



OPEN

## Identification and verification of microtubule associated genes in lung adenocarcinoma

YuHui Wei, CaiZhen Yang, JinMei Wei, WenTao Li, YuanWen Qin & GuangNan Liu 

Associated with high morbidity and mortality, lung adenocarcinoma (LUAD) is lacking in effective prognostic prediction and treatment. As chemotherapy drugs commonly used in clinics, microtubule-targeting agents (MTAs) are limited by high toxicity and drug resistance. This research aimed to analyze the expression profile of microtubule-associated genes (MAGs) in LUAD and explore their therapy efficiency and impact on prognosis. Key MAGs were identified as novel molecular targets for targeting microtubules. The LUAD project in The Cancer Genome Atlas (TCGA) database was used to identify differently expressed MAGs. On the one hand, a microtubule-related prognostic signature was constructed and validated, and its links with clinical characteristics and the immune microenvironment were analyzed. On the other hand, hub MAGs were obtained by a protein–protein interaction (PPI) network. Following the expression of hub MAGs, patients with LUAD were classified into two molecular subtypes. A comparison was made of the differences in half-maximal drug inhibitory concentration (IC50) and tumor mutational burden (TMB) between groups. In addition, the influence of MAGs on the anticancer efficacy of different therapies was explored. MAGs, which were included in both the prognosis signature and hub genes, were considered to have great value in prognosis and targeted therapy. They were identified by quantitative real-time polymerase chain reaction (qRT-PCR). A total of 154 differently expressed MAGs were discovered. For one thing, a microtubule-related prognostic signature based on 14 MAGs was created and identified in an external validation cohort. The prognostic signature was used as an independent prognostic factor. For another, 45 hub MAGs were obtained. In accordance with the expression profile of 45 MAGs, patients with LUAD were divided into two subtypes. Distinct differences were observed in TMB and IC50 values of popular chemotherapy and targeted drugs between subtypes. Finally, five genes were included in both the prognosis signature and hub genes, and identified by qRT-PCR. A microtubule-related prognosis signature that can serve as an independent prognostic factor was constructed. Microtubule subtype influenced the efficacy of different treatments and could be used to guide therapy selection. In this research, five key MAGs, including MYB proto-oncogene like 2 (MYBL2), nucleolar and spindle-associated protein 1 (NUSAP1), kinesin family member 4A (KIF4A), KIF15 and KIF20A, were verified and identified. They are promising biomarkers and therapeutic targets in LUAD.

Lung cancer has become one leading cause of cancer-related mortality<sup>1</sup>, and its most common subtype is lung adenocarcinoma (LUAD). Tumor-node-metastasis (TNM)-based staging is the most widely used strategy for predicting survival in the clinic. However, the prognosis of lung cancer differs substantially among patients of the same stage. Nowadays, surgery, radiotherapy, chemotherapy, immunotherapy and molecular-targeted therapy are the main treatments for LUAD. Multiple therapies bring new opportunities to patients with LUAD. Nevertheless, valid markers are needed to guide treatment selection. In addition, blockades still exist, including drug resistance<sup>2</sup>, big side effects<sup>3</sup> and low response rates. Hence, effective biomarkers and therapeutic targets remain urgently needed by LUAD patients. It is necessary to further explore the molecular mechanisms of abnormal microtubule activities and provide evidence for the development of molecularly targeted drugs.

Microtubules which are hollow tubes assembled by tubulins play a crucial role in multiple cellular processes as parts of the cytoskeleton<sup>4</sup>. They have become one of the effective antitumor targets because tumor cell proliferation can be inhibited by interfering with microtubule assembly. Multiple microtubule-targeting agents (MTAs) have been synthesized and employed as first-line anticancer drugs like taxane and colchicine. They significantly prolong the survival of patients. The feasibility of microtubules as anti-tumor targets is unquestionable.

Department of Respiratory and Critical Care, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China. ✉email: gnliu63@hotmail.com

Many studies have been conducted on prognostic signatures like epithelial–mesenchymal transition (EMT)-related signatures<sup>5</sup> and hypoxia-related signatures<sup>6</sup>. However, very little research has investigated the value of microtubules in prognosis and guiding treatment.

In this research, the expression and clinical significance of microtubule-associated genes (MAGs) in LUAD were explored systematically. The gene-expression profiling and clinical features of 489 LUAD samples were obtained from The Cancer Genome Atlas (TCGA) database. MAGs were found from the gene set enrichment analysis (GSEA) database. Based on differently expressed MAGs, a microtubule-related prognostic signature was constructed and identified in a validation cohort. The prognostic value of MAGs in LUAD was highlighted. Unlike general prognosis signatures, the hub genes of differently expressed MAGs were further identified. According to the expression levels of hub MAGs, 489 patients were divided into two groups. A comparison was made of tumor mutational burden (TMB) and the IC50 values of popular chemotherapy and targeted drugs between the two groups. In addition, their relationships with the tumor immune microenvironment were analyzed. MAGs, which were included in both the prognosis signature and hub genes, were considered to be related to both prognosis and chemotherapy sensitivity in LUAD and served as potential targets. Finally, five MAGs were verified in total by quantitative real-time polymerase chain reaction (qRT-PCR).

The value of MAGs in prognosis and guiding treatment was assessed to identify and verify key MAGs. It will provide a foundation for the establishment of biomarkers and the development of molecular MTAs.

## Materials and methods

### Data collection and process

Ribonucleic acid (RNA) sequencing profiles (including both fragments per kilobase of exon model per million mapped fragments (FPKM) and COUNTS values), somatic variation and accompanying clinical data were collected from the TCGA-LUAD project<sup>7</sup> as training cohorts.

Transcript expression data were obtained from 594 samples and clinical information was obtained from 585 patients in total. After the exclusion of those without survival time, stage or metastasis, 489 LUAD samples with reasonably completed clinical information were selected for the next analysis. GSE50081, a validation set, was retrieved from the Gene Expression Omnibus (GEO) database<sup>8</sup> using the package TCGA biolinks<sup>9</sup>.

To find microtubule-related ontology gene sets, “microtubule” was utilized as a search phrase in the Molecular Signatures Database (MSigDB)<sup>10</sup>. All the genes engaged in these pathways were regarded as MAGs.

### Enrichment analysis of differently expressed genes

The DESeq2 package<sup>11</sup> was used for comparing expression profiles between tumor and normal samples. Genes with FC > 1.5 and p < 0.05 were viewed as differently expressed genes (DEGs). Up- and down-regulation genes were presented on the volcano map. GSEA was employed for the detection of differently activated biological processes using the clusterProfiler<sup>12</sup> package. As reference gene sets, the biological processes ontology derived from the gene ontology resource were employed. The pathways whose adjusted p-values are under 0.05 were regarded as differently activated pathways.

### Construction of microtubule-associated prognostic signature

Differently expressed MAGs were defined by taking the intersection of DEGs and MAGs. Univariate and multivariate Cox regressions were performed to determine independent prognostic MAGs. After univariate Cox regression analysis, differently expressed MAGs were subjected to performing multivariate Cox regression. Independent prognostic genes were defined as multivariate analysis with P < 0.05. After that, the least absolute shrinkage and selection operator (LASSO), a machine learning approach to selecting variables in multicollinear variables, was employed to squeeze genes. Finally, a Cox proportional hazards model was created. The above procedures were carried out by use of the R package glmnet<sup>13</sup>. To ensure repeatability, the random number seed was set as 3. The best lambda value was obtained by cross-validation 1,000 times. At last, the best genes for the construction of the microtubule-associated prognostic signature were chosen. The signature was constructed using the following formula:

$$\text{Riskscore} = \sum_{n=1}^{\infty} (\text{Coefficient}(i) * X_i)$$

The coefficient was presented from the weight of multivariate analysis. Time-dependent receiver operating characteristic (ROC) curves were utilized for evaluating the effectiveness of the signature and created by using the R package time ROC.

### Establishment and evaluation of a nomogram

For ease in clinical application, a nomogram combining multiple clinical indicators was established. The discrimination ability to predict survival outcomes was assessed by the concordance index (C index) and decision curve analysis (DCA). In addition, the calibration plots of 1, 3 and 5 years were created to testify uniformity power.

### Prognostic value of the signature

The model was used to derive and group the risk scores of each patient. The patients were separated into two groups based on their median risk scores. The Kaplan–Meier (KM) technique was used to assess the capability of categorizing death risks.

To further investigate the heterogeneity between the two groups, the differences in clinical features were compared. Clinical features included age, gender and stages T, N, M. T, N and M stand for original tumor diameter,

distant metastasis and lymph node metastasis, respectively. The information containing letters TX, MX and NX was deemed unclear and not taken into account in statistical analysis.

### Immune landscape analysis

Immunotherapy is an important part of anti-cancer therapy. This research was aimed at comparing the immune microenvironment to see if any variation existed in the immune landscape. Cibersort and xcell were chosen as prediction algorithms because of quantifying immune infiltration based on different principles. Cibersort is a deconvolution-based algorithm, whereas xcell calculates the proportions of cells based on marker genes by single sample GSEA (ssGSEA)<sup>14</sup>. The proportion of 22 common immune cells was quantified by Cibersort, and helper T cells were quantified by xcell. The aforesaid algorithms were deployed on TIMER2.0<sup>15</sup>, an open-access portal that integrates multiple immune infiltration techniques.

Immune checkpoint inhibitors (ICIs) have been utilized as one of the main treatments for LUAD with satisfactory effect. The expression levels of programmed death protein 1 (PD-1), programmed death ligand 1 (PD-L1) and lymphocyte activation gene-3 (LAG3) were examined as common immune checkpoints, to predict the response of ICIs.

### External validation of the signature

The signature was validated by selecting GEO datasets as external cohorts. In the validation cohort, risk scores were calculated, and the patients were separated into low- and high-score groups. The prognostic value of the signature was examined by performing KM analysis and time-dependent ROC curves. Immune cell infiltration was also measured and used to investigate its links with the immune microenvironment. The procedures of these analyses were the same as those in training cohorts.

### Protein–protein interaction network

The STRING database is an integrative database for the consolidation and prediction of protein interaction. Based on the database, the correlations of differently expressed MAGs were described and then displayed as a protein–protein interaction (PPI) network in Cytoscape. Molecular complex detection (MCODE) is a kind of strong software for locating key subnetworks in a huge interactive network according to the relationship between edges and nodes. Subnetworks were screened by MCODE, and the genes found there were designated as hub MAGs.

### Identification of subgroups using consensus clustering

Based on hub MAGs, unsupervised clustering was performed to identify subgroups using the R package ConsensusClusterPlus. The Partitioning Around Medoids (PAM) algorithm was used as a clustering algorithm, and sampling was performed 1000 times with a sub-sampling ratio of 0.8. The seed for the random number generator was 123,456. The optimal K value was obtained while the cumulative distribution function (CDF) reached a near-maximal value and remained constant. Heatmaps were visually inspected to identify the optimal number of gene clusters. T-distributed stochastic neighbor embedding (tSNE) analysis was applied, to ensure the accuracy of classification. Then, survival outcomes between clusters were compared by using survival analysis.

### Comparison of biological pathways and therapeutic effect

Difference-in-difference and GSEA analyses described above were performed to probe into the biological function of hub MAGs. Then, TMB and IC50 were used as indicators to predict diverse therapeutic effects. TMB was defined to be the number of somatic mutations in every 1 Mb exon. Patients with higher TMB are more likely to benefit from immunotherapy. The R package maftools were used for analyzing somatic mutations and calculating the TMB of each patient. The IC50 value was used to evaluate and compare the response to a variety of treatments. Lower IC50 indicated higher drug sensitivity. The IC50 of commonly used drugs including targeted and chemotherapy drugs was estimated using R package pRRophetic.

### Identification of differently expressed hub MAGs

After the intersection of DEGs and hub MAGs was taken, five differentially expressed hub MAGs were obtained: nucleolar and spindle-associated protein 1 (NUSAP1), KIF15, MYBL2, KIF20A and KIF4A. Multiple immunohistochemistry results of proteins in various tissues were found in the Human Protein Atlas (HPA) database, which provided a wealth of resources for protein-level gene verification. Matching protein expression in normal lung and LUAD was downloaded and compared to verify the expression of differently expressed hub MAGs in LUAD.

### Cell culture

A549 and BEAS-2B were purchased from NEWGAINBIO company and cultured in an incubator with 5% carbon dioxide (CO<sub>2</sub>) at 37 °C. Then, 1640 (Gibco, the United States of America (USA)) with 10% fetal bovine serum (FBS; BI, China) was used as the culture medium of A549 cells. Dulbecc's modified Eagle's medium (DMEM, Gibco, USA) with 10% FBS (BI, China) was used as the culture medium of BEAS-2B cells.

### qRT-PCR

After more than 80% cell fusion rates, the cells were treated with TRIzol Reagent (Takara, Japanese) to isolate total RNA. Next, complementary deoxyribonucleic acid (cDNA) was synthesized by use of the PrimeScript™ RT reagent Kit with genomic DNA (gDNA) Eraser (Takara, Japanese) once RNA purity was determined. On the StepOnePlus real-time PCR system (StepOnePlus; Applied Biosystems), 2× RealStar Fast SYBR qPCR Mix

(Genstar, China) was used to perform RT-qPCR.  $\beta$ -Actin (ACTB) was regarded as a reference gene to normalize gene expression. All specific primers were synthesized by Sangon Biotechnology (Shanghai, China).

### Statistical analysis

R 4.1.2 and GraphPad Prism version 8.0 (GraphPad Software, Inc., La Jolla, California (CA), USA) were used to perform all the above procedures. Wilcoxon rank-sum test was used to compare the differences between low- and high-score groups, including clinical characteristics, immune checkpoint expression, TMB and IC50. Survival curves were drawn by the KM analysis and then compared by the log-rank test. Student's t-test was used to evaluate the differences in the expression levels of MAGs between A549 and BEAS-2B. The inspection level was set as 0.05.

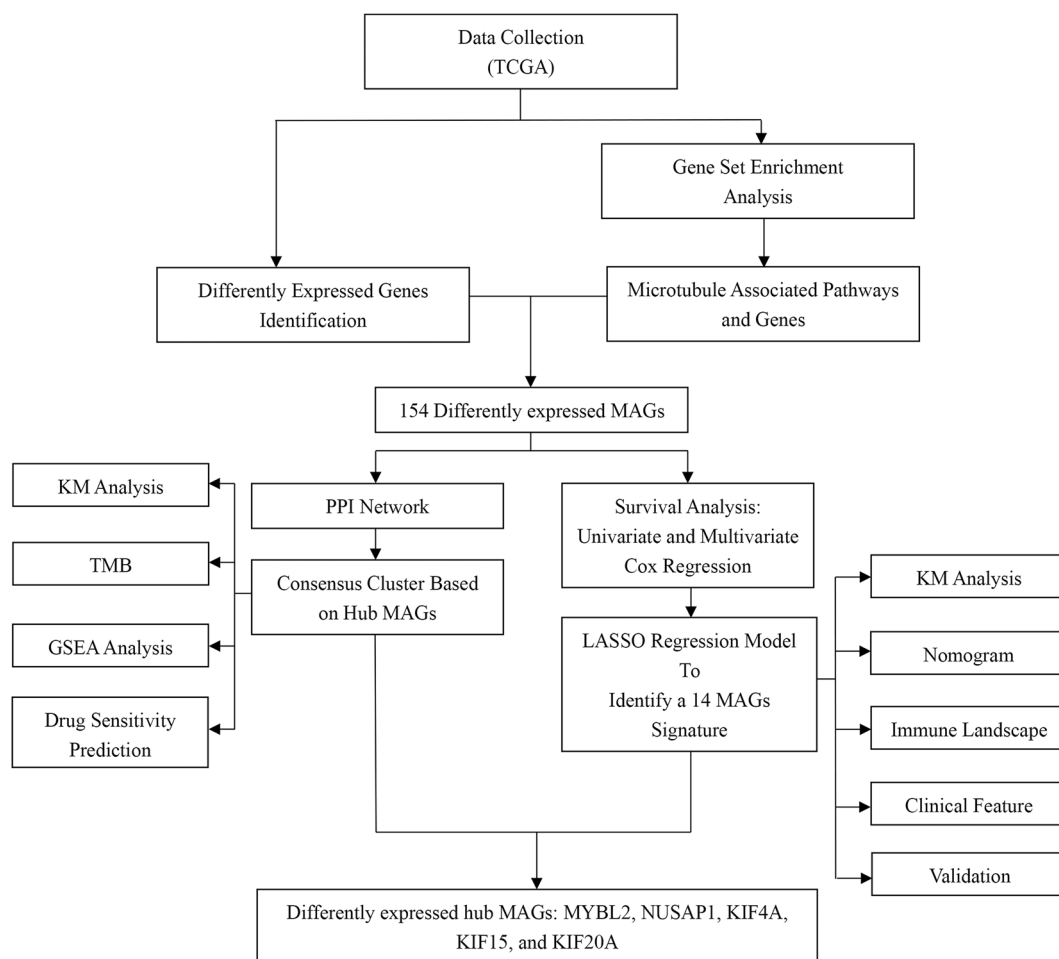
### Ethics approval and consent to participate

Both TCGA and GEO are public databases. Users can download relevant data for free to conduct research and post relevant articles. Based on open-source data, this study required no ethical approval.

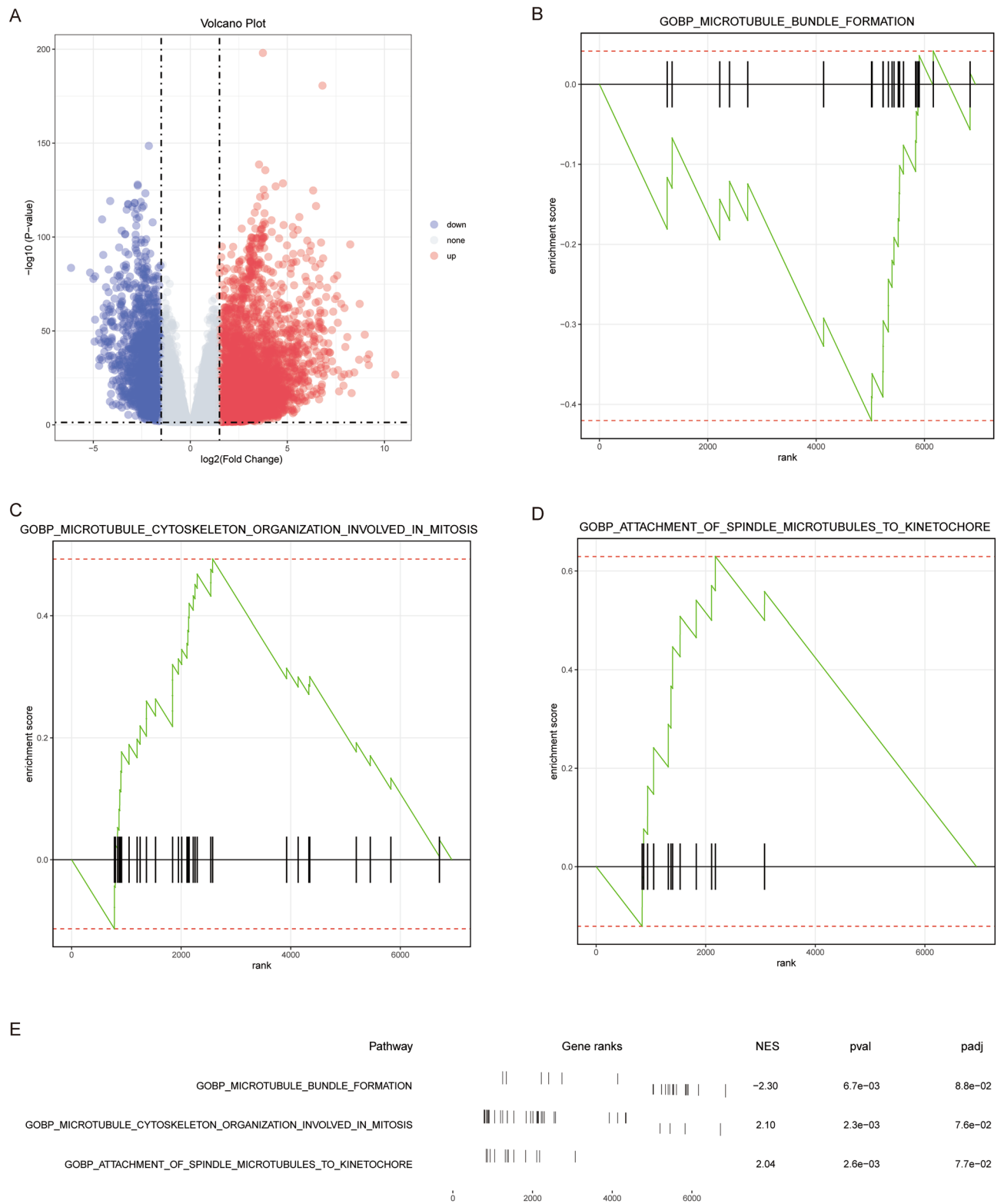
## Result

### Identification and enrichment analysis of DEGs

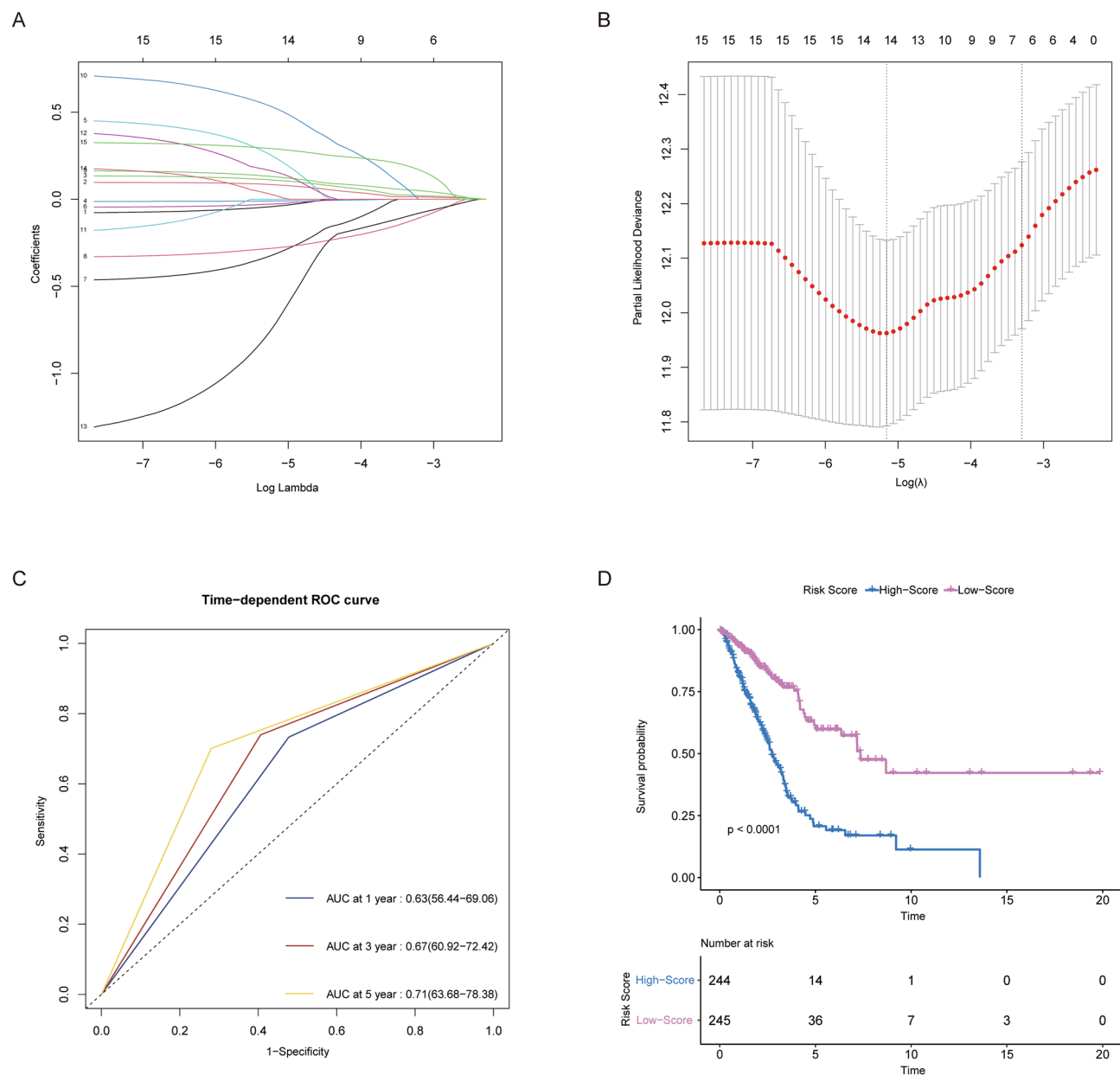
The diagram of the experiment workflow is presented in Fig. 1. After difference-in-difference analysis, 6940 genes ( $FC > 1.5$ ,  $P < 0.05$ ) were obtained. DEGs were displayed on the volcano map (Fig. 2A). Here, GSEA analysis was performed to explore the biological processes involved in tumorigenesis. Surprisingly, the GSEA analysis revealed three microtubule-related pathways as follows: GOBP\_MICROTUBULE\_BUNDLE\_FORMATION, GOBP\_MICROTUBULE\_CYTOSKELETON\_ORGANIZATION\_INVOLVED\_IN\_MITOSIS and GOBP\_ATTACHMENT\_OF\_SPINDLE\_MICROTUBULES\_TO\_KINETOCHORE. GOBP\_MICROTUBULE\_BUNDLE\_FORMATION was significantly enriched in normal samples, while others were significantly enriched in the tumor. Details of pathways are displayed in Fig. 2B–E. The role of microtubules in LUAD was highlighted and worthy of future investigation.



**Figure 1.** This is the workflow visualization of this research.



**Figure 2.** Genes and pathways differentially expressed in LUAD. **(A)** DEGs in the training cohort (n = 489). **(B–D)** Differently expressed microtubule-related pathways. **(E)** A merged plot of GSEA analysis. LUAD: lung adenocarcinoma; DEGs: differentially expressed genes; GSEA: gene set enrichment analysis.



**Figure 3.** Construction of microtubule-associated prognostic signature. (A,B) LASSO analysis discovered MAGs for constructing a signature. (C) Time-dependent ROC curves and AUC. (D) Results of the KM analysis of overall survival for low- and high-score groups. LASSO: the least absolute shrinkage and selection operator; MAGs: microtubule-associated genes; ROC: receiver operating characteristic; AUC: area under the curve; KM analysis: Kaplan–Meier analysis.

By searching in the MSigDB database, 30 microtubule-related ontology gene sets were discovered. After deduplication, 956 genes that participated in these pathways were considered MAGs. After the intersection of DEGs and MAGs was taken, 154 differently expressed MAGs were selected for further analysis.

### Construction of microtubule-associated prognostic signature

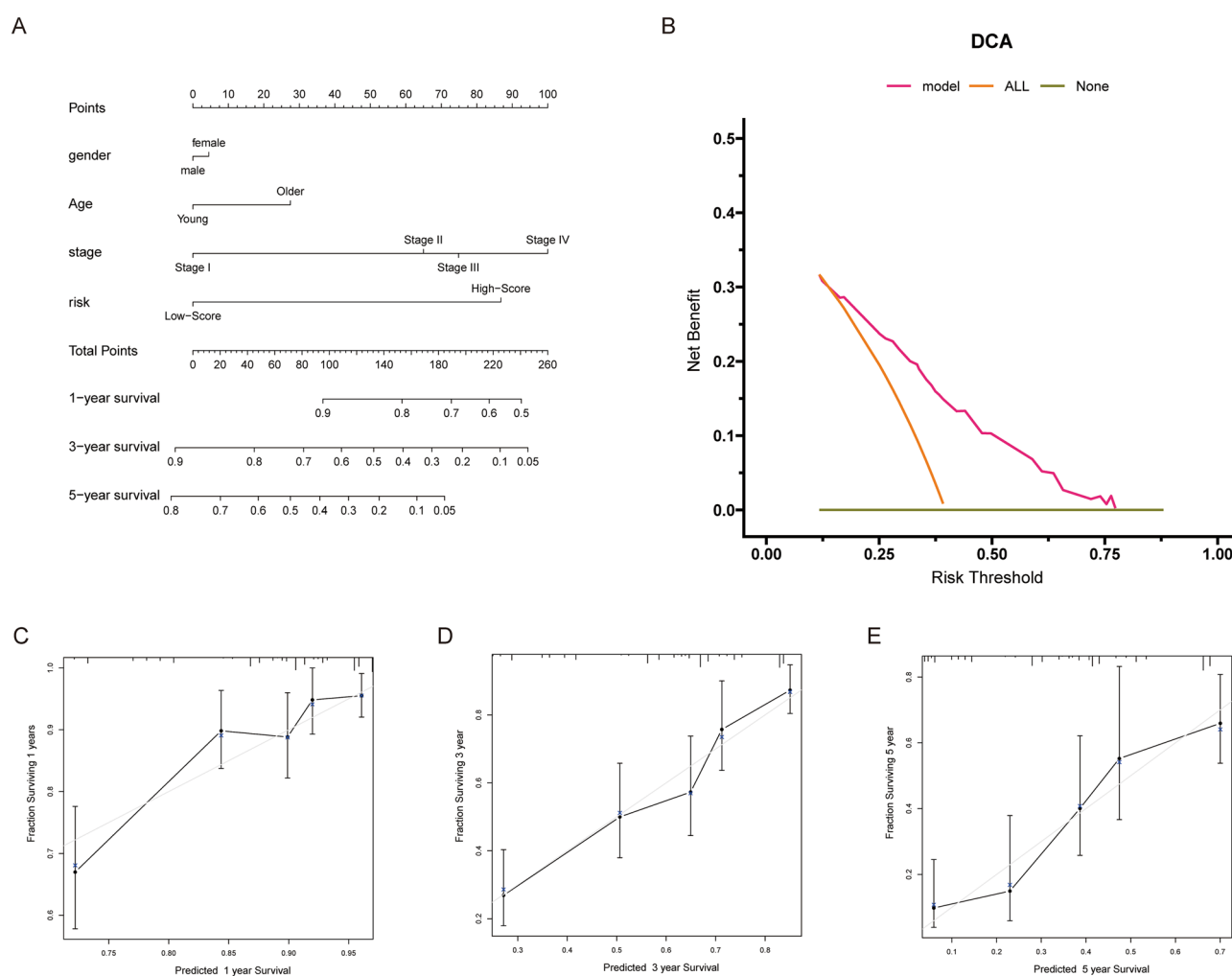
A total of 15 genes were identified as independent prognostic MAGs and selected for the building of prognostic models. Then, 14 genes were screened by LASSO regression to build a microtubule-associated risk signature (Fig. 3A,B). The formula is as follows:

$$\begin{aligned} \text{Riskscore} = & (-0.039459074148011 * \text{TUBB3}) + (0.0826175098637292 * \text{KIF4A}) \\ & + (0.109104945255747 * \text{KIF20A}) + (-0.0100318589132996 * \text{MYBL2}) \\ & + (0.23439864802325 * \text{SGO1}) + (-0.0262042220273366 * \text{NUSAP1}) \\ & + (-0.310922864156712 * \text{KIF15}) + (-0.278759908343141 * \text{NCKAP5}) \\ & + (0.128416494502307 * \text{CDK5R1}) + (0.519865071940523 * \text{FSIP2}) \\ & + (0.155416212889263 * \text{DNAI2}) + (-0.703695256275053 * \text{DNAH12}) \\ & + (0.0193141341488732 * \text{ZBBX}) + (0.28855536069112 * \text{NLRP5}) \end{aligned}$$

Then, the area under the ROC curve (AUC) was calculated to preliminarily test efficacy. Time-dependent curves are shown in Fig. 3C, which demonstrated strong predictive capacity for 1, 3 and 5 years of survival (AUC=0.63 for 1 year of survival, 0.67 for 3 years of survival and 0.67 for 5 years of survival).

### Establishment and evaluation of a nomogram

In Fig. 4A, a nomogram that includes four parameters (gender, age, stage and signature risk) was established. Then, several quantitative metrics were used to evaluate the signature, including c-index, and ROC and DCA curves. The c-index, which ranges between 0.5 and 1, is a regularly used metric for signature prediction accuracy. The signature provided a c-index of 0.717, which indicated its great performance in prediction accuracy. DCA curves are a simple method of measuring the clinical utility of different models. As seen in DCA curves, the curve representing the signature did not cross the two extreme curves, which also demonstrated a significant level of net benefit over a broad range of risk thresholds. As a result, the signature may provide clinical benefits for patients (Fig. 4B). All of the above metrics indicated adequate discriminatory capacity. According to the



**Figure 4.** Establishment and evaluation of a nomogram. (A) Nomogram plot for predicting 1, 3 and 5-year overall survival. (B) DCA curve used for evaluating discrimination ability. (C–E) Calibration curves for 1, 3 and 5 years to testify uniformity power. DCA: decision curve analysis.

		High-score	Low-score	P
Sample		244	245	
Gender	Female (%)	123 (50.4)	141 (57.6)	0.135
Age	Age < 60 (%)	85 (35.4)	67 (28.0)	0.101
M (%)	M0	172 (70.5)	154 (62.9)	0.348
	M1	11 (4.5)	6 (2.4)	
	M1a	0 (0.0)	2 (0.8)	
	M1b	3 (1.2)	2 (0.8)	
	MX	58 (23.8)	81 (33.1)	
N (%)	N0	145 (59.4)	173 (70.6)	0.030
	N1	52 (21.3)	40 (16.3)	
	N2	41 (16.8)	27 (11.0)	
	N3	2 (0.8)	0 (0.0)	
	NX	4 (1.6)	5 (2.0)	
T (%)	T1	27 (11.1)	37 (15.1)	0.334
	T1a	19 (7.8)	27 (11.0)	
	T1b	25 (10.2)	30 (12.2)	
	T2	82 (33.6)	76 (31.0)	
	T2a	40 (16.4)	39 (15.9)	
	T2b	14 (5.7)	9 (3.7)	
	T3	24 (9.8)	20 (8.2)	
	T4	12 (4.9)	5 (2.0)	
	TX	1 (0.4)	2 (0.8)	
Stage (%)	Stage I	3 (1.2)	2 (0.8)	0.046
	Stage IA	49 (20.1)	79 (32.2)	
	Stage IB	67 (27.5)	66 (26.9)	
	Stage II	1 (0.4)	0 (0.0)	
	Stage IIA	24 (9.8)	26 (10.6)	
	Stage IIB	36 (14.8)	32 (13.1)	
	Stage IIIA	42 (17.2)	27 (11.0)	
	Stage IIIB	8 (3.3)	2 (0.8)	
	Stage IV	14 (5.7)	11 (4.5)	

**Table 1.** Clinical features of high- and low-score groups.

calibration curves of 1, 3 and 5 years, the calculated and actual parameters were near on plots (Fig. 4C–E), which confirmed the effectiveness of the signature in prediction consistency.

### Worse clinical outcomes of the high-score group

After the classification of patients by median risk scores, KM analysis revealed a significant difference between the two groups. The high-score group experienced a faster reduction in survival chances over time (Fig. 3D).

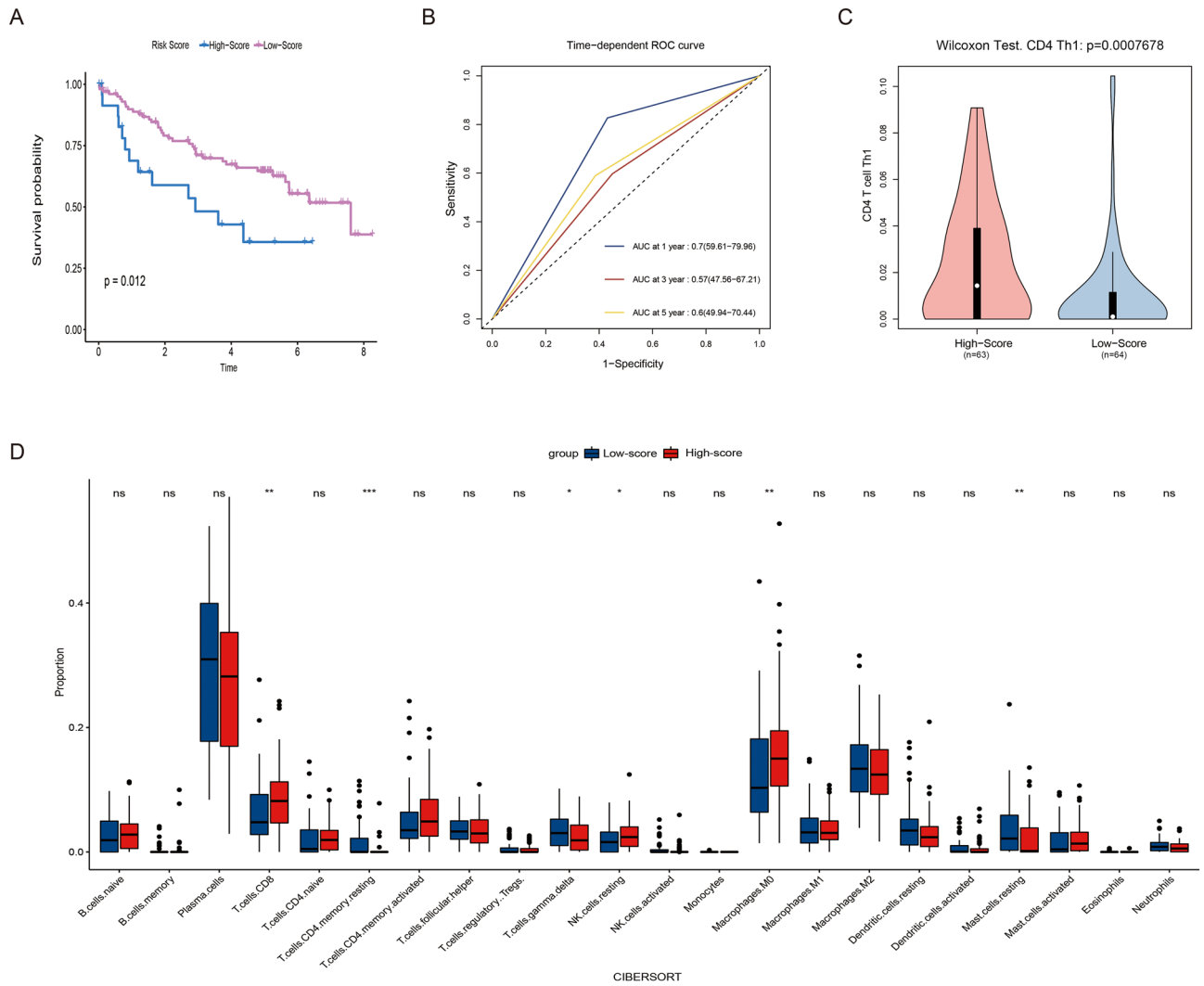
To broaden the applicable scope of the signature, the relationship between the signature and clinical traits was explored. According to the results of clinical analysis, significant differences were found between groups in N ( $P = 0.03$ ) and stage ( $P = 0.046$ ). Patients with higher risk scores are prone to develop node metastasis and have a higher stage (Table 1).

### Immune landscape comparison

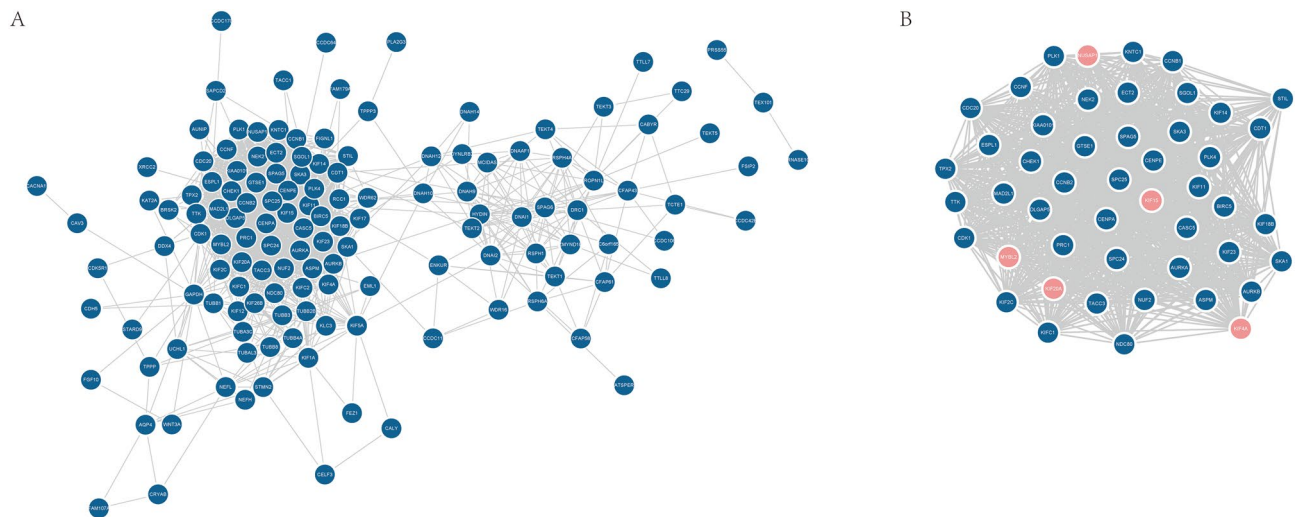
The high-score group had higher expression levels of PD-L1 and LAG3 compared with the low-score one (Fig. 5A). Immunotherapy has a greater possibility of helping patients with greater immune checkpoint levels. It indicated that patients in the high-score group may benefit from ICIs. Proportions of the cluster of CD4+ T Th1 and CD4+ Th2 cells were estimated by xcell and depicted in Fig. 5B,C. The immune infiltration heterogeneity of the two groups was investigated using the CIBERSORTABS algorithm (Fig. 5D). With the exception of helper T, mast and natural killer (NK) cells, as well as macrophages, most immune cells showed no significant difference. The high score group was more abundant in M0 macrophages and resting NK, CD4+ Th1 and CD4+ Th2 cells. Mast cells and M2 macrophages were observed to infiltrate more in the low-score group. Immune infiltration imbalance may be involved in the prognostic difference between low- and high-score groups. It may affect the response of ICIs by influencing the expression of immune checkpoints.



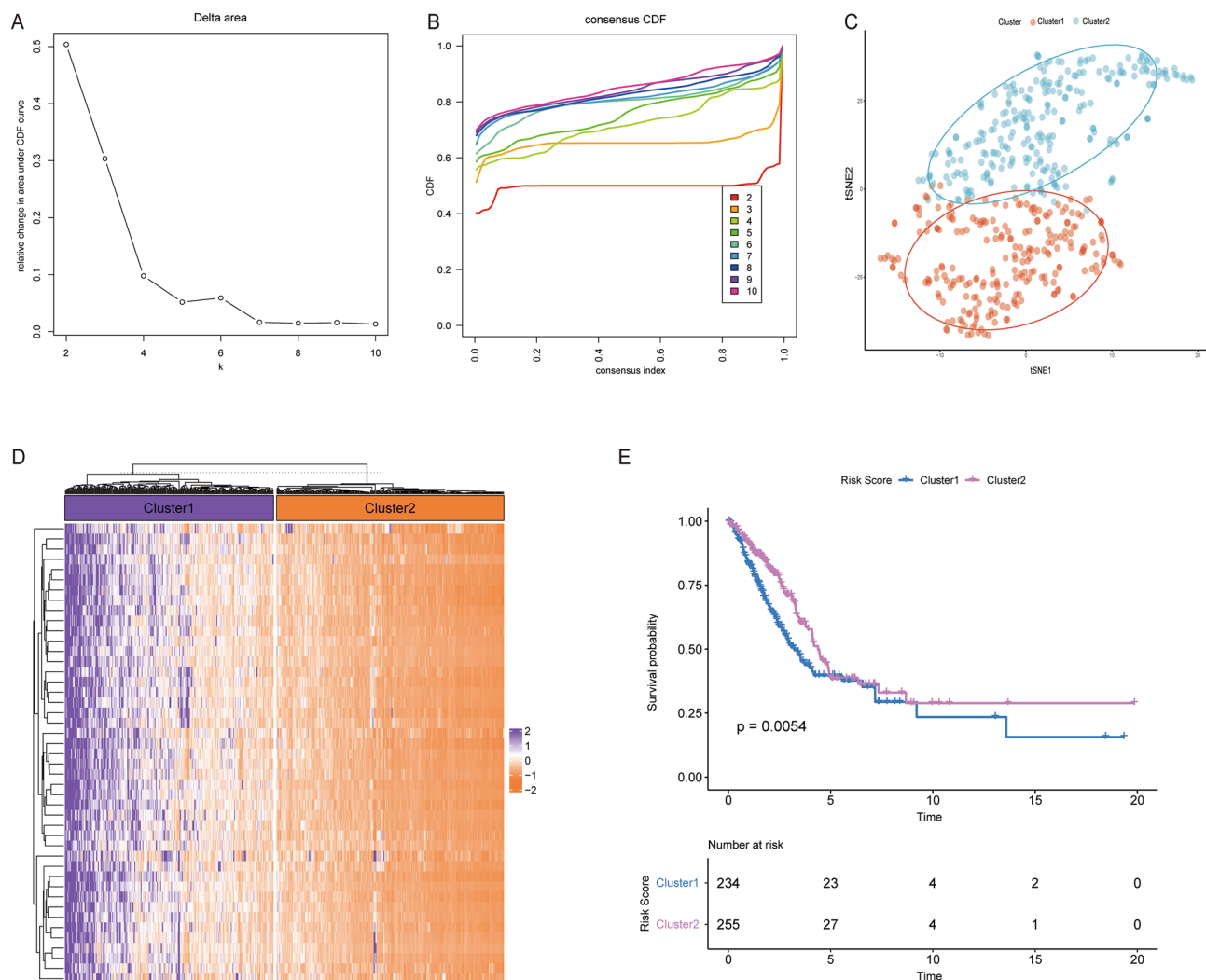




**Figure 6.** External validation of the signature in GSE50081 (n=127). **(A)** Results of the KM analysis of overall survival for low- and high-score groups. **(B)** Time-dependent ROC curves and AUC. **(C)** Infiltration of CD4+ Th1 cells estimated by xcell. **(D)** Infiltration of 22 common immune cells estimated by cibersort.



**Figure 7.** **(A)** PPI network of 154 differently expressed MAGs. **(B)** 48 microtubule-associated proteins screened by MCODE. MCODE: molecular complex detection.



**Figure 8.** Identification of subgroups using consensus clustering. (A) Changes in the area under the CDF curve. (B) CDF curve. (C) tSNE analysis of clusters. (D) Heatmap of clusters. (E) KM analysis between clusters. CDF: consensus clustering distribution function.

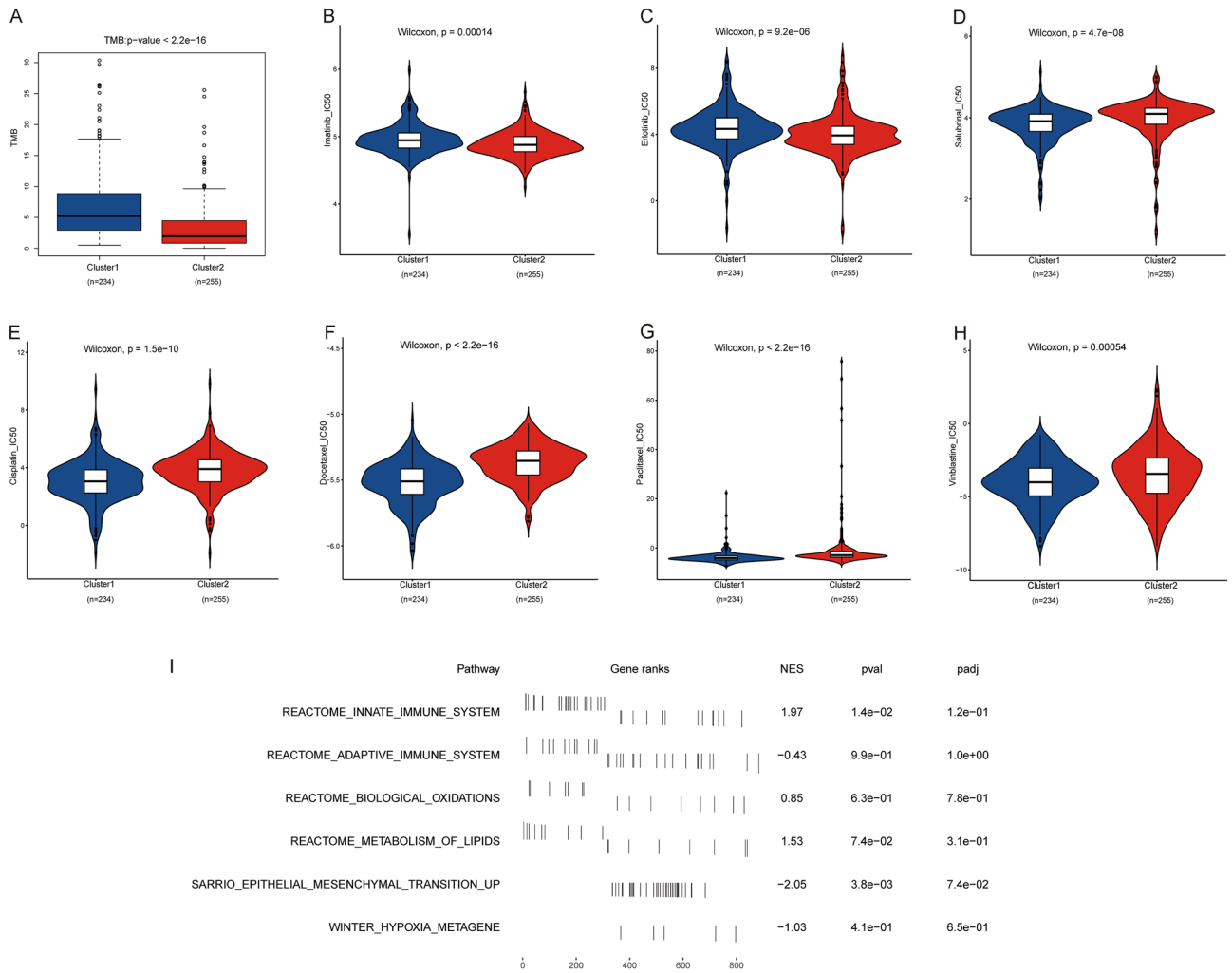
### Biological pathways and therapeutic effects

TMB was assessed as a measure of immunotherapy response. With more mutation-derived antigens, T cells tend to recognize neoantigens and kill cancerous cells<sup>16</sup>. As a widely accepted immunotherapy biomarker, patients with higher TMB are more inclined to gain more benefits from ICB therapy. Patients in cluster1 had significantly higher TMB than those in cluster2, which indicated that immunotherapy would benefit them more (Fig. 9A). The IC50 of commonly used chemotherapeutic drugs is shown in Fig. 9E–H. Patients in cluster1 had a lower IC50 of chemotherapeutic drugs incorporating cisplatin, docetaxel, paclitaxel and vinblastine, which indicated better chemosensitivity. IC50 values varied depending on the targeted medicines (Fig. 9B–D). Patients with a high expression level of hub MAGs may find it difficult to benefit from erlotinib. Erlotinib is used for advanced non-small cell lung cancer (NSCLC) that has failed chemotherapy. On the other hand, salubrinal would more benefit hub MAGs.

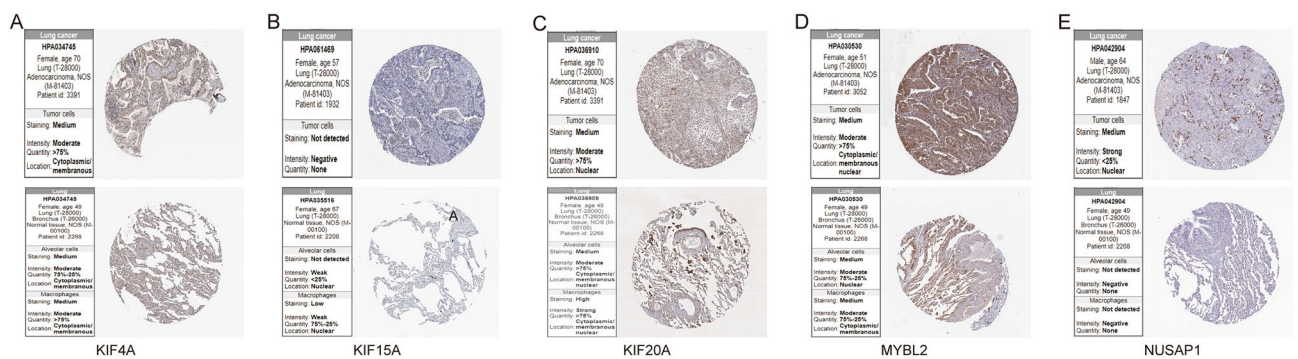
As indicated by the GSEA analysis, hub MAGs were involved in multiple oncogenic pathways. Microtubule-associated pathways were the top-ranked pathways in the GSEA-enriched pathways, which confirmed the association between the functions of hub MAGs and microtubules. Besides, other cancer-associated pathways were also found in the GSEA-enriched pathways. Some of the enrichment results are shown in Fig. 9I. Except for microtubule-associated pathways, hub MAGs are also engaged in other critical pathways including immune, oxidations, metabolism, EMT and hypoxia. They may be implicated in other cancer-related pathways to promote the malignant phenotype of cancer.

### Verification of differently expressed hub MAGs

The expression of differentially expressed hub MAGs was verified in the HPA database and experiment. It was shown that all of the five genes were significantly overexpressed in LUAD (Fig. 10A–E). Compared with BEAS-2B cells, they experienced a significant increase in A549 cells (Fig. 11).



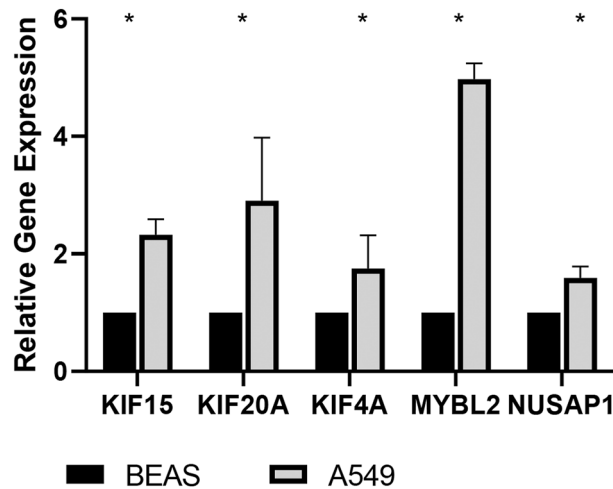
**Figure 9.** Comparison of biological pathways and therapeutic effects between clusters. (A) TMB values in two clusters. (B–D) Comparison of IC50 in Imatinib (B), Erlotinib (C) and Salubrinal (D). (E–H) Comparison of IC50 in Cisplatin (E), Docetaxel (F), Paclitaxel (G) and Vinblastine (H).



**Figure 10.** Immunohistochemical analysis of five MAGs in the HPA database. (A) KIF4A. (B) KIF15. (C) KIF20A. (D) MYBL2. (E) NUSAP1.

**Discussion**

In this research, it was found that microtubule-associated biological processes were differently expressed significantly between tumor and normal samples. Microtubules are a well-known anti-cancer target. However, the clinical application of MTAs is restricted by toxicity and drug resistance. The exploration of novel molecular targets with high efficiency is necessary. Previous research mainly focused on a single gene, but little literature described the characteristics of all MAGs. Singharajkomron<sup>17</sup> investigated the prognostic function of microtubule-associated proteins in lung cancer and distinguished six genes as biomarkers. Consistent with this, KIF4A



**Figure 11.** Validation by qRT-PCR. The relative expression levels of five MAGs were detected by qRT-PCR in A549 and BEAS-2B cells. N = 3. \* represents  $P < 0.05$ . \*\* represents  $P < 0.01$ .

was also identified as a biomarker in this research. However, their methods of subsequent analysis mainly focused on prognostic value of a single gene. In contrast, this study was more inclined to make an overall analysis of MAGs. In addition, the value of MAGs in guiding treatment was further assessed in this research.

Several genes in the model had already been proven as independent prognostic factors. Tubulin  $\beta$  3 (TUBB3) was overexpressed in various cancers and linked to the poor response of MTAs<sup>18</sup>. It has been discovered to be involved in the postoperative survival of NSCLC patients<sup>19</sup>.

In general, patients with high scores developed more advanced clinical stages and were more likely to develop lymphatic metastasis. Cancer metastasis is a complex process which involves both tumor cell phenotype and the tumor microenvironment (TME)<sup>20</sup>. Lymph node and blood metastasis are the most common metastatic routes. Tubulin, a part of the cytoskeleton, works with actin to create impetus through contraction. Depending on this physical process, microtubules promote tumor metastasis through cellular division, cell shape and migration<sup>21</sup>. It has been discovered that TUBB3 inhibits migration by interfering with microtubule dynamics in malignant melanoma cells<sup>22</sup>. It could explain the malignant phenotype of the high-score group.

Immunotherapy is an emerging and promising therapy for NSCLC patients. However, the relationships between the immune microenvironment and microtubules still require further investigation. As complex surroundings that consist of immune and intestinal cells, TME exerts an influence on tumor regression and therapeutic response<sup>23,24</sup>. Tumor immune infiltration differed between low- and high-score groups, which was validated in the external cohort. Patients with high-risk scores demonstrated a higher infiltration of resting NK and activated CD4+ T cells. NK cells are the only innate immune cells to kill tumor cells directly. Resting and activating NK cells are the adverse status of NK cells. The activation of resting NK cells is regulated more strictly with less cytotoxicity<sup>25</sup>. Given the higher infiltration of resting NK cells in the high-score group, immunotherapy based on the activation of NK cells might be a promising therapy. CD4+ T cells are perceived to exert an anti-tumor effect mainly by helping CD8+ cytotoxic T lymphocytes. Other inhibitor factors may account for no significant difference in CD8+ T cell counts between groups.

The comparison of immune checkpoint expression levels showed that the high-score group expressed higher levels of PD-L1 and LAG3. This indicates that patients with a high score may benefit from immunotherapy.

Based on 154 differently expressed MAGs, hub genes were screened by a PPI network. Then, two distinct molecular subtypes were identified based on the expression characteristics of 45 hub MAGs. The survival outcomes and therapy responses of these two subgroups differed significantly.

GSEA analysis revealed that hub MAGs may be involved in multiple oncogenic pathways. Su<sup>26</sup> discovered that the protein regulator of cytokinesis 1 (PRC1) mediated the recruitment of tumor-associated macrophages and regulatory cells (Tregs), which facilitated immunosuppression and angiogenesis in double-negative prostate cancer. Xu<sup>27</sup> reported that the overexpression of epithelial cell transforming sequence 2 (ECT2) would promote the polarization of M2 macrophages in hepatocellular carcinoma. It has been identified that TPX2 level is correlated with EMT in hepatocellular carcinoma and cholangiocarcinoma<sup>28,29</sup>. STIL has been reported recently to enhance metastasis in lung cancer by EMT and hypoxia<sup>30</sup>.

Then, the relationship between hub MAGs and various therapeutic responses was explored. Patients in cluster1 would benefit from immunotherapy and chemotherapy. On the contrary, patients in cluster2 seemed more suitable for targeted therapy. Hub MAGs are involved in not only chemotherapy sensitivity but also immunotherapy response and targeted therapy, which may serve as a reference for therapy selection.

Surprisingly, five genes were included in both the prognostic signature and hub MAGs. They are not only critical in the aberrant activities of microtubules but also likely to be involved in other carcinogenic processes. Hence, the five genes might be potential prognostic factors and therapeutic targets.

The pan-cancer analysis indicated that MYBL2 was observed to be an overexpressed oncogene in multiple cancers. The underlying mechanism might include the regulation of the immune microenvironment and cell

proliferation and spread<sup>31</sup>. Its ability to promote cell proliferation has been experimentally verified in NSCLC<sup>32</sup>. It not only promoted cancer metastasis via EMT but also regulated cancer therapy resistance through the prolongation of cell survival<sup>33</sup>. NUSAP1 is involved in microtubule organization and binding, and its promoting effect on cancer malignant phenotype has been explored in multiple studies<sup>34</sup>. The immune infiltration analysis of hepatocellular carcinoma showed that NUSAP1 might regulate the immune microenvironment by influencing CD4+ T cells<sup>35</sup>.

Kinesins are a type of motor protein that moves along microtubules, which supports a large number of essential functions including cell division and transport. It has been found that kinesins play a key role in cancer progression. Their expression has been reported to link to prognosis, immunological response, drug assistance and other oncogene pathways. The importance of kinesins in the mitogen-activated protein kinase (MAPK) pathway has been studied extensively, which is a complex cascade reaction in essential cell activities<sup>36</sup>.

It has been identified that KIF4A, KIF15 and KIF20A as members of the kinesin family are risk factors with prognostic value in LUAD<sup>37–39</sup>. The overexpression of these genes promotes cancer by not just directly controlling tumor cells, but also affecting the tumor immune microenvironment. In bladder cancer, KIF4A potentiated an immunosuppressive microenvironment by increasing the recruitment of myeloid-derived suppressor cells<sup>40</sup>. KIF15 functioned in macrophage polarization, and its decreased expression resulted in the enhancement of macrophage migration<sup>41</sup>. Furthermore, it was worth noting that KIF20A was linked to CD4+ T-cell response, which was consistent with the above-described immune analysis. In immunotherapy-treated patients with head-and-neck malignant tumors, KIF20A-specific Th1-cell responses were observed<sup>42</sup>. In addition, these five genes were reported to be involved in multiple oncogenic pathways. KIF4A, for example, was observed to accelerate angiogenesis in renal cell carcinoma cells<sup>43</sup> and enhance drug resistance in LUAD<sup>44</sup>.

To summarize, these five genes were closely related to multiple cancers. They were key factors in tumor progression and influenced it both directly and indirectly. These five genes might be effective therapeutic targets in the future.

As is known to all, little literature has discussed the prognostic value of microtubules before. It is the first comprehensive analysis of MAGs in LUAD. The findings may provide new insight into the role of microtubules in the progression, prognosis and treatment of LUAD. Of course, limitations also exist in this research. The expression of hub MAGs in cells was verified. Further validation in clinical samples and further mechanistic investigation are needed.

## Conclusion

A microtubule-related prognosis signature which can serve as an independent prognostic factor was constructed. MAGs were proved to be important factors that regulated therapy response. Five key MAGs were verified as novel prognostic markers and therapeutic targets. In this study, the values of MAGs in prognosis and guiding therapy were explored, which provided evidence for the prognosis and treatment of LUAD.

## Data availability

The datasets presented in this study can be found in the TCGA-LUAD project (<https://portal.gdc.cancer.gov>) and GSE50081 (<https://www.ncbi.nlm.nih.gov/geo/>).

Received: 9 June 2023; Accepted: 17 September 2023

Published online: 26 September 2023

## References

- Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249. <https://doi.org/10.3322/caac.21660> (2021).
- Wu, S. G. & Shih, J. Y. Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer. *Mol. Cancer* **17**, 38. <https://doi.org/10.1186/s12943-018-0777-1> (2018).
- Suresh, K., Naidoo, J., Lin, C. T. & Danoff, S. Immune checkpoint immunotherapy for non-small cell lung cancer: Benefits and pulmonary toxicities. *Chest* **154**, 1416–1423. <https://doi.org/10.1016/j.chest.2018.08.1048> (2018).
- Kavallaris, M. Microtubules and resistance to tubulin-binding agents. *Nat. Rev. Cancer* **10**, 194–204. <https://doi.org/10.1038/nrc2803> (2010).
- Li, F. *et al.* Identifying the EMT-related signature to stratify prognosis and evaluate the tumor microenvironment in lung adenocarcinoma. *Front. Genet.* <https://doi.org/10.3389/fgene.2022.1008416> (2022).
- Chen, J., Fu, Y., Hu, J. & He, J. Hypoxia-related gene signature for predicting LUAD patients' prognosis and immune microenvironment. *Cytokine* <https://doi.org/10.1016/j.cyto.2022.155820> (2022).
- Tomczak, K., Czerwinska, P. & Wiznerowicz, M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol.* **19**, A68–77. <https://doi.org/10.5114/wo.2014.47136> (2015).
- Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Res* **41**, D991–995. <https://doi.org/10.1093/nar/gks1193> (2013).
- Colaprico, A. *et al.* TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71. <https://doi.org/10.1093/nar/gkv1507> (2016).
- Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004> (2015).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
- Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287. <https://doi.org/10.1089/omi.2011.0118> (2012).
- Engelbrechtsen, S. & Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **11**, 123. <https://doi.org/10.1186/s13148-019-0730-1> (2019).
- Finotello, F. & Trajanoski, Z. Quantifying tumor-infiltrating immune cells from transcriptomics data. *Cancer Immunol. Immunother.* **67**, 1031–1040. <https://doi.org/10.1007/s00262-018-2150-z> (2018).

15. Sturm, G. *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445. <https://doi.org/10.1093/bioinformatics/btz363> (2019).
16. Chan, T. A. *et al.* Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. *Ann. Oncol.* **30**, 44–56. <https://doi.org/10.1093/annonc/mdy495> (2019).
17. Singharajkomron, N. *et al.* Evaluating the expression and prognostic value of genes encoding microtubule-associated proteins in lung cancer. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms232314724> (2022).
18. Person, F. *et al.* Prevalence of betaIII-tubulin (TUBB3) expression in human normal tissues and cancers. *Tumour Biol.* **39**, 1010428317712166. <https://doi.org/10.1177/1010428317712166> (2017).
19. Sun, S. *et al.* Prognostic significance of the mRNA expression of ERCC1, RRM1, TUBB3 and TYMS genes in patients with non-small cell lung cancer. *Exp. Ther. Med.* **10**, 937–941. <https://doi.org/10.3892/etm.2015.2636> (2015).
20. Suhail, Y. *et al.* Systems biology of cancer metastasis. *Cell Syst.* **9**, 109–127. <https://doi.org/10.1016/j.cels.2019.07.003> (2019).
21. Hall, A. The cytoskeleton and cancer. *Cancer Metastasis Rev.* **28**, 5–14. <https://doi.org/10.1007/s10555-008-9166-3> (2009).
22. Altonsy, M. O. *et al.* Beta3-tubulin is critical for microtubule dynamics, cell cycle regulation, and spontaneous release of microvesicles in human malignant melanoma cells (A375). *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms21051656> (2020).
23. Hinshaw, D. C. & Shevde, L. A. The tumor microenvironment innately modulates cancer progression. *Cancer Res.* **79**, 4557–4566. <https://doi.org/10.1158/0008-5472.CAN-18-3962> (2019).
24. Wu, T. & Dai, Y. Tumor microenvironment and therapeutic response. *Cancer Lett.* **387**, 61–68. <https://doi.org/10.1016/j.canlet.2016.01.043> (2017).
25. Bryceson, Y. T., March, M. E., Ljunggren, H. G. & Long, E. O. Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion. *Blood* **107**, 159–166. <https://doi.org/10.1182/blood-2005-04-1351> (2006).
26. Su, W. *et al.* The polycomb repressor complex 1 drives double-negative prostate cancer metastasis by coordinating stemness and immune suppression. *Cancer Cell* **36**, 139–155.e110. <https://doi.org/10.1016/j.ccell.2019.06.009> (2019).
27. Xu, D. *et al.* ECT2 overexpression promotes the polarization of tumor-associated macrophages in hepatocellular carcinoma via the ECT2/PLK1/PTEN pathway. *Cell Death Dis.* **12**, 162. <https://doi.org/10.1038/s41419-021-03450-z> (2021).
28. Zou, Z. *et al.* TPX2 level correlates with cholangiocarcinoma cell proliferation, apoptosis, and EMT. *Biomed. Pharmacother.* **107**, 1286–1293. <https://doi.org/10.1016/j.biopha.2018.08.011> (2018).
29. Liang, B. *et al.* TPX2 level correlates with hepatocellular carcinoma cell proliferation, apoptosis, and EMT. *Dig. Dis. Sci.* **60**, 2360–2372. <https://doi.org/10.1007/s10620-015-3730-9> (2015).
30. Ma, Q. *et al.* Hypoxia promotes chemotherapy resistance by down-regulating SKA1 gene expression in human osteosarcoma. *Cancer Biol. Ther.* **18**, 177–185. <https://doi.org/10.1080/15384047.2017.1294285> (2017).
31. Chen, X. *et al.* Pan-cancer analysis indicates that MYBL2 is associated with the prognosis and immunotherapy of multiple cancers as an oncogene. *Cell Cycle* **20**, 2291–2308. <https://doi.org/10.1080/15384101.2021.1982494> (2021).
32. Xiong, Y. C., Wang, J., Cheng, Y., Zhang, X. Y. & Ye, X. Q. Overexpression of MYBL2 promotes proliferation and migration of non-small-cell lung cancer via upregulating NCAPH. *Mol. Cell Biochem.* **468**, 185–193. <https://doi.org/10.1007/s11010-020-03721-x> (2020).
33. Musa, J., Aynaud, M. M., Mirabeau, O., Delattre, O. & Grunewald, T. G. MYBL2 (B-Myb): A central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. *Cell Death Dis.* **8**, e2895. <https://doi.org/10.1038/cddis.2017.244> (2017).
34. Xu, Z. *et al.* NUSAP1 knockdown inhibits cell growth and metastasis of non-small-cell lung cancer through regulating BTG2/PI3K/Akt signaling. *J. Cell Physiol.* **235**, 3886–3893. <https://doi.org/10.1002/jcp.29282> (2020).
35. Zhu, W., Xu, J., Chen, Z. & Jiang, J. Analyzing roles of NUSAP1 from clinical, molecular mechanism and immune perspectives in hepatocellular carcinoma. *Front. Genet.* **12**, 689159. <https://doi.org/10.3389/fgene.2021.689159> (2021).
36. Liang, Y. J. & Yang, W. X. Kinesins in MAPK cascade: How kinesin motors are involved in the MAPK pathway?. *Gene* **684**, 1–9. <https://doi.org/10.1016/j.gene.2018.10.042> (2019).
37. Qiao, Y. *et al.* Increased KIF15 expression predicts a poor prognosis in patients with lung adenocarcinoma. *Cell Physiol. Biochem.* **51**, 1–10. <https://doi.org/10.1159/000495155> (2018).
38. Zhao, X. *et al.* Overexpression of KIF20A confers malignant phenotype of lung adenocarcinoma by promoting cell proliferation and inhibiting apoptosis. *Cancer Med.* **7**, 4678–4689. <https://doi.org/10.1002/cam4.1710> (2018).
39. Zhang, L. *et al.* A novel PHD-finger protein 14/KIF4A complex overexpressed in lung cancer is involved in cell mitosis regulation and tumorigenesis. *Oncotarget* **8**, 19684–19698. <https://doi.org/10.18632/oncotarget.14962> (2017).
40. Lin, N. *et al.* KIF4A promotes tumor progression of bladder cancer via CXCL5 dependent myeloid-derived suppressor cells recruitment. *Sci. Rep.* **12**, 6015. <https://doi.org/10.1038/s41598-022-10029-x> (2022).
41. de Vries, S. *et al.* P23 acts as functional RBP in the macrophage inflammation response. *Front. Mol. Biosci.* **8**, 625608. <https://doi.org/10.3389/fmolb.2021.625608> (2021).
42. Tomita, Y. *et al.* Identification of promiscuous KIF20A long peptides bearing both CD4+ and CD8+ T-cell epitopes: KIF20A-specific CD4+ T-cell immunity in patients with malignant tumor. *Clin. Cancer Res.* **19**, 4508–4520. <https://doi.org/10.1158/1078-0432.CCR-13-0197> (2013).
43. Liu, G. *et al.* The kinesin motor protein KIF4A as a potential therapeutic target in renal cell carcinoma. *Investig. New Drugs* **38**, 1730–1742. <https://doi.org/10.1007/s10637-020-00961-y> (2020).
44. Pan, L. N., Zhang, Y., Zhu, C. J. & Dong, Z. X. Kinesin KIF4A is associated with chemotherapeutic drug resistance by regulating intracellular trafficking of lung resistance-related protein. *J. Zhejiang Univ. Sci. B* **18**, 1046–1054. <https://doi.org/10.1631/jzus.B1700129> (2017).

## Acknowledgements

TCGA and GEO databases were acknowledged for providing their platforms and contributors for uploading their meaningful datasets.

## Author contributions

The study was designed by Y.H.W. The data were analyzed by Y.H.W. and C.Z.Y. Article writing was conducted by J.M.W., W.T.L. and Y.W.Q. equally. The manuscript was revised by G.N.L. All the authors listed have made a direct, substantial and intellectual contribution to the work and approved it for publication.

## Funding

This work gained support from the Guangxi Clinical Medical Research Center for Respiratory Diseases (No. 2022AC04005) and the National Nature Science Foundation of China (No. 82060003).

## Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to G.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023