Translational Oncology 66 (2026) 102699

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Integrative bioinformatics identifies NSCLC biomarkers associated with LPS metabolism and circadian disruption

Yu Chen [a,1], Yifan Zhao [b,1], Qinghua Kong [c,1], Yabin Li [d], Ping Jiang [a], Lichao Fan [a,*]

[a] Department of Respiratory and Critical Care Medicine, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China
[b] Department of Pharmacy, Suzhou Kowloon Hospital, Shanghai Jiao Tong University School of Medicine, Suzhou, China
[c] Department of Respiratory and Critical Care Medicine, Dahua Hospital, Shanghai, China
[d] School of Life Science and Technology, Tongji University, Shanghai, China

ARTICLE INFO

Keywords:
Non-small cell lung cancer (NSCLC)
Lipopolysaccharide (LPS) metabolism
Circadian rhythm disruption
Tumor immune microenvironment
Targeted therapy

ABSTRACT

Purpose: Non-small cell lung carcinoma (NSCLC) is a leading cause of cancer-related mortality worldwide, highlighting the urgent need for early detection and targeted therapies. While lipopolysaccharide (LPS) metabolism and circadian rhythm disruption are emerging as important factors in cancer progression, their specific roles in NSCLC remain poorly understood.

Methods: We integrated multiple GEO datasets to identify NSCLC-associated differentially expressed genes (DEGs). Weighted gene co-expression network analysis (WGCNA) identified key gene modules, followed by functional enrichment analysis. A hybrid machine learning approach combining Lasso regression and random forest was used to identify hub genes. Immune infiltration analysis evaluated associations with the tumor microenvironment, and diagnostic performance was validated in an independent cohort. Functional roles of the candidate gene CACNA2D2 were assessed through gain- and loss-of-function experiments in A549 cells, evaluating viability, proliferation, migration, and invasion.

Results: We identified 889 significant DEGs enriched in inflammatory and immune-related pathways. WGCNA revealed the magenta module as highly associated with NSCLC, involved in angiogenesis and extracellular matrix organization. Machine learning identified nine hub genes (CACNA2D2, ASPA, LRRN3, ABCA6, TNFSF12, AHNAK, TACC1, ID4, TSLP) showing excellent diagnostic performance (AUC: 0.832–0.906). These genes correlated significantly with immune cell infiltration patterns. Functional validation established CACNA2D2 as a tumor suppressor, where its depletion enhanced malignant phenotypes while overexpression suppressed them.

Conclusion: This study identifies a novel gene signature linked to LPS metabolism and circadian disruption in NSCLC, with validated diagnostic utility and implications for tumor immune regulation. CACNA2D2 emerges as a key tumor suppressor, offering insights for early detection and targeted therapy development.

Introduction

With persistently high mortality and increasing incidence rates, lung cancer represents a critical global health burden among all malignancies [1,2]. Non-small cell lung carcinoma (NSCLC) accounts for roughly 85 % of all diagnosed lung malignancies, with the remaining 15 % comprising small cell lung cancer (SCLC). The majority NSCLC subtype carries a generally poor prognosis [3,4]. Despite advances in targeted therapies and immunotherapy, most NSCLC patients present with advanced-stage disease due to diagnostic challenges, leading to suboptimal treatment outcomes [5,6]. This demonstrates the pressing demand for new molecular markers and personalized therapeutic interventions to boost survival.

Emerging evidence suggests that lipopolysaccharide (LPS)-related genes (LRGs) and circadian rhythm-related genes (CRGs) play critical roles in lung cancer pathogenesis and progression [7]. LPS metabolism not only regulates the immune response and tumor microenvironment, but also influences tumor immune escape and inflammatory responses,

* Corresponding author.
E-mail address: 1801041@tongji.edu.cn (L. Fan).
[1] These authors contributed equally.

https://doi.org/10.1016/j.tranon.2026.102699
Received 25 November 2025; Received in revised form 20 January 2026; Accepted 3 February 2026

thereby promoting cancer progression [8,9]. LPS enhances tumor cell invasiveness and metastasis by upregulating oncogenes and triggering local inflammation, particularly in NSCLC. In addition, LPS-related genes may further promote tumor growth and proliferation by inducing immune responses, altering cellular metabolism and regulating cytokine expression.

Emerging evidence suggests that the functional synergy between LPS-mediated inflammatory signaling and circadian pacing represents a critical 'two-hit' model in NSCLC development [10,11]. While LPS-TLR4 activation is fundamental to pulmonary microenvironmental remodeling, its signaling intensity is strictly governed by the circadian clock's transcriptional control over inflammatory effectors [12,13]. Physiological circadian gating ensures that inflammatory flares remain transient and localized. Conversely, circadian disruption in NSCLC patients eliminates this temporal oversight, shifting the balance toward a constitutive inflammatory state [14]. The convergence of dysregulated LPS metabolism and circadian desynchrony promotes a pro-tumorigenic environment characterized by persistent cytokine secretion, which facilitates immune evasion and accelerates the malignant transition of lung epithelia.

Meanwhile, circadian rhythm disruption (CRD) has been strongly linked to the development of various cancers [15,16]. Dysregulation of circadian rhythms, that orchestrate essential cellular functions such as mitotic progression, DNA damage response, and RNA biosynthesis, may affect these mechanisms, thereby promoting tumor development [17, 18]. Abnormal circadian rhythms may alter the growth pattern, division rate, and immune escape ability of tumor cells, which in turn exacerbate the malignant progression of cancer [19]. Although preliminary studies have explored the roles of LPS metabolism and circadian rhythms in lung cancer, systematic studies are still insufficient, and the specific mechanisms of how these genes interact with each other and act on tumors have not been clarified. Thus, systematic exploration of LPS metabolic pathways and CRG functions in NSCLC promises to reveal fundamental oncogenic mechanisms, also provide new strategies for early diagnosis, prognostic assessment and personalized treatment. By identifying key biomarkers of lung cancer, more precise therapeutic targets can be provided to clinics, enhancing therapeutic efficacy and improving patient survival prognosis

To systematically investigate the roles of LPS metabolism and circadian rhythm disruption in NSCLC, this investigation combines computational biology and artificial intelligence approaches to pinpoint essential gene networks and central regulators associated with NSCLC advancement. We further elucidate the biological functions of these hub genes and assess their diagnostic potential. By characterizing the expression patterns of LPS- and circadian-related genes and their influence on the tumor microenvironment, our findings provide novel insights for early detection and precision therapy in NSCLC.

Methods

1. Data Sourcing and Preprocessing

Gene expression data were extracted from multiple microarray datasets available in the Gene Expression Omnibus (GEO) repository. The search strategy utilized combinations of the following terms: "Non-small cell lung cancer," "NSCLC," "lung adenocarcinoma," and "lung squamous cell carcinoma," restricted to "Homo sapiens."

Specifically including: GSE103512 [20],GSE116959 [21], GSE140797 [22],GSE101929 [23],GSE40791 [24],GSE49644 [25], GSE54495 [26],GSE18842 [27]. The datasets were derived from multiple experimental platforms, with detailed platform specifications provided in Table 1. The datasets of GSE103512, GSE116959, GSE140797, GSE101929, GSE40791, and GSE49644 were employed as the training set to identify hub genes associated with NSCLC; GSE54495 and GSE18842 were used as the testing set to verify the differentiation ability of the hub genes.

Table 1
Summary of datasets analyzed in this study.

Datasets	Platform	Control	Disease	
GSE103512	GPL13158	9	60	Training datasets
GSE116959	GPL17077	11	57	
GSE140797	GPL13497	7	7	
GSE101929	GPL570	/	66	
GSE40791	GPL570	100	94	
GSE49644	GPL570	9	9	
GSE54495	GPL570	13	17	Test datasets
GSE18842	GPL570	45	46	

To ensure data quality, genes with zero or negative expression values were removed during the preprocessing stage. Additionally, mean expression values were computed across duplicate probes for each gene to reduce technical variability. Subsequently, the gene expression matrices of all samples were normalized to reduce the impact of technical errors on the analysis results. To evaluate sample distribution between healthy controls and lung cancer groups while addressing inter-dataset variability, we applied the ComBat algorithm for batch effect correction. This approach minimized non-biological differences between datasets, ensuring the reliability of subsequent analyses.

By collating the published literature, 6571 LRGs and 2091 CRGs (STable 1) were screened, which provided data for subsequent functional enrichment analysis and screening of key genes [28,29].

2. Identification of Differentially Expressed Genes (DEGs) in NSCLC

Following data normalization and batch effect correction, the identification of differentially expressed genes within the training dataset was performed utilizing the limma (v3.50.0) and surrogate variable analysis (v3.44.0) packages. DEGs were selected according to thresholds of $|\log_2 FC| > 1$ and an adjusted $P < 0.05$. For visualization, we employed ggplot2 (v3.3.6) to generate volcano plots highlighting significant up-/down-regulated genes, while pheatmap (v1.0.12) was used to construct clustered heatmaps displaying expression patterns across samples.

3. Identifying of key modules based on WGCNA

To elucidate the co-expression networks and identify disease-relevant gene modules, Weighted Gene Co-expression Network Analysis (WGCNA) was performed on the identified DEGs. Prior to network construction, a rigorous pre-filtering step was implemented to remove genes with low expression variance, ensuring the focus remained on biologically informative features. Sample quality was assessed using the goodSamplesGenes function. To ensure cohort homogeneity, hierarchical clustering of samples was conducted, and an outlier threshold of $h = 5$ was applied to eliminate abnormal samples.We determined the optimal soft-thresholding power (β) using the pickSoftThreshold function to satisfy the scale-free topology criterion. Based on a signed R^2 threshold of 0.8, a power of $\beta = 14$ was selected to construct the adjacency matrix. This matrix was subsequently transformed into a Topological Overlap Matrix (TOM) to quantify gene-gene co-expression proximity and mitigate technical noise.Module identification was performed using the cutreeDynamic algorithm with the following parameters: a minimum module size (minModuleSize) of 10 and a tree-cutting sensitivity (deepSplit) of 2. To simplify the network and consolidate biologically redundant clusters, similar modules were merged using a height cutoff of 0.35, which corresponds to a correlation coefficient of 0.65 between module eigengenes (MEs).

The association between identified modules and the NSCLC phenotype was quantified by calculating the correlation between MEs and clinical traits. Key modules strongly associated with NSCLC were prioritized through a combined analysis of Module Membership (MM)-the correlation between individual genes and their respective Mes-and Gene Significance (GS)-the correlation between gene expression and

clinical status. Modules exhibiting the highest significance were selected for subsequent hub gene identification and functional enrichment analysis to uncover the molecular mechanisms driving NSCLC progression.

4. Functional enrichment analysis

To systematically investigate the biological significance and mechanistic contributions of key genes in NSCLC pathogenesis, we performed comprehensive functional enrichment analysis using clusterProfiler (v4.12.6). Gene Ontology (GO) analysis systematically characterized DEGs across three domains: Biological Proces (BP), Molecular Functions (MF), and Cellular Components (CC). Parallel KEGG pathway analysis identified significantly enriched signaling pathways, revealing their potential involvement in NSCLC pathogenesis.

To evaluate genome-wide expression profiles, we conducted Gene Set Enrichment Analysis (GSEA) using the Hallmark gene set collection (v2023.1) from the Molecular Signatures Database (MSigDB) as the reference. Furthermore, Gene Set Variation Analysis (GSVA) was employed to quantify pathway activity at the individual sample level, transforming the gene-level expression matrix into a gene-set-level enrichment matrix.Based on the GSVA (v1.44.2) package, we utilized the gsvaParam function with the following prioritized settings: the Gaussian cumulative distribution function (kcdf = "Gaussian") was applied for the kernel density estimation, suitable for normalized expression data. The extension parameter was set to 1, and the maxDiff parameter was set to TRUE to calculate the enrichment scores as the difference between the largest and smallest vertical distances in the random walk. Gene sets were filtered to include those with a minSize of 1.

To identify significant functional pathways associated with NSCLC, the resulting GSVA score matrix was subjected to differential analysis using the limma package. A linear model was fitted to the scores, and empirical Bayes statistics were applied to determine the t-values and significance. Pathways with $P < 0.05$ and an absolute t-value threshold were considered significantly enriched, revealing the dynamic functional landscape across different pathological states.

5. Machine learning analytics

To identify NSCLC-related hub genes and develop diagnostic models, our model architecture incorporated Lasso regression to enhance feature interpretability while minimizing overfitting. Random Forest (RF) and XGBoost for capturing non-linear relationships, and Support Vector Machine (SVM) for classification, all models were subjected to 10-fold cross-validation during both training and validation phases using the training dataset to ensure robustness and generalizability.

System performance was holistically evaluated through receiver operating characteristic (ROC) curve analysis and area under the curve (AUC) metrics. The optimal algorithm (e.g., Elastic Net or XGBoost) demonstrating the highest AUC was selected to identify hub genes, which were subsequently validated in an external cohort datasets. To characterize the genomic distribution of these hub genes, we employed the circlize package to generate chromosomal location maps, visualizing their spatial organization across the genome.

To investigate the functional relevance of hub genes, we systematically constructed a protein-protein interaction (PPI) network using GeneMANIA to parse the regulatory mechanisms of key genes and provide potential biomarkers for NSCLC early diagnosis and targeted treatment.

6. Consensus clustering analysis

For further exploration of the molecular heterogeneity of NSCLC, we used the ConsensusClusterPlus (V1.58.0) in this study for consensus clustering analysis (Consensus Clustering). Employing gene expression signatures of prioritized core genes, with the k-means clustering (k-means) method adopted, the similarity between samples was measured by Euclidean Distance. To ensure result stability, the maximum cluster number (maxK) was set to 9, and 80 % of the samples (pItem=0.8) were randomly selected, with the procedure iterated 1000 times.

The effect under different number of clusters was evaluated by consensus clustering heat map with Cumulative Distribution Function curve. The optimal clustering number, determined through the analysis, was used to classify distinct molecular subtypes of NSCLC. Box line plots were generated to examine the expression variations of key genes across these subtypes, uncovering potential molecular characteristics of NSCLC.

7. XGBoost-based hub genes modeling and interpretive analysis

To evaluate the predictive importance of hub genes, we implemented an XGBoost model (v1.7.8.1) with SHAP analysis for interpretability. The data were randomly partitioned into training (70 %) and test sets (30 %). Model hyperparameters included: maximum tree depth (max_depth = 3), learning rate ($\eta = 0.2$), and 100 boosting iterations (nrounds = 100). Model interpretability was achieved through SHAP value estimation, revealing the directional impact of individual genes on predictions.

To assess the model performance, we used ROC, calibration curves, and Decision Curve Analysis (DCA). Additionally, SHAP analysis was conducted to produce Beeswarm plots, illustrating the individual contributions of each core gene to classification prediction and their mechanisms of action.

8. Immune infiltration analysis

Quantification of tumor-infiltrating immune cells was performed via CIBERSORT analysis, enabling proportional estimation of 22 immune cell phenotypes characteristic of the NSCLC microenvironment. Distributions of immune cell infiltration levels were statistically compared between NSCLC and control cohorts using boxplot representations. Additionally, we performed correlation analysis to assess associations between hub genes and immune cells, visualizing these relationships through heatmaps and network diagrams to characterize the NSCLC immune microenvironment.

9. ROC analysis and core gene validation

In order to quantify the diagnostic accuracy of the prioritized hub genes in NSCLC, we performed ROC analysis using the pROC package (v1.18.0) in R. The AUC was calculated to quantify the predictive performance of each core gene. These findings were subsequently validated in independent datasets (GSE54495 and GSE18842).

To facilitate the clinical application of the hub genes' diagnostic performance in NSCLC, a nomogram model was developed based on their expression levels. This predictive model serves dual purposes: forecasting NSCLC risk in patients while quantifying each factor's diagnostic contribution. Furthermore, calibration curves were generated to evaluate the consistency between the predicted probabilities of NSCLC and the actual outcomes.

10. Cell Culture

A549 cells were obtained from iCell Bioscience Inc. in Shanghai, China. Cell maintenance was performed using DMEM with 10 % FBS and 1 % penicillin-streptomycin antibiotic cocktail, cultured at 37 °C under 5 % CO_2 humidified atmosphere.

11. Cell Transfection

Sh-lncRNA and its negative control (Sh-NC) were synthesized by

Ribobio (Guangdong, China). Cells in 6-well plates (70–80 % confluent) were transfected using Lipofectamine 2000 (Invitrogen, USA): plasmid/siRNA was diluted in Opti-MEM (5 min), mixed with Lipofectamine 2000 (20 min), then added to cells. Transfected cells were harvested after 48 h for experiments.

12. Quantitative Real-Time PCR

Total RNA isolation was achieved with TRIzol (Invitrogen), reverse transcription carried out using Primescript RT Reagent (TaKaRa), and quantitative PCR performed on StepOne Plus (Applied Biosystems) with SYBR® Premix Ex Taq™ (TaKaRa). The amplification protocol comprised a 5-minute 95 °C hot start, then 40 cycles of three-step amplification (15 s at 95 °C, 30 s at 58 °C, and 30 s at 74 °C per cycle). Data were normalized to β-actin (ΔΔCt method). Primer

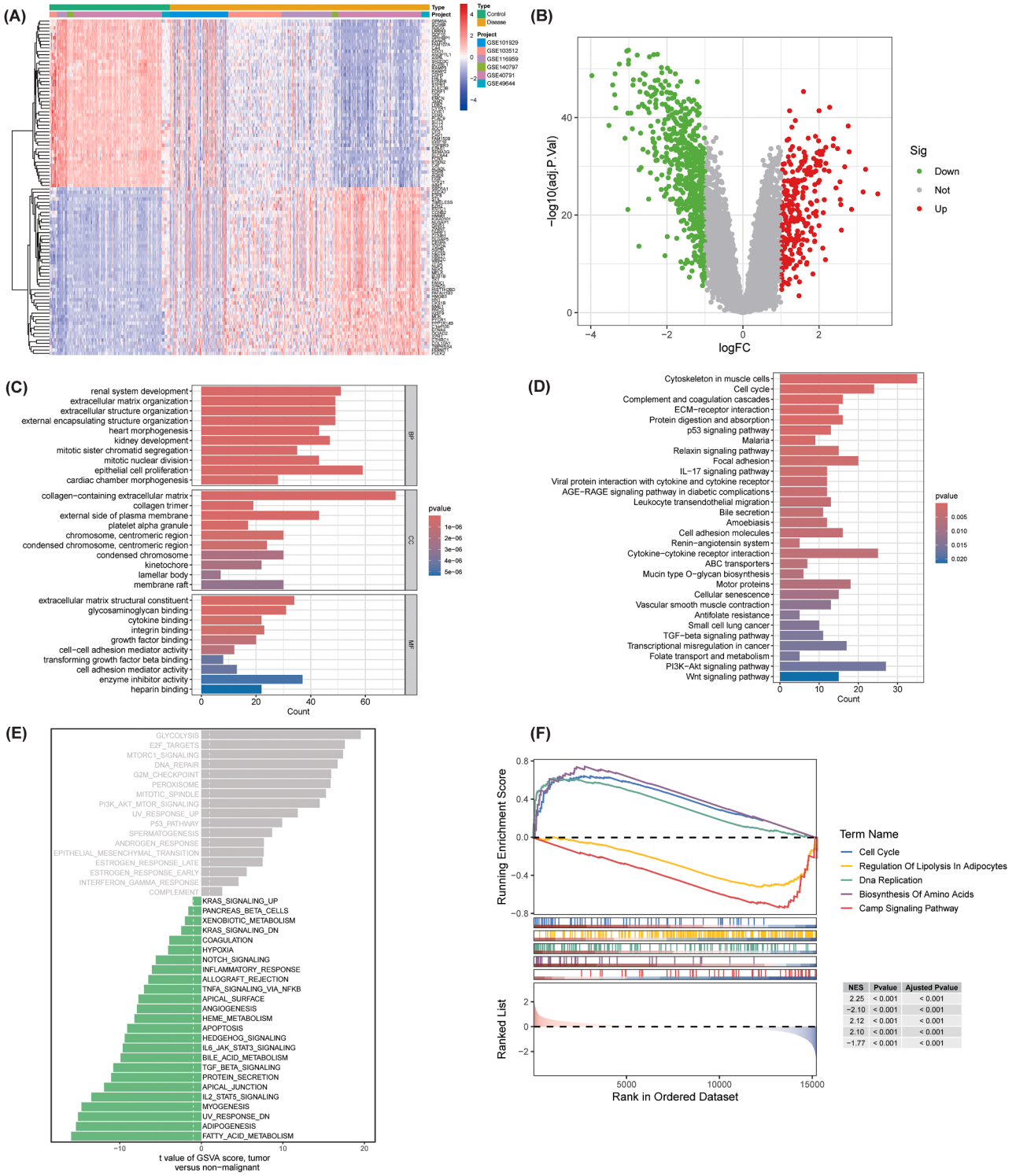

Fig. 1. Differentially Expressed Genes (DEGs) in Non-Small Cell Lung Cancer (NSCLC). (A) Volcano plot showing the differential expression of genes between NSCLC samples (Disease) and healthy controls (Control), with genes identified by $|\log_2 FoldChange| > 1$ and $P < 0.05$. (B) Heatmap of the top 100 DEGs. (C) Gene Ontology (GO) enrichment analysis of DEGs. (E) GSVA of significant enrichment of NSCLC samples. (F) GSEA highlighting the most enriched pathways in NSCLC.

sequences: (1) CACNA2D2: F 5′-CAGAAGTTGGTGGAGAAGGT-3′, R 5′-ATGTCAGGAGCTGCTTCAG-3′; (2) β-actin: F 5′-AGGAAGGACCTG-TATGCCAACA-3′, R 5′-GCGCGGTGATCTCTTTCTG-3′.

13. Western Blot (WB)

Protein extraction was precisely executed employing RIPA lysis buffer. Proteins (50–80 μg) were separated by SDS-PAGE, transferred to nitrocellulose membranes (Absin, China), and probed with primary antibodies: anti-CACNA2D2 (1:2000, 22,145-1-AP) and anti-β-actin (1:1000, 66,009-1-Ig; loading control) for 2 h (RT) + overnight (4 °C). Membranes were blocked for 3 h in 5 % non-fat milk prior to incubation with secondary antibodies, with subsequent infrared imaging performed using an Odyssey scanner (LI-COR, USA).

14. Transwell migration assay

Cells were trypsinized (0.05 % trypsin/0.02 % EDTA), resuspended in RPMI-10 % FBS (1×10^5 cells/mL), and loaded into 96-well chemotaxis chambers with collagen-coated (20 μg/mL, 6 h) 8-μm filters. After 12 h migration (37 °C) toward RPMI-10 % FBS or test medium, non-migrated cells were removed by swabbing. Migrated cells on filter undersides were Hoechst 33,342-stained (10 μM) and counted (100×, 3 fields/filter).

15. Statistical analysis

Shapiro-Wilk test assessed data normality; non-normal data used nonparametric tests. Group comparisons employed *t*-test/Mann-Whitney *U test* ($p < 0.05$). Inter-group comparisons were performed using the Kruskal-Wallis nonparametric test followed by Dunn's multiple comparison procedure.

Results

1. DEGs Identification and Functional Enrichment Analysi

Following batch effect correction, differential expression analysis between Disease and Control samples in the training set identified 889 significant DEGs ($|\log_2 FC| > 1$, $P < 0.05$), including 271 genes showing increased expression and 618 genes with decreased expression. The heatmap clearly demonstrated the overall trend of gene expression in NSCLC samples (Fig. 1A), indicating that numerous genes displayed significant variations in expression levels in lung cancer tissues.

To characterize DEG expression patterns, the top 100 DEGs were selected for visualization in a volcano plot (Fig. 1B). Cluster analysis showed the top 100 DEGs clearly distinguished NSCLC from controls, confirming their significant differential expression. This result further validated the potential biological significance of DEGs in NSCLC.

To delineate the biological significance of DEGs, we conducted functional enrichment analysis to identify overrepresented pathways and molecular functions. GO term analysis identified the BP category as being markedly enriched for pathways such as epithelial cell proliferation and mitotic nuclear division, indicating that these genes may be crucial for tumor cell proliferation and division (Fig. 1C). In terms of CC, DEGs were mainly distributed in structures such as collagen-containing extracellular matrix and condensed chromosome, reflecting the fact that these genes might play key roles in remodeling of extracellular matrix and regulation of chromosome structure. MF analysis demonstrated that DEGs were significantly associated with extracellular matrix structural components, cytokine binding, and integrin binding, suggesting their potential involvement in NSCLC progression through the regulation of immune responses, extracellular matrix interactions, and cell signaling. Additionally, KEGG pathway analysis further demonstrated significant enrichment of DEGs in key inflammatory pathways, particularly the IL-17, Wnt, and PI3K-Akt signaling pathways (Fig. 1D). These pathways were critically involved in mediating inflammatory responses, immune modulation, and tumor microenvironment remodeling, indicating their potential roles in NSCLC pathogenesis. The prominent enrichment of IL-17 and PI3K-Akt signaling pathways particularly highlighted their significance in NSCLC development.

GSVA validated DEG enrichment in FATTY_ACID_METABOLISM, BILE_ACID_METABOLISM, and IL2_STAT5_SIGNALING (Fig. 1E). GSEA further revealed top pathways: Cell Cycle, Regulation of Lipolysis Adipocytes, DNA Replication, Amino Acid Biosynthesis, and cAMP Signaling. (Fig. 1F). The pathway enrichment results underscored the pivotal involvement of inflammatory response and immune regulation in NSCLC pathogenesis, offering important mechanistic clues for further investigation.

2. WGCNA analysis and key module enrichment analysis results

For further resolution of the potential regulatory networks of DEGs in NSCLC samples, WGCNA was adopted to construct and analyze modules on DEGs. Through sample clustering and soft threshold screening, β=14 was determined to be the most appropriate soft threshold for generating a biologically meaningful weighted network where adheres to scale-free network properties (Fig. 2A, B). Based on the dynamic tree cutting method, multiple gene modules were finally identified and distinguished by different colors (Fig. 2C).

Module-trait correlation analysis between modules and clinical characteristics revealed that the magenta module (MEmagenta, $R = 0.71$, $P = 1e\text{-}65$), on the other hand, had the strongest association with the Disease group (Fig. 2D). Furthermore, GS and MM analyses further suggested a significant correlation between genes in the MEbrown module and the clinical phenotype of tumor progression ($R = 0.89$, $P = 2.6e\text{-}186$) (Fig. 2E). The levels of expression genes within the MEmagenta were illustrated by a heat map (Fig. 2F).

Enrichment analysis for the MEmagenta module's 542 genes identified statistically significant biological process terms related to vasculature development regulation, angiogenesis, and positive regulation of vasculature development. CC analysis showed enrichment in collagen-containing extracellular matrix and actin filament bundles, while MF analysis identified actin binding, cytokine binding, and transmembrane receptor protein kinase activity (Fig. 2G). KEGG pathway analysis further demonstrated involvement in TGF-β signaling, cAMP signaling, and muscle cell cytoskeleton pathways (Fig. 2H), suggesting the MEmagenta module's critical roles in NSCLC immune response, metastasis, and proliferation.

Through WGCNA analysis, several key gene modules related to the biological process of NSCLC, which provided an important basis for subsequent biomarker screening and therapeutic target development.

3. Machine learning analysis and integration of LPS-related and circadian rhythm genes.

For further exploration on the potential roles of LRGs and CRGs in lung cancer development, the MEmagenta, which was closely associated with NSCLC progression, was first identified by WGCNA analysis, which was executed on intersection analysis with LPS and CRGs. The Venn diagram (Fig. 3A) revealed that 59 genes in the MEmagenta had significant intersection with LPS-related genes and circadian rhythm genes, indicating that these genes may jointly play key roles in immune regulation and tumor progression in NSCLC.

For establishment of a diagnostic model related to NSCLC, it was adopted that in this study various machine learning algorithms for the analysis, including Lasso regression, RF, Extreme Gradient Boosting (XGBoost), Elasticity Networks (Enet), Partial Least Squares Regression (plsRglm), Generalized Augmented Regression Model, Naive Bayes, Stepwise Generalized Linear Model (Stepglm), Linear Discriminant Analysis (LDA), and SVM. The above algorithms were systematically analyzed under a ten-fold cross-validation framework using a total of 59

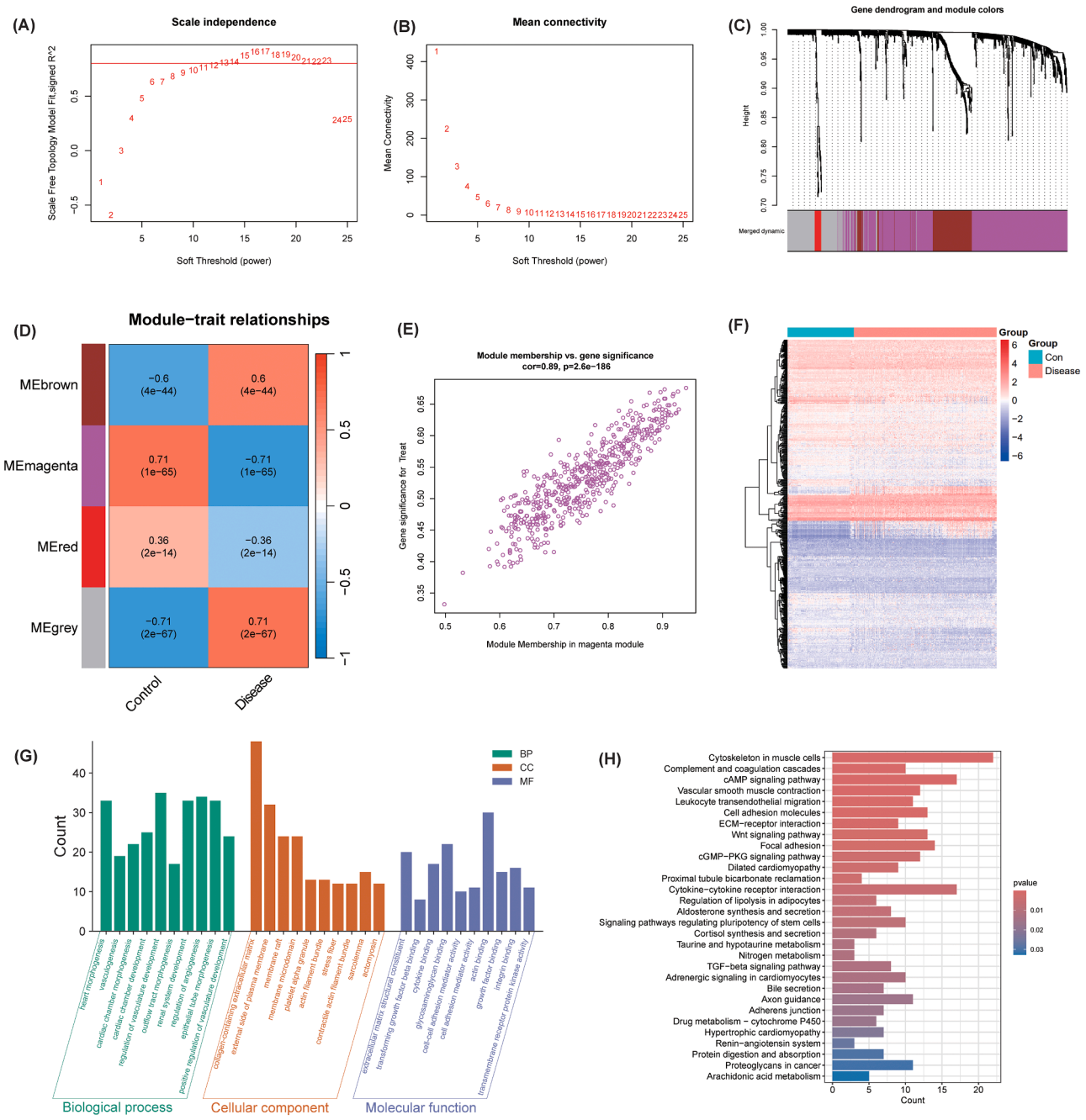

Fig. 2. WGCNA analysis and key module enrichment results. (A-B) Sample clustering and scale-free topology model fit analysis determined $\beta = 14$ as the optimal soft threshold. (C) Gene dendrogram and module colors, where different modules are marked with distinct colors. (D) Module-trait correlation analysis identified the MEmagenta as significantly associated with the Disease group ($R = 0.71$, $P = 2e\text{-}67$). (E) Gene significance (GS) and module membership (MM) analysis. (F) Heatmap illustrating the expression levels of genes within the MEmagenta. (G) GO enrichment analysis of MEmagenta genes. (H) KEGG pathway analysis of MEmagenta genes.

genes available in LRGs, CRGs, and MEturquoise in both the training and validation sets, and finally plsRglm was selected as the best diagnostic model based on the performance of the model (Fig. 3B). The selection of plsRglm over traditional models like RF or SVM was primarily due to its specialized capacity to handle the high multicollinearity inherent in transcriptomic datasets. While RF may lose structured co-expression information through random feature partitioning, and SVM is often sensitive to noise in high-dimensional spaces where the number of predictors vastly exceeds observations, plsRglm extracts orthogonal latent components that maximize the covariance between gene expression and the NSCLC phenotype. This ensures both robust predictive stability and enhanced generalizability across heterogeneous cohorts.

For further identification of the hub genes in the optimal diagnostic model, twenty genes closely associated with NSCLC progression were first screened by Lasso regression analysis (Fig. 3C, D), with the importance of these genes in the model determined. The contribution of these genes was then further evaluated by an RF approach with the top 20 genes selected according to IMPORTANCE (Fig. 3E, F). The intersection results of the Lasso regression and RF analyses (Fig. 3G) identified nine hub genes in the best diagnostic model. The expression differences of these 9 hub genes were also demonstrated using volcano plots, showing markedly lower expression levels in Disease compared to Control groups (Fig. 3H). Notably, our initial screening categorized the lead hub gene, CACNA2D2, as both an LRG and a CRG, suggesting its

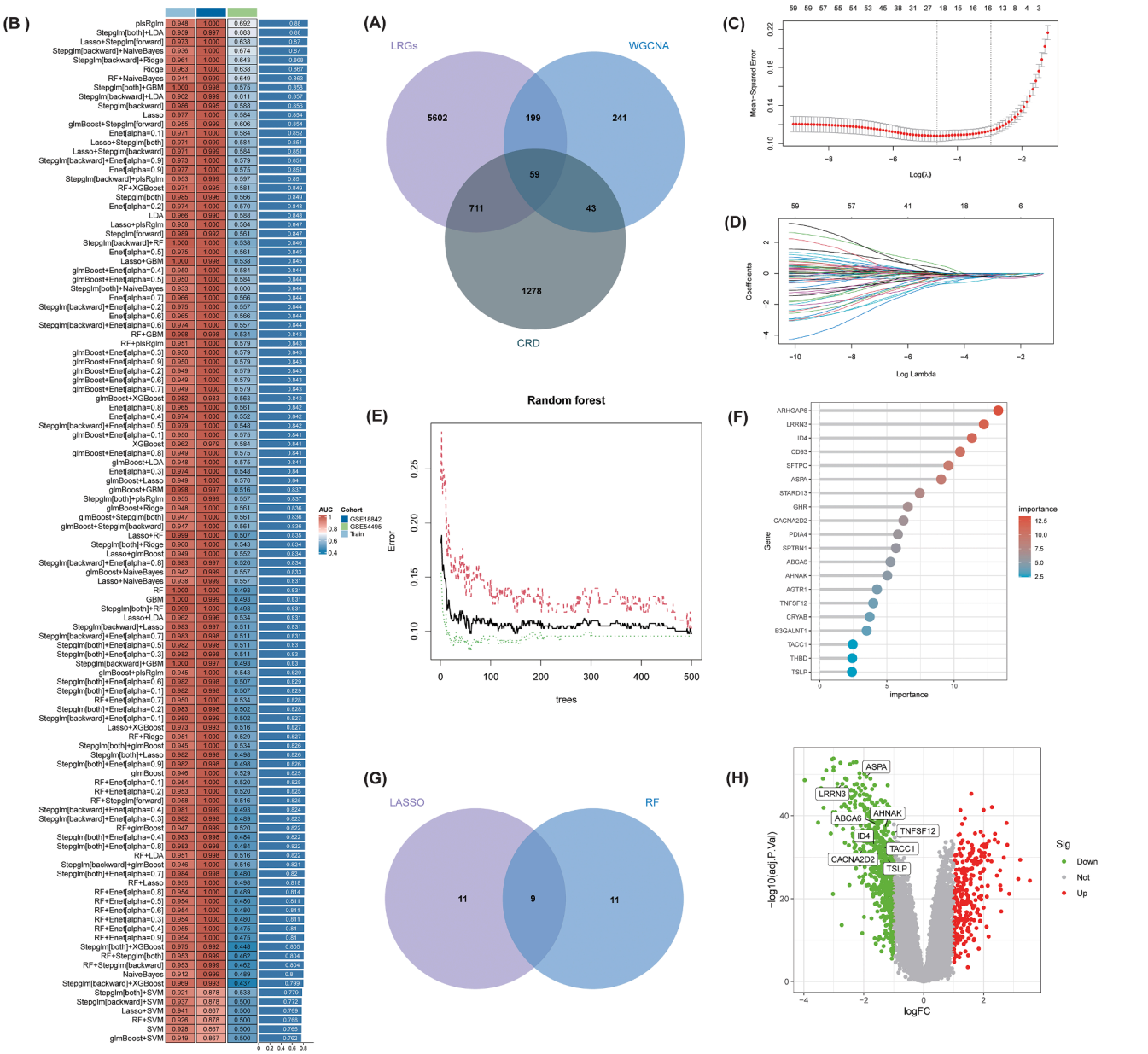

Fig. 3. Machine learning analysis and integration of LPS-related and circadian rhythm genes. (A) Venn diagram showing the intersection of 59 genes among LPS-related genes (LRGs), circadian rhythm genes (CRGs), and the MEmagenta identified in WGCNA. (B) Performance comparison of multiple machine learning models for NSCLC diagnosis using ten-fold cross-validation. (C-D) Lasso regression analysis identified 20 key genes significantly associated with NSCLC progression. (E-F) Random forest analysis ranked gene importance and selected the top 20 genes. (G) Intersection analysis of Lasso regression and random forest results identified nine hub genes crucial for the diagnostic model. (H) Volcano plot showing differential expression of the nine hub genes.

dual role in metabolic and temporal regulation. The expression differences of these 9 hub genes were also demonstrated using volcano plots, showing markedly lower expression levels in Disease compared to Control groups.

4. Functional validation and characterization of hub genes

For further validation on the diagnostic value and biological functions of the nine hub genes screened in NSCLC, it was performed that receiver operating characterization (ROC) curve analysis, PPI network construction, gene correlation analysis, chromosomal localization, and GO BP enrichment analysis.

Firstly, the training set data were subjected to ROC analysis of nine hub genes, revealing their classification accuracy between NSCLC and normal populations (Fig. 4A). The results revealed that these genes' AUC values ranged from 0.832 to 0.906, with ASPA and LRRN3 having the highest AUC (AUC=0.906), indicating their strong diagnostic value.

Secondly, the construction of PPI network revealed the interaction relationship among hub genes (Fig. 4B), which further supported their potential roles in the development of NSCLC. Additionally, significant co-expression patterns emerged from correlation analysis of the core gene set (Fig. 4C), and it was found that some of the genes exhibited significant co-expression patterns in the NSCLC samples, suggesting their potential involvement in the same signaling pathway or regulatory network. Meanwhile, these genes were distributed on multiple chromosomes, indicating that they might affect NSCLC progression through multiple different genetic mechanisms (Fig. 4D).

Finally, we performed GO enrichment analysis (Fig. 4E). The results displayed that these genes were primarily associated with BP such as GO:0044,291 (cell-cell contact zone) and GO:0042,975 (peroxisome

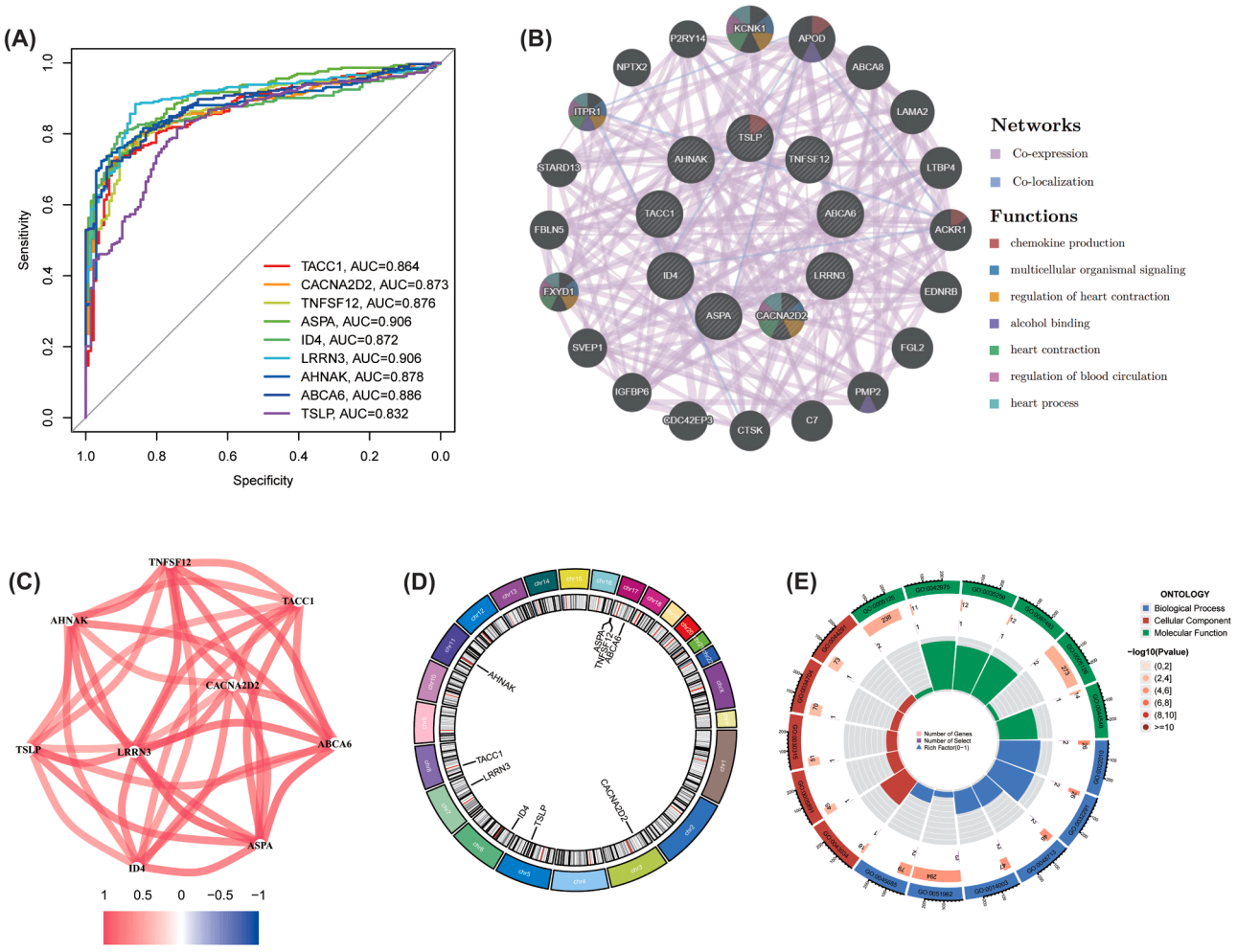

Fig. 4. Functional validation and characterization of hub genes. (A) ROC curve analysis for the nine hub genes. (B) PPI network illustrating the interactions among the hub genes. (C) Correlation analysis of hub gene expression in NSCLC samples. (D) Chromosomal distribution of the hub genes. (E) GO enrichment analysis of 9 hub genes.

proliferator activated receptor binding).

5. Consensus clustering and SHAP interpretability analysis.

Using expression profiles of the 9 hub genes (CACNA2D2, ASPA, LRRN3, ABCA6, TNFSF12, AHNAK, TACC1, ID4, TSLP), CC profiling was performed on all samples. The clustering results showed that the best cluster consistency of the sample groupings was achieved when the number of clusters was 2 ($k = 2$) (Fig. 5A). According to the clustering results, all samples were categorized into Group A and Group B, comparative analysis revealed marked expression variation of hub genes across clusters ($P < 0.001$) (Fig. 5B, C). Expression profiling showed divergent clustering behavior, with samples robustly separating into Group A and Group B cohorts. For assessment of the potential value of the 9 key genes in classification prediction, the XGBoost algorithm was used for modeling. Cross-validation results declared that the XGBoost model exhibited excellent prediction performance (AUC Train= 1.000, Test = 0.914) in both the training and test sets (Fig. 5D). Additionally, the clinical utility of the model was evaluated by DCA, which showed that the XGBoost model had a high net gain under different risk thresholds, further validating the reliability and clinical application potential of the model (Fig. 5E).

For further discussion on the specific contribution of key genes to the model prediction results, an interpretive analysis of the XGBoost model was conducted using SHAP analysis. The results of SHAP values showed that TSLP, ID4, and TACC1 were the most important features affecting the model prediction (Fig. 5F). SHAP analysis further illustrated the relationship between hub gene expression levels and their contribution to model predictions, providing theoretical support for understanding the underlying molecular mechanisms of the disease.

6. Immune infiltration analysis

Immune cell infiltration levels of the two groups were assessed by CIBERSORT algorithm (Fig. 6A). Comparative analysis uncovered statistically significant disparities in immune cell frequencies among the groups. Significant immune cell redistribution was observed, with Macrophages M0/M1 and activated Dendritic cells demonstrating infiltration increases, while resting CD4+ memory T cells and resting mast cells displayed substantial depletion. These alterations indicate a pivotal role of immune cell infiltration in driving lung cancer development and advancement.

To assess hub gene-immune cell associations, we analyzed correlations between the 9 core genes' expression and immune infiltration levels (Fig. 6B). Immune correlation analysis revealed distinct patterns among hub genes (all $P < 0.001$): ABCA6 showed positive associations with CD4+ memory resting T cells, monocytes and resting mast cells, the associations are negative with plasma cells and M0/M1 macrophage phenotypes (Fig. 6C); TSLP positively correlated with monocytes and CD4+ memory T cells but negatively with Tregs and plasma cells

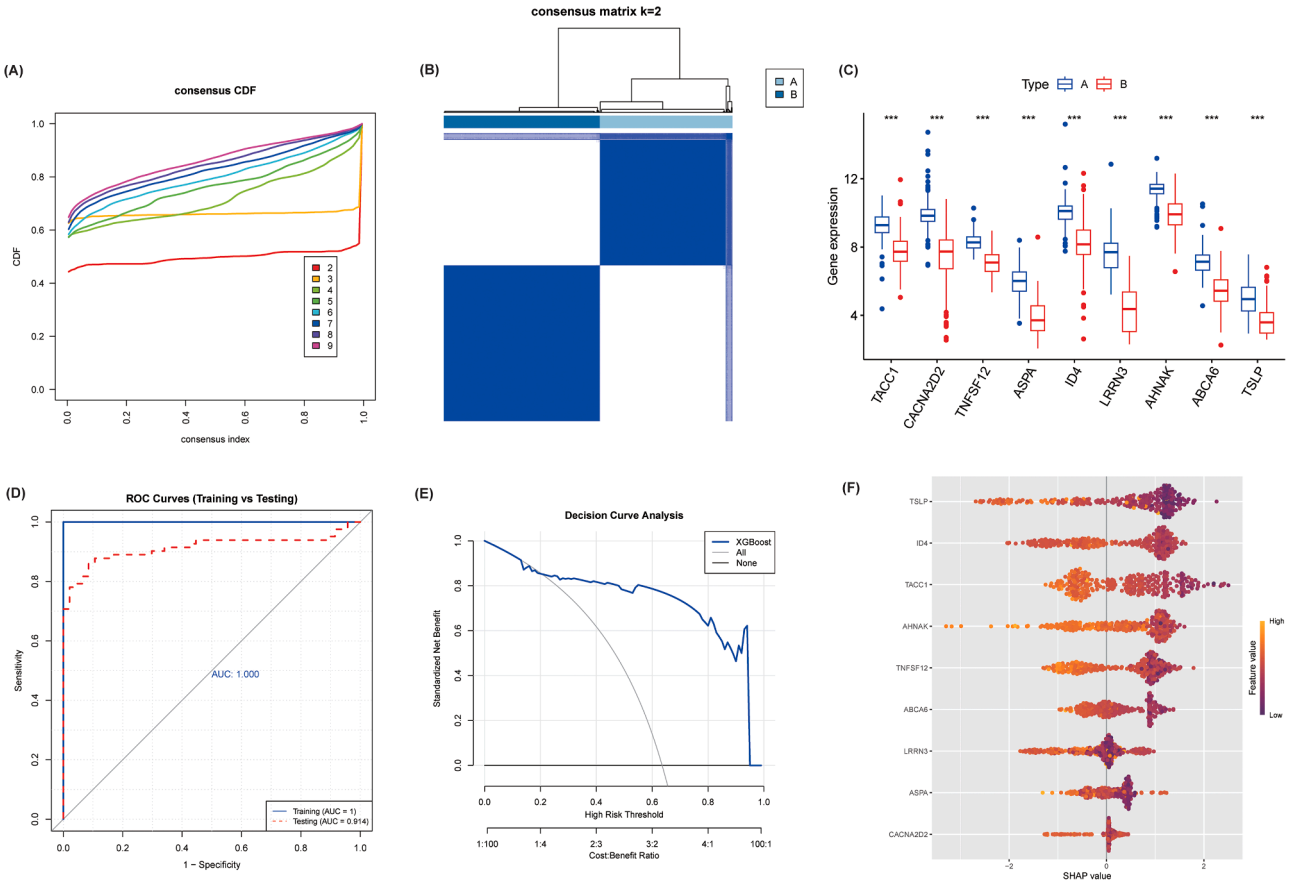

Fig. 5. Consensus clustering and SHAP interpretability analysis based on hub genes. (A) Consensus clustering analysis based on the expression of hub genes, showing optimal clustering stability when $k = 2$. (B-C) Heatmap and boxplot demonstrating significant differences in hub gene expression between the two clusters (A and B) ($P < 0.001$). (D) ROC curves of the XGBoost classification model. (E) DCA evaluating the clinical utility of the XGBoost model. (F) SHAP analysis of 9 hub genes in the XGBoost model.

(Fig. 6D); AHNAK associated with monocytes and resting mast cells but inversely with plasma cells (Fig. 6E); LRRN3 expression displayed significant positive associations with monocytes, resting CD4+ memory T cells and resting mast cells, but inverse relationships with plasma cells and Tregs (Fig. 6F); ID4 related to monocytes and activated CD4+ memory T cells but opposed macrophages M1 (Fig. 6G); ASPA demonstrated co-expression patterns with monocytes and resting CD4+ memory T cell subsets in the immune microenvironment, whereas it exhibits negative associations with Tregs and plasma cells (Fig. 6H); TNFSF12 linked to resting CD4+ memory T cells, and macrophages M2 but conflicted with activated CD4+ memory T cells (Fig. 6I); CACNA2D2 associated with monocytes, resting CD4+ memory T cells and resting mast cells correlate positively, in contrast to the inverse correlations observed with macrophages M1 and activated CD4+ memory T cells (Fig. 6J); and TACC1 connected to monocytes and resting CD4+ memory T cells while opposing plasma cells and Tregs (Fig. 6K). These findings demonstrate hub genes' intricate immunomodulatory roles in NSCLC, offering novel insights for tumor microenvironment regulation and precision therapy development.

7. Expression validation of hub genes and type construction of nomogram models

Nine hub genes' expression patterns were validated in two independent cohorts (GSE54495 and GSE18842) (Fig. 7A, B). ROC curve evaluation demonstrated strong predictive performance for the majority of genes across both validation cohorts, among them CACNA2D2, TNFSF12, ID4, LRRN3, AHNAK and ABCA6 had AUC values close to or reaching 1.000, suggesting very high differentiation ability, while the AUC value of TSLP was relatively low, pointing that its predictive efficacy might be affected by other factors.

A nomogram was developed using eight consistently validated genes (CACNA2D2, ASPA, LRRN3, ABCA6, TNFSF12, AHNAK, TACC1, ID4) to predict patient disease risk (Fig. 7C). To evaluate the calibration of the model, a calibration plot was constructed (Fig. 7D). The curve demonstrates that the predicted probabilities align reasonably well with the actual outcomes, despite minor variations common in clinical cohorts. The results displayed that genes such as CACNA2D2, TNFSF12, and ID4 contributed more, further validating their key role in disease prediction. The consistent downregulation of CACNA2D2 across multiple independent human tissue cohorts provided a high-confidence clinical basis for its identification as a tumor suppressor, warranting subsequent functional investigation in the NSCLC cell model.

8. CACNA2D2 inhibits tumorigenesis and metastasis

Driven by its unique role at the intersection of LPS metabolism and circadian regulation, along with its superior stability in clinical datasets, we prioritized CACNA2D2 for experimental confirmation over other identified hub genes. As a master regulator of Ca^{2+} homeostasis, CACNA2D2 offers a direct mechanistic link to the conceptual framework of this study. CACNA2D2 has been demonstrated to be a tumor suppressor, and its downregulation promotes tumorigenesis. To investigate the role of CACNA2D2, we conducted experiments involving both overexpression and knockdown of this gene. WB analysis confirmed the successful establishment of both overexpression and knockdown

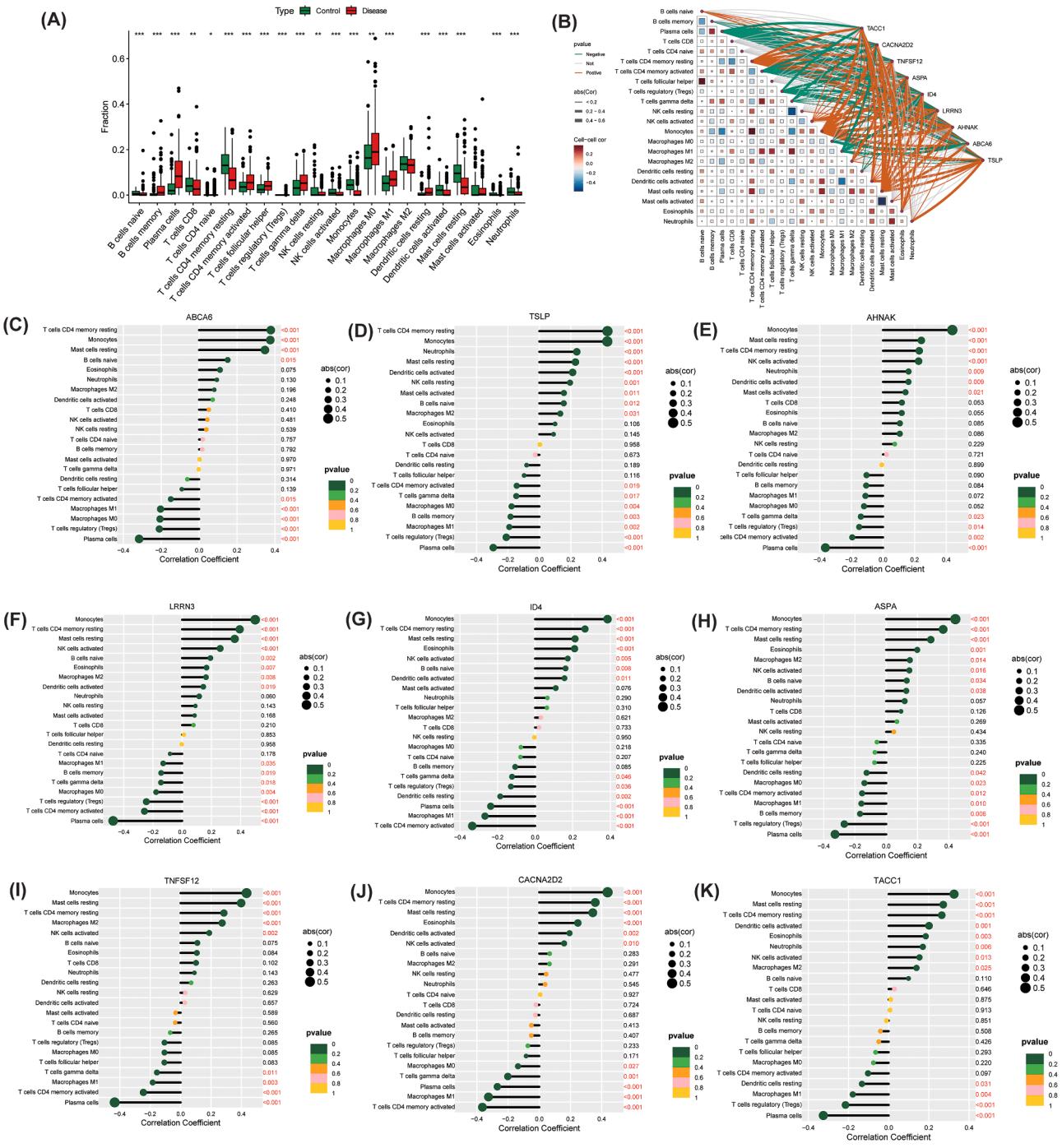

Fig. 6. Immune infiltration analysis and correlation with hub genes in Lung Cancer. (A) Immune cell infiltration levels in disease and control groups analyzed using the CIBERSORT algorithm. (B-K) Correlation analysis between the expression levels of nine hub genes and immune cell infiltration.

constructs (Fig. 8A-C). Cell counting kit-8 (CCK-8) assay results indicated that cell viability increased following si-CACNA2D2 treatment, while it decreased after CACNA2D2-OE (Fig. 8D). Compared to the control group, live/dead staining results demonstrated that the number of viable cells increased after si-CACNA2D2 treatment, whereas it decreased following CACNA2D2-OE (Fig. 8E). To further elucidate the relationship between CACNA2D2 and tumor cell development and metastasis, we examined its association with clonogenicity, migration, and invasion. It was observed that si-CACNA2D2 treatment enhanced clonogenicity, migration, and invasion capabilities, while CACNA2D2-OE treatment resulted in a reduction of these abilities (Fig. 8F-K).

9. Prognostic Significance of CACNA2D2 as a Key Hub Gene

To further explore the prognostic significance of the core hub gene beyond its diagnostic utility, we performed an overall survival (OS) analysis using the GEPIA2 database. In the TCGA-LUAD cohort, patients with low expression of CACNA2D2 exhibited significantly shorter survival times compared to those with high expression (Log-rank $P = 0.022$) (Supplementary Figure 2). This prognostic validation in a large-scale clinical cohort underscores the critical role of CACNA2D2 in lung cancer progression and suggests its potential as a prognostic indicator for patient outcomes.

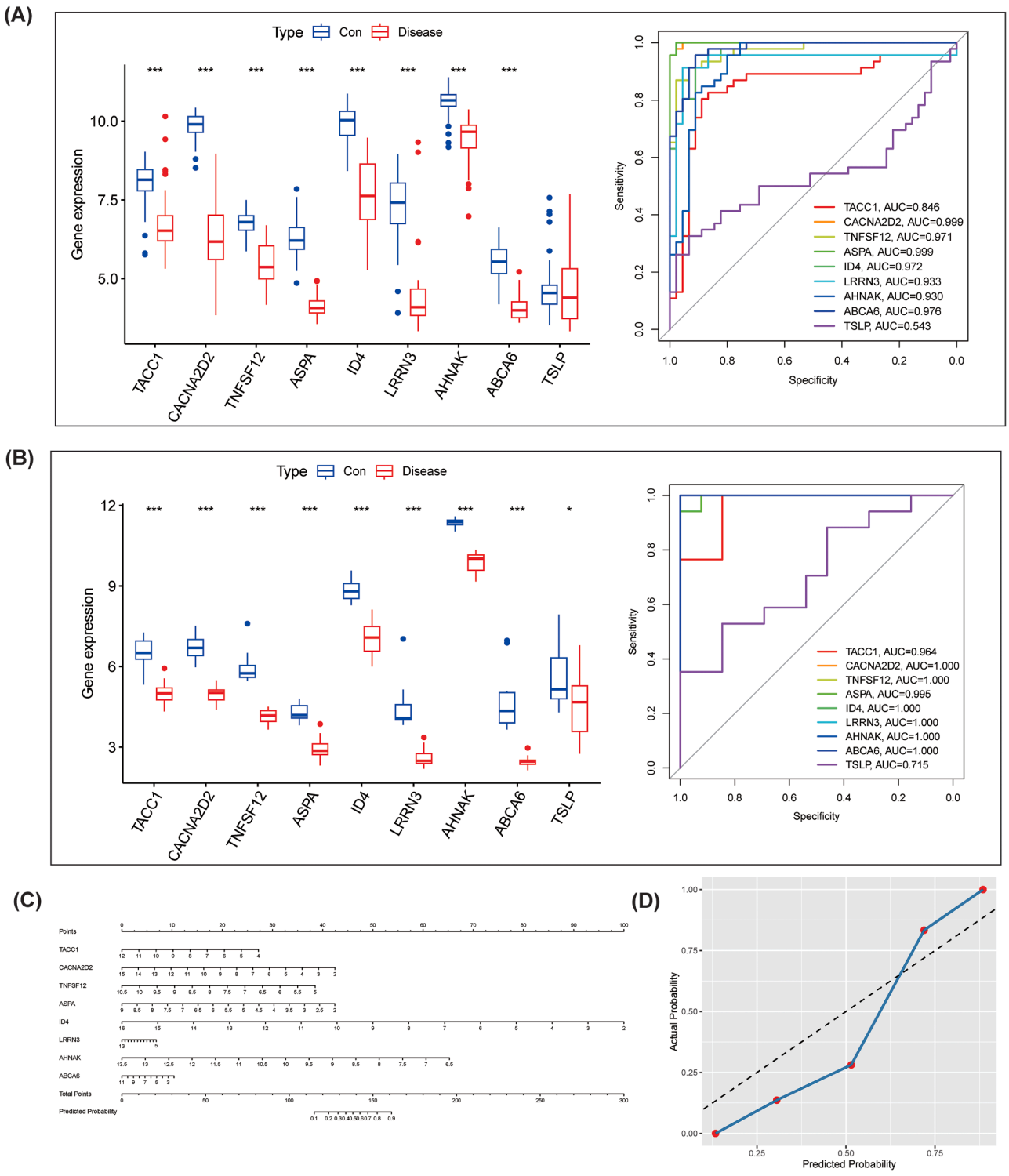

Fig. 7. Validation of hub gene expression and construction of a nomogram model. (A, B) Validation of the expression levels of nine hub genes in two independent datasets (GSE54495 and GSE18842).(C) A nomogram model constructed based on eight genes (CACNA2D2, ASPA, LRRN3, ABCA6, TNFSF12, AHNAK, TACC1, and ID4) that exhibit the same significant expression trend in the validation datasets. (D) Calibration curve of the nomogram model.

Discussion

NSCLC is a malignancy with exceptionally high case fatality rates worldwide, and its high incidence and poor prognosis make it an urgent public health problem [30,31]. Despite some progress in targeted therapy and immunotherapy in recent years, the problems of difficult early diagnosis and limited therapeutic effects remain prominent. Consequently, identifying novel biomarkers is essential for improving early diagnosis and refining targeted therapy strategies. Through integrated bioinformatics and machine learning analysis of LPS metabolism and circadian rhythm pathways, we identified novel NSCLC-associated biomarkers, offering both mechanistic insights into lung cancer pathogenesis and potential diagnostic/therapeutic targets.

The significant enrichment of LRGs and CRGs in NSCLC suggested

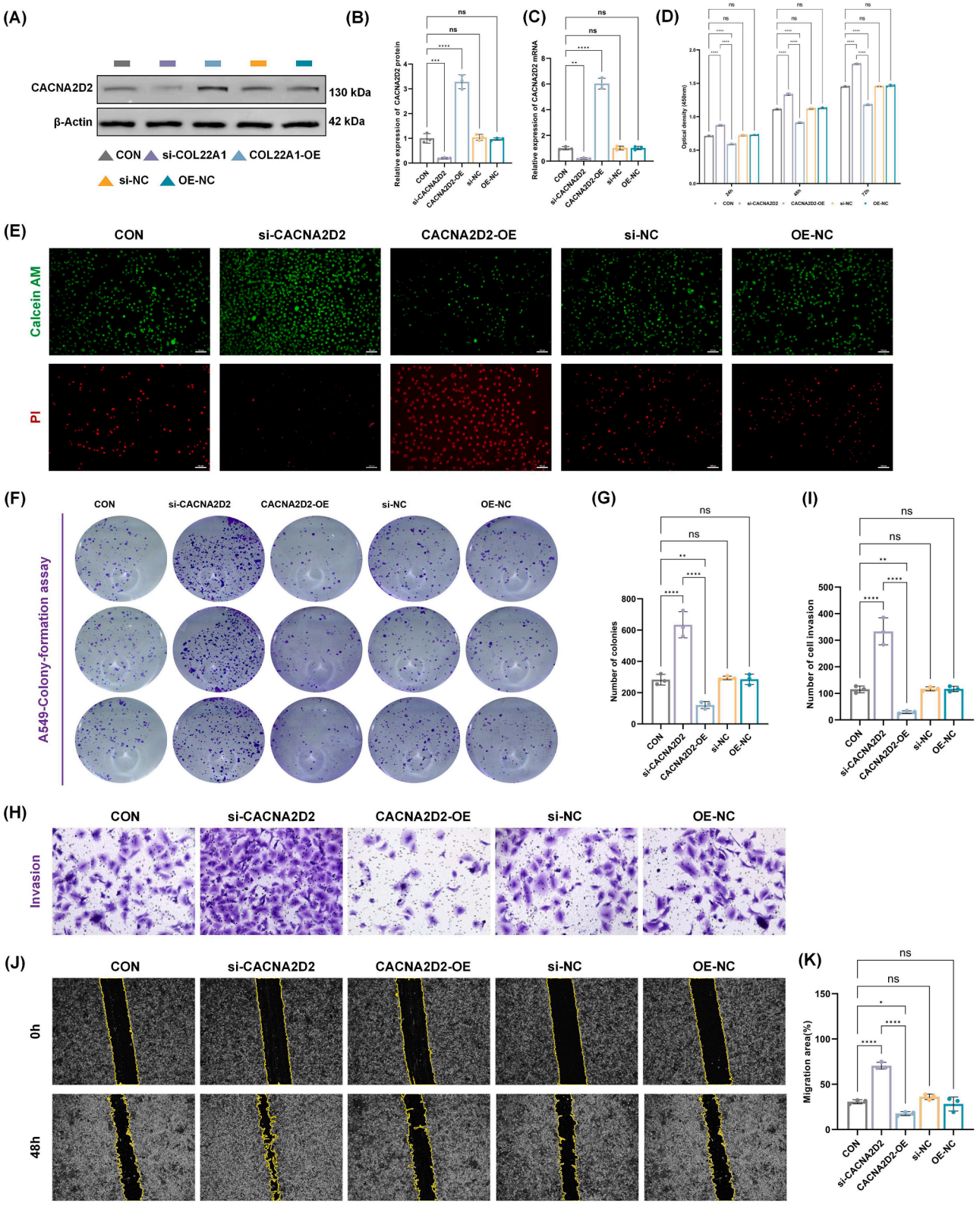

Fig. 8. CACNA2D2 inhibits tumorigenesis and metastasis. (A, B) WB assay for CACNA2D2 expression. (C) PCR detection of CACNA2D2 expression. (D) CCK-8 results at different times. (E) Live/dead staining assessment of tumor cell viability. Green cells are stained with calcein AM, and red cells are stained with propidium iodide. (F, G) The colony formation test (using crystal violet staining) is used to detect the colony-forming ability of tumor cells. (H, I) The Transwell assay is used to evaluate the invasive ability of tumor cells. (J, K) Cell scratch assay. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

their potential role in tumor progression. It has been shown that LPS induces an inflammatory response through activation of Toll-like receptors (TLRs), which promotes immune escape from the tumor microenvironment and tumor cell invasiveness [32,33]. The LRGs identified in this study were markedly overrepresented in the IL-17 and PI3K-Akt signaling pathways, further supporting the mechanism by which LPS promoted tumor progression by modulating inflammatory and immune responses in NSCLC. Previous studies have revealed that LPS advanced epithelial-mesenchymal transition and invasive capacity of NSCLC cells through transcriptional activation and functional enhancement of hypoxia-inducible factor 1α (HIF-1α), further revealing the role of LPS in NSCLC to promote tumor progression by regulating inflammation-related microenvironment and tumor hypoxia mechanisms [34].

Systematic reviews confirm circadian misalignment as a validated promoter of diverse cancer development, such effects promoted oncogenesis by dysregulating fundamental biological processes: cell cycle checkpoints, DNA repair fidelity, and metabolic homeostasis [30,31]. The significant enrichment of CRGs in this study suggested that circadian rhythm disruption might further exacerbate the malignant progression of NSCLC by altering the growth pattern and immune escape ability of tumor cells. It has been declared that the expression profiles of CRGs were closely associated with the prognosis of NSCLC, and that they further reveal the critical role of circadian rhythm disruption in the immune microenvironment and tumor progression of NSCLC by disrupting the metabolism and cellular functions of the cancer cells, further mediated by impeding anti-tumor immune cell trafficking while facilitating inflammatory cell infiltration [35].

Through machine learning methods, 9 hub genes screened in this study (CACNA2D2, ASPA, LRRN3, ABCA6, TNFSF12, AHNAK, TACC1, ID4, and TSLP), exhibited high predictive performance in NSCLC diagnosis. Among them, CACNA2D2 and LRRN3 had the highest AUC values, suggesting that they might serve as potential diagnostic markers for NSCLC. The tumor suppressive function of CACNA2D2 has been validated in diverse malignancies, and its low expression has been associated with enhanced tumor aggressiveness and metastatic ability [36,37]. LRRN3 plays a crucial role in immune regulation, with CACNA2D2 expression levels demonstrating significant correlations with tumor-infiltrating immune cell populations [35,36]. These findings align with the immune infiltration analysis in this study, suggesting that hub genes may be key regulators of the tumor immune microenvironment.

Our computational and experimental findings collectively substantiate this integrated conceptual framework. The prioritized hub genes, most notably CACNA2D2 and TNFSF12, function as critical molecular nodes at the intersection of calcium-dependent signaling and immunoregulatory cascades. For instance, CACNA2D2 is fundamental to cellular Ca^{2+} homeostasis—a physiological requirement for both the maintenance of circadian oscillations and the modulation of inflammatory activation following LPS exposure [38,39]. Consequently, the marked downregulation of CACNA2D2 identified in our cohorts likely signifies a dual regulatory failure: the collapse of circadian-mediated calcium rhythmicity and a compromised cellular capacity to buffer LPS-induced inflammatory stress. This synergistic disruption culminates in the aggressive malignant phenotypes observed in our A549 functional assays, highlighting CACNA2D2 as a pivotal rheostat in the circadian-inflammation axis in NSCLC.

To advance our understanding of CACNA2D2′s involvement in tumor invasiveness and metastatic progression, we conducted a series of cellular experiments. As an auxiliary subunit of calcium channels, CACNA2D2 plays a crucial role in regulating cellular calcium signaling and cell cycle progression, reinforcing its classification as a validated tumor suppressor across various malignancies2. Our functional studies corroborate this paradigm: the loss of CACNA2D2 promotes oncogenesis, as evidenced by increased cell viability, clonogenicity, migration, and invasion following si-CACNA2D2 treatment. Conversely, the CACNA2D2-OE suppresses malignant phenotypes, leading to reduced proliferation, survival, and metastatic potential. The expression of CACNA2D2 is linked to heightened tumor invasiveness. The concordance observed across various assays, including viability (CCK-8), survival (live/dead), clonogenicity, migration, and invasion, demonstrates that CACNA2D2 serves to constrain fundamental hallmarks of cancer progression. Mechanistically, its tumor-suppressive function likely arises from dysregulated calcium homeostasis and disrupted cell cycle control, which in turn affects critical oncogenic pathways. These findings reveal CACNA2D2′s dual role as a crucial controller of tumor aggressiveness and a prospective treatment target. While our in vitro assays focused on the A549 cell line, the profound biological effects observed upon both knockdown and restoration of CACNA2D2 expression align with the marked downregulation observed in our high-throughput tissue analysis. This "gain-and-loss" approach effectively confirms CACNA2D2 as a functional tumor suppressor in NSCLC, consistent with previous reports comparing its basal expression in malignant versus normal lung epithelial cell lines.

Importantly, the clinical relevance of our findings is reinforced by the survival analysis, which demonstrates that low CACNA2D2 expression is a significant predictor of poor overall survival in LUAD patients. This prognostic association implies that the disruption of calcium-mediated homeostasis—likely linked to the LPS-circadian axis—has tangible consequences for patient longevity. By bridging bioinformatics-derived diagnostic potential with clinical survival data, our study provides a more comprehensive perspective on the role of CACNA2D2 as a key tumor suppressor in the NSCLC landscape.

Immune profiling revealed significantly elevated infiltration of M1/M0 macrophages and activated dendritic cells, coupled with reduced CD4+ resting memory T cells and mast cells, in NSCLC samples compared to controls. This suggested that there might be a significant immunosuppressive state in the immune microenvironment of NSCLC, which was consistent with the role of tumor-associated macrophages (TAMs) in promoting tumor growth and immune escape that has been reported in previous studies [40,41]. Hub gene expression profiles demonstrated significant correlations with immune infiltration patterns, suggesting their regulatory involvement in tumor-immune microenvironment modulation. The expression pattern of ABCA6 revealed contrasting immune interactions: positively coordinated with resting CD4+ memory T cell infiltration but inversely related to M1 macrophage abundance, suggesting that it might play a role in suppressing tumor-associated inflammatory responses [42,43]. Beyond ABCA6, other components of the 9-gene signature exhibit distinct yet synergistic roles in defining the immune landscape. LRRN3 and AHNAK are intricately involved in T-cell receptor (TCR) signaling and lymphocyte activation; their downregulation is consistent with the T-cell exhaustion and the observed reduction in CD^{4+} memory T cells [44,45]. TNFSF12 (TWEAK) and TSLP serve as pivotal inflammatory rheostats; TNFSF12 modulates macrophage recruitment and pro-inflammatory cytokine production, while TSLP bridges innate and adaptive immunity, often being dysregulated in the lung to promote pro-tumorigenic inflammation [46,47]. Furthermore, metabolic-immune gatekeepers such as ASPA and ID4 link biochemical homeostasis to the TME, potentially influencing the energetic state of infiltrating leukocytes and the differentiation of regulatory T cells.

The results of this study align with established literature while also breaking new ground. For instance, the roles of LPS and CRGs in NSCLC have been partially reported [48], but this study systematically identified potential biomarkers for these genes in NSCLC and revealed their interaction mechanisms by integrating multi-omics data and machine learning approaches. Moreover, this study linked LPS metabolism and circadian rhythm genes to the tumor immune microenvironment for the first time, providing a new perspective for understanding the immune escape mechanism in NSCLC.

Although this study provided valuable findings, there were still some limitations. Firstly, the data analysis was based on public datasets,

which might have batch effects and sample heterogeneity, and experimental validation was needed to further confirm the functions of the hub genes in the future. Secondly, this study mainly focused on the changes in gene expression levels without addressing the mechanisms at the levels of epigenetic regulation and protein modification, which can be analyzed more comprehensively in the future by combining multi-omics data. Future functional studies employing in vitro and in vivo approaches will be essential to fully characterize the mechanistic roles of these hub genes in NSCLC progression.

Conclusion

In this study, the roles of LPS metabolism and CRGs in NSCLC were deeply explored. Several potential biomarkers, including the tumor suppressor CACNA2D2, were successfully screened by combining bioinformatics analysis and machine learning technology. Functional validation experiments demonstrated the critical role ofCACNA2D2 in suppressing NSCLC malignancy, where its knockdown enhanced tumor cell proliferation, clonogenicity, migration, and invasion while its overexpression inhibited these oncogenic processes. This work provided a stronger theoretical foundation for further clinical applications. Hub gene characterization through both functional studies and immune landscape quantification clarified their pathogenic roles in shaping immunosuppressive microenvironments, offering new insights for early diagnosis and novel targeted therapy of NSCLC. Future studies will further validate the clinical utility of these biomarkers, with an emphasis on the therapeutic potential of CACNA2D2, thereby promoting advancements in treatment strategies for NSCLC.

Funding

This study was supported by the National Natural Science Foundation of China (81800063); Three-year action plan for strengthening the construction of the public health system in Shanghai (GWVI-11.2-YQ40) and Outstanding Young Talent Program of Shanghai Pulmonary Hospital (2023).

CRediT authorship contribution statement

Yu Chen: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yifan Zhao:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Qinghua Kong:** Writing – review & editing, Supervision, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Yabin Li:** Writing – review & editing, Validation, Supervision, Methodology, Data curation, Conceptualization. **Ping Jiang:** Writing – review & editing, Validation, Supervision, Software, Resources, Formal analysis, Data curation, Conceptualization. **Lichao Fan:** Writing – review & editing, Writing – original draft, Visualization, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2026.102699.

References

[1] I. Soerjomataram, F. Bray, Planning for tomorrow: global cancer incidence and the role of prevention 2020-2070, Nat. Rev. Clin. Oncol. 18 (2021) 663–672, https://doi.org/10.1038/s41571-021-00514-z.
[2] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA. *Cancer J. Clin.* 71 (2021) 209–249, https://doi.org/10.3322/caac.21660.
[3] Y. Yin, D. Li, M. He, J. Wang, The identification of hub biomarkers and pathways in lung cancer and prognostic evaluation, Transl. Cancer Res. 11 (2022) 2622–2635, https://doi.org/10.21037/tcr-22-245.
[4] W. Li, J.-B. Liu, L.-K. Hou, F. Yu, J. Zhang, W. Wu, X.-M. Tang, F. Sun, H.-M. Lu, J. Deng, et al., Liquid biopsy in lung cancer: significance in diagnostics, prediction, and treatment monitoring, Mol. Cancer 21 (2022) 25, https://doi.org/10.1186/s12943-022-01505-z.
[5] A. Jachowski, M. Marcinkowski, J. Szydłowski, O. Grabarczyk, Z. Nogaj, Ł. Marcin, A. Pławski, P.P. Jagodziński, B.K. Słowikowski, Modern therapies of nonsmall cell lung cancer, J. Appl. Genet. 64 (2023) 695–711, https://doi.org/10.1007/s13353-023-00786-4.
[6] F. Wang, Q. Su, C. Li, Identidication of novel biomarkers in non-small cell lung cancer using machine learning, Sci. Rep. 12 (2022) 16693, https://doi.org/10.1038/s41598-022-21050-5.
[7] N. Bao, X. Zhang, C. Lin, F. Qiu, G. Mo, A scoring model based on bacterial lipopolysaccharide-related genes to predict prognosis in NSCLC, Front. Genet. 15 (2024) 1408000, https://doi.org/10.3389/fgene.2024.1408000.
[8] T. Xu, R. Deng, X. Li, Y. Zhang, M.-Q. Gao, RNA-seq analysis of different inflammatory reactions induced by lipopolysaccharide and lipoteichoic acid in bovine mammary epithelial cells, Microb. Pathog. 130 (2019) 169–177, https://doi.org/10.1016/j.micpath.2019.03.015.
[9] D. Yi, Y. Hou, L. Wang, D. Zhao, B. Ding, T. Wu, H. Chen, Y. Liu, P. Kang, G. Wu, Gene expression profiles in the intestine of lipopolysaccharide-challenged piglets, Front. Biosci. Landmark Ed. 21 (2016) 487–501, https://doi.org/10.2741/4404.
[10] C.-H. Liu, Z. Chen, K. Chen, F.-T. Liao, C.-E. Chung, X. Liu, Y.-C. Lin, P. Keohavong, G.D. Leikauf, Y.P. Di, Lipopolysaccharide-mediated chronic inflammation promotes tobacco carcinogen-induced lung cancer and determines the efficacy of immunotherapy, Cancer Res. 81 (2021) 144–157, https://doi.org/10.1158/0008-5472.CAN-20-1994.
[11] H. Xu, J. Zhou, F. Ye, Y. Gao, Serum lipopolysaccharide associated with new-onset atrial fibrillation in patients with non-small-cell lung cancer a retrospective observational study, Front. Surg. 11 (2024), https://doi.org/10.3389/fsurg.2024.1404450.
[12] I. Guijarro-Muñoz, M. Compte, A. Álvarez-Cienfuegos, L. Álvarez-Vallina, L. Sanz, Lipopolysaccharide activates toll-like receptor 4 (TLR4)-mediated NF-κb signaling pathway and proinflammatory response in Human pericytes, J. Biol. Chem. 289 (2014) 2457–2468, https://doi.org/10.1074/jbc.M113.521161.
[13] P. Wang, X. Han, B. Mo, G. Huang, C. Wang, LPS enhances TLR4 expression and IFN-γ production via the TLR4/IRAK/NF-κb signaling pathway in rat pulmonary arterial smooth muscle cells, Mol. Med. Rep. 16 (2017) 3111–3116, https://doi.org/10.3892/mmr.2017.6983.
[14] N.G. Casanova, R.L. De Armond, S. Sammani, X. Sun, B. Sun, C. Kempf, C. Bime, J. G.N. Garcia, S. Parthasarathy, Circadian disruption dysregulates lung gene expression associated with inflammatory lung injury, Front. Immunol. 15 (2024) 1348181, https://doi.org/10.3389/fimmu.2024.1348181.
[15] C. Fu, L. Sun, C. Feng, T. Zhou, Y. Bi, A prognostic model of lung adenocarcinoma constructed based on circadian rhythm genes and its potential clinical significance, Front. Oncol. 15 (2025) 1464578, https://doi.org/10.3389/fonc.2025.1464578.
[16] O. Ogunlusi, A. Ghosh, M. Sarkar, K. Carter, H. Davuluri, M. Chakraborty, K. Eckel-Mahan, A. Keene, J.S. Menet, D. Bell-Pedersen, et al., Rhythm is essential: unraveling the relation between the circadian clock and cancer, Crit. Rev. Oncol. Hematol. 208 (2025) 104632, https://doi.org/10.1016/j.critrevonc.2025.104632.
[17] B. Zhao, E. Nepovimova, Q. Wu, The role of circadian rhythm regulator PERs in oxidative stress, immunity, and cancer development, Cell Commun. Signal. CCS 23 (2025) 30, https://doi.org/10.1186/s12964-025-02040-2.
[18] M. Motiei, R. Abu-Dawud, A. Relógio, C. Assaf, Circadian rhythms in haematological malignancies: therapeutic potential and personalised interventions, EBioMedicine 110 (2024) 105451, https://doi.org/10.1016/j.ebiom.2024.105451.
[19] Y. Zhang, Q. Zhang, R. Liu, D. Zhang, G. Hu, X. Chen, Circadian disruption in cancer and regulation of cancer stem cells by Circadian clock genes: an updated review, Cancer Lett. 611 (2024) 217391, https://doi.org/10.1016/j.canlet.2024.217391.
[20] J. Brouwer-Visser, W.-Y. Cheng, A. Bauer-Mehren, D. Maisel, K. Lechner, E. Andersson, J.T. Dudley, F. Milletti, Regulatory T-cell genes drive altered immune microenvironment in adult solid cancers and allow for immune contextual patient subtyping, Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. 27 (2018) 103–112, https://doi.org/10.1158/1055-9965.EPI-17-0461.
[21] L. Moreno Leon, M. Gautier, R. Allan, M. Ilié, N. Nottet, N. Pons, A. Paquet, K. Lebrigand, M. Truchi, J. Fassy, et al., The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress, Oncogene 38 (2019) 7146–7165, https://doi.org/10.1038/s41388-019-0935-y.
[22] J. Liang, H. Li, J. Han, J. Jiang, J. Wang, Y. Li, Z. Feng, R. Zhao, Z. Sun, B. Lv, et al., Mex3a Interacts with LAMA2 to promote lung adenocarcinoma metastasis via PI3K/AKT pathway, Cell Death. Dis. 11 (2020) 614, https://doi.org/10.1038/s41419-020-02858-3.

[23] K.A. Mitchell, A. Zingone, L. Toulabi, J. Boeckelman, B.M. Ryan, Comparative transcriptome profiling reveals coding and noncoding RNA differences in NSCLC from African Americans and European Americans, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 23 (2017) 7412–7425, https://doi.org/10.1158/1078-0432.CCR-17-0527.
[24] Y. Zhang, O. Foreman, D.A. Wigle, F. Kosari, G. Vasmatzis, J.L. Salisbury, J. van Deursen, P.J. Galardy, USP44 Regulates centrosome positioning to prevent aneuploidy and suppress tumorigenesis, J. Clin. Invest. 122 (2012) 4362–4374, https://doi.org/10.1172/JCI63084.
[25] Y. Sun, A. Daemen, G. Hatzivassiliou, D. Arnott, C. Wilson, G. Zhuang, M. Gao, P. Liu, A. Boudreau, L. Johnson, et al., Metabolic and transcriptional profiling reveals pyruvate dehydrogenase Kinase 4 as a mediator of epithelial-mesenchymal transition and drug resistance in tumor cells, Cancer Metab. 2 (2014) 20, https://doi.org/10.1186/2049-3002-2-20.
[26] J.-C.J. Tsay, Z. Li, T.-A. Yie, F. Wu, L. Segal, A.K. Greenberg, E. Leibert, M. D. Weiden, H. Pass, J. Munger, et al., Molecular characterization of the peripheral airway field of cancerization in lung adenocarcinoma, PLoS. One 10 (2015) e0118132, https://doi.org/10.1371/journal.pone.0118132.
[27] A. Sanchez-Palencia, M. Gomez-Morales, J.A. Gomez-Capilla, V. Pedraza, L. Boyero, R. Rosell, M.E. Fárez-Vidal, Gene expression profiling reveals novel biomarkers in nonsmall cell lung cancer, Int. J. Cancer 129 (2011) 355–364, https://doi.org/10.1002/ijc.25704.
[28] L. He, Y. Fan, Y. Zhang, T. Tu, Q. Zhang, F. Yuan, C. Cheng, Single-cell transcriptomic analysis reveals circadian rhythm disruption associated with poor prognosis and drug-resistance in lung adenocarcinoma, J. Pineal Res. 73 (2022) e12803, https://doi.org/10.1111/jpi.12803.
[29] J. Gong, B. Shi, P. Yang, A. Khan, T. Xiong, Z. Li, Unveiling immune infiltration characterizing genes in hypertrophic cardiomyopathy through transcriptomics and bioinformatics, J. Inflamm. Res. 17 (2024) 3079–3092, https://doi.org/10.2147/JIR.S454446.
[30] B.C. Bade, C.S. Dela Cruz, Lung Cancer 2020: epidemiology, etiology, and prevention, Clin. Chest Med 41 (2020) 1–24, https://doi.org/10.1016/j.ccm.2019.10.001.
[31] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, CA. *Cancer J. Clin.* 70 (2020) 7–30, https://doi.org/10.3322/caac.21590.
[32] R. Afroz, E.M. Tanvir, M. Tania, J. Fu, M.A. Kamal, M.A. Khan, LPS/TLR4 pathways in breast cancer: insights into cell signalling, Curr. Med. Chem. 29 (2022) 2274–2289, https://doi.org/10.2174/0929867328666210811145043.
[33] M. Genin, F. Clement, A. Fattaccioli, M. Raes, C. Michiels, M1 and M2 macrophages derived from THP-1 cells differentially modulate the response of cancer cells to etoposide, BMC. Cancer 15 (2015) 577, https://doi.org/10.1186/s12885-015-1546-9.
[34] N. Yang, Y. Liang, P. Yang, F. Ji, Propofol suppresses LPS-induced nuclear accumulation of HIF-1α and tumor aggressiveness in non-small cell lung cancer, Oncol. Rep. 37 (2017) 2611–2619, https://doi.org/10.3892/or.2017.5514.
[35] M. Li, Z. Chen, T. Jiang, X. Yang, Y. Du, J. Liang, L. Wang, J. Xi, M. Lin, M. Feng, Circadian rhythm-associated clinical relevance and tumor microenvironment of non-small cell lung cancer, J. Cancer 12 (2021) 2582–2597, https://doi.org/10.7150/jca.52454.
[36] M. Warnier, M. Roudbaraki, S. Derouiche, P. Delcourt, A. Bokhobza, N. Prevarskaya, P. Mariot, CACNA2D2 Promotes tumorigenesis by stimulating cell proliferation and angiogenesis, Oncogene 34 (2015) 5383–5394, https://doi.org/10.1038/onc.2014.467.
[37] X. Kang, F. Kong, K. Huang, L. Li, Z. Li, X. Wang, W. Zhang, X. Wu, LncRNA MIR210HG promotes proliferation and invasion of non-small cell lung cancer by upregulating methylation of CACNA2D2 promoter via binding to DNMT1, OncoTargets Ther 12 (2019) 3779–3790, https://doi.org/10.2147/OTT.S189468.
[38] S. Haddad, C. Ablinger, R. Stanika, M. Hessenberger, M. Campiglio, N.J. Ortner, P. Tuluc, G.J. Obermair, A biallelic mutation in CACNA2D2 associated with developmental and epileptic encephalopathy affects calcium channel-dependent as well as synaptic functions of A2δ-2, J. Neurochem. 169 (2025) e16197, https://doi.org/10.1111/jnc.16197.
[39] J. Kim, S.G. Jeon, H.-R. Jeong, H. Park, J.-I. Kim, H.-S.L- Hoe, Type Ca2+ channel inhibition rescues the LPS-induced neuroinflammatory response and impairments in spatial memory and dendritic spine formation, Int. J. Mol. Sci. 23 (2022) 13606, https://doi.org/10.3390/ijms232113606.
[40] S. Ostrand-Rosenberg, P. Sinha, D.W. Beury, V.K. Clements, Cross-talk between myeloid-derived suppressor cells (MDSC), macrophages, and dendritic cells enhances tumor-induced immune suppression, Semin. Cancer Biol. 22 (2012) 275–281, https://doi.org/10.1016/j.semcancer.2012.01.011.
[41] A. Sandström Gerdtsson, M. Knulst, J. Botling, A. Mezheyeuski, P. Micke, S. Ek, Phenotypic characterization of spatial immune infiltration niches in non-small cell lung cancer, Oncoimmunology. 12 (2023) 2206725, https://doi.org/10.1080/2162402X.2023.2206725.
[42] M. Pasello, A.M. Giudice, C. Cristalli, M.C. Manara, C. Mancarella, A. Parra, M. Serra, G. Magagnoli, F. Cidre-Aranaz, T.G.P. Grünewald, et al., ABCA6 Affects the malignancy of ewing sarcoma cells via cholesterol-guided inhibition of the IGF1R/AKT/MDM2 axis, Cell. Oncol. Dordr. Neth. 45 (2022) 1237–1251, https://doi.org/10.1007/s13402-022-00713-5.
[43] M. Pasello, A.M. Giudice, C. Cristalli, M.C. Manara, C. Mancarella, A. Parra, M. Serra, G. Magagnoli, F. Cidre-Aranaz, T.G.P. Grünewald, et al., ABCA6 Affects the malignancy of ewing sarcoma cells via cholesterol-guided inhibition of the IGF1R/AKT/MDM2 axis, Cell. Oncol. Dordr. Neth. 45 (2022) 1237–1251, https://doi.org/10.1007/s13402-022-00713-5.
[44] J.P. Chou, C.M. Ramirez, J.E. Wu, R.B. Effros, Accelerated aging in HIV/AIDS: novel biomarkers of senescent Human CD8+ T cells, PLoS. One 8 (2013) e64702, https://doi.org/10.1371/journal.pone.0064702.
[45] D. Matza, A. Badou, K.S. Kobayashi, K. Goldsmith-Pestana, Y. Masuda, A. Komuro, D. McMahon-Pratt, V.T. Marchesi, R.A. Flavell, A scaffold protein, AHNAK1 is required for calcium signalling during T cell activation, Immunity. 28 (2008) 64–74, https://doi.org/10.1016/j.immuni.2007.11.020.
[46] S. Solouki, W. Huang, J. Elmore, C. Limper, F. Huang, A. August, TcR signal strength and antigen affinity regulate CD8+ memory T cells, J. Immunol. Baltim. Md 1950 205 (2020) 1217–1227, https://doi.org/10.4049/jimmunol.1901167.
[47] C. Jiang, Z. Zhou, G. Shu, G. Yin, M. Wang, C. Jiang, Z. Zhou, G. Shu, G. Yin, M. Wang, Over-expression of TNFRSF12A promotes immune suppression and facilitates angiogenesis in triple-negative breast cancer, Biology. (Basel) 14 (2025), https://doi.org/10.3390/biology14111513.
[48] M.S. Rabby, M.M. Islam, S. Kumar, M. Maniruzzaman, M.A.M. Hasan, Y. Tomioka, J. Shin, Identification of potential biomarkers for lung cancer using integrated bioinformatics and machine learning approaches, PLoS. One 20 (2025) e0317296, https://doi.org/10.1371/journal.pone.0317296.