wang, F. Zhou and Y. Zhou, Identifying biomarkers of papillary renal cell carcinoma associated with pathological stage by weighted gene co-expression network analysis, *Oncotarget* **8** (2017), 27904–27914.
- [11] L. Yuan, L. Chen, K. Qian, G. Qian, C.L. Wu, X. Wang and Y. Xiao, Co-expression network analysis identified six hub genes in association with progression and prognosis in human clear cell renal cell carcinoma (ccRCC), *Genom Data* **14** (2017), 132–140.
- [12] W.C. Chou, A.L. Cheng, M. Brotto and C.Y. Chuang, Visual gene-network analysis reveals the cancer gene co-expression in human endometrial cancer, *BMC Genomics* **15** (2014), 300.
- [13] H. Okayama, T. Kohno, Y. Ishii, Y. Shimada, K. Shiraishi, R. Iwakawa, K. Furuta, K. Tsuta, T. Shibata, S. Yamamoto, S. Watanabe, H. Sakamoto, K. Kumamoto, S. Takenoshita, N. Gotoh, H. Mizuno, A. Sarai, S. Kawano, R. Yamaguchi, S. Miyano and J. Yokota, Identification of Genes Upregulated in ALK-Positive and EGFR/KRAS/ALK-Negative Lung Adenocarcinomas, *Cancer Research* **72** (2012), 100–111.
- [14] Y. Zhang, O. Foreman, D.A. Wigle, F. Kosari, G. Vasmatazis, J.L. Salisbury, J. van Deursen and P.J. Galaray, USP44 regulates centrosome positioning to prevent aneuploidy and suppress tumorigenesis, *Journal of Clinical Investigation* **122** (2012), 4362–4374.
- [15] L. Gautier, L. Cope, B.M. Bolstad and R.A. Irizarry, Affy – analysis of Affymetrix GeneChip data at the probe level, *Bioinformatics* **20** (2004), 307–315.
- [16] M.E. Ritchie, B. Phipson, D. Wu, Y.F. Hu, C.W. Law, W. Shi and G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Research* **43** (2015).
- [17] S. Horvath and J. Dong, Geometric Interpretation of Gene Co-expression Network Analysis, *Plos Computational Biology* **4** (2008).
- [18] M.J. Mason, G.P. Fan, K. Plath, Q. Zhou and S. Horvath, Signed weighted gene co-expression network analysis of transcriptional regulation in murine embryonic stem cells, *Bmc Genomics* **10** (2009).
- [19] A.M. Yip and S. Horvath, Gene network interconnectedness and the generalized topological overlap measure, *BMC Bioinformatics* **8** (2007), 22.
- [20] 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 and C. von Mering, STRING v10: protein-protein interaction networks, integrated over the tree of life, *Nucleic Acids Research* **43** (2015), D447–D452.
- [21] G. Dennis, B.T. Sherman, D.A. Hosack, J. Yang, W. Gao, H.C. Lane and R.A. Lempicki, DAVID: Database for annotation, visualization, and integrated discovery, *Genome Biology* **4** (2003).
- [22] V.K. Mootha, C.M. Lindgren, K.F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstrale, E. Laurila, N. Houstis, M.J. Daly, N. Patterson, J.P. Mesirov, T.R. Golub, P. Tamayo, B. Spiegelman, E.S. Lander, J.N. Hirschhorn, D. Altshuler and L.C. Groop, PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes, *Nat Genet* **34** (2003), 267–73.
- [23] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander and J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, *Proc Natl Acad Sci U S A* **102** (2005), 15545–50.
- [24] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res* **13** (2003), 2498–504.
- [25] S. Peters, A.A. Adjei, C. Gridelli, M. Reck, K. Kerr, E. Felip and E.G.W. Group, Metastatic non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, *Ann Oncol* **23**(Suppl 7) (2012), vii56–64.
- [26] P.A. Russell, Z. Wainer, G.M. Wright, M. Daniels, M. Conron and R.A. Williams, Does Lung Adenocarcinoma Subtype Predict Patient Survival? A Clinicopathologic Study Based on the New International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society International Multidisciplinary Lung Adenocarcinoma Classification, *Journal of Thoracic Oncology* **6** (2011), 1496–1504.
- [27] Q.H. Zhou, Individualized Surgical Treatment for Locally Advanced Non-Small Cell Lung Cancer Based on Molecular Biomarkers., *Journal of Thoracic Oncology* **8** (2013), S284–S284.
- [28] L.Y. Zhang, H.Y. Shao, Y.J. Huang, F. Yan, Y.J. Chu, H. Hou, M. Zhu, C.H. Fu, F. Aikhionbare, G.W. Fang, X. Ding and X.B.A. Yao, PLK1 Phosphorylates Mitotic Centromere-associated Kinesin and Promotes Its Depolymerase Activity, *Journal of Biological Chemistry* **286** (2011), 3033–3046.
- [29] H. Shao, Y. Huang, L. Zhang, K. Yuan, Y. Chu, Z. Dou, C. Jin, M. Garcia-Barrio, X. Liu and X. Yao, Spatiotemporal dynamics of Aurora B-PLK1-MCAK signaling axis orchestrates kinetochore bi-orientation and faithful chromosome segregation, *Sci Rep* **5** (2015), 12204.
- [30] C.T. Friel and J. Howard, The kinesin-13 MCAK has an unconventional ATPase cycle adapted for microtubule depolymerization, *Embo Journal* **30** (2011), 3928–3939.
- [31] C.Q. Wang, F.G. Xiang, Y.J. Li, X.M. Xing, N. Wang, J.H. Chi and W.J. Yu, Relation between the expression of mitotic centromere-associated kinesin and the progression of squamous cell carcinoma of the tongue, *Oral Surgery Oral Medicine Oral Pathology Oral Radiology* **117** (2014), 353–360.
- [32] L. Bie, G. Zhao, Y.P. Wang and B. Zhang, Kinesin family member 2C (KIF2C/MCAK) is a novel marker for prognosis in human gliomas, *Clinical Neurology and Neurosurgery* **114** (2012), 356–360.