BMC Cancer

RESEARCH Open Access

# Integrative analysis of a novel signature incorporating metabolism and stemness-related genes for risk stratification and assessing clinical outcomes and therapeutic responses in lung adenocarcinoma

Wanrong Zheng[1†], Chuchu Zhou[2,3†], Zixin Xue[2,3†], Ling Qiao[2], Jianjun Wang[1] and Feng Lu[2,3*]

## Abstract

**Background** Metabolism and stemness-related genes (msRGs) are critical in the development and progression of lung adenocarcinoma (LUAD). Nevertheless, reliable prognostic risk signatures derived from msRGs have yet to be established.

**Methods** In this study, we downloaded and analyzed RNA-sequencing and clinical data from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases. We employed univariate and multivariate Cox regression analyses, along with least absolute shrinkage and selection operator (LASSO) regression analysis, to identify msRGs that are linked to the prognosis of LUAD and to develop the prognostic risk signature. The prognostic value was evaluated using Kaplan-Meier analysis and log-rank tests. We generated receiver operating characteristic (ROC) curves to evaluate the predictive capability of the prognostic signature. To estimate the relative proportions of infiltrating immune cells, we utilized the CIBERSORT algorithm and the MCPCOUNTER method. The prediction of the half-maximal inhibitory concentration (IC50) for commonly used chemotherapy drugs was conducted through ridge regression employing the "pRRophetic" R package. The validation of our analytical findings was performed through both in vivo and in vitro studies.

**Results** A novel five-gene prognostic risk signature consisting of S100P, GPX2, PRC1, ARNTL2, and RGS20 was developed based on the msRGs. A risk score derived from this gene signature was utilized to stratify LUAD patients into high- and low-risk groups, with the former exhibiting significantly poorer overall survival (OS). A nomogram was

†Wanrong Zheng, Chuchu Zhou and Zixin Xue contributed equally to this work.

*Correspondence:
Feng Lu
Lufeng@henu.edu.cn

Full list of author information is available at the end of the article

constructed incorporating the risk score and other clinical characteristics, showcasing strong capabilities in estimating the OS rates for LUAD patients. Furthermore, we observed notable differences in the infiltration of various immune cell subtypes, as well as in responses to immunotherapy and chemotherapy, between the low-risk and high-risk groups. Results from gene set enrichment analysis (GSEA) and in vitro studies indicated that the prognostic signature gene ARNTL2 influenced the prognosis of LUAD patients, primarily through the activation of the PI3K/AKT/mTOR signaling pathway.

**Conclusions** Utilizing this gene signature for risk stratification could help with clinical treatment management and improve the prognosis of LUAD patients.

**Keywords** Lung adenocarcinoma, Metabolism and stemness-related genes signature, Immune infiltration, Bioinformatics

## Introduction

Lung cancer is recognized as the most lethal form of malignant tumor in both the United States and China, with around 85% of cases categorized as non-small cell lung cancer (NSCLC) [1, 2]. Currently, lung adenocarcinoma (LUAD) represents the predominant subtype of NSCLC, maintaining a consistently high incidence rate. In recent years, despite advancements in treating LUAD through improvements in radiotherapy, chemotherapy, surgery, and targeted therapies, patient outcomes remain largely disappointing. Over the last ten years, immune checkpoint blockade immunotherapy aimed at targets such as PDCD1, CD274, and CTLA-4, among other immune regulatory components, has emerged as a promising option for LUAD patients [3]. Nevertheless, a significant number of patients exhibit resistance to these immunotherapy agents [4]. Hence, there is a pressing need to investigate more effective methods for prognostic evaluation and to discover dependable biomarkers capable of predicting the success of anti-immune checkpoint therapies, as well as to stratify those who would benefit from treatment. Ultimately, this could lead to the creation of highly personalized management and treatment strategies for individuals with LUAD. Additionally, in today's digital age, the field-effect transistor (FET) has developed into a versatile device with a wide range of applications across various domains [5–7]. They have garnered considerable attention recently because of their benefits, including rapid sample analysis, label-free identification, a broad dynamic range, and economical fabrication techniques, especially on flexible materials, which exceed the efficiency of traditional methods [8–11]. Nevertheless, when it comes to analyzing or detecting complicated real-world samples, the existence of various interfering agents makes it challenging to simultaneously identify multiple biomarkers or physiological indicators. Consequently, the screening and identification of particular biomarkers hold significant importance.

Metabolic reprogramming is the major hallmark of tumor development. Numerous recent investigations indicate that abnormal metabolism is linked to poor clinical outcomes in a variety of tumor types, including lung adenocarcinoma (LUAD) [12]. During cancer progression, tumor cells frequently alter the metabolic pathways involved in lipid production, glycolytic processes, oxidative phosphorylation (OXPHOS), glutaminolysis, and mitochondrial functions, ensuring an adequate supply of energy, redox balance, and material necessary for their proliferation and spreading [13].

Cancer stem cells (CSCs), a distinct subpopulation of tumor cells, have garnered significant interest as targets for cancer therapy, owing to their abilities for self-renewal and differentiation into multiple cell types, which play a crucial role in promoting tumor growth and diversity. Compared to regular cancer cells, CSCs exhibit greater aggressiveness, thereby facilitating tumor invasion and the spread of cancer to other parts of the body [14, 15]. Malta et al. discovered characteristics of stemness linked to oncogenic dedifferentiation through a comprehensive analysis of various cancers utilizing a machine learning approach [16]. They established stem cell indices to assess the stemness of each tumor sample within The Cancer Genome Atlas (TCGA) database, revealing that these indices could effectively predict metastatic occurrences and provide insights into intratumoral heterogeneity. Notably, the mRNA expression-based stemness index (mRNAsi) can measure cancer stemness by examining the transcriptomic data of cancer specimens [17–19]. In the study of lung adenocarcinoma (LUAD), mRNAsi was employed to evaluate LUAD cases from the TCGA, and the analysis revealed that mRNAsi was significantly elevated in cancer patients. The mRNAsi scores corresponded with clinical stage progression, demonstrating that the low mRNAsi cohort exhibited a superior overall survival rate over a span of 5 years in the majority of LUAD cases [20]. In studies focusing on pancreatic cancer (PDAC), it was observed that patients displaying high mRNAsi expression had a considerably shorter overall survival time compared to those with low mRNAsi expression [21]. Recent developments in metabolomics reveal that the metabolic reprogramming of CSCs plays a crucial role in satisfying energy requirements, preserving

stemness, and supporting cancer proliferation and invasion [22, 23]. However, so far, the exact mechanisms and roles of the interactions between metabolism and stemness in the advancement and prognosis of LUAD are not yet fully understood. Moreover, there is a noticeable absence of prognostic models that incorporate metabolism and stemness-related genes (msRGs) to forecast outcomes for patients with LUAD.

In this study, we hypothesize that there is a prognostic relevance of the interplay between metabolism and stemness in LUAD. Through comprehensive bioinformatics approaches, including univariate Cox regression, LASSO regression, and multivariate Cox regression analyses, we developed a novel prognostic risk signature for LUAD that incorporates genes related to cancer metabolism and stemness. This risk signature was then validated using a merged GEO database comprising seven independent external cohorts. Following this, we confirmed the expression of the identified signature genes using GEPIA and HPA datasets, as well as RT-PCR analyses of clinically matched tissue samples. Additionally, we conducted functional enrichment and somatic mutation analyses to investigate the underlying mechanisms contributing to survival variances across different risk subtypes. Finally, we assessed the relationship between risk scores and levels of immune infiltration, responses to immunotherapy, and sensitivities to chemotherapeutic agents. We also preliminarily examined the potential molecular mechanism by which the prognostic risk gene ARNTL2 influences the prognosis of LUAD through both in vivo and in vitro experimental approaches. Consequently, our research offers new perspectives on personalized treatment strategies and prognostic evaluations for LUAD patients, focusing on the interaction between metabolism and stemness.

## Materials and methods

### Data collection and preprocessing

The expression profiles of mRNA and clinical data were retrieved from both TCGA (https://portal.gdc.cancer.gov/) and GEO (Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/). The microarray datasets for GSE11969, GSE13213, GSE41271, GSE42127, GSE50081, GSE68465, and GSE72094 underwent preprocessing using the affy package in R, which involved background adjustment, normalization, and log2 transformation [24]. To address potential batch effects and other undesired variations among the seven databases, we utilized the surrogate variable analysis (SVA) package from Bioconductor [25]. Subsequently, these seven datasets were combined to form an external validation cohort. The metabolism-related genes and the mRNA expression-based stemness index (mRNAi) for each sample were sourced from earlier studies, respectively [16, 26, 27].

### Screening and functional enrichment analysis of MsRGs

Clustering via unsupervised non-negative matrix factorization (NMF) was conducted on the expression profiles of metabolism-related genes utilizing the NMF package, which is based on the TCGA database. The relationship between all candidate genes and OS was assessed through the use of the "survival" package in R. The optimal cluster value is identified as the point at which the correlation coefficient begins to decline. Analyses for OS and progression-free survival (RFS) among the identified clusters were carried out using the 'survival' and 'survminer' packages in R. The Limma package was employed to examine the differentially expressed genes (DEGs) across LUAD subtypes, based on mRNA expression data concerning the previously mentioned metabolism-related genes, with thresholds set at $|\log FC| \geq 1$ and $P < 0.05$ for the identification of differential genes. Additonally, LUAD patients were categorized based on the median stemness index, and stemness-related genes were identified through differential expression analysis between the two groups using the Limma package. The overlap of differentially expressed metabolism-associated genes and stemness-related genes resulted in the identification of metabolism-stemness-related genes (msRGs). Next, to further investigate the functional significance of the candidate msRGs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted using the R package clusterProfiler [28]. Then, the STRING online database (https://string-db.org) was utilized to assess the pr otein-protein interaction (PPI) network of the candidate msRGs, applying a confidence score threshold of $> 0.4$ as the cut-off criterion.

### Gene set enrichment analyses (GSEA)

Curated collections from version 7.4 of the Molecular Signatures Database served as the target sets for the GSEA analysis, which was conducted using the GSEA software version 4.2.1 (http://www.gsea-msigdb.org/gsea). The entire transcriptome of tumor samples was utilized in the GSEA, and only gene sets that met the criteria of $P < 0.05$ and FDR, $q < 0.05$ were considered statistically meaningful [26].

### MsRGs prognostic risk model construction

Initially, univariate Cox regression was carried out using differentially expressed msRGs (DE-msRGs) to identify prognosis-related genes in the TCGA_LUAD cohort. Subsequently, to refine the list of candidate genes and mitigate the dimensionality of the extensive dataset, least absolute shrinkage and selection operator (LASSO) regression was employed through the R package 'glmnet'. Following this, the selected genes underwent multivariate Cox regression analysis to filter out additional candidate

genes and create a msRGs-related prognostic risk signature aimed at predicting survival outcomes for LUAD patients. The risk score was computed using the equation: Risk score = ∑(Coefi * Expi). Here, Coefi denotes the regression coefficient obtained from the multivariate Cox regression analysis, while Expi indicates the expression levels of the identified prognostic risk genes. Based on the risk score formula, the median risk score was established as the cut-off point, leading to the categorization of LUAD patients into low- and high-risk groups. The model's sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve (AUC) were evaluated. The GEO merged dataset, which comprises seven GSE datasets, served as the external validation cohort.

#### Nomogram model construction and evaluation

Drawing from the findings of the multivariate Cox regression analysis along with clinical case parameters, a predictive nomogram model was additionally developed to quantitatively estimate the likelihood of OS at 1, 3, and 5 years for patients with LUAD utilizing the 'rms' and 'survival' packages in R [29]. Calibration curves were produced to evaluate the model's precision. Furthermore, decision curve analysis (DCA) was conducted to assess the net clinical benefits offered by the nomogram model [30].

#### Immune infiltration analysis

The online database GSCA (https://guolab.wchscu.cn/GSCA/#/) was utilized to determine the relationship between prognostic risk genes and the levels of immune cell infiltration in LUAD tissues. We investigated immune cell infiltration in LUAD samples through three distinct methods. Initially, we applied the Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) algorithm [31], which evaluates each tumor sample for stromal, immune, and tumor purity levels. This analysis was conducted utilizing the R package named "estimate." Subsequently, to gain a deeper insight into the specific types of immune cell infiltration within each tumor, we employed the Cell Type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) algorithm [32] to quantitatively assess the proportions of 22 different immune cell types in LUAD samples. Additionally, we implemented the Microenvironment Cell Populations (MCP)-counter algorithm [33] using the "MCP-counter" R package for further analysis. The relationship between the risk score and immune cell abundance was evaluated using the Spearman correlation test, applying a cutoff threshold of $P < 0.05$.

#### Prediction of immunotherapy and chemotherapeutic response

The expression levels of immune checkpoint genes and members of the class I human leukocyte antigen (HLA) family were compared between high-risk and low-risk groups. Somatic mutation data differences between these groups were analyzed and visualized using waterfall charts generated by the 'maftool' R package. To assess the predictive value of risk scores on the immunotherapy response of LUAD patients, both the Tumor Immune Dysfunction and Exclusion (TIDE) score and Tumor Mutation Burden (TMB) score were utilized. TIDE scores for LUAD patients were obtained from the TIDE database (http://tide.dfci.harvard.edu/) [34]. Additionally, the half-maximal inhibitory concentration (IC50) of chemotherapeutic drugs for each LUAD patient was assessed using the 'pRRophetic' R package, based on data from the Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/) [35]. The drug sensitivity was determined by comparing the IC50 values between patients with low- and high-risk scores, where a lower IC50 value indicated greater sensitivity.

#### Collection of paired clinical lung adenocarcinoma tissues

Tissues from paired LUAD and adjacent normal lung samples were sourced from both Huaihe Hospital of Henan University and Puyang Hospital of Traditional Chinese Medicine, China. Informed consent was obtained from all patients prior to the collection of samples. The Ethics Committee of the Medical School at Henan University, China, approved this study (HUSOM-2018-282). All procedures conducted in this research adhered to the established guidelines provided by the approval.

#### Cell culture and stable transfection of ShRNA

Cell culture and plasmid construction were performed as previously studies [36]. Details of relevant contents are described in the Supplementary Materials and Methods.

#### Transwell assay and in vivo lung metastasis

Cell invasion and migration were assessed using both transwell assays and in vivo analysis of lung metastasis [36]. Transwell filters (Costar) with pore sizes of 8.0 μm were treated with Matrigel (BD Biosciences), diluted at a ratio of 1:6 in serum-free medium. A total of $2.5 \times 10^5$ cells in 200 μl of serum-free media were introduced into the upper chamber coated with Matrigel. The lower chamber was filled with 500 μl of complete medium. Following a 36-hour incubation period, cells remaining in the upper chamber were gently removed using cotton swabs, while those that migrated through the filters into the lower wells were fixed using 3.7% formaldehyde and stained with 1% crystal violet (Sigma). The number

of cells in 3 randomly chosen fields (×200 magnification) from each well was counted. For lung metastasis assay, five-week-old female BALB/C nude mice were procured from Weitong Lihua Animal Co. (Beijing, China) and kept in a designated pathogen-free environment at our institution. A549 cells that were persistently infected with lentiviruses containing either shARNTL2 or control shRNA were harvested for the analysis of pulmonary metastasis. A single-cell suspension with $2 \times 10^6$ cells in 200 μL of PBS was injected into the tail veins of BALB/C nude mice ($n = 5$). All mice were euthanized by injecting excessive pentobarbital sodium [37, 38], and lung samples were taken six weeks post-injection. The experimental protocol received approval from the Animal Care and Research Committee of Henan University.

#### Real-time qPCR and western blot analysis

Total RNA was isolated from frozen matched tissues utilizing Trizol reagent (Takara, Dalian, China) following the instructions provided by the manufacturer. The extracted RNAs underwent reverse transcription (RR036A, Takara) to synthesize cDNA. Real-time PCR was conducted using the TB GreenTM premix Ex TaqTM (RR420A, Takara) and the ABI Prism 7900 System, adhering to these procedures: the denaturation phase occurred for 10 s at 95 °C, the annealing phase lasted for 20 s at 60 °C, the extension phase continued for 30 s at 72 °C, and a total of 40 cycles were completed. The gene expression levels were determined using the $2^{-\Delta\Delta Ct}$ method, with GAPDH serving as the endogenous control. The primer sequences are provided in Table S1. Each independent experiment was conducted a minimum of three times. For the western blot analysis, frozen tissues or cell lines were lysed on ice utilizing RIPA lysis buffer containing 500 mM NaCl, 50 mM Tris pH 8.0, 1 mM EDTA, 1% NP-40, and 1× cocktail of protease inhibitors (Roche, Lewes, UK). Protein concentrations were measured in accordance with the manufacturer's guidelines (Pierce, Rockford, IL). Equal amounts of protein from each sample were separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequently transferred onto polyvinylidene fluoride (PVDF) membranes. Following this, the membranes were incubated with the specified primary antibodies and horseradish peroxidase (HRP)-conjugated secondary antibodies, in accordance with the recommended protocols. Finally, signals were visualized using the enhanced chemiluminescence (ECL) detection kit (Amersham Pharmacia Biotech, Inc., NJ) and a chemiluminescent imaging system (Tanon-5200). The targeted bands presented in the main figures are derived from the corresponding regions of the uncropped original immunoblots after antibody hybridization (Additional file 4). Immunoblots were conducted in two independent experiments. Further details regarding the relevant content have been previously described [36, 39]. Information on the antibodies utilized can be found in Table S2.

#### Statistical analysis

The continuous data are represented as the mean ± standard deviation (SD) unless specified otherwise. For survival analysis, the Kaplan-Meier method along with log-rank tests were utilized. Relationships between datasets were evaluated using Spearman's correlation. Categorical variables were analyzed using either Fisher's exact test or Chi-square tests. All statistical analyses were carried out using R software (version 4.4.0) or the SangerBox platform (http://sangerbox.com/), while GraphPad was used for handling experimental data. All $p$ *values* $< 0.05$ were deemed statistically significant.

## Results

#### Identification of metabolism and stemness-related genes (msRGs) in LUAD cohort

From the TCGA_LUAD data, 1,466 metabolism-related genes were extracted (Table S3), resulting in a transcriptome matrix of a total of 1,411 genes (Table S4). Using the expression profiles of these 1,411 candidate genes, the TCGA_LUAD dataset was classified into two distinct groups by NMF with the optimal value of k determined through a comprehensive analysis of the Residual Sum of Squares (RSS), Cophenetic Correlation Coefficient, Silhouette Score, and consensus matrix classification (Fig. 1A and B, Fig. S1A, and Table S5). To validate the performance of NMF, T-SNE, PCA, and UMAP analyses were conducted, which supported the classification into two subgroups: cluster 1 (C1, n = 200) and cluster 2 (C2, n = 268) (Fig. S1B-D). Furthermore, survival analysis revealed that LUAD patients in subtype C2 experienced a longer median OS ($P < 0.001$), progression-free interval (PFI) ($P < 0.01$), and disease specific survival (DSS) ($P < 0.001$), whereas disease-free interval (DFI) did not show a significant difference between the two subgroups (Fig. 1C-F). Differential analysis of the two subtypes was conducted using the 'limma' package in R, yielding a total of 1,024 metabolism-related differentially expressed genes (DEGs) (Fig. 1G and Table S6). Subsequently, 614 stemness-related genes were identified based on the median of the stemness index through the limma package (Fig. 1H and Table S7). Finally, 415 candidate genes, referred to as metabolism-stemness-related genes (msRGs), were generated by intersecting the differential genes related to metabolism with the stemness-related genes (Fig. 1I).

#### Functional enrichment and genetic mutation landscape of MsRGs

We then assessed the expression levels of 415 msRGs in both normal and tumor samples utilizing the

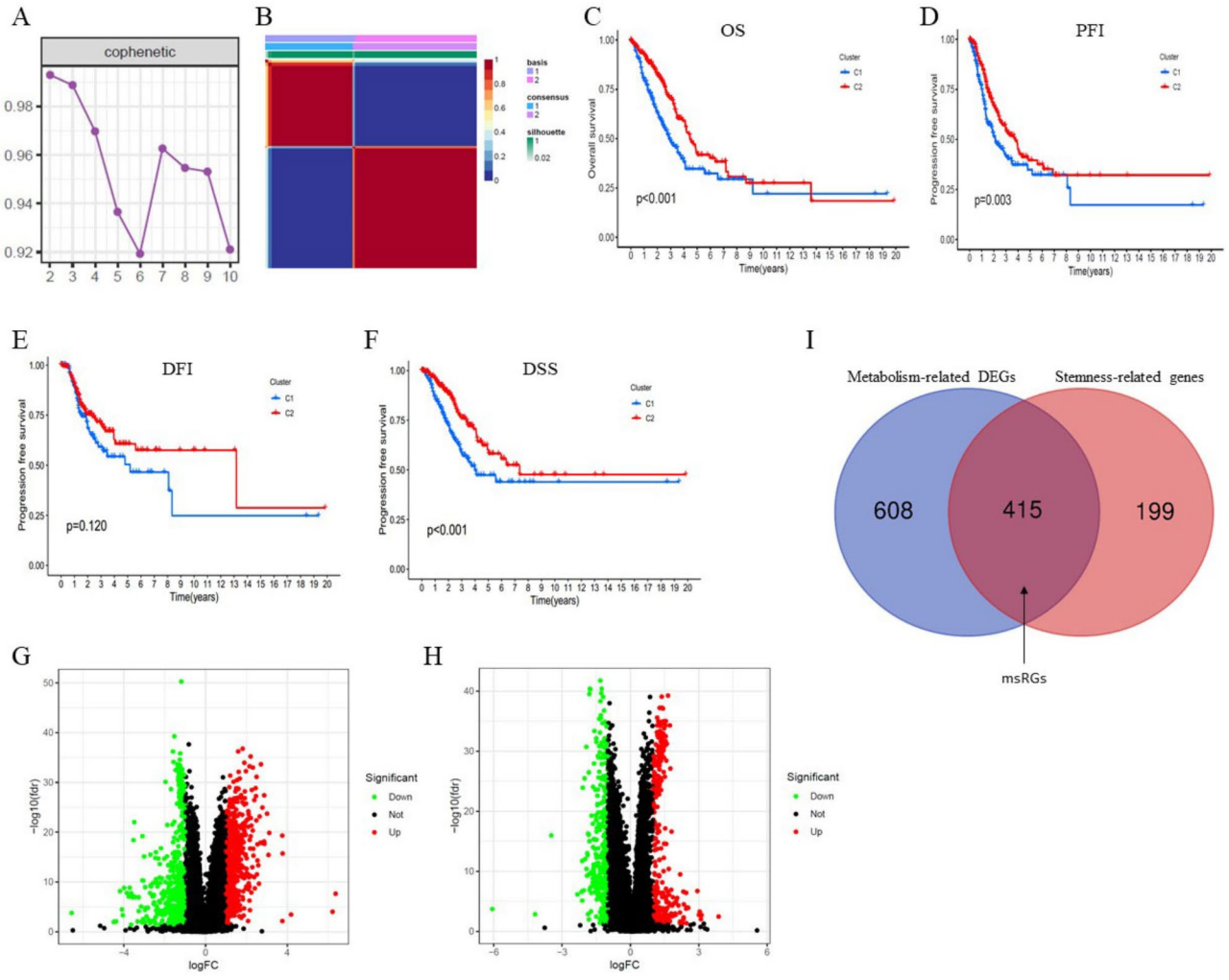


**Fig. 1** Identification of metabolism and stemness-related genes (msRGs). **A** The cophenetic coefficient for clusters k = 2 to 10 indicates that the most significant cointegration correlation coefficient is observed in cluster k = 2. **B** The scheme that partitions the samples into two subgroups demonstrates optimal performance in consensus clustering. **C-F** Kaplan-Meier plots of overall survival (OS) (**C**), progression-free interval (PFI) (**D**), disease-free internval (DFI) (**E**), and disease specific survival (DSS) (**F**) for the two metabolism subgroups of LUAD patients, as derived from the TCGA database. **G**, **H** Volcano plots illustrating the expression of metabolism-related DEGs (**E**) and stemness-related genes (**F**) based on the TCGA database. **I** Venn diagrams depict the overlaps between metabolism-related DEGs and stemness-related genes. DEGs, differentially expressed genes

TCGA_LUAD database. Following this, we identified 301 differentially expressed msRGs (DE-msRGs), comprising 188 upregulated genes and 113 downregulated genes (Fig. 2A and Table S8). To explore the functional roles of DE-msRGs in LUAD, we conducted gene set enrichment analysis (GSEA). Findings indicated that the GO biological processes (BP) were predominantly enriched in nuclear division, organelle fission, chromosome segregation, and mitotic nuclear division. Notable cellular components (CC) related to these processes included the spindle, chromosomal region, and condensed chromosome, among others. The molecular function (MF) terms mainly involved tubulin binding, microtubule binding, and cytoskeletal motor activity (Fig. 2B). KEGG pathway analysis revealed that the target genes were significantly implicated in several cancer-related pathways, including the cell cycle, oocyte meiosis, motor proteins, progesterone-mediated oocyte maturation, and cellular senescence (Fig. 2C). To further examine the involvement of DE-msRGs in the initiation and progression of LUAD, we analyzed the regulatory relationships among DE-msRGs and constructed a protein-protein interaction (PPI) network with the STRING database. This network diagram included 214 nodes and 4305 edges, highlighting CDK1, CCNB1, KIF23, KIF11, CCNA2, BUB1B, MCM10, CDC45, CDC20, and TTK as the ten primary core genes (Fig. 2D). Finally, we utilized the TCGA_LUAD mutation database to conduct an analysis of the genetic mutation background for 62 out of 301 DE-msRGs genes that exhibited ten or more mutation cases via GSCA online

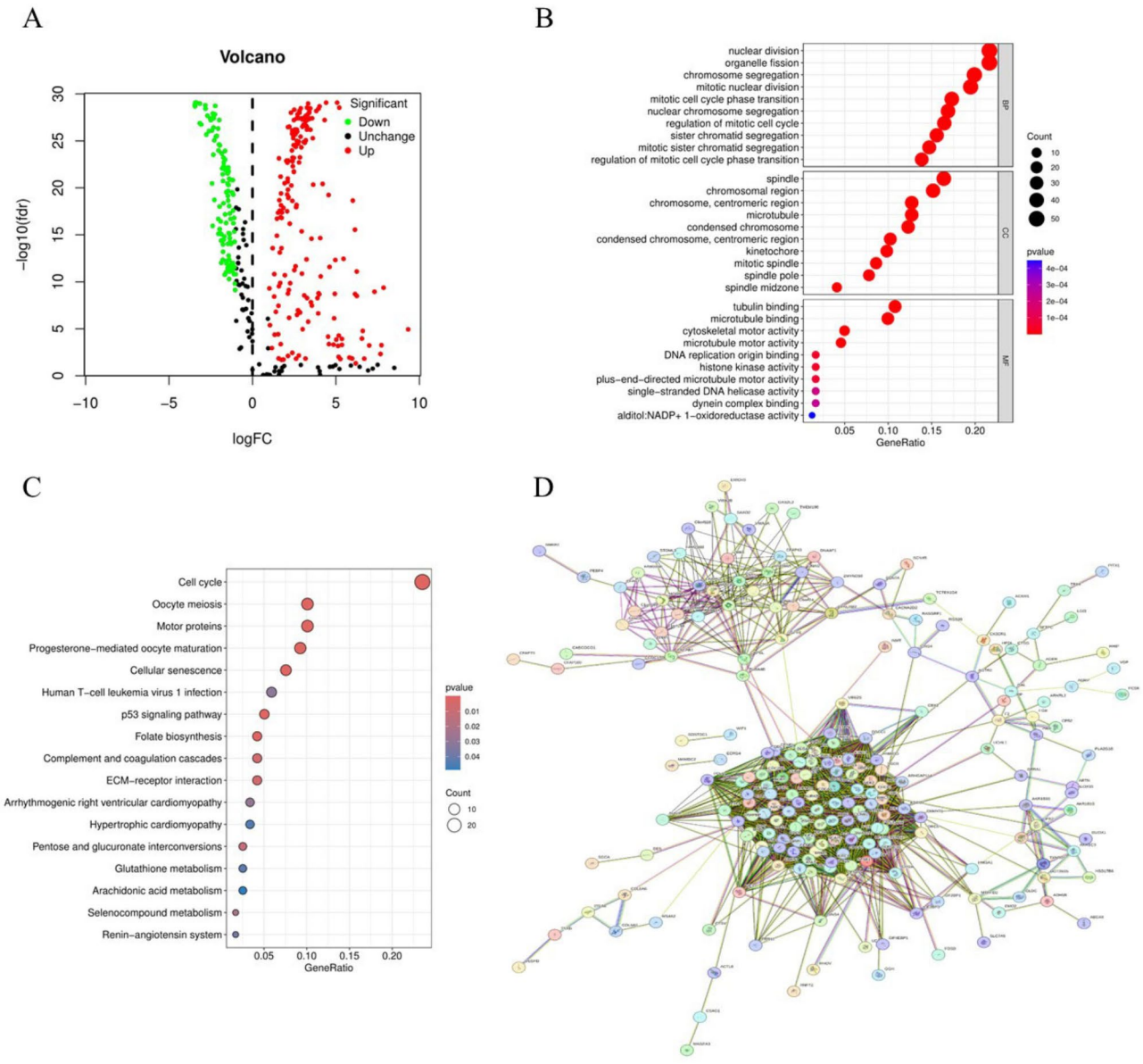


**Fig. 2** Functional enrichment analysis of metabolism and stemness-related DEGs and the establishment of protein-protein interaction networks. **A** Volcano plot illustrating the metabolism and stemness-related DEGs. **B** Representative results from gene ontology (GO) analyses of DEGs sourced from the TCGA database. **C** Representative analyses of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways related to DEGs in the TCGA database. **D** The protein-protein interaction network for metabolism and stemness-related DEGs, constructed using the STRING database. DEGs, differentially expressed genes

(https://guolab.wchscu.cn/GSCA/#/mutation) (Table S9 and Fig. S2A-D). Among the 401 LUAD patients documented in the TCGA database, 266 individuals (66.33%) were identified as having genetic mutations, with ERICH3 showing the highest mutation rate at 21%. This was followed by ASPM, FCGBP, ITGA8, COL6A6, CENPF, SLIT3, POLQ, SCN7A, and MKI67. We also assessed the CNV frequencies of 62 DE-msRGs in LUAD, noticing that ACKR1 had the most significant amplification frequency at 72.09%. In contrast, BUB1B and GLDC both demonstrated notable CNV loss frequencies of 53.68% and 53.29%, respectively (Table S10). From this analysis, we observed that msRGs in LUAD tissues exhibited significant transcriptional and genetic changes, which may play a role in the oncogenesis of LUAD.

### Establishment and validation of a novel msRGs-related gene prognostic model for LUAD

To investigate the possible predictive significance of the 301 DE-msREs in LUAD, we conducted a univariate Cox regression analysis. When Hazard Ratio (HR) ≠ 1, with $p < 0.05$, we identified 81 genes from the TCGA

database as prognostic markers influencing the OS of LUAD patients (Table S11). Next, we utilized LASSO regression with 10-fold cross-validation to ascertain the optimal lambda values derived from the minimum partial likelihood deviance, correlating with 9 out of the 81 prognostic genes that showed a strong association with OS (Fig. 3A and B and Table S12). In our analysis, we used the TCGA_LUAD database as the training cohort ($n = 464$) alongside a merged GEO database serving as the validation cohort ($n = 1489$). By performing multivariate Cox regression analysis, we pinpointed an optimal gene signature comprising five metabolism-stemness-related DEGs (msrDEGs) (S100P, GPX2, PRC1, ARNTL2, and RGS20), along with their respective coefficients for LUAD (Fig. 3C). Subsequently, we formulated a prognostic risk model as follows: risk score = (0.082×expression value of S100P + 0.078×expression value of GPX2 + 0.575×expression value of PRC1 + 0.322×expression value of ARNTL2 + 0.372×expression value of RGS20). The risk score for each patient in the TCGA database was computed. Principal component analysis (PCA) indicated that patients with LUAD could be categorized into two molecular subgroups according to risk scores in both the TCGA_LUAD cohort and the GEO-merged cohort (Fig. S3A, 3B). We then employed the "survminer" R software package to derive the median cut-off point, which allowed us to separate patients from the TCGA_LUAD database into high- and low-risk groups based on their individual risk scores. Kaplan-Meier analysis indicated that the OS for the high-risk group was significantly poorer than that of the low-risk group ($P < 0.001$) (Fig. 3D). Figure 3E illustrated that individuals in the high-risk group were more likely to express the risk genes. In addition, we found that the five-gene prognostic signature exhibited higher AUC values in the time-dependent ROC analysis for both 1-year and 3-year survival compared to each of the previously mentioned risk prognostic genes within the TCGA cohort (Fig. 3F and Fig. S3C). To further validate the predictive capability of the five-gene prognostic signature, we incorporated the GEO-merged cohort ($n = 1489$) as an external validation group to assess the results obtained from the TCGA training cohort. Aligning with the training cohort outcomes (Fig. 3G-I and Fig. S3D), the KM curves for the validation cohort consistently indicated that the high-risk group had a poorer prognosis than the low-risk group (Fig. 3G). The time-dependent ROC analysis revealed that the AUCs for 1-year, 3-year, and 5-year OS in the GEO-merged cohort were 0.724, 0.699, and 0.672, respectively (Fig. 3I). In summary, the five-gene prognostic risk signature demonstrated strong performance in predicting the OS of patients with LUAD.

#### Relationship between prognostic risk signature and clinicopathologic characteristics in LUAD

Next, we analyzed the differences in risk scores based on various stratified features to investigate the association between the prognostic risk signature and clinical and pathological characteristics. The violin plots illustrated that LUAD patients with advanced pathological stage ($P < 0.001$), lymph node metastasis ($P < 0.001$), higher T stage ($P < 0.05$), and those who are female ($P < 0.001$) exhibited higher risk scores in the TCGA_LUAD cohort (Fig. 4A-E). Furthermore, an analysis of the TCGA database revealed that LUAD patients with advanced pathological stage ($P < 0.001$), lymph node metastasis ($P < 0.001$), and higher T stage ($P < 0.001$) experienced shorter OS as determined by stratified survival analysis (Fig. 4F-J). Additionally, we further explored the independence of the prognostic risk signature and clinicopathological factors through univariate and multivariate Cox regression analyses based on the TCGA cohort. Univariate Cox regression analysis demonstrated that the clinical features of pathological stage (HR = 1.586, $P < 0.001$), T stage (HR = 1.712, $P < 0.001$), N stage (HR = 2.473, $P < 0.001$), and risk score (HR = 1.805, $P < 0.001$) were significantly correlated with survival rates in LUAD. In contrast, multivariate Cox regression analysis indicated that tumor pathological stage (HR = 1.314, $P < 0.05$), N stage (HR = 1.580, $P < 0.05$), and risk score (HR = 1.765, $P < 0.001$) were independent predictors of OS in LUAD patients (Fig. 4K and L). These results suggest the potential for integrating risk score, N stage, and tumor pathological stage for prognosis stratification in LUAD.

#### Validation of prognostic risk genes

To further validate the prognostic risk model, we initially examined the correlation between the expression levels of five prognostic risk genes and the survival rate of patients with LUAD. The Kaplan-Meier curves revealed that high expression of S100P, PRC1, ARNTL2, and RGS20 genes was significantly associated with shorter OS in LUAD patients, while high expression of GPX2 may suggest a poor prognosis in LUAD patients, although this was not statistically significant (HR = 1.19, $P = 0.15$) (Fig. 5A). Subsequently, we investigated the expression patterns of these five risk genes using databases and RT-qPCR. GEPIA analysis indicated that mRNA levels of S100P, GPX2, PRC1, and ARNTL2 were markedly elevated in LUAD compared to normal samples. The mRNA level of RGS20 was also higher in LUAD, but not significantly different (Fig. 5B). RT-qPCR analysis of eight paired clinical LUAD tissue samples showed increased mRNA expression of these five prognostic risk genes in LUAD tissues compared to control tissues, consistent with the GEPIA results except for RGS20 (Fig. 5C). Furthermore, immunohistochemical analyses from the Human Protein Atlas

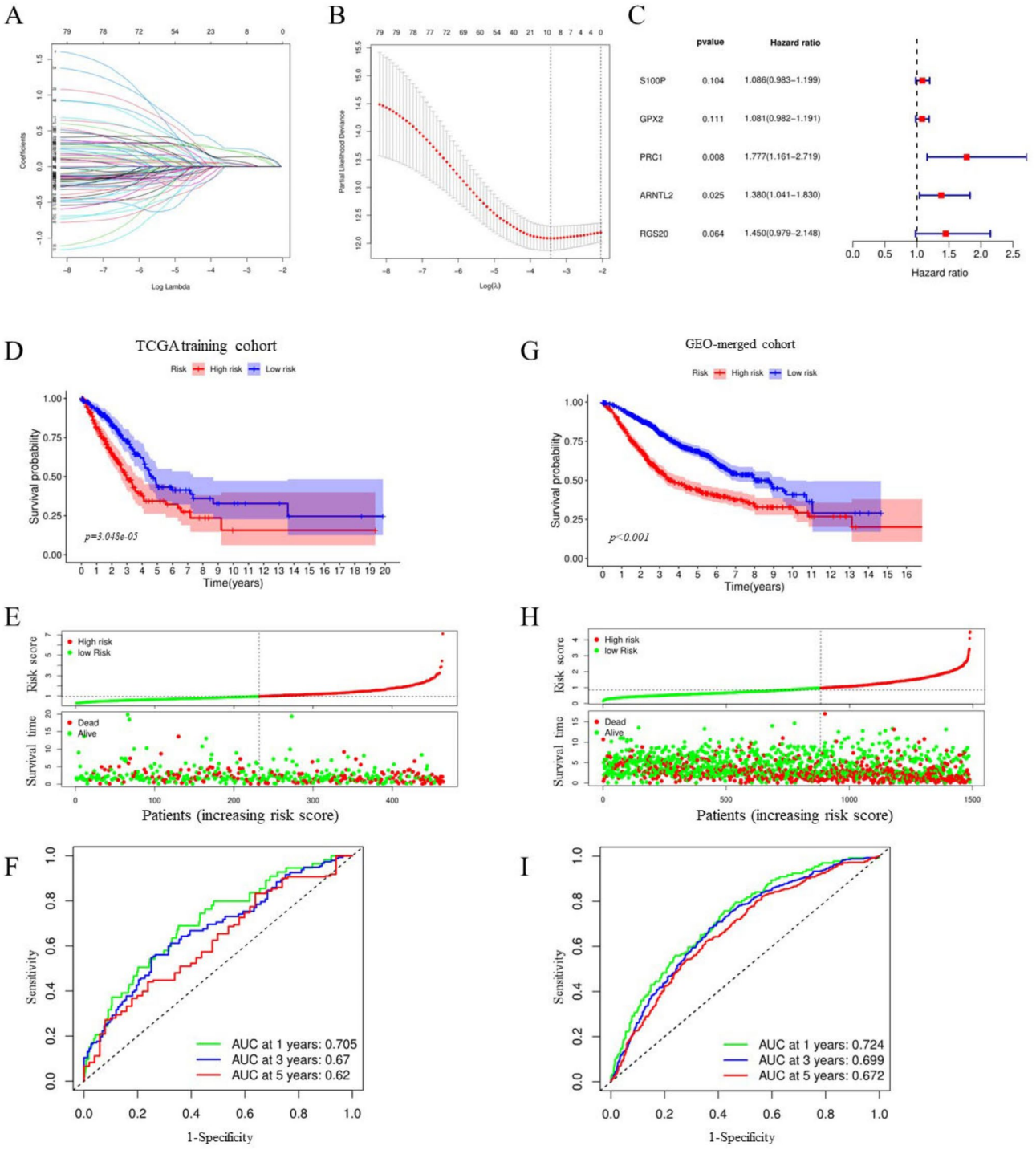


**Fig. 3** The metabolism and stemness-related gene signature and prognostic classifier in the LUAD cohort. **A** LASSO coefficient analysis of the metabolism and stemness-related DEGs is presented. The dotted lines in the graph indicate the values selected through 3-fold cross-validation. **B** A three-fold cross-validation is performed to determine the tuning parameter in the LASSO model, with partial likelihood deviation values plotted against log(λ), and the error bars representing standard error (SE). **C** A forest plot displays the hazard ratios for five metabolism and stemness-related prognostic DEGs obtained from multivariate Cox regression analyses. **D**, **G** Kaplan-Meier plot analyses are shown for the TCGA cohort (**D**) and the GEO-merged cohort (**G**). **E**, **H** The risk distribution among patients is depicted in the training cohort (**E**) and the GEO-merged cohort (**H**). **F**, **I** The ROC curves of the risk signature are illustrated for the training cohort (**F**) and the GEO-merged cohort (**I**). DEGs refers to differentially expressed genes

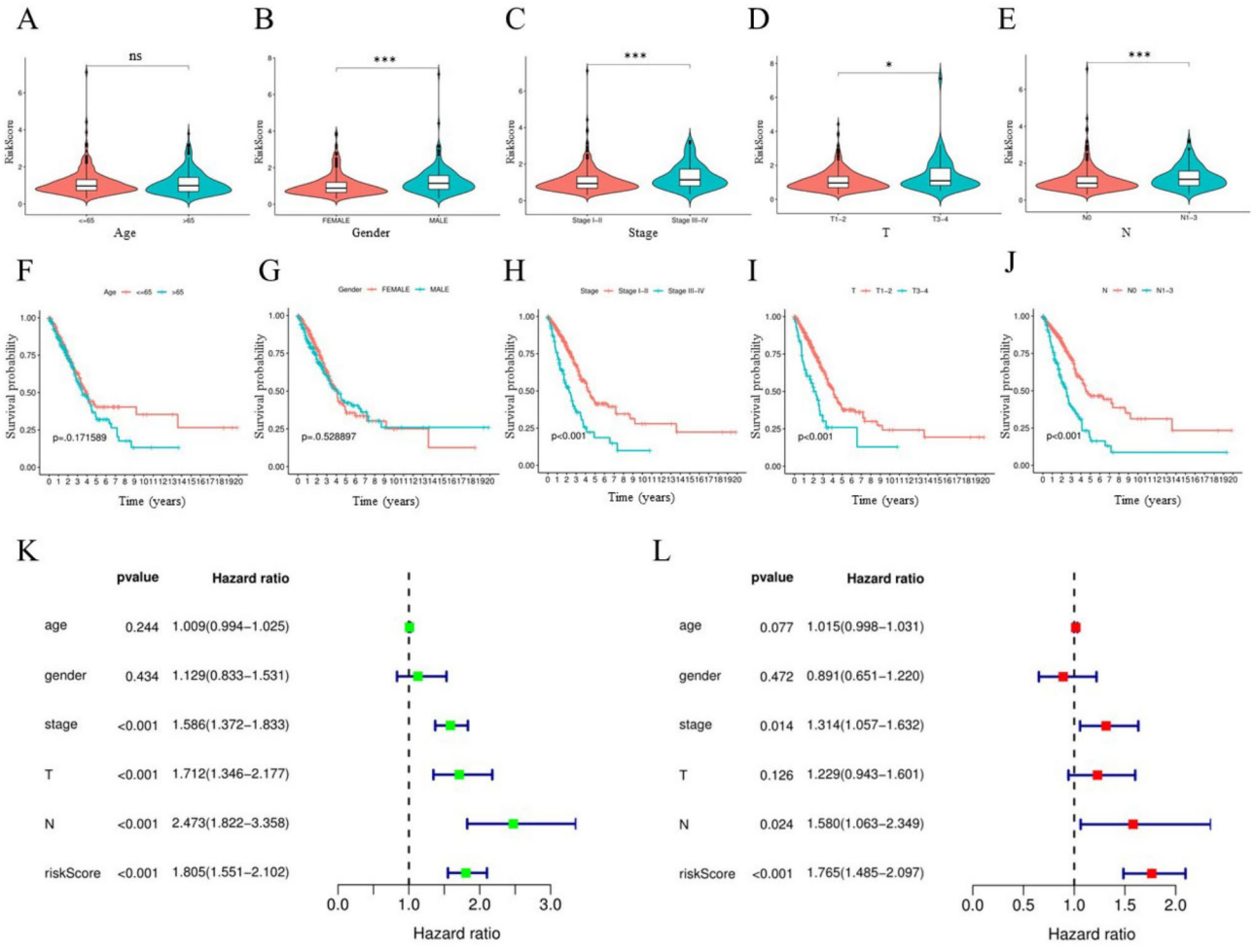


**Fig. 4** The correlation between the prognostic risk signature and clinicopathological characteristics in the TCGA_LUAD cohort. **A-E** Comparison of the risk score across different subgroups stratified by clinicopathological characteristics, including age (**A**), gender (**B**), pathological stage (**C**), T stage (**D**), and N stage (**E**). **F-J** Kaplan-Meier curves depicting the probability of OS stratified by the same clinicopathological characteristics: age (**F**), gender (**G**), pathological stage (**H**), T stage (**I**), and N stage (**J**). **K**, **L** Cox regression analyses for univariate (**K**) and multivariate (**L**) models, incorporating age, gender, pathological stage, T stage, N stage, and risk score as factors. OS, overall survival; T, tumor size; N, lymph node metastasis; ns, no significance. Statistical significance is indicated as * $P < 0.05$ and *** $P < 0.001$

(HPA) database demonstrated high staining intensity of S100P, GPX2, and PRC1 in LUAD tissues, contrasting with low intensity or absence of staining in normal tissues, while ARNTL2 and RGS20 did not exhibit significant differences (Fig. 5D).

### Development and validation of the prognostic prediction nomogram

In accordance with the current clinical context and the results of independent prognostic analyses, we developed a composite nomogram that integrates the risk score, pathological stage, T stage, N stage, and age to predict the OS of LUAD patients at 1, 3, and 5 years, based on the TCGA_LUAD cohort (Fig. 6A). The calibration curve for patient survival prediction demonstrated strong performance for the nomogram when compared to an ideal model (Fig. 6B). The area under the curve (AUC) values for 1-year, 3-year, and 5-year predictions of the nomogram exceeded those of other clinical variables (Fig. 6C-E). Additionally, we employed decision curve analysis (DCA) to assess the model's effectiveness. The standardized net benefit confirmed the predictive value of the nomogram for OS at 1 year and 3 years when compared to individual clinical variables (Fig. 6F-H).

### Association of risk subtypes with tumor immune microenvironment

Based on the close correlation between the expression of risk prognostic genes and the levels of infiltration by multiple immune cells, as observed in the GSCA database, we found that ARNTL2 expression showed the strongest positive correlation with nTreg cell infiltration levels ($R = 0.45$, $P < 0.001$), while exhibiting the most substantial negative correlation with Th17 cells ($R$=-0.32, $P < 0.001$)

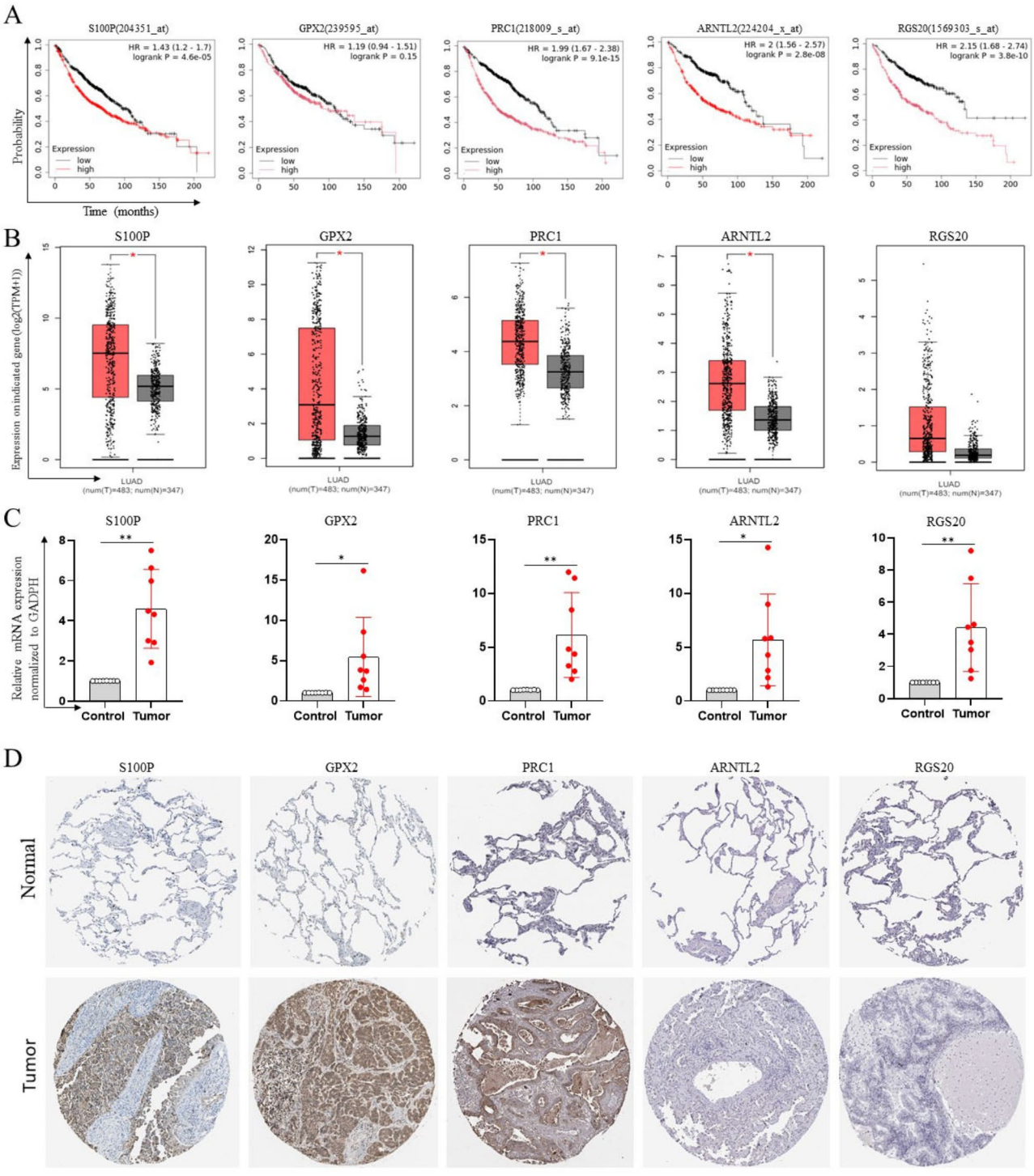


**Fig. 5** The expression patterns of prognostic risk genes and their correlation with OS in patients with LUAD. **A** The Kaplan-Meier plot depicts the relationship between the expression levels of S100P, GPX2, PRC2, ARNTL2, and RGS20 and OS in LUAD patients. **B** The expression patterns of S100P, GPX2, PRC2, ARNTL2, and RGS20 in LUAD and normal samples are presented based on data from the GEPIA database. **C** RT-qPCR analysis of S100P, GPX2, PRC2, ARNTL2, and RGS20 was conducted using matched clinical tissues ($n = 8$). **D** Immunohistochemical analysis of S100P, GPX2, PRC2, ARNTL2, and RGS20 was performed on LUAD and normal tissue samples sourced from the Human Protein Atlas (HPA) database. HR, hazard ratio; OS, overall survival; LUAD, lung adenocarcinoma. Statistical significance is indicated as $^{*}P < 0.05$; $^{**}P < 0.01$

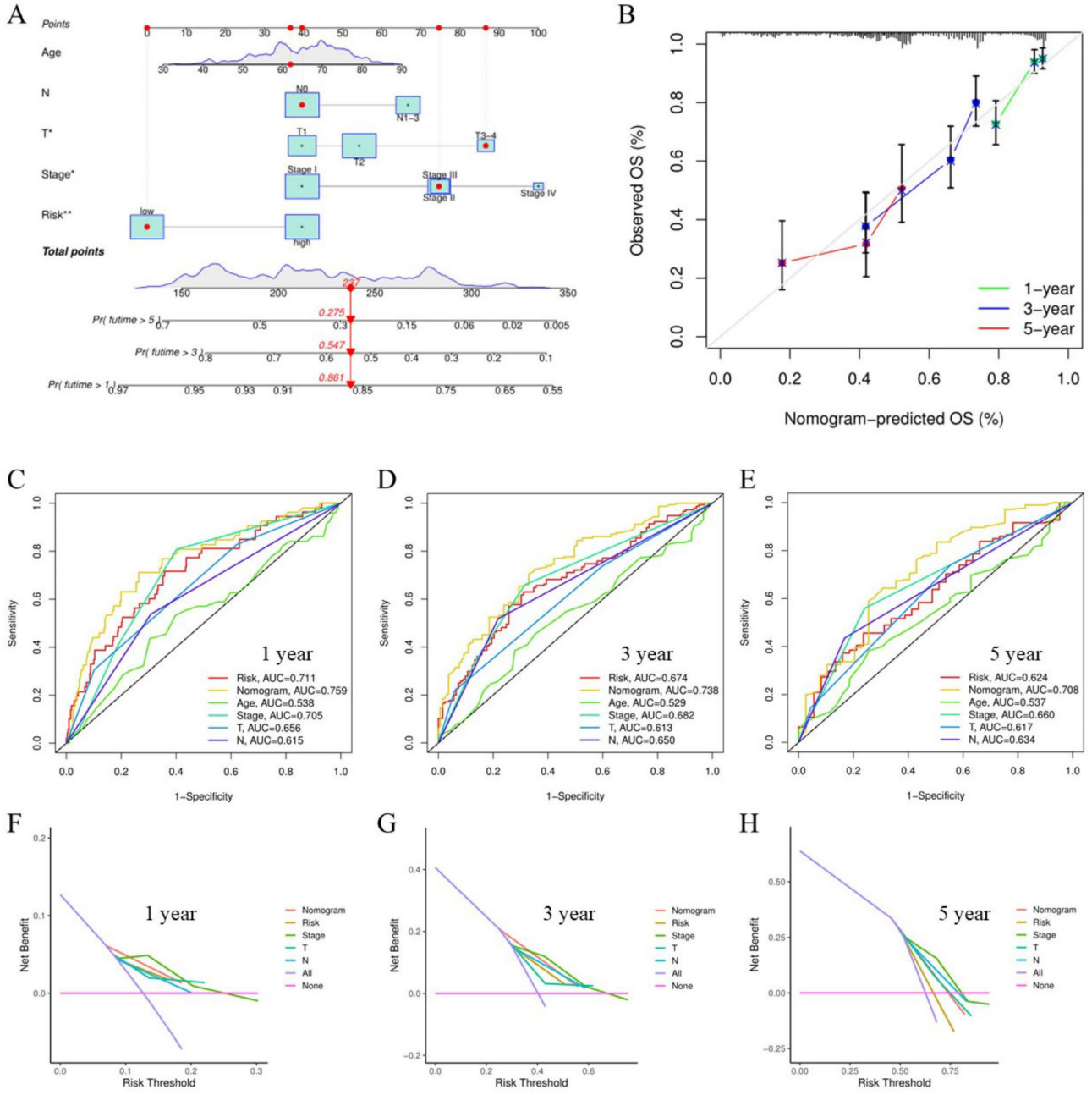


**Fig. 6** Construction and validation of a nomogram utilizing the TCGA database. **A** The nomogram developed predicts the probabilities of OS at 1, 3, and 5 years. The red line exemplifies the method for prognostic prediction. **B** Calibration curves demonstrate the nomogram's performance in predicting 1-, 3-, and 5-year OS by comparing observed and predicted outcomes. **C-E** ROC curve analyses assess the predictive efficiency of the nomogram for 1-, 3-, and 5-year OS based on the TCGA database. **F-H** Decision curve analysis (DCA) is employed to evaluate the model's effectiveness. OS, overall survival; AUC, area under the curve; ROC, receiver operating characteristic

(Fig. 7A). GPX2 expression demonstrated the highest positive correlation with Th17 cells ($R = 0.17$, $P < 0.001$) and the most significant negative correlation with macrophages ($R$=-0.28, $P < 0.001$) (Fig. S4A). Additionally, PRC1 expression levels correlated positively with nTreg cells at $R = 0.61$ ($P < 0.001$), whereas a negative correlation was noted with NKT cells ($R$=-0.43, $P < 0.001$) (Fig. S4B). RGS20 expression showed the most significant positive correlation with exhausted T cells ($R = 0.30$, $P < 0.001$), and the strongest negative correlation with Th17 cells ($R$=-0.24, $P < 0.001$) (Fig. S4C). Lastly, S100P expression exhibited a notable positive correlation with nTreg cells ($R = 0.33$, $P < 0.001$) and the highest negative correlation with macrophages ($R$=-0.30, $P < 0.001$) (Fig. S4D).

Next, we further investigated the relationship between the risk score and immunological infiltrates using the

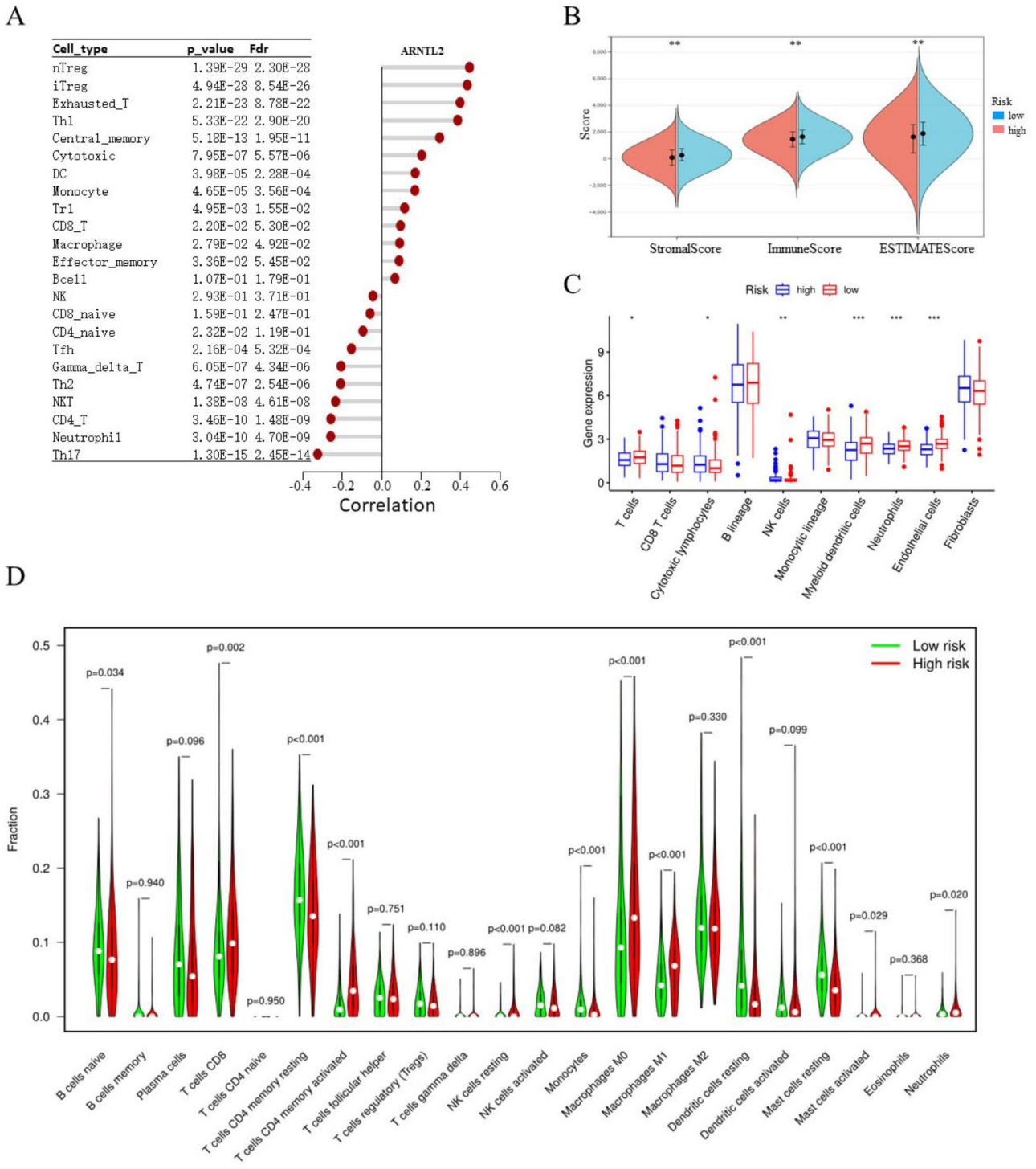


**Fig. 7** The infiltration of immune cells in high-risk and low-risk groups. **A** The correlation analysis of the prognostic risk gene ARNTL2 expression and immune cell infiltration was conducted using the GSCA database. **B** The ESTIMATE algorithm was employed to compare the ESTIMATE score, immune score, and stromal score between the high-risk and low-risk groups. **C**, **D** The levels of immune cell infiltration in LUAD patients were compared between high-risk and low-risk groups using the MCPcounter (C) and CEBERSORT (D) algorithms. Statistical significance is indicated as $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

ESTIMATE algorithm. The findings revealed that the low-risk group exhibited higher ESTIMATE, stromal, and immune scores compared to the high-risk group (Fig. 7B, $P < 0.001$), indicating a greater abundance of immune and stromal cells in LUAD patients with the low-risk signature. Additionally, we estimated the abundance of two stromal cells and eight immune-related cells in each sample using the MCP-counter algorithm. When comparing the high-risk group with the low-risk group, the box plot of immune cell infiltration fraction showed significant increases in T cells, myeloid dendritic cells, neutrophils, and endothelial cells in the low-risk group, whereas cytotoxic lymphocytes and NK cells exhibited marked decreases (Fig. 7C).

Subsequently, we also employed the CIBERSORT method to further investigate the relative proportion of 22 tumor-infiltrating immune cells. As shown in Fig. 7D, there were significant differences in immune cell infiltration between the high- and low-risk groups. The low-risk group exhibited a relatively high proportion of B cells naïve, plasma cells, T cells CD4 memory resting, monocytes, dendritic cells resting, and mast cells resting, while showing low infiltration of T cells CD8, T cells CD4 memory activated, NK cells resting, and macrophages M0 and M1.

#### Association of risk subtypes with immunotherapies and chemotherapeutic responses in LUAD

Recently, some LUAD patients have been found to benefit significantly from immunotherapy, especially immune checkpoint inhibitors. Currently, various biomarkers, such as the expression levels of immune checkpoint proteins, class I human leukocyte antigen (HLA) family members, and tumor mutational burden (TMB), have shown potential in predicting the response to immunotherapy. Therefore, we first compared the expression of different immune checkpoints between the high- and low-risk groups. Figure 8A showed that the high-risk group had significantly higher expression levels of TNFSF4, CD276, LAG3, and ARHGEF5 compared to the low-risk group. On the other hand, the low-risk group exhibited higher expression levels of GEM, CD27, BTLA, and CD47. Next, we examined the expression of human leukocyte antigen (HLA) class I family members. This is because HLA participates in neoantigen presentation and the cytolytic activity of T cells by presenting intracellular peptides on the cell surface. HLA deficiencies impair the ability of cells to present neoantigens and may lead to immune tolerance. As expected, excluding HLA-A, -B, -C, -H, and -G, the low-risk group exhibited a significantly higher level of gene expression for various members of the HLA family, including HLA-L, HLA-J, HLA-F, HLA-E, HLA-DRB1, HLA-DRB5, HLA-DRB6, HLA-DRA, HLA-DQB1, HLA-DQB2, HLA-DQA1, HLA-DQA2, HLA-DPB1, HLA-DPB2, HLA-DPA1, HLA-DOA, HLA-DOB, HLA-DMA, and HLA-DMB, compared to the high-risk group (Fig. 8B). Next, through an extensive mutation analysis of the TCGA_LUAD database, we explored the mutation profiles of LUAD patients categorized into high-risk and low-risk groups. As illustrated in Fig. 8C, the most prevalent variant classification was missense mutation, with a significantly higher proportion of mutations in genes such as TP53, TTN, MUC16, CSMD3, and RYR2 observed in the high-risk group compared to the low-risk group. The tumor mutational burden (TMB) in the high-risk group was also significantly elevated relative to that in the low-risk group ($P < 0.001$) (Fig. 8D). Furthermore, we assessed the potential response to immunotherapy in each patient using the tumor immune dysfunction and exclusion (TIDE) algorithm, revealing that patients in the high-risk group (90%, 209/232) were more likely to respond to immunotherapy than those in the low-risk group (73%, 169/232) (Fig. 8E and F).

In recent decades, chemotherapy has emerged as a central approach in cancer treatment. However, the heterogeneity of tumors has posed a challenge, as the response to the same chemotherapy can vary among patients. We utilized the R package pRRophetic to calculate inhibitory concentration (IC50) of indicated drugs and analyze the correlation between risk scores and clinical responses to chemotherapy drugs. We selected two targeted drugs (gefitinib and axitinib) and six chemotherapy drugs (cisplatin, paclitaxel, gemcitabine, doxorubicin, etoposide, and docetaxel) that have been commonly used in the treatment of lung adenocarcinoma. Based on the Cancer Genome Project (CGP) database, differences were observed in the drug response between the high-risk and low-risk groups. Specifically, the chemotherapy drug cisplatin, paclitaxel, gemcitabine, doxorubicin, etoposide, and docetaxel exhibited a high drug response in the high-risk group. Conversely, in the low-risk group, the targeted drug axitinib showed a favorable response (Fig. S5A-H). Overall, the risk score-related msRGs developed in our study demonstrate considerable potential for predicting prognosis and the benefits of immunotherapy and chemotherapy.

#### The prognostic risk gene ARNTL2 affects the prognosis of LUAD patients and involves the activation of the PI3K/AKT/mTOR signaling pathway

Aryl hydrocarbon receptor nuclear translocator like 2 (ARNTL2) serves as a transcription factor linked to the prognosis and metastasis of various tumors and is one of the PAS superfamily [40, 41]. Previous research has indicated that ARNTL2 correlates with poor survival rates and levels of immune infiltration in breast cancer, LUAD, and clear cell renal cell carcinoma [42–44].

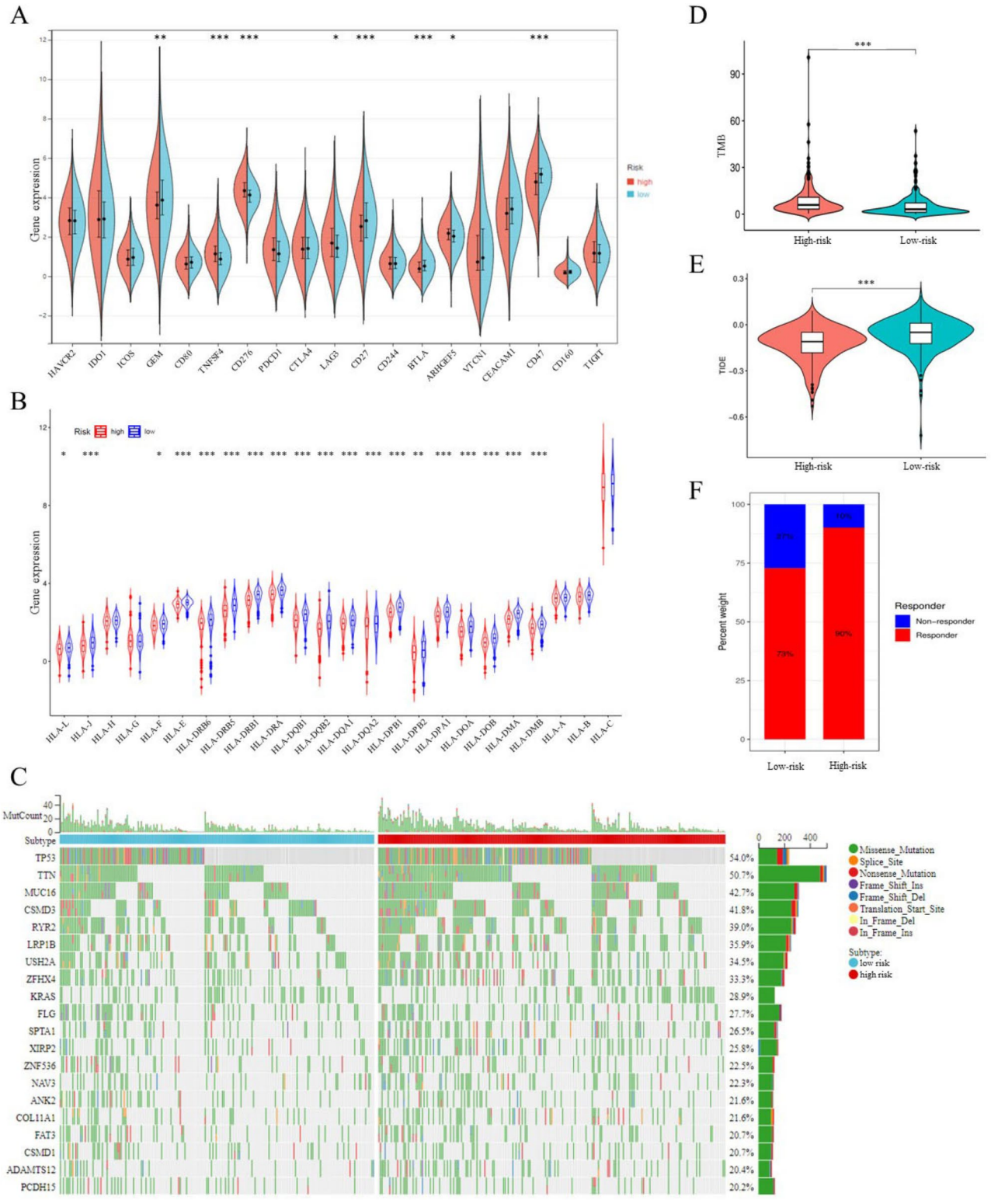


**Fig. 8** Analysis of immune checkpoints, HLA, TMB, and TIDE. **A** It illustrates the differences in immune checkpoint gene expression between high- and low-risk groups. **B** This panel compares the expression levels of HLA members across the two risk groups. **C** Mutation mapping of LUAD patients highlights the top 20 genes with the highest mutation frequencies, differentiated by risk group. **D** A comparison of tumor mutation burden (TMB) among distinct risk groups is shown. **E**, **F** These panels display the TIDE values and immunotherapy response results for LUAD patients categorized by risk group. HLA, human leukocyte antigen; TMB, tumor mutation burden; TIDE, tumor immune dysfunction and exclusion. Statistical significance is indicated as $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

Our observations also highlighted a significant impact of ARNTL2 on the OS of LUAD patients (Fig. 5A). Furthermore, analyses from multiple independent GEO databases (GSE68465, GSE37745, GSE30219, GSE31210, and GSE67639) reinforced the notion that ARNTL2 acts as a negative regulatory factor for survival among patients with LUAD, utilizing the LOGpc online tool (https://bioinfo.henu.edu.cn/DatabaseList.jsp) (Fig. S6A-E). Nonetheless, the molecular mechanisms underlying the impact of ARNTL2 on malignant biological behaviors and the prognosis of LUAD are currently unclear. Therefore, ARNTL2 was chosen for further investigation.

We first investigated the potential roles of ARNTL2 in individual LUAD cells utilizing the CancerSEA database (http://biocc.hrbmu.edu.cn/CancerSEA/). Our analysis revealed that the functions of ARNTL2 were primarily linked to angiogenesis, differentiation, epithelial-mesenchymal transition (EMT), invasion, metastasis, quiescence, and stemness (Fig. 9A). Research by Kim KT (EXP0066) indicated a positive correlation between elevated ARNTL2 expression and metastasis, angiogenesis, and quiescence, with Spearman coefficients of 0.72, 0.62, and 0.48, respectively ($P<0.001$) (Fig. 9B). Moreover, tumor samples from patient-derived xenografts (LC-PT-45) also showed enrichment in functions related to DNA damage, metastasis, and invasion (Fig. 9C). Through GSEA analysis, we identified several crucial regulatory genes linked to the PI3K/AKT/mTOR signaling pathway that were consistently enriched in LUAD patients demonstrating high ARNTL2 expression across the TCGA_LUAD, GSE68465, and GSE31210 databases, suggesting a positive relationship between ARNTL2 expression and the PI3K/AKT/mTOR signaling pathway (Fig. 9D). Subsequently, we assessed the protein levels of ARNTL2 in the A549, H1373, H1573, H1299, and BEAS-2B cell lines using Western blotting. Notably, ARNTL2 was significantly overexpressed in A549 and H1299 cells compared to normal lung tissues and the BEAS-2B cell line (Fig. 9E). To silence endogenous ARNTL2, we performed lentiviral transfection of shRNA targeted at ARNTL2 in A549 and H1299 cells. Our results demonstrated that ARNTL2 knockdown significantly reduced the migration of both cell lines in transwell assays, compared to the control group (Fig. 9F, G). Furthermore, an in vivo lung metastasis assay indicated that mice with A549/sh2ARNTL2 exhibited fewer pulmonary metastasis nodules and weight of lungs than the mock control (Fig. S6F and Fig. 9H-J). Additionally, analysis using Western blot demonstrated that the knockdown of ARNTL2 resulted in reduced levels of p-PI3K$^{Tyr458}$, p-AKT$^{Ser473}$, and p-mTOR$^{Ser2448}$ in A549 and H1299 cell lines (Fig. 9K). Prior research indicated a strong correlation between the activation of PI3K/AKT/mTOR pathway and the malignant transformation of various tumors, along with the metastatic behavior of cancer cells [45–47]. Consequently, we investigated the expression levels of crucial proteins linked to epithelial-mesenchymal transition (EMT). Our findings revealed that the silencing of ARNTL2 in A549 and H1299 cells significantly lowered the expression of N-cadherin and Vimentin, while simultaneously elevating the levels of E-cadherin and γ-catenin (Fig. 9K). These results imply that ARNTL2 might influence several biological characteristics of lung adenocarcinoma cells through the activation of the PI3K/AKT/mTOR signaling pathway.

## Discussion

Metabolic reprogramming has emerged as a critical characteristic of cancer cells in recent times. Once believed to be solely a consequence of accelerated cell division, new evidence has changed this perspective by showing that metabolic reprogramming can actually promote tumor formation [48]. Cancer stem cells (CSCs) represent a distinct subset of tumor cells endowed with stem-like features that enable them to persist despite conventional therapeutic approaches, which contributes to the development of metastatic disease and tumor relapse [49, 50]. Given the diverse prognoses among patients with LUAD and the interdependent relationship between metabolic reprogramming and CSCs in the tumor microenvironment, creating an effective classification system to stratify patients by varying risk levels based on msRGs would be advantageous. Furthermore, such a system could facilitate a thorough evaluation of prognosis, immune status, and therapeutic responses in LUAD patients. This study marks the first systematic examination of the expression levels of 415 msRGs in LUAD tissues. Through a multistep selection process, a prognostic risk model including five DE-msRGs was developed that is significantly linked to OS, responses to immunotherapy and chemotherapy, as well as the tumor microenvironment of LUAD patients, potentially serving as a valuable tool for assessing the effectiveness of immunotherapy and chemotherapy in this context.

In recent years, a growing body of research has demonstrated that metabolic reprogramming significantly influences the growth and function of CSCs [51, 52]. As reported by Park et al., elevated levels of pyruvate kinase M2 (PKM2), a vital enzyme within the glycolytic metabolic pathway, frequently correlate with tumor invasion and unfavorable patient outcomes across various cancers. This relationship arises from PKM2's involvement in glycolytic metabolic reprogramming, which enhances the survival and proliferation of tumor stem cells [53]. In a study conducted in 2022, Luo et al. utilized genetic mouse mammary tumor models alongside human breast cancer samples to identify that the histone reader ZMYND8 was uniquely expressed in breast cancer stem cells (BCSCs).

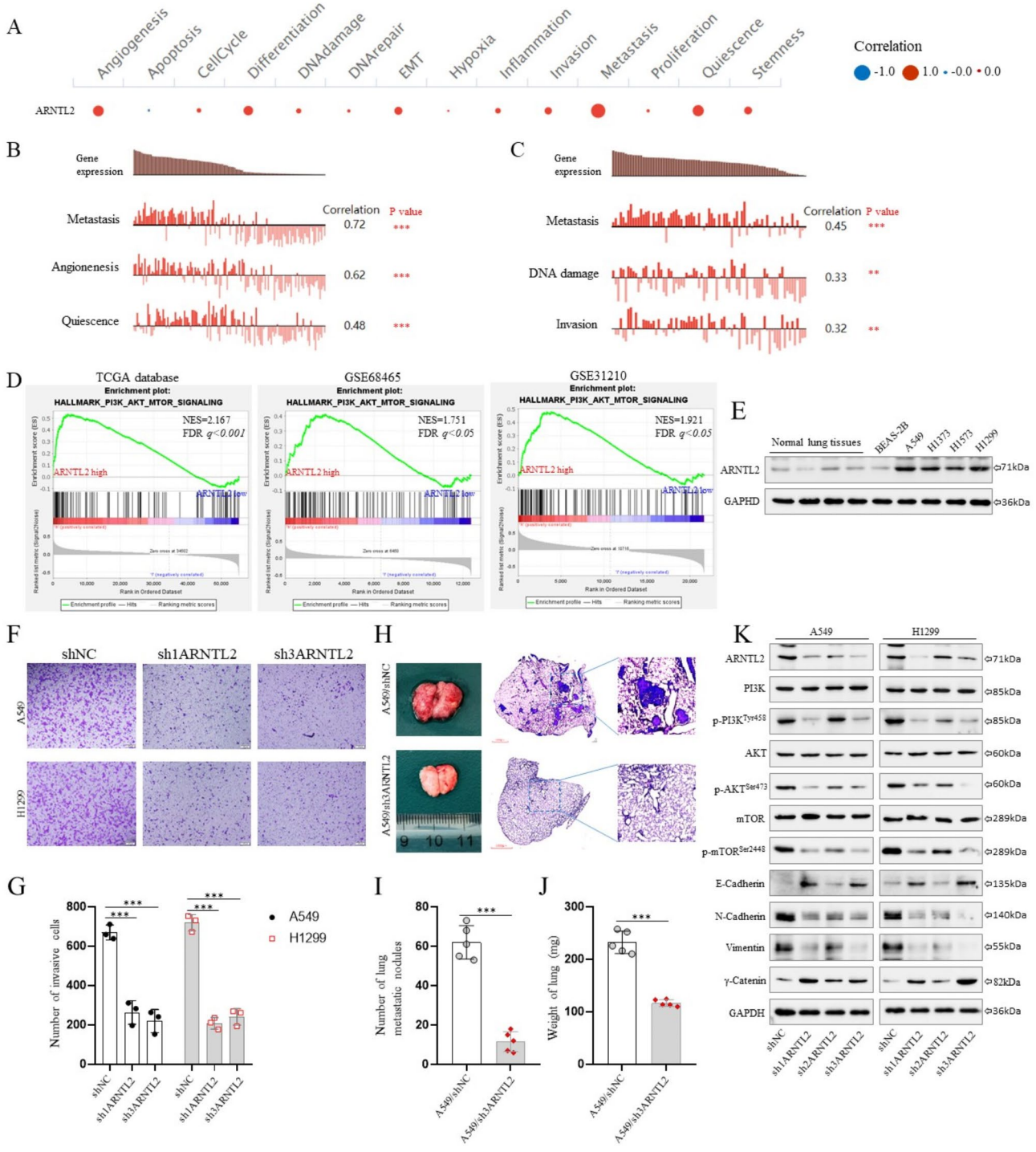


**Fig. 9** Functional analysis and validation of ARNTL2. **A** The correlation of ARNTL2 with functional state in LUAD based on the CancerSEA database is illustrated through an interactive bubble chart. **B**, **C** Detailed functional correlations are presented for the LUAD chip (**B**) and a patient-derived xenograft model (**C**). **D** Enrichment analysis was conducted using Gene Set Enrichment Analysis (GSEA) to compare high and low expression levels of ARNTL2 across the TCGA, GSE68465, and GSE31210 databases. **E** Western blot analysis was performed to assess ARNTL2 protein expression levels in normal lung tissues and the indicated cell lines. **F**, **G** A Transwell assay was utilized to evaluate the invasive ability of A549 and H1299 cells following ARNTL2 silencing, with a scale bar of 100 μm. **H-J** The impact of ARNTL2 silencing on invasion and metastasis in A549 cells was analyzed using a lung metastasis model. The metastatic tumor lesions in each mouse lung were assessed through Hematoxylin and eosin (H&E) staining (**H**). Representative images of H&E-stained lung sections from mice injected intravenously with the indicated cells are shown (**H**). The number of lung metastatic nodules (**I**) and the lung weight of mice (**J**) in the indicated groups were measured and analyzed. **K** Western blot analysis was conducted to evaluate the expression levels of the indicated proteins. The target bands are derived from the corresponding region of the original blot images. $^{**}p < 0.01$, $^{***}p < 0.001$

This protein facilitates the epithelial-mesenchymal transition (EMT), supports BCSC preservation and self-renewal, and contributes to oncogenic transformation by elevating cholesterol biosynthesis and oxidation while inhibiting cholesterol efflux and the catabolism of 27-Hydroxycholesterol (27-HC). This imbalance results in the accumulation of 27-HC within BCSCs, ultimately initiating breast tumors [54]. Furthermore, findings by Lv et al. indicated that specifically knocking down HIF-1α to target hypoxia not only restrains cell proliferation and spheroid formation but also reduces the expression of CSC-related genes and diminishes the activity of the Wnt/β-catenin signaling pathway [55]. Despite these insights, there remains limited understanding of metabolism-stemness-related genes (msRGs) and their specific contributions in LUAD. In the current study, we conducted a comprehensive bioinformatics analysis using the TCGA_LUAD dataset to assess metabolism-related and stemness-related genes. We identified a total of 81 genes, including ABCA8, ADH1B, CX3CR1, OIP5, among others, that exhibited significant associations with OS in LUAD patients. Following this, a five-gene prognostic risk signature comprising S100P, GPX2, PRC1, ARNTL2, and RGS20 was created through LASSO regression and multivariate Cox regression analysis. To the best of our knowledge, this represents the first prognostic gene signature linked to DE-msRGs in LUAD. This five-gene model may offer a novel approach for assessing patients with LUAD, aiding in prognosis prediction as well as influencing decisions regarding immunotherapy and chemotherapy treatments.

Previous studies have underscored the critical importance of risk signature genes identified in this research across multiple cancer types. S100P, which consists of 95 amino acids and belongs to the S100 protein family, is found to be overexpressed in various solid tumors [56–58]. Experimental findings indicate that elevated levels of S100P can facilitate cancer progression and reduce patient survival rates through its specific functions in cell proliferation, angiogenesis, and metastasis [59, 60]. GPX2 (Glutathione peroxidase 2), an enzyme that acts as an antioxidant, is predominantly expressed in the gastrointestinal system, particularly within the esophagus and liver [61]. This gene is also found to be upregulated in several cancers, including colorectal, gastric, esophageal, and breast cancer [62–66]. A recent study revealed that GPX2 contributes to the epithelial-mesenchymal transition (EMT) and invasion of non-small cell lung cancer (NSCLC) cells by activating the PI3K/AKT/mTOR/Snail signaling pathway through the elimination of reactive oxygen species (ROS) [67]. PRC1 (protein regulator of cytokinesis-1) acts as a factor associated with microtubules and plays a role in cytokinesis. Its overexpression has been shown to significantly enhance both the proliferation and metastasis of hepatocellular carcinoma cells, correlating with early recurrence and unfavorable patient outcomes by modulating the oncogenic effects of the Wnt signaling pathway [68]. Conversely, silencing PRC1 has been found to markedly decrease monolayer colony formation and inhibit the proliferation, invasion, and migratory capacities of gastric carcinoma cells [69]. Additionally, another risk gene, ARNTL2 (aryl hydrocarbon receptor nuclear translocator like 2), functions as a biologically relevant partner of circadian and hypoxia factors [70, 71]. Increasing evidence indicates that ARNTL2 is closely linked with poor prognosis in diverse tumors, contributing to the activation of the PI3K/AKT signaling pathway as well as the TGF/BETA pathway, while also inhibiting both apoptosis and ferroptosis [72–74]. Moreover, RGS20, a crucial regulator of neuronal G protein-coupled receptor signaling pathways within the brain, belongs to the RGS subfamily. Elevated expression of RGS20 has been linked to nodal metastasis in triple-negative breast cancer and promotes NSCLC cell growth by initiating autophagy through the suppression of PKA-Hippo signaling [75, 76]. These findings align with the results of the current study, where S100P, GPX2, PRC1, and ARNTL2 were shown to be upregulated in LUAD tissues. Furthermore, the overexpression of S100P, PRC1, ARNTL2, and RGS20 demonstrated a significant correlation with poor clinical outcomes in LUAD patients. However, the above-mentioned signature genes have seldom been investigated regarding their roles in the context of combined metabolism and cancer stemness. In the present research, a new prognostic risk model featuring five genes, developed through extensive bioinformatics analysis and experimental investigations, may offer insights for novel molecular subtyping and prognostic evaluation of LUAD, as well as for the personalized treatment and management of patients in the future.

The signaling pathway involving phosphatidylinositol 3-kinase (PI3K), AKT, and the mammalian target of rapamycin (mTOR) is crucial for cellular survival and growth, often showing disruptions in malignant diseases [77]. This pathway is influenced by a variety of upstream signaling proteins and, in turn, regulates numerous downstream effectors through interactions with different compensatory signaling pathways, notably the RAF/MEK/ERK pathway [78, 79]. Currently, the limited clinical effectiveness of existing targeted therapies, along with the challenges posed by tumour heterogeneity in various cancer types, highlights the necessity for a deeper understanding of the regulatory mechanisms governing the PI3K/AKT/mTOR pathway. This knowledge is essential for formulating effective personalized treatment approaches. In our study, we found that heightened expression of ARNTL2 was significantly associated with the activation of the PI3K/AKT/mTOR pathway,

as demonstrated by GSEA analysis. The knockdown of ARNTL2 resulted in a decrease in the phosphorylation levels of PI3K, AKT, and mTOR, while promoting the expression of E-cadherin and γ-catenin, and diminishing levels of N-cadherin and Vimentin in H1299 and A549 cell lines. Our results indicate that the activation of the PI3K/AKT/mTOR signaling pathway may be a key mechanism in the advancement of tumors characterized by ARNTL2 overexpression. However, it is important to acknowledge the limitations of the current work. Our research predominantly relies on retrospective analyses of publicly available databases, which may introduce bias, and features limited in vitro and in vivo experiments. If circumstances permit in the future, the subsequent integration of single-cell sequencing, spatiotemporal sequencing, large-scale tissue chip analysis, and serological detection of prognostic risk genes will aid in the validation and enhancement of the stability and reliability of risk models. Furthermore, the robustness and reproducibility of this five-gene risk signature warrant further validation in extensive prospective real-world studies, utilizing predictive algorithms in FET biosensors [80–82].

## Conclusion

In this research, we successfully developed a novel risk stratification model that correlates with metabolism and stemness-related genes, derived from comprehensive analysis and experimental research. This model has the potential to enhance clinical classification and management of lung adenocarcinoma, thereby improving patient prognosis.

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13984-6.

**Additional File 1**: **Table S1**. Oligo sequences used in quantitative real-time PCR. **Table S2**. Antibodies used in this study. **Table S3**. 1466 metabolism-related genes. **Table S4**. Expression matrix of 1411 motebolism-related genes based on TCGA_LUAD. **Table S5**. The results of NMF analysis based on 1411 metabolism-related genes. **Table S6**. Differentially expressed genes between subtype C1 and C2. **Table S7**. Stemness index-related differentially expressed genes. **Table S8**. 301 differentially expressed metabolism and stemness-related genes based on TCGA_LUAD database. **Table S9**. 62 DE-msRGs genes with mutation cases greater than 10 in the TCGA_LUAD database. **Table S10**. Copy number variation (CNV) amplifications and deletions of 62 DE-msRGs in LUAD patients. **Table S11**. Results of univariate Cox regression analyses from 301 DE-msRGs. **Table S12**. Results of LASSO Cox regression based on TCGA_LUAD database.

**Additional File 2**: **Fig. S1**. Developing the metabolism-related molecular patterns. **Fig. S2**. Landscape of genetic variations of metabolism and stemness-related genes (msRGs) in lung adenocarcinoma (LUAD). **Fig. S3**. Principal component analysis (PCA) and receive operating characteristic (ROC) curve analysis based on the TCGA cohort and GEO-merged cohort. **Fig. S4**. Correlation analysis between prognostic risk genes and immune-infiltrating levels based on GSCA database. **Fig. S5**. The anticipated reactions of patients in the TCGA database to targeted treatments and conventional chemotherapy across various risk categories were analyzed. **Fig. S6**. The Kaplan-Meier plot depicts the relationship between the expression levels of ARNTL2 and OS in LUAD patients, and an in vivo lung metastasis assay.

**Additional File 3** Supplementary materials and methods.

**Additional File 4** Original blots of Fig. 9.

### Acknowledgements

No.

### Author contributions

Feng Lu conceived and supervised this study. Wanrong Zheng, Chuchu Zhou, and Zixin Xue participated in the experimental data collection; Ling Qiao and Jianjun Wang provided technical assistance; Feng Lu drafted and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

### Funding

This work was supported by the National Natural Science Foundation of China (No: 81372147) and Henan University support grant CX3070A0780502.

### Data availability

Publicly available datasets were analyzed in this study. These data can be found here: https://portal.gdc.cancer.gov/ and http://www.ncbi.nlm.nih.gov/geo/; TCGA_OV, GSE11969, GSE13213, GSE41271, GSE42127, GSE50081, GSE68465, and GSE72094.

## Declarations

### Ethics approval and consent to participate

Our research adheres to the Declaration of Helsinki. The studies involving human participants were reviewed and approved by the Ethics Committee of Medical School of Henan University, China (HUSOM-2018-282). Informed consent was obtained from all subjects involved in the study. The animal study was reviewed and approved by the Ethic Committee of Medical School of Henan University.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

[1]Department of Medical Oncology, Huaihe Hospital of Henan University, Kaifeng, China
[2]Department of Immunology, School of Basic Medical Sciences, Henan University, Kaifeng, China
[3]Joint National Laboratory for Antibody Drug Engineering, Henan University, Kaifeng, China

Received: 4 September 2024 / Accepted: 20 March 2025
Published online: 01 April 2025

### References

1. Lovly CM. Expanding horizons for treatment of Early-Stage lung cancer. N Engl J Med. 2022;386:2050–1.
2. Henderson LM, Su IH, Rivera MP, Pak J, Chen X, Reuland DS, et al. Prevalence of lung cancer screening in the US, 2022. JAMA Netw Open. 2024;7:e243190.
3. Wang Q, Tang Z, Li C, Li X, Su C. Evaluating distinct KRAS subtypes as potential biomarkers for immune checkpoint inhibitor efficacy in lung adenocarcinoma. Front Immunol. 2023;14:1297588.
4. Zavitsanou AM, Pillai R, Hao Y, Wu WL, Bartnicki E, Karakousi T, et al. KEAP1 mutation in lung adenocarcinoma promotes immune evasion and immunotherapy resistance. Cell Rep. 2023;42:113295.
5. Das A, Rewari S, Kanaujia BK, Deswal SS, Gupta RS. Analytical investigation of a triple surrounding gate germanium source metal–oxide–semiconductor

field-effect transistor with step graded channel for biosensing applications. Int J Numer Model Electron Networks Devices Fields 2023; 36.
6. Bhardwaj A, Solay LR, Kumar N, Amin SI, Singh A, Raj B, et al. Doping-less TFET based common source amplifier implementation and behaviour analysis under symmetric and asymmetric conditions. Silicon. 2022;14:12251–60.
7. Das A, Rewari S, Kanaujia BK, Deswal SS, Gupta RS. Numerical modeling of a dielectric modulated surrounding-triple-gate germanium-source MOSFET (DM-STGGS-MOSFET)-based biosensor. J Comput Electron. 2023;22:742–59.
8. Nguyen TT, Nguyen CM, Huynh MA, Vu HH, Nguyen TK, Nguyen NT. Field effect transistor based wearable biosensors for healthcare monitoring. J Nanobiotechnol. 2023;21:411.
9. Mao S, Chang J, Pu H, Lu G, He Q, Zhang H, et al. Two-dimensional nanomaterial-based field-effect transistors for chemical and biological sensing. Chem Soc Rev. 2017;46:6872–904.
10. Mandal N, Pakira V, Samanta N, Das N, Chakraborty S, Pramanick B, et al. PSA detection using label free graphene FET with coplanar electrodes based microfluidic point of care diagnostic device. Talanta. 2021;222:121581.
11. Hwang MT, Heiranian M, Kim Y, You S, Leem J, Taqieddin A, et al. Ultrasensitive detection of nucleic acids using deformed graphene channel field effect biosensors. Nat Commun. 2020;11:1543.
12. Sun Z, Zhang R, Zhang X, Sun Y, Liu P, Francoeur N, et al. LINE-1 promotes tumorigenicity and exacerbates tumor progression via stimulating metabolism reprogramming in non-small cell lung cancer. Mol Cancer. 2022;21:147.
13. Sun L, Zhang H, Gao P. Metabolic reprogramming and epigenetic modifications on the path to cancer. Protein Cell. 2022;13:877–919.
14. Zhou H, Tan L, Liu B, Guan XY. Cancer stem cells: recent insights and therapies. Biochem Pharmacol. 2023;209:115441.
15. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. J Biomed Sci. 2018;25:20.
16. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. Cell. 2018;173:338–e354315.
17. Zhang Y, Liu D, Li F, Zhao Z, Liu X, Gao D, et al. Identification of biomarkers for acute leukemia via machine learning-based stemness index. Gene. 2021;804:145903.
18. Tan J, Zhu H, Tang G, Liu H, Wanggou S, Cao Y, et al. Molecular subtypes based on the stemness index predict prognosis in glioma patients. Front Genet. 2021;12:616507.
19. Wang Z, Yang L, Huang Z, Li X, Xiao J, Qu Y, et al. Identification of prognosis biomarkers for High-Grade serous ovarian cancer based on stemness. Front Genet. 2022;13:861954.
20. Zhang Y, Tseng JT, Lien IC, Li F, Wu W, Li H. mRNAsi index: machine learning in mining lung adenocarcinoma stem cell biomarkers. Genes (Basel) 2020; 11.
21. Tang R, Liu X, Wang W, Hua J, Xu J, Liang C, et al. Identification of the roles of a stemness index based on mRNA expression in the prognosis and metabolic reprograming of pancreatic ductal adenocarcinoma. Front Oncol. 2021;11:643465.
22. Sun L, Suo C, Li ST, Zhang H, Gao P. Metabolic reprogramming for cancer cells and their microenvironment: beyond the Warburg effect. Biochim Biophys Acta Rev Cancer. 2018;1870:51–66.
23. El Hout M, Cosialls E, Mehrpour M, Hamaï A. Crosstalk between autophagy and metabolic regulation of cancer stem cells. Mol Cancer. 2020;19:27.
24. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy–analysis of affymetrix genechip data at the probe level. Bioinformatics. 2004;20:307–15.
25. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The Sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28:882–3.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–50.
27. He Z, Wang C, Xue H, Zhao R, Li G. Identification of a Metabolism-Related risk signature associated with clinical prognosis in glioblastoma using integrated bioinformatic analysis. Front Oncol. 2020;10:1631.
28. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16:284–7.
29. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. J Natl Cancer Inst. 1998;90:766–71.
30. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak. 2006;26:565–74.
31. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
32. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7.
33. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218.
34. Zheng H, Liu H, Ge Y, Wang X. Integrated single-cell and bulk RNA sequencing analysis identifies a cancer associated fibroblast-related signature for predicting prognosis and therapeutic responses in colorectal cancer. Cancer Cell Int. 2021;21:552.
35. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41:D955–961.
36. Lu F, Zhou J, Chen Q, Zhu J, Zheng X, Fang N, et al. PSMA5 contributes to progression of lung adenocarcinoma in association with the JAK/STAT pathway. Carcinogenesis. 2022;43:624–34.
37. Xie S, Jin Y, Wang J, Li J, Peng M, Zhu X. DOCK1 regulates the malignant biological behavior of endometrial cancer through c-Raf/ERK pathway. BMC Cancer. 2024;24:296.
38. Ren G, Li H, Hong D, Hu F, Jin R, Wu S, et al. LINC00955 suppresses colorectal cancer growth by acting as a molecular scaffold of TRIM25 and Sp1 to inhibit DNMT3B-mediated methylation of the PHIP promoter. BMC Cancer. 2023;23:898.
39. Yang Z, Zhu J, Yang T, Tang W, Zheng X, Ji S, et al. Comprehensive analysis of the lncRNAs-related immune gene signatures and their correlation with immunotherapy in lung adenocarcinoma. Br J Cancer. 2023;129:1397–408.
40. Mazzoccoli G, Pazienza V, Panza A, Valvano MR, Benegiamo G, Vinciguerra M, et al. ARNTL2 and SERPINE1: potential biomarkers for tumor aggressiveness in colorectal cancer. J Cancer Res Clin Oncol. 2012;138:501–11.
41. Ha NH, Long J, Cai Q, Shu XO, Hunter KW. The circadian rhythm gene Arntl2 is a metastasis susceptibility gene for Estrogen Receptor-Negative breast cancer. PLoS Genet. 2016;12:e1006267.
42. Wang X, Li Y, Fu J, Zhou K, Wang T. ARNTL2 is a prognostic biomarker and correlates with immune cell infiltration in Triple-Negative breast cancer. Pharmgenomics Pers Med. 2021;14:1425–40.
43. Zhang H, Shan G, Jin X, Yu X, Bi G, Feng M, et al. ARNTL2 is an indicator of poor prognosis, promotes epithelial-to-mesenchymal transition and inhibits ferroptosis in lung adenocarcinoma. Transl Oncol. 2022;26:101562.
44. Wang S, Ma X, Ying Y, Sun J, Yang Z, Li J, et al. Upregulation of ARNTL2 is associated with poor survival and immune infiltration in clear cell renal cell carcinoma. Cancer Cell Int. 2021;21:341.
45. Li Y, Zhou Y, Huang M, Wang Z, Liu D, Liu J, et al. DHCR7 promotes tumorigenesis via activating PI3K/AKT/mTOR signalling pathway in bladder cancer. Cell Signal. 2023;102:110553.
46. Shi L, Wang X, Guo S, Gou H, Shang H, Jiang X, et al. TMEM65 promotes gastric tumorigenesis by targeting YWHAZ to activate PI3K-Akt-mTOR pathway and is a therapeutic target. Oncogene. 2024;43:931–43.
47. Mabuchi S, Kuroda H, Takahashi R, Sasano T. The PI3K/AKT/mTOR pathway as a therapeutic target in ovarian cancer. Gynecol Oncol. 2015;137:173–9.
48. Sebastian C. Tracking down the origin of cancer: metabolic reprogramming as a driver of stemness and tumorigenesis. Crit Rev Oncog. 2014;19:363–82.
49. Hua Z, White J, Zhou J. Cancer stem cells in TNBC. Semin Cancer Biol. 2022;82:26–34.
50. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature. 2001;414:105–11.
51. Chong KH, Chang YJ, Hsu WH, Tu YT, Chen YR, Lee MC et al. Breast cancer with increased drug resistance, invasion ability, and cancer stem cell properties through metabolism reprogramming. Int J Mol Sci 2022; 23.
52. Yao X, Li W, Li L, Li M, Zhao Y, Fang D, et al. YTHDF1 upregulation mediates hypoxia-dependent breast cancer growth and metastasis through regulating PKM2 to affect Glycolysis. Cell Death Dis. 2022;13:258.
53. Park JH, Pyun WY, Park HW. Cancer metabolism: phenotype, signaling and therapeutic targets. Cells 2020; 9.
54. Luo M, Bao L, Chen Y, Xue Y, Wang Y, Zhang B, et al. ZMYND8 is a master regulator of 27-hydroxycholesterol that promotes tumorigenicity of breast cancer stem cells. Sci Adv. 2022;8:eabn5295.
55. Lv Z, Liu RD, Chen XQ, Wang B, Li LF, Guo YS, et al. HIF–1α promotes the stemness of oesophageal squamous cell carcinoma by activating the Wnt/β–catenin pathway. Oncol Rep. 2019;42:726–34.

56. Guerreiro Da Silva ID, Hu YF, Russo IH, Ao X, Salicioni AM, Yang X, et al. S100P calcium-binding protein overexpression is associated with immortalization of human breast epithelial cells in vitro and early stages of breast cancer development in vivo. Int J Oncol. 2000;16:231–40.
57. Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med. 2002;8:816–24.
58. Hamada S, Satoh K, Hirota M, Kanno A, Ishida K, Umino J, et al. Calcium-binding protein S100P is a novel diagnostic marker of cholangiocarcinoma. Cancer Sci. 2011;102:150–6.
59. Basu GD, Azorsa DO, Kiefer JA, Rojas AM, Tuzmen S, Barrett MT, et al. Functional evidence implicating S100P in prostate cancer progression. Int J Cancer. 2008;123:330–9.
60. Diederichs S, Bulk E, Steffen B, Ji P, Tickenbrock L, Lang K, et al. S100 family members and trypsinogens are predictors of distant metastasis and survival in early-stage non-small cell lung cancer. Cancer Res. 2004;64:5564–9.
61. Chu FF, Doroshow JH, Esworthy RS. Expression, characterization, and tissue distribution of a new cellular selenium-dependent glutathione peroxidase, GSHPx-GI. J Biol Chem. 1993;268:2571–6.
62. Liu D, Sun L, Tong J, Chen X, Li H, Zhang Q. Prognostic significance of glutathione peroxidase 2 in gastric carcinoma. Tumour Biol. 2017;39:1010428317701443.
63. Murawaki Y, Tsuchiya H, Kanbe T, Harada K, Yashima K, Nozaka K, et al. Aberrant expression of selenoproteins in the progression of colorectal cancer. Cancer Lett. 2008;259:218–30.
64. Lei Z, Tian D, Zhang C, Zhao S, Su M. Clinicopathological and prognostic significance of GPX2 protein expression in esophageal squamous cell carcinoma. BMC Cancer. 2016;16:410.
65. Suzuki S, Pitchakarn P, Ogawa K, Naiki-Ito A, Chewonarin T, Punfa W, et al. Expression of glutathione peroxidase 2 is associated with not only early hepatocarcinogenesis but also late stage metastasis. Toxicology. 2013;311:115–23.
66. Naiki-Ito A, Asamoto M, Hokaiwado N, Takahashi S, Yamashita H, Tsuda H, et al. Gpx2 is an overexpressed gene in rat breast cancers induced by three different chemical carcinogens. Cancer Res. 2007;67:11353–8.
67. Peng F, Xu Q, Jing X, Chi X, Zhang Z, Meng X, et al. GPX2 promotes EMT and metastasis in non-small cell lung cancer by activating PI3K/AKT/mTOR/Snail signaling axis. FASEB Bioadv. 2023;5:233–50.
68. Chen J, Rajasekaran M, Xia H, Zhang X, Kong SN, Sekar K, et al. The microtubule-associated protein PRC1 promotes early recurrence of hepatocellular carcinoma in association with the Wnt/β-catenin signalling pathway. Gut. 2016;65:1522–34.
69. Zhang B, Shi X, Xu G, Kang W, Zhang W, Zhang S, et al. Elevated PRC1 in gastric carcinoma exerts oncogenic function and is targeted by Piperlongumine in a p53-dependent manner. J Cell Mol Med. 2017;21:1329–41.
70. Ikeda M, Yu W, Hirai M, Ebisawa T, Honma S, Yoshimura K, et al. cDNA cloning of a novel bHLH-PAS transcription factor superfamily gene, BMAL2: its mRNA expression, subcellular distribution, and chromosomal localization. Biochem Biophys Res Commun. 2000;275:493–502.
71. Lebailly B, Langa F, Boitard C, Avner P, Rogner UC. The circadian gene Arntl2 on distal mouse chromosome 6 controls thymocyte apoptosis. Mamm Genome. 2017;28:1–12.
72. Ge W, Wang Y, Quan M, Mao T, Bischof EY, Xu H, et al. Activation of the PI3K/AKT signaling pathway by ARNTL2 enhances cellular Glycolysis and sensitizes pancreatic adenocarcinoma to erlotinib. Mol Cancer. 2024;23:48.
73. Wang Z, Liu T, Xue W, Fang Y, Chen X, Xu L, et al. ARNTL2 promotes pancreatic ductal adenocarcinoma progression through TGF/BETA pathway and is regulated by miR-26a-5p. Cell Death Dis. 2020;11:692.
74. Wang T, Wang K, Zhu X, Chen N. ARNTL2 upregulation of ACOT7 promotes NSCLC cell proliferation through Inhibition of apoptosis and ferroptosis. BMC Mol Cell Biol. 2023;24:14.
75. Li Q, Jin W, Cai Y, Yang F, Chen E, Ye D, et al. Regulator of G protein signaling 20 correlates with clinicopathological features and prognosis in triple-negative breast cancer. Biochem Biophys Res Commun. 2017;485:693–7.
76. Ding X, Li X, Jiang Y, Li Y, Li H, Shang L, et al. RGS20 promotes non-small cell lung carcinoma proliferation via autophagy activation and Inhibition of the PKA-Hippo signaling pathway. Cancer Cell Int. 2024;24:93.
77. Yu L, Wei J, Liu P. Attacking the PI3K/Akt/mTOR signaling pathway for targeted therapeutic treatment in human cancer. Semin Cancer Biol. 2022;85:69–94.
78. Polivka J Jr., Janku F. Molecular targets for cancer therapy in the PI3K/AKT/mTOR pathway. Pharmacol Ther. 2014;142:164–75.
79. Badoiu SC, Greabu M, Miricescu D, Stanescu S, II, Ilinca R, Balan DG et al. PI3K/AKT/mTOR dysregulation and reprogramming metabolic pathways in renal cancer: crosstalk with the VHL/HIF axis. Int J Mol Sci 2023; 24.
80. Das A, Rewari S, Kanaujia BK, Gupta RS. Recent technological advancement in surrounding gate MOSFET for biosensing Applications - a synoptic study. Silicon. 2021;14:5133–43.
81. Bhardwaj A, Kumar P, Raj B, Anand S. Design and performance analysis of a GAA electrostatic doped negative capacitance vertical nanowire tunnel FET. J Electron Mater. 2023;52:3103–11.
82. Bhardwaj A, Kumar P, Raj B, Anand S. Design and performance optimization of dopingless vertical nanowire TFET using gate stacking technique. J Electron Mater. 2022;51:4005–13.


## Publisher's note