Electronic Journal of Biotechnology 82 (2026) 100714

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

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

Research article

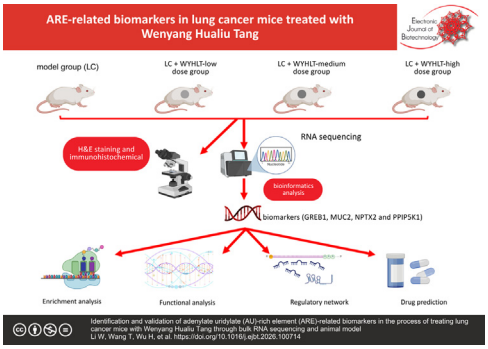
Identification and validation of adenylate uridylate (AU)-rich element (ARE)-related biomarkers in the process of treating lung cancer mice with Wenyang Hualiu Tang through bulk RNA sequencing and animal model ☆

Wen-xin Li [a,1], Teng-yu Wang [a,b,c,1], Hui-ting Wu [a,b,c,*], Ming-feng Xiong [a,b,c,*]

[a] Preventive Center of Traditional Chinese Medicine, The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang, the People's Republic of China
[b] China-Japan Friendship Jiangxi Hospital, National Regional Center for Respiratory Medicine, Nanchang, the People's Republic of China
[c] Ten Layer Pulse Research Institute, Nanchang University, Nanchang, the People's Republic of China

GRAPHICAL ABSTRACT

Identification and validation of adenylate uridylate (AU)-rich element (ARE)-related biomarkers in the process of treating lung cancer mice with Wenyang Hualiu Tang through bulk RNA sequencing and animal model



ARTICLE INFO

Article history:
Received 11 August 2025
Accepted 10 March 2026
Available online 6 May 2026

Keywords:
Adenylate uridylate-rich element
Animal model

ABSTRACT

Background: Lung cancer (LC) remains the leading cause of cancer-related mortality. While the traditional formula Wenyang Hualiu Tang (WYHLT) has shown efficacy against LC, the role of adenylate uridylate (AU)-rich element (ARE)-related genes (AREGs) in this disease and its treatment is unexplored. This study aimed to elucidate the mechanism by which AREGs are involved in WYHLT treatment of LC.
Results: Transcriptome sequencing of a mouse LC model with WYHLT intervention (low, medium, high dose) identified four key biomarkers: GREB1, MUC2, NPTX2, and PPIP5K1. Enrichment analysis linked MUC2, NPTX2, and PPIP5K1 to translation. A constructed regulatory network revealed interactions

Abbreviations: AREGs, adenylate uridylate-rich element-related genes; AU, adenylate uridylate; AUBPs, Adenylate-Uridylate-rich elements binding proteins; CRC, colorectal cancer; DEGs, differentially expressed genes; FC, fold change; GO, Gene Ontology; GSEA, gene set enrichment analysis; H&E, Hematoxylin and eosin; HCC, hepatocellular carcinoma; IARC, International Agency for Research on Cancer; IHC, immunohistochemical; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC, lung cancer; NSCLC, non-small-cell lung cancer; PPI, protein-protein interaction; WYHLT, Wenyang Hualiu Tang.

☆ Audio abstract available in Supplementary material.

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

* Corresponding authors.
E-mail addresses: 2213856748@qq.com (H.T. Wu), 15270889352@163.com (M.F. Xiong).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ejbt.2026.100714

Biomarkers
Bulk RNA sequencing
Circular RNAs (circRNAs)
Long non-coding RNAs (lncRNAs)
Lung cancer
MicroRNAs (miRNAs)
Regulatory network analysis
Transcriptomics
Wenyang Hualiu Tang (WYHLT)

involving three biomarkers (PPIP5K1, GREB1, NPTX2), 16 long non-coding RNAs (lncRNAs), and 10 circular RNAs (circRNAs), with lncRNA Malat1 targeting four microRNAs (miRNAs). Drug prediction analysis identified seven drugs targeting MUC2 and one targeting NPTX2, with PD-98059 showing the highest binding affinity for MUC2. Reverse transcription quantitative polymerase chain reaction (RT-qPCR) validation confirmed that WYHLT intervention increased PPIP5K1 and MUC2 expression and decreased GREB1 and NPTX2 expression compared to the LC model.

Conclusions: This study identifies GREB1, MUC2, NPTX2, and PPIP5K1 as potential biomarkers involved in WYHLT treatment of LC. The findings provide new insights into the molecular mechanisms of WYHLT and novel directions for LC therapeutic research.

How to cite: Li W, Wang T, Wu H, et al. Identification and validation of adenylate uridylate (AU)-rich element (ARE)-related biomarkers in the process of treating lung cancer mice with Wenyang Hualiu Tang through bulk RNA sequencing and animal model. Electron J Biotechnol 2026;82. https://doi.org/10.1016/j.ejbt.2026.100714.

1. Introduction

Lung cancer (LC) is one of the leading causes of cancer-related mortality worldwide [1]. An estimated 2 million new cases and 1.76 million deaths are reported annually [2]. Based on histopathological analysis, lung cancer was classified into four major histological subtypes: small-cell lung cancer, squamous cell carcinoma, adenocarcinoma, and large-cell carcinoma. The latter three, collectively known as non-small-cell lung cancer (NSCLC), account for 80% of all lung cancer cases [3]. According to the latest statistical data released by the International Agency for Research on Cancer (IARC), the numbers of new cases and deaths from lung cancer are continuously increasing each year [4]. Although traditional treatment modalities, including surgery, immunotherapy, chemotherapy, and radiotherapy, can extend patients' survival to a certain extent, they are accompanied by several limitations. For example, drug resistance in conventional therapies is one of the reasons for treatment failure in gastric cancer [5] and also includes limited side effect profiles [6]. The use of traditional Chinese medicine in the treatment of tumors is on the rise because of its low toxicity and minimal side effects. When combined with modern therapies, it not only mitigated the adverse reactions of radiotherapy and chemotherapy but also significantly eased the economic burden on patients' families, boasting reliable efficacy [7,8].

Wenyang Hualiu Tang (WYHLT) is a classic Chinese herbal formula based on Yang-He Tang. It is composed of *Astragalus membranaceus* (Huangqi), *Vaccaria segetalis* (Wangbuliuxing), *Angelica sinensis* (Danggui), *Zingiber officinale* (Ganjjiang), *Atractylodes macrocephala* (Baizhu), *Poria cocos* (Fuling), *Aucklandia lappa* (Muxiang), and *Ligusticum chuanxiong* (Chuanxiong) [9]. Studies have shown that it possesses various biological functions and exerts antitumor effects by inhibiting tumor cell proliferation and promoting apoptosis [10]. Simultaneously, it can regulate inflammatory signaling pathways to achieve anti-inflammatory efficacy [11]. Traditionally, it has been used to treat malignant tumors such as non-small cell lung cancer, breast cancer, and liver cancer [12,13], as well as various types of chronic abscesses characterized by yang deficiency with cold congelation, phlegm-stasis intermingling, and toxin retention [14].

Rooted in the fundamental theory of traditional Chinese medicine, it aligned with the treatment approaches for lung cancer [15]. Moreover, in clinical practice, it was validated that this formula could impede high-risk pulmonary nodules and exhibited a remarkable curative effect on lung cancer [16]. The formula designated Radix Aconiti Lateralis Preparata and *A. membranaceus* as the principal drugs. Additionally, Rhizoma Zingiberis and Semen Sinapis Albae functioned to warm the lungs, strengthen the spleen, and replenish qi to resolve phlegm [17]. The combination of *A. macrocephala* and *P. cocos* served to dry dampness and strengthen the spleen. Honey-fried *Glycyrrhiza uralensis* harmonized the properties of various drugs and reinforced qi while resolving phlegm. *V. segetalis*, in combination with *A. sinensis* and *L. chuanxiong*, excelled at promoting blood circulation, unblocking the meridians, regulating qi, and dissipating nodules, effectively ameliorating the syndrome of blood stasis and coagulation. The integrated use of these herbs fulfilled the functions of tonifying the lung, strengthening the spleen, promoting blood circulation to remove blood stasis, and resolving phlegm to dissipate nodules. Modern research also demonstrated that numerous yang-warming traditional Chinese medicines could play an anti-lung-cancer role by regulating RNA [18]. For instance, atractylenolide I could inhibit PDK1 (protein) and the *lncRNA HOTAIR*, resulting in a reduction in the expression of the downstream effector EZH2. Its combination with the EGFR-TKI erlotinib showed a synergistic effect and offered an additional novel strategy to improve the fight against lung cancer [19]. The activation of IRE1α, TRAF2, p-ASK1, and p-JNK in A549 cells treated with tetramethylpyrazine could be inhibited by knocking down IRE1α with siRNA [20]. Despite the remarkable achievements of WYHLT in lung cancer treatment research, the precise mechanisms of how it regulates the expression of key genes and the profound connections between these gene-expression changes and the alterations in tumor-cell biological behaviors remain to be further explored. The identification of molecular biomarkers has become crucial for early diagnosis and prognosis prediction. Supporting this, epigenetic markers have been demonstrated to effectively forecast disease risk, underscoring the potential of transcriptomic biomarkers in cancer research.

While the antitumor efficacy of WYHLT has been documented, its underlying molecular mechanisms, particularly regarding the comprehensive regulation of gene expression networks, remain largely uncharted. Beyond transcriptional control, post-transcriptional regulation of mRNA stability and translation represents a critical layer in determining cellular protein abundance and function, especially in cancer. In this context, adenylate uridylate (AU)-rich elements (AREs) emerge as pivotal cis-acting regulators. AREs, found in the 3′-untranslated regions of 8–10% of human transcripts, are recognized by a family of RNA-binding proteins known as ARE-binding proteins (AUBPs) [21,22,23,24]. This interaction profoundly influences the stability, degradation, and translation efficiency of mRNAs encoding key players in tumorigenesis, such as cytokines, inflammatory mediators, and oncoproteins [22,23,24]. Critically, dysregulation of this ARE-dependent post-transcriptional control has been implicated in the initiation and progression of various cancers, including colorectal cancer (CRC) and hepatocellular carcinoma (HCC), by modulating the expression

of oncogenes and tumor suppressors [25]. Therefore, investigating the role of ARE-related genes (AREGs) in lung cancer and their potential modulation by WYHLT represents a promising strategy to uncover a more upstream and network-level regulatory mechanism underlying the formula's therapeutic effects. This approach bridges the multi-target characteristics of TCM with contemporary molecular oncology focused on RNA-based regulation.

During the preliminary phase of this study, an animal model of LC treated with WYHLT was established, and transcriptome sequencing was performed to obtain self-sequencing data. AREGs were retrieved from public databases. Using the self-sequencing data, bioinformatics methods (differential analysis), including differential expression analysis, were employed to obtain biomarkers associated with AU-rich elements in the treatment of WYHLT for LC. Furthermore, through functional enrichment analysis, molecular regulatory network construction, drug prediction, and other means, the potential relationships among LC, WYHLT, and AU-rich elements were further explored in-depth. These investigations provided a robust theoretical foundation for the effective diagnosis and treatment of LC.

2. Material and methods

2.1. Cell cultivation

The mouse lung cancer cell lines, Lewis lung carcinoma (LLC, [LL/2; LLC1]), were purchased from Pricella (Wuhan, China). The cells are cultured in a 37°C, 5% CO_2 incubator with DMEM medium, supplemented with 10% FBS and 1% penicillin/streptomycin. Finally, LLC cells in the logarithmic growth phase were collected for the next step in constructing the lung cancer (LC) model.

2.2. Establishment of a subcutaneous tumor mouse model of LC and intervention with WYHLT

In total, 20 male C57BL/6J mice, aged 5 weeks, were obtained from Beijing Sibeifu Bio-Technology Co., Ltd. (Production License No.: SCXK (Beijing) 2019-0010; use License No.: SYXK (Dian) K2022-0006). Firstly, 20 mice were adaptively fed for 1 week and then randomly divided into 4 groups, namely the model group (LC), and the intervention groups (LC + WYHLT-low dose, LC + WYHLT-medium dose, LC + WYHLT-high dose), with 5 mice in each group. A total of 120 μL of cells (LLC cells) at a concentration of 1 × 106 cells/mL were inoculated into the right axilla of each C57 mouse in each group to construct the LC model. After tumor formation, mice in LC group (as LC group) were administered saline by gavage, 200 μl per mouse per day, for 14 d; mice in LC + WYHLT-low dose group (as drug-low group) were administered WYHLT by gavage, 7.0625 g/kg per mouse per day, for 14 d; mice in LC + WYHLT-medium dose group (as drug-medium group) were administered WYHLT by gavage, 14.1250 g/kg per mouse per day, for 14 d; mice in LC + WYHLT-high dose group (as drug-high group) were administered WYHLT by gavage, 28.2500 g/kg per mouse per day, for 14 d. Throughout the experiment, the body weight of mice in 4 groups was monitored and recorded.

Finally, after 14 d of intervention, we euthanized all mice in the 4 groups (5 mice per group, totaling 20 mice), removed the LC tumor tissues from each group, and recorded and compared the weight and volume of the tumor tissues in each group. A portion of LC tumor tissue samples was stored at –80°C. The other portion of LC tumor tissue samples was fixed in 4% paraformaldehyde for histological examination.

2.3. Hematoxylin and eosin (H&E) staining

To compare the histopathological morphology of LC tumor tissues among the 4 groups, H&E staining was performed, following several key steps. Initially, LC tumor tissue samples from 4 groups were fixed in 4% paraformaldehyde for 24 to 48 h. The fixed samples were subsequently dehydrated using alcohol, cleared with dimethylbenzene, infiltrated with paraffin, and embedded. Then, the tissues were cut into 5-μm thick slices. For H&E staining, the tissue slices were incubated at 64°C for 1 h, dewaxed in dimethylbenzene and rehydrated with alcohol. Counterstaining with hematoxylin was performed for 5 min, and next, the slides were stained with eosin for 5–10 s. Finally, the samples were dehydrated, cleared, and mounted, after which whole-slide scan analysis was conducted.

2.4. Immunohistochemical (IHC) analysis

Lung cancer is often accompanied by the infiltration of macrophages. To clarify the regulatory effects of WYHLT on CD68 and Ly6G after treatment, IHC analysis was performed, involving several key steps. Initially, LC tumor tissues among the 4 groups underwent decalcification treatment before being fixed in 4% paraformaldehyde for 24–48 h. The fixed samples were then dehydrated with alcohol, cleared with dimethylbenzene, infiltrated with paraffin, and embedded. Then, the tissues were cut into 5 μm thick slices.

For the IHC assay, the tissue slices were incubated at 64°C for 1 h, dewaxed in dimethylbenzene, rehydrated with alcohol, and subjected to antigen retrieval. Endogenous peroxidase activity was blocked using 3% H_2O_2. The samples were incubated with 5% bovine serum albumin (BSA) at 37°C for blocking and then treated with primary antibodies (Table S1) diluted in 2% BSA. Next, 100 μL of reaction enhancer was applied, and the samples were treated with anti-mouse IgG polymer. Diaminobenzidine (DAB) was used for coloration, and counterstaining with hematoxylin was performed for 5 min. Finally, the samples were dehydrated, cleared, and mounted, after which whole-film scan analysis was conducted.

2.5. RNA sequencing and data preprocessing

In this study, a total of 20 LC tumor tissue samples were obtained from 4 groups for RNA sequencing. Total RNA was extracted from 20 LC tumor tissue samples using TRIzol (Invitrogen, CA, USA). Next, the quantity and integrity of the total RNA were assessed. Samples with concentrations > 50 ng/μL, RIN values > 7.0, OD 260/280 > 1.8, and total RNA > 1 μg were deemed suitable for downstream experiments. Then, two rounds of purification with Dynabeads Oligo (dT) 25-61005 (Thermo Fisher, CA, USA) were used to isolate Poly(A) RNA from 1 μg of total RNA. Subsequently, the Magnesium RNA Fragmentation Module (NEB, cat. e6150, USA) was used to fragment the poly(A) RNA at 94°C for 5–7 min. SuperScript™ II Reverse Transcriptase (Invitrogen, cat. 1896649, USA) was used to reverse-transcribe the cleaved RNA fragments into cDNA. Furthermore, Polymerase Chain Reaction (PCR) amplification was performed. The insert size of the final cDNA library averaged 300 ± 50 bp. Sequencing was carried out on the Illumina NovaSeq 6000 platform using the paired-end 150 bp (PE150) sequencing mode. After sequencing, low-quality reads were filtered out using the Fastp software (https://github.com/OpenGene/fastp). The presence of AT/GC separation was assessed using Phred software (https://www.phrap.org/phred-phrapconsed.html). Importantly, we employed HISAT2 (https://ccb.jhu.edu/software/hisat2) to align the reads with the reference genome of *Mus musculus* GRCm38. Finally, the mapped reads of each specimen were put together via StringTie (https://ccb.jhu.edu/software/stringtie) with default settings. Then, all transcriptomes originating from every sample were combined to establish a more inclusive transcriptome by means of gffcompare (https://github.com/gpertea/gffcompare/). Following the generation of the

final transcriptome, StringTie was utilized to ascertain the expression levels of mRNAs through the computation of FPKM (FPKM = [-total exon fragments/mapped reads (millions) × exon length(kB)]). Principal component analysis (PCA) was then performed on the 4 groups of sequencing data using the FPKM.

2.6. Screening of differentially expressed genes (DEGs) and candidate genes

In differential expression analysis, we used the DESeq2 R package. To control for potential technical variation introduced by sequencing the samples in two separate batches, we included the sequencing batch as a covariate in the design formula of DESeq2. Specifically, the model design for differential comparisons was ~ sequencing_batch + group, where group represents different treatment groups (e.g., LC model group, WYHLT intervention group). The "DESeq2″ package (v 1.42.0) [26] was used to perform 4 separate differential expression analyses, aiming to identify 3 DEGs with $|\log_2$Fold Change (FC)| > 1 and $p < 0.05$. Specifically, DEGs-L were identified by comparing case samples from the drug-low group with samples from the LC group. DEGs-M were determined by comparing case samples from the drug-medium group with samples from the LC group. DEGs-H were found by comparing case samples from the drug-high group with samples from the LC group. Meanwhile, to respectively illustrate the results of each DEGs, the volcano maps were plotted via "ggplot2″ package (v 3.4.1) [27] and the DEGs were sorted according to log2(fold change)FC, with the up- and down-regulated top 10 genes in the volcano plot were labeled. The heatmaps were drawn via "pheatmap" package (v 1.0.12) [28] to show up- and down-regulated top10 genes in DEGs.

Following this, the "VennDiagram" package (v 1.7.3) [29] was employed to identify up-regulated candidate genes by overlapping up-regulated genes in DEGs-L, up-regulated genes in DEGs-M and up-regulated genes in DEGs-H. The down-regulated candidate genes were obtained using the same method, by intersecting the down-regulated genes in DEGs-L, down-regulated genes in DEGs-M and down-regulated genes in DEGs-H. Finally, the up-regulated candidate genes and down-regulated candidate genes were combined to obtain candidate genes.

2.7. Functional analysis and protein–protein interaction (PPI) network construction of candidate genes

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were employed to reveal the biological processes ($p < 0.05$) of candidate genes via "clusterProfiler" (v 4.7.1.003) [30]. Subsequently, for GO, the enriched pathways were ranked from highest to lowest according to the number of genes they contained, and the top 3 pathways were presented. For KEGG, the top 5 pathways with the most prominent enrichment results were selected and presented. A PPI network (interaction scores = 0.4) was created by the STRING online site (https://string-db.org) and displayed through "Cytoscape" (v 3.9.1) [31].

2.8. Screening of biomarkers related to AU-rich element and correlation analysis

A total of 4883 AU-rich element-related genes (AREGs) were acquired from ARED (https://brp.kfshrc.edu.sa/ared) (Table S2). To identify potential biomarkers that are both differentially expressed in response to WYHLT treatment and functionally related to AU-rich element-mediated post-transcriptional regulation, we intersected the list of 44 candidate genes (obtained in Section 2.6) with the curated list of 4883 AREGs. This overlap was performed using the "VennDiagram" package (v 1.7.3) [29]. Genes appearing in both lists were defined as ARE-related biomarkers for further analysis. This approach aligns with established bioinformatics workflows for integrating differential expression data with functional gene sets to prioritize candidate biomarkers [32].

Additionally, to clarify the correlation between biomarkers, Spearman correlation coefficients (r) among each biomarker in all samples were calculated via "psych" (v 2.1.6) [33] ($|r| > 0.3$ and $p < 0.05$).

2.9. Functional enrichment analysis

To understand the role of biomarkers in biological processes, gene set enrichment analysis (GSEA) was carried out on all samples in the transcriptome sequencing dataset. Firstly, the "m2.cp.v2023.1.Mm.symbols.gmt" was obtained from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb) to serve as a background gene set. Subsequently, Spearman correlation coefficients among biomarkers and genes in the transcriptome sequencing dataset were calculated via "psych" (v 2.1.6) [33]. The correlation coefficients were then sorted, and GSEA was conducted using the "clusterProfiler" package (v 4.7.1.003) [30] (p adj < 0.05). The pathways were sorted according to p values from smallest to largest, and the top 10 pathways were chosen for display.

Additionally, to investigate potential functional similarities between the biomarkers, the correlation analysis was carried out with the "OmicCircos" package (v 1.40.0) [34]. At last, the GeneMANIA database (https://genemania.org/) was employed to forecast genes related to biomarker function and explore the closely related functions.

2.10. Chromosomal localization and subcellular analyses

In order to identify the location of biomarkers on mouse chromosomes, the "RCircos" package (v 1.2.2) [35] was used to visualize the chromosomal distribution of biomarkers.

Additionally, to explore the distribution of biomarkers within subcells, the FASTA files for the biomarkers were first downloaded from the Gene database (https://www.ncbi.nlm.nih.gov/gene/). Subsequently, subcellular localization analysis was implemented using the mRNALocater database (https://www.csbio.sjtu.edu.cn/bioinf/mRNALocater/).

2.11. Development of molecular regulatory networks

To examine the molecular mechanisms regulating biomarkers, based on miRanda database (https://mirtoolsgallery.tech/mirtoolsgallery/node/1055) and miRDB database (https://mirdb.org/) in "multiMiR" package (v 0.98.0.2) [36], the microRNAs (miRNAs) corresponding to each biomarker were obtained. Then, the shared miRNAs from the 2 databases were chosen as the key miRNAs. Subsequently, we based on the key miRNAs, separately predicted the associated lncRNAs and circRNAs using the StarBase database (https://rnasysu.com/encori/), and generated miRNA-lncRNA and miRNA-circRNA interaction pairs. Finally, on the basis of the aforementioned biomarkers, miRNAs, lncRNA and circRNA, the biomarkers-miRNA-lncRNA/circRNA regulatory network was drawn using "Cytoscape" software (v 3.9.1) [31].

2.12. Drug prediction analysis and molecular docking

In order to explore potential drugs that could regulate biomarkers, the DGIdb (https://dgidb.org/) was used to search for drugs associated with biomarkers. The predicted drugs were subsequently loaded into "Cytoscape" software to visualize the biomarkers-drugs network. Moreover, in order to ascertain the interaction between the predicted drugs and biomarkers, molecu-

lar docking could be performed to determine the most promising drugs. Firstly, we searched the Protein Data Bank (PDB, https://www.rcsb.org) and the PubChem database (https://pubchem.ncbi.nlm.nih.gov) to obtain the protein structure data of biomarkers and the structural data of predicted drugs, respectively. Then, the CB-Dock2 online website (https://cadd.labshare.cn/cb-dock2/php/index.php) was used to add non-polar hydrogens to the three-dimensional structures and calculate charges for molecular docking simulation. A docking score under −1.2 kcal/mol pointed to the fact that the drug had a high binding affinity with the target. Finally, the molecular docking result with the highest binding energy was selected and visualized using Pymol software.

2.13. Expression level analysis and RT-qPCR of biomarkers

In our transcriptome sequencing dataset, based on the biomarker expression in differential analysis, the expression results of biomarkers in the model group (LC group) and the intervention groups (drug-low, drug-medium and drug-high groups) were displayed using a violin plot ($p < 0.05$).

RT-qPCR was then performed to confirm the expression of the biomarkers. Specifically, a sum of 20 LC tumor tissues from 4 groups was used. RNA extraction from the 20 samples was performed using TRIzol (Ambion, Austin, USA). Then, reverse transcription to cDNA was carried out using the SureScript First-Strand cDNA Synthesis Kit (Servicebio, Wuhan, China). Primer sequences for PCR could be found in Table S3. RT-qPCR was conducted using the 2xUniversal Blue SYBR Green qPCR Master Mix (Servicebio, Wuhan, China). M-GAPDH served as an internal reference gene. The relative quantification of biomarkers was determined using the 2-ΔΔCT method [37]. Finally, Graphpad Prism (v 5.0.0) [38] was employed to plot and calculate the p-value ($p < 0.05$).

2.14. Statistical analysis

Statistical analysis was carried out using R (v 4.2.2). Difference analysis between groups was executed via the Wilcoxon test ($p < 0.05$). Differences between PCR experimental categories were obtained through a *t*-test.

3. Results

3.1. Development and assessment of mouse models for LC

Differences in tumor volume and weight between the LC group and each treatment group were analyzed using independent two-sample t-tests (or Wilcoxon rank-sum tests where appropriate). Overall differences among all groups were assessed using one-way ANOVA or Kruskal-Wallis test, followed by post-hoc pairwise comparisons. Effect sizes were calculated using Cohen's d. A *p*-value < 0.05 was considered statistically significant. In the LC animal model, compared to the LC group, the mice in drug-low, drug-medium and drug-high groups exhibited a marked decrease in body weight (Fig. 1A). More importantly, a notable and dose-dependent decrease in both the volume and weight of LC tumor tissues was observed (Fig. 1B–C, Fig. S1A–C). Specifically, compared to the LC group (mean tumor volume: 802.46 ± 60.73 mm^3, mean ± SEM), the average tumor volumes in the drug-high, drug-medium, and drug-low groups were 302.70 ± 69.83 mm^3, 397.98 ± 37.90 mm^3, and 529.99 ± 45.10 mm^3, respectively. These represent reductions of 62.28%, 50.41%, and 33.95% relative to the LC control group (Fig. S1D). Statistical analysis using independent t-tests revealed that all WYHLT-treated groups showed significantly smaller tumor volumes compared to the LC group (drug-high vs. LC: $p = 0.0013$; drug-medium vs. LC: $p = 0.0005$; drug-low vs. LC: $p = 0.0177$). The effect sizes, quantified by Cohen's d, were large for all comparisons (drug-high: d = −3.045; drug-medium: d = −3.511; drug-low: d = −1.881; Fig. S1E). Furthermore, one-way ANOVA followed by post-hoc tests confirmed significant overall differences among the four groups (Kruskal-Wallis $\chi^2 = 17.331$, $p = 0.0006$). These findings suggested that WYHLT effectively inhibited tumor growth in a mouse model of LC in a dose-dependent manner.

HE staining revealed that the cancer cells in the LC group exhibited significantly enlarged nuclei, an altered nucleocytoplasmic ratio, and irregularly shaped cell depressions. In contrast, drug-low, drug-medium and drug-high groups displayed a reduction in the number of cancer cells, with more intact cell morphology and lower degrees of differentiation (Fig. 1D). These findings suggested that treatment with WYHLT significantly improved tissue development in the LC model.

IHC showed that in the LC group, both CD68 and Ly6G proteins were highly expressed. However, in the LC tumor tissues of drug-medium and drug-high groups, both CD68 and Ly6G protein levels were significantly reduced (Fig. 1E–F). These results indicated that treatment with WYHLT exerted an inhibitory effect on the development of lung cancer.

3.2. Identification of 44 candidate genes

First and foremost, high-quality reads were obtained after filtering the transcriptome data. Additionally, the excellent base quality in the detected data further demonstrated its accuracy (Table S4). Comparison with the reference dataset revealed that exonic regions accounted for the highest proportion in the distribution map of reads from all samples, exonic regions accounted for the highest proportion (Fig. S2A). Meanwhile, the overall distribution of gene expression levels across all samples was largely consistent, and the dispersion within the sample groups was relatively smaller than that between the groups (Figs. S2B–C). These findings suggest that the transcriptome sequencing data were of high quality and appropriate for further analysis.

Differential expression analysis identified a total of 6484 DEGs-L (3510 up-regulated and 2974 down-regulated) between the drug-low group and the LC group (Fig. 2A–B), 2735 DEGs-M (1198 up-regulated and 1537 down-regulated) between the drug-medium group and the LC group (Fig. 2C–D) and 1202 DEGs-H (572 up-regulated and 630 down-regulated) between the drug-high group and the LC group (Fig. 2E–F). Subsequently, the 3510 DEGs-L-UP, 1198 DEGs-M-UP and 572 DEGs-L-UP were overlapped, as well as the 2974 DEGs-L-DOWN, 1537 DEGs-M-DOWN and 630 DEGs-L-DOWN (Fig. 2G–H). This overlap resulted in the identification of 44 candidate genes, containing 6 up-regulated candidate genes and 38 down-regulated candidate genes. Moreover, enrichment analysis revealed that 44 candidate genes were significantly associated with 415 GO terms, which included 333 biological processes (BPs), 28 cellular components (CCs), and 54 molecular functions (MFs). Notable GO terms encompassed ''heart contraction'' (BP), ''myofibril'' (CC), and ''actin binding'' (MF) (Fig. 3A, Table S5). Additionally, the candidate genes showed significant enrichment in 15 KEGG pathways, including ''cytoskeleton in muscle cells'', ''motor proteins'' and so on (Fig. 3B, Table S6). To further understand the protein interaction relationship of the 44 candidate genes, after the removal of the isolated protein, a PPI network including 71 reciprocal relationships of 32 proteins was created, showing that genes such as Scgb1a1 interacted closely with others like Bpifb1 and Lamp3 within the network (Fig. 3C).

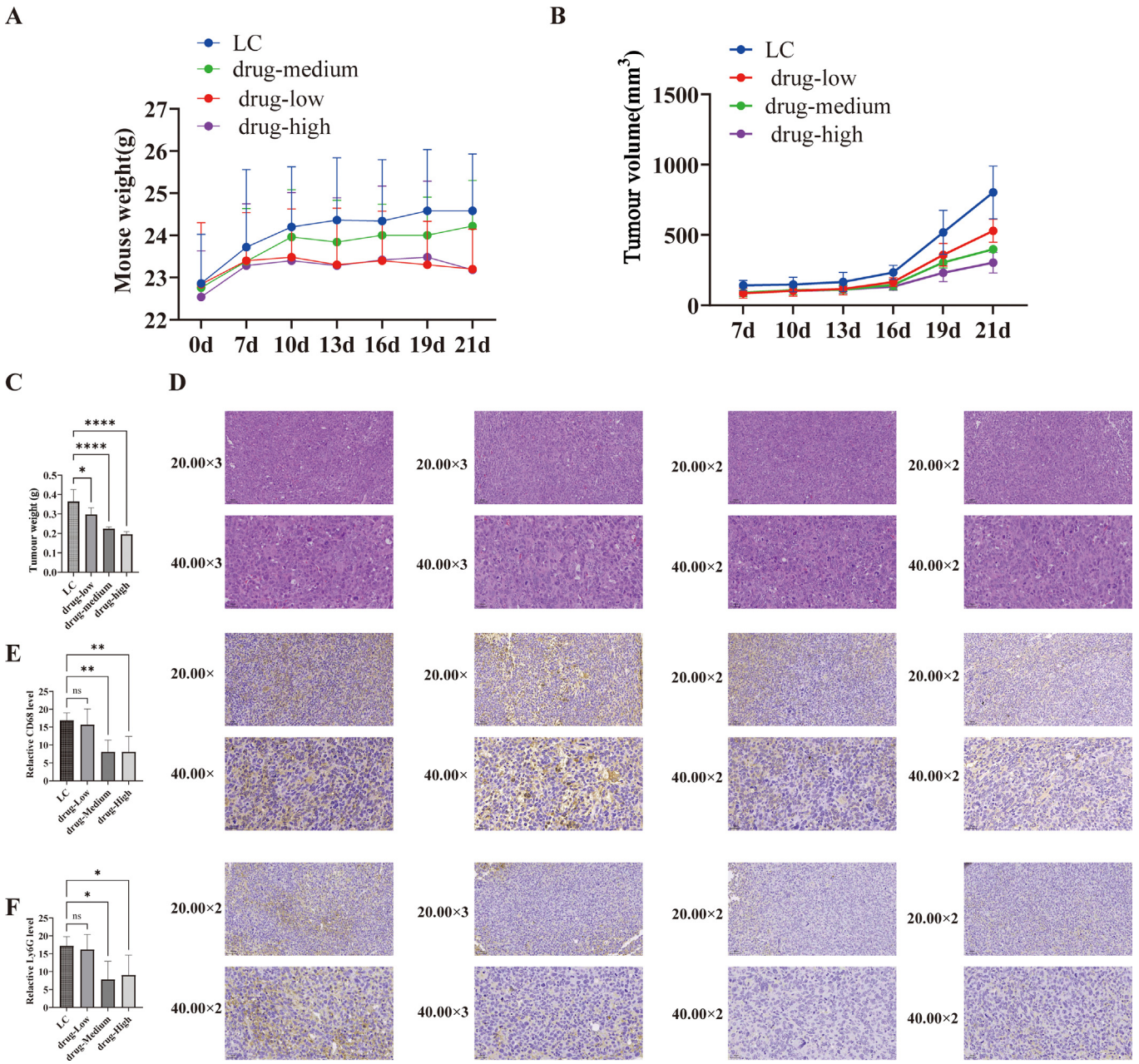


Fig. 1. Experimental analysis of body weight, tumor volume, and histological characteristics in LC, drug-low, drug-medium, and drug-high groups. (A) Changes in body weight in the LC, drug-low, drug-medium, and drug-high groups. (B) Changes of tumor volume for LC tumor tissues in LC, drug-low, drug-medium, and drug-high groups. Data are presented as mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ vs. LC group (independent *t*-test). (C) Tumor weight in the LC, drug-low, drug-medium, and drug-high groups. (D) Hematoxylin and eosin (H&E) staining of tumor tissues from the LC, drug-low, drug-medium, and drug-high groups. (E) Immunohistochemistry (IHC) analysis of CD68 protein expression in tumor tissues from the LC, drug-low, drug-medium, and drug-high groups. (F) IHC analysis of Ly6G protein expression in tumor tissues from the LC, drug-low, drug-medium, and drug-high groups. For each group with 5 samples, the statistical testing methods need to be provided by the laboratory. Statistical analysis was performed using Prism GraphPad. All experimental data are expressed as mean ± standard deviation (Mean ± SD). Comparisons of quantitative data between two groups were conducted using the *T*-test, while comparisons among multiple groups were performed using one-way ANOVA. A *p*-value < 0.05 was considered statistically significant; $p < 0.01$ was considered highly statistically significant; and $p < 0.001$ was considered extremely statistically significant.

3.3. Chosen GREB1, MUC2, NPTX2 and PPIP5K1 as biomarkers

This study ultimately identified 4 biomarkers associated with AU-rich element from 44 candidate genes, namely GREB1, MUC2, NPTX2 and PPIP5K1 (Fig. 4A). In addition, Spearman correlation analysis revealed a strong negative correlation between MUC2 and NPTX2 ($r = -0.57$, $p < 0.01$), and a notable positive correlation between GREB1 and NPTX2 ($r = 0.57$, $p < 0.01$) (Fig. 4B). To ensure the robustness of the research results, we accounted for potential technical variations introduced by the sequencing of samples in two batches by including "sequencing batch" as a fixed effect in the DESeq2 model (~ sequencing_batch + group) for reanalysis. The supplementary analysis confirmed that the four previously identified key biomarkers (GREB1, MUC2, NPTX2, PPIP5K1) exhibited consistent directions of differential expression across all intergroup comparisons (i.e., the up-/down-regulation trends following WYHLT intervention compared to the LC model group) and still met the criteria for significant differences ($|log2FC| > 1$, $p < 0.05$). This indicates that the core findings of this study are not affected by sequencing batches, and the conclusions are reliable.

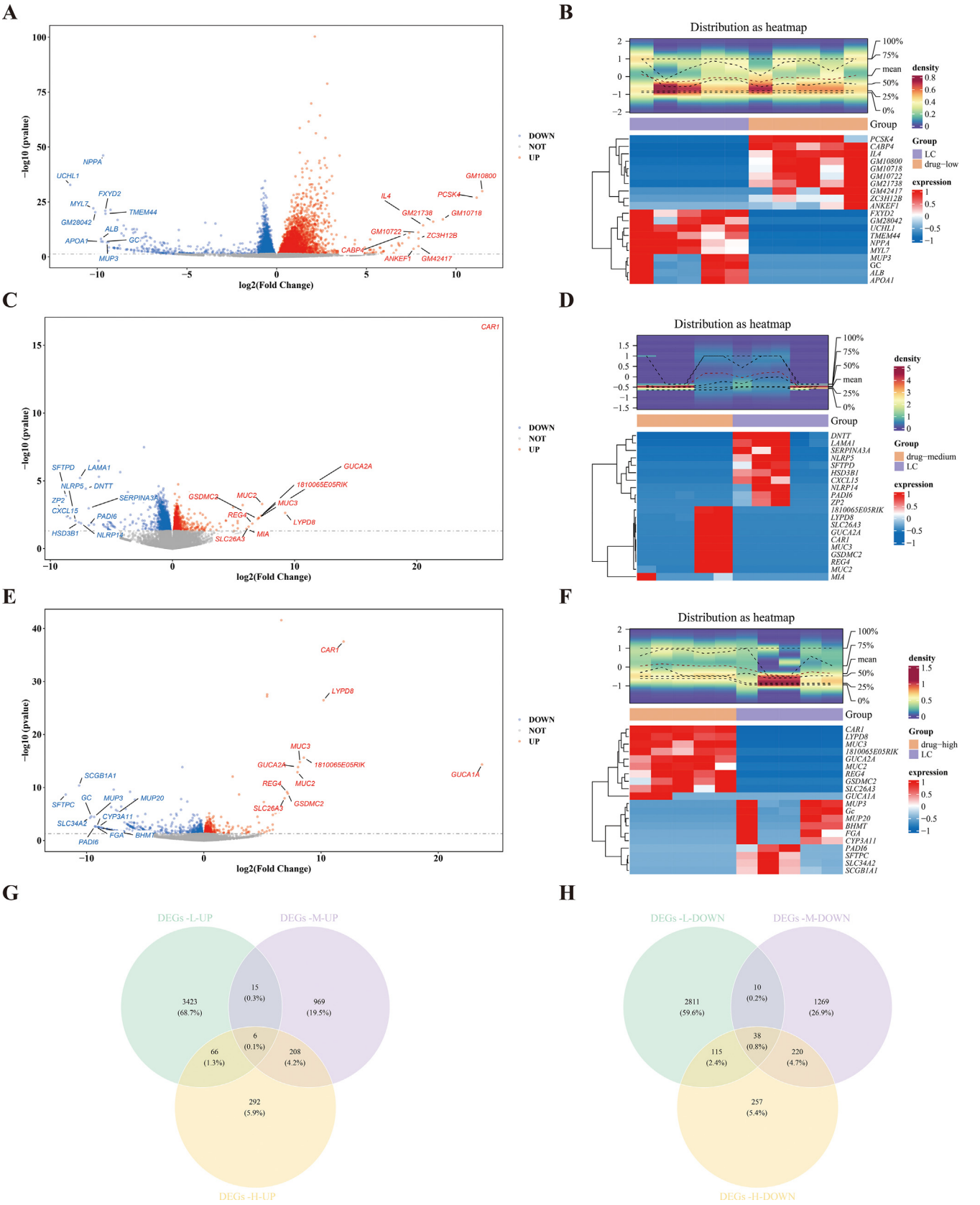


Fig. 2. Differential expression analysis and overlap of differentially expressed genes (DEGs) between drug-low, drug-medium, drug-high, and LC groups. (A) Volcano plot of DEGs between drug-low group and LC group ($|\log_2 \text{foldchange(FC)}| > 1.0$, $p < 0.05$). (B) Heatmap of the top 10 up- and down-regulated DEGs between the drug-low group and the LC group. (C) Volcano plot of DEGs between drug-medium group and LC group ($|\log_2 FC| > 1.0$, $p < 0.05$). (D) Heatmap of the top 10 up- and down-regulated DEGs between the drug-medium group and the LC group. (E) Volcano plot of DEGs between drug-high group and LC group ($|\log_2 FC| > 1.0$, $p < 0.05$). (F) Heatmap of the top 10 up- and down-regulated DEGs between the drug-high group and the LC group. (G) Venn diagram illustrating the overlap of up-regulated DEGs across drug-low, drug-medium, and drug-high groups compared to the LC group. (H) Venn diagram illustrating the overlap of down-regulated DEGs across drug-low, drug-medium, and drug-high groups compared to the LC group. The statistical test method for the volcano plot and heatmap is the Wald test from DESeq2 (based on the negative binomial distribution), while no statistical test was applied for the Venn diagram. Each group contains five samples.

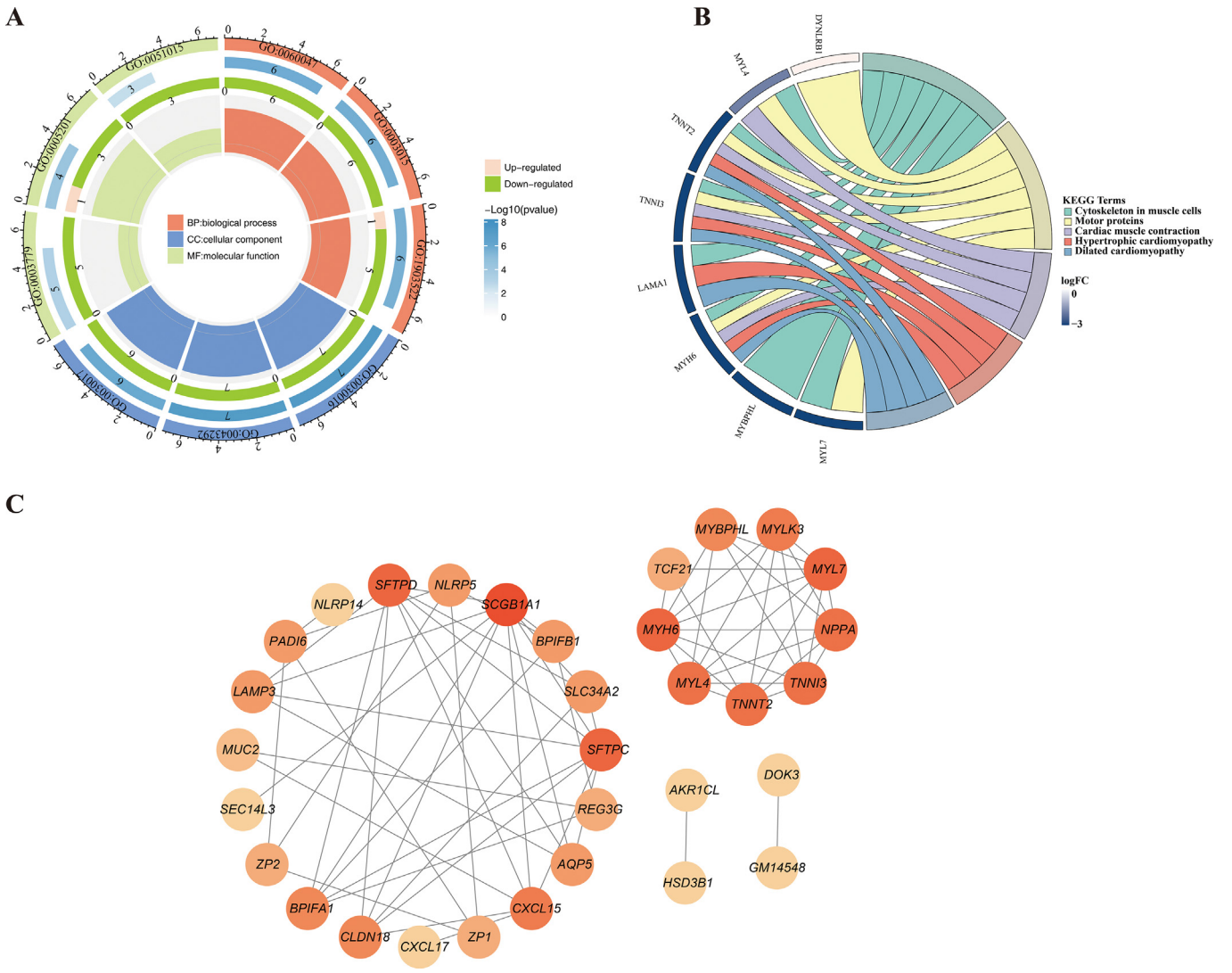


Fig. 3. Functional enrichment and protein–protein interaction (PPI) network analysis of candidate genes. (A) A circular plot of Gene Ontology (GO) enrichment analysis for candidate genes. The figure displays the top three enriched terms (belonging to Biological Process-BP, Cellular Component-CC, and Molecular Function-MF, respectively) filtered based on *p*-value ($p < 0.05$). The radial length of each sector corresponds to the number of genes enriched in that term. The enrichment significance of each term is indicated by its *p*-value and annotated in the figure (or represented by color intensity). (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of candidate genes, highlighting significant pathways ($p < 0.05$). (C) PPI network of candidate genes. The PPI analysis did not involve statistical testing and only provided a confidence score of 0.4. For enrichment analysis, the statistical method used was the hypergeometric distribution test (also known as Fisher's exact test) implemented in the clusterProfiler package, with FDR correction applied. The analysis was conducted using all samples ($n = 20$).

3.4. Exploring the pathways through which biomarkers might influence the development of LC

To investigate the potential biological functions and pathways involved in the progression of lung cancer for the four identified biomarkers (GREB1, MUC2, NPTX2, and PPIP5K1), positive enrichment indicates that genes positively correlated with the biomarker are overrepresented in the pathway, whereas negative enrichment indicates that genes negatively correlated with the biomarker are overrepresented. GSEA revealed that GREB1 was significantly enriched in 26 pathways, such as developmental biology (Fig. 5A, Table S7), MUC2 was closely related to 265 biological pathways, like translation (Fig. 5B, Table S8), NPTX2 was closely related to 55 biological pathways, like cytoplasmic ribosomal proteins (Fig. 5C, Table S9), and genes most negatively correlated with PPIP5K1 were enriched for pathways such as mRNA splicing (Fig. 5D, Table S10). These results indicate that MUC2, NPTX2, and PPIP5K1 are all closely associated with the 'translation' process, suggesting that Wenyang Hualiu Tang may exert its therapeutic effects by modulating protein synthesis metabolism in lung cancer cells. Among them, 3 biomarkers (MUC2, NPTX2 and PPIP5K1) were associated with translation. Subsequently, the correlation analysis between biomarkers revealed that the functional similarity score of GREB1 and other biomarkers was the highest, indicating that it might have functional similarities with other genes (Fig. 5E). Through GeneMANIA analysis, 20 genes related to the function of the biomarkers were identified. These genes interact with the biomarkers through 7 different mechanisms, including complement binding, opsonin binding and so on (Fig. 5F). These analyses identified biomarkers and their related pathways, providing insight into their potential involvement in LC progression.

3.5. Identified the chromosomal and subcellular locations of the biomarkers

Chromosomal localization analysis revealed that GREB1 was located on chromosome 12, MUC2 on chromosome 7, NPTX2 on chromosome 5, and PPIP5K1 on chromosome 2 (Fig. 6A). The chro-

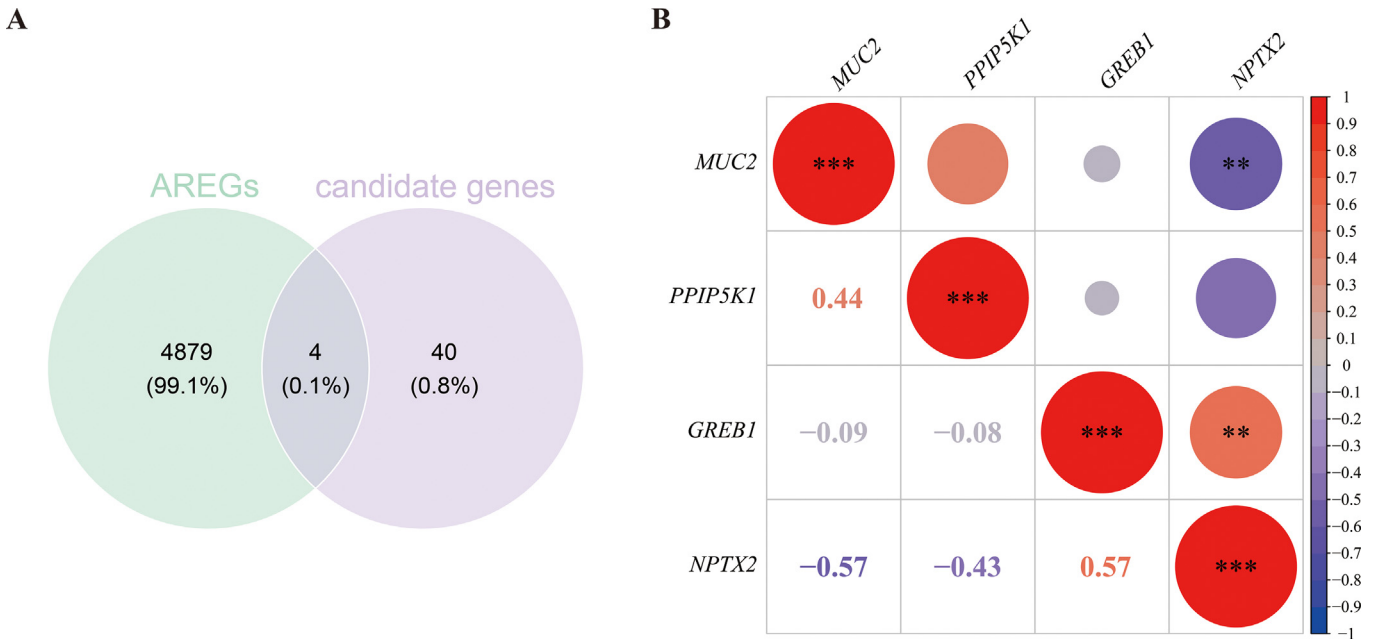


Fig. 4. Identification and correlation analysis of biomarkers. (A) Identification of four biomarkers (GREB1, MUC2, NPTX2, and PPIP5K1). The green circle represents adenylate uridylate-rich element-related genes, with a total of 4883 genes. The purple circle represents candidate genes, containing 44 genes. There are 4 overlapping genes in the intersection region, accounting for 0.1%. (B) Spearman correlation analysis between biomarkers. The statistical test method used for correlation analysis is the Spearman rank correlation test, with all samples included (n = 20). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mosomal localization of these biomarkers (Fig. 6A) provides a genomic context that may inform future studies on genetic regulation or biomarker association analysis. Additionally, subcellular localization analysis provided new directions for the functional study of these biomarkers. GREB1 and PPIP5K1 were predominantly localized in the cell nucleus, while MUC2 and NPTX2 were mainly found in the cytoplasm (Fig. 6B).

3.6. Exploring the regulation mechanism of the biomarkers

Subsequently, to explore the potential regulatory mechanisms of these biomarkers, we constructed a biomarker-miRNA-lncRNA/circRNA regulatory network. The core components of this network can be summarized as follows (Table S11, Fig. 7):

(1) Biomarkers and their targeted miRNAs: A total of 7 key miRNAs were predicted for the three biomarkers (PPIP5K1, GREB1, NPTX2). Among them, 5 miRNAs target PPIP5K1 (mmu-miR-743a-3p, mmu-miR-122-5p, mmu-miR-761, mmu-miR-450b-3p, mmu-miR-743b-3p), 1 miRNA targets GREB1 (mmu-miR-345-3p), and 1 miRNA targets NPTX2 (mmu-miR-96-5p).
(2) Competitive endogenous RNA (ceRNA) interactions: Based on the aforementioned miRNAs, we further predicted their upstream regulatory lncRNAs and circRNAs. The network includes a total of 16 lncRNAs and 10 circRNAs.
(3) Key hub molecules: Notably, one lncRNA (Malat1) simultaneously targets 4 miRNAs (mmu-miR-345-3p, mmu-miR-743a-3p, mmu-miR-743b-3p, mmu-miR-96-5p), suggesting that Malat1 may serve as a core ceRNA by sponging these miRNAs to regulate the expression of downstream biomarkers. Additionally, one circRNA (Etv3) is associated with two miRNAs (mmu-miR-450b-3p and mmu-miR-761).

These findings reveal a complex regulatory network involving miRNAs, lncRNAs, and circRNAs, which may collectively influence the progression of lung cancer.

3.7. Various drugs targeting biomarkers for LC treatment

Given the potential critical role of the aforementioned biomarkers in lung cancer, we further utilized drug databases to predict potential therapeutic agents that may target these markers, aiming to provide candidate compounds for translational research. Subsequently, we constructed a biomarker-drug network encompassing 7 drugs (fluorouracil, GPI-0100, dextran sulfate sodium, chemopreventive agent, PD-98059, recombinant interferon and decitabine) targeting MUC2 and a drug (zinpentraxin alfa) targeting NPTX2 (Fig. 8A). Molecular docking revealed that PD-98059 exhibits the highest binding affinity (−7.8 kcal/mol) with the MUC2 protein, suggesting its high theoretical feasibility as a lead compound targeting MUC2. As shown in the molecular docking results, the docking scores of MUC2 and 3 drugs (decitabine, fluorouracil and PD-98059) were all lower than −1.2 kcal/mol, indicating a high affinity between the biomarker and the drugs; among them, the binding energy between MUC2 and PD-98059 was the highest (binding energy = −7.8 kcal/mol) (Table 1, Fig. 8B).

3.8. Validation of biomarker expression levels in LC animal models

In our transcriptome sequencing dataset, all 3 intervention groups (drug-low, drug-medium and drug-high groups) showed a significant increase in the expression levels of PPIP5K1 and MUC2 compared to the LC groups ($p < 0.05$). In contrast, GREB1 and NPTX2 expression levels were significantly diminished in all 3 intervention groups ($p < 0.05$) (Fig. 9A). Furthermore, RT-PCR analysis of LC tissue samples from the animal model validated the expression of these biomarkers ($p < 0.05$). The expression trends of most genes across different groups were consistent with the bioinformatics predictions. Unfortunately, the drug-low group showed higher MUC2 expression and lower NPTX2 expression compared to the LC group, but the differences were not significant (Fig. 9B).

4. Discussion

This study is the first to systematically investigate the expression changes and regulatory networks of ARE-related genes in a

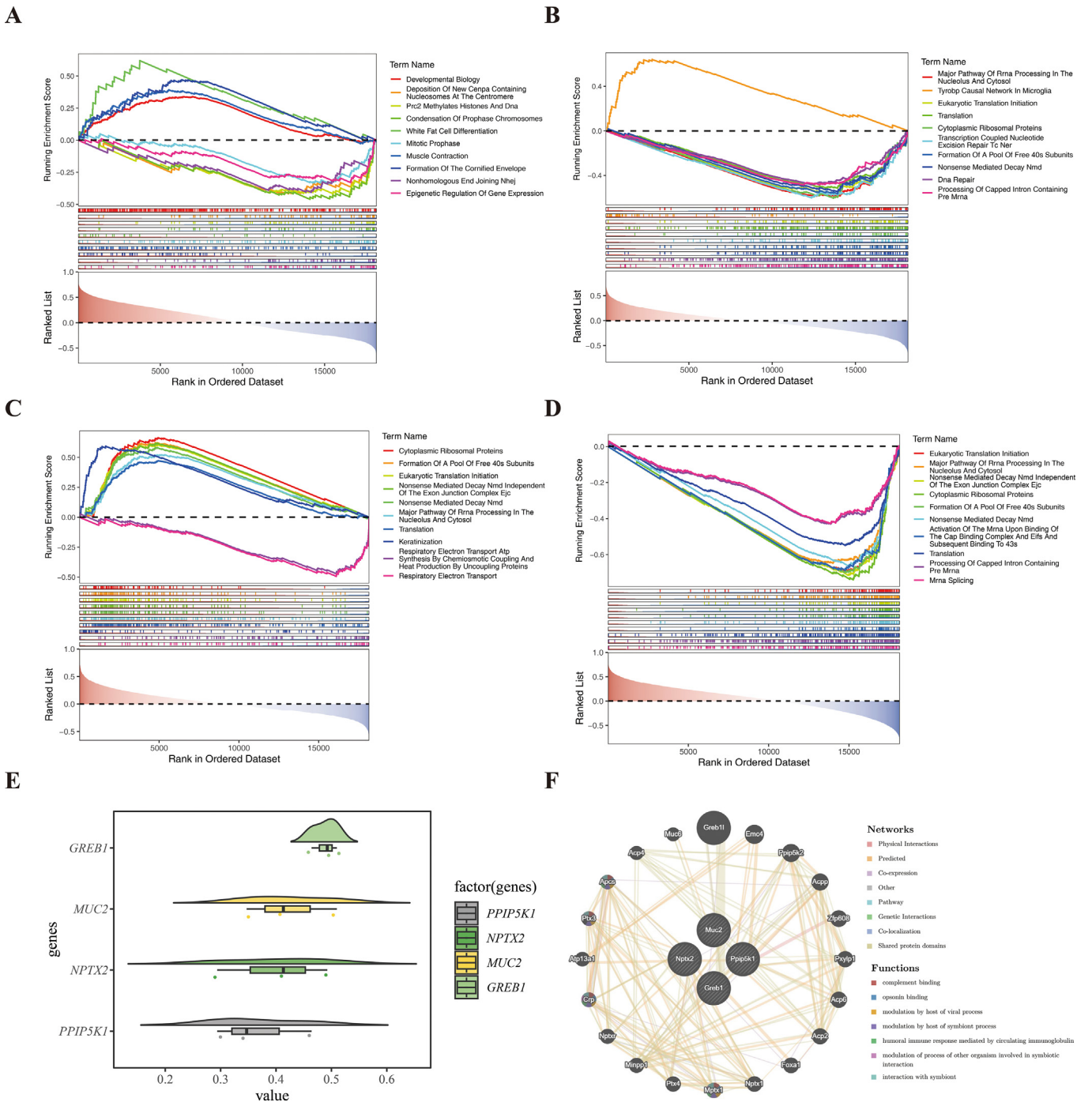


Fig. 5. Gene set enrichment analysis (GSEA) and associated pathways of biomarkers. (A) GSEA GREB1; (B) GSEA of MUC2; (C) GSEA of NPTX2; (D) GSEA of PPIP5K1; (E). Functional similarity analysis between biomarkers. (F). GeneMANIA analysis of the biomarkers. GSEA employs the pre-ranked gene set enrichment analysis (GSEA) method by calculating the Spearman rank correlation coefficient between each hub gene and all other genes across the genome. Based on this, genes are ranked, and permutation tests are used to evaluate the significance of enrichment of specific functional gene sets at the top or bottom of the ranked list, thereby identifying biological pathways significantly associated with the expression patterns of hub genes. Similarity analysis is based on the GO semantic similarity measure using the Wang algorithm. This method quantifies functional similarity by comparing the topological structure and semantic relationships of GO terms corresponding to genes in the GO directed acyclic graph. It integrates the comprehensive similarity of three ontologies—biological process, cellular component, and molecular function—and finally derives the overall functional similarity score for gene pairs through the geometric mean of the scores from these three dimensions. GeneMANIA utilizes a Gaussian Random Field Classifier for prediction, and the prediction results are evaluated using a hypergeometric distribution test.

model system of lung cancer intervention with a traditional Chinese medicine compound (Wenyang Huailiu Tang). This marks the introduction of the modern molecular biological mechanism of ARE-mediated post-transcriptional regulation into the interpretation of the efficacy of traditional Chinese medicine compounds, establishing a novel research connection. In the preliminary phase of this study, an animal model of LC intervened by WYHLT was successfully established. The low-dose, medium-dose, and high-dose drug-intervention groups were meticulously subjected to differential analysis in comparison with the model group. This analysis yielded a total of 44 candidate genes. These candidate genes were then crossed with 4883 AREGs, ultimately leading to the identifica-

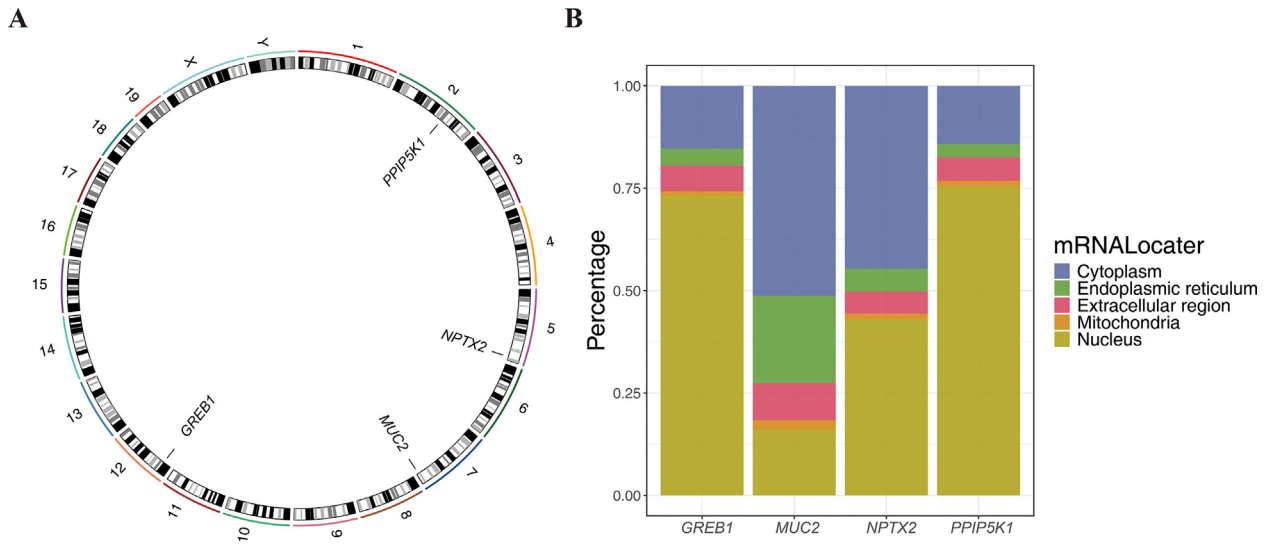


Fig. 6. Chromosomal and subcellular localization of biomarkers. (A) Chromosomal localization of GREB1, MUC2, NPTX2, and PPIP5K1. (B) Subcellular localization of GREB1, MUC2, NPTX2, and PPIP5K1.

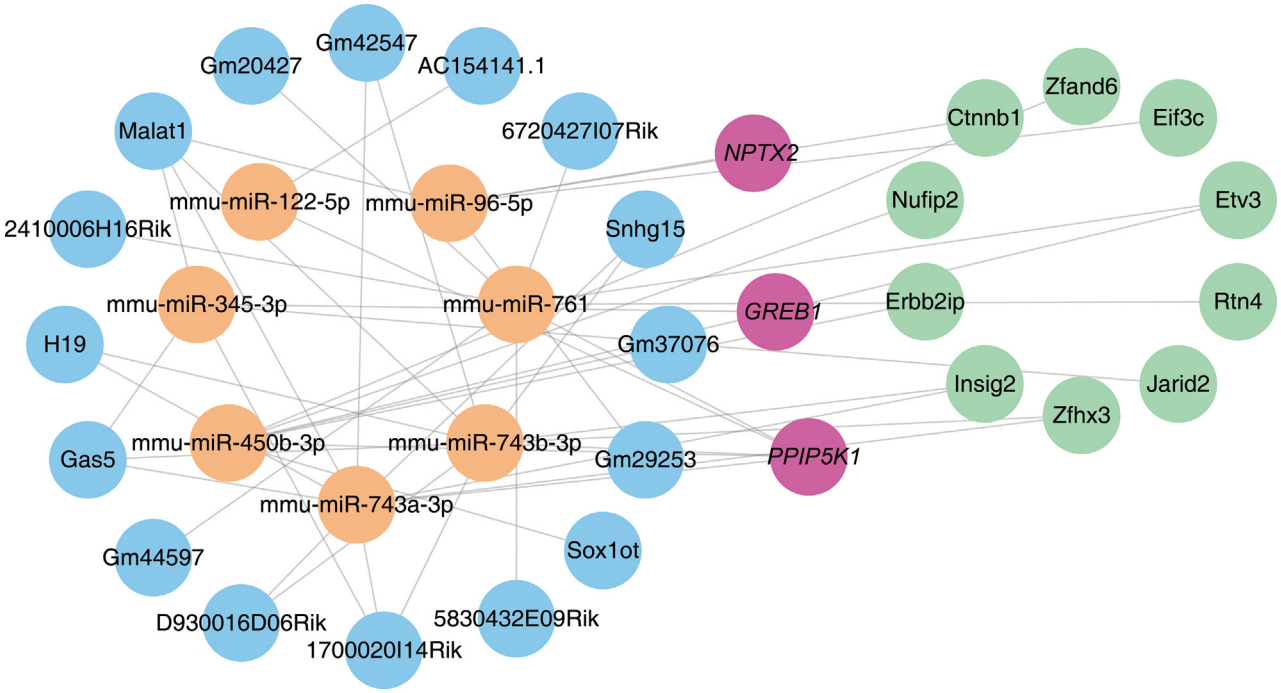


Fig. 7. Biomarker-miRNA-lncRNA/circRNA regulatory network.

tion of four biomarkers: GREB1, MUC2, NPTX2, and PPIP5K1. Quantitative detection was performed by RT-qPCR, and the expression trends obtained were consistent with those of previous studies. Furthermore, the regulatory mechanisms of AREGs in LC were comprehensively analyzed based on transcriptome data. The discovery of these biomarkers has undeniably opened up a novel direction for lung cancer research and holds significant potential value.

Our research found that by identifying four specific AREGs (GREB1, MUC2, NPTX2, and PPIP5K1) as key mediators through which WYHLT exerts its effects, we have advanced the current understanding of the pathogenesis of lung cancer and the mechanism of action of traditional Chinese medicine. This expands the known functions of AREGs, extending their relevance from fields such as CRC and hepatocellular carcinoma (HCC) to the drug-modulable treatment of lung cancer. The significance of these biomarkers in cancer has been supported by the literature. For example, studies have shown that GREB1 mutations may mediate primary resistance to cisplatin treatment and influence the clinical prognosis of LUAD patients [39]. MUC2 is a gene that encodes a mucin. Mucins are high molecular weight glycoproteins composed of polysaccharides and proteins that are abundant on various mucosal surfaces of the human body. Among them, mucins (MUC1–MUC24) form a family of glycoproteins involved in cell signaling and barrier protection and have been implicated in the progression of numerous malignancies, including gastric, pancreatic, ovarian, breast and lung cancer [40]. MUC2 is negative in all primary lung cancers, but is positive in less than half of lung metas-

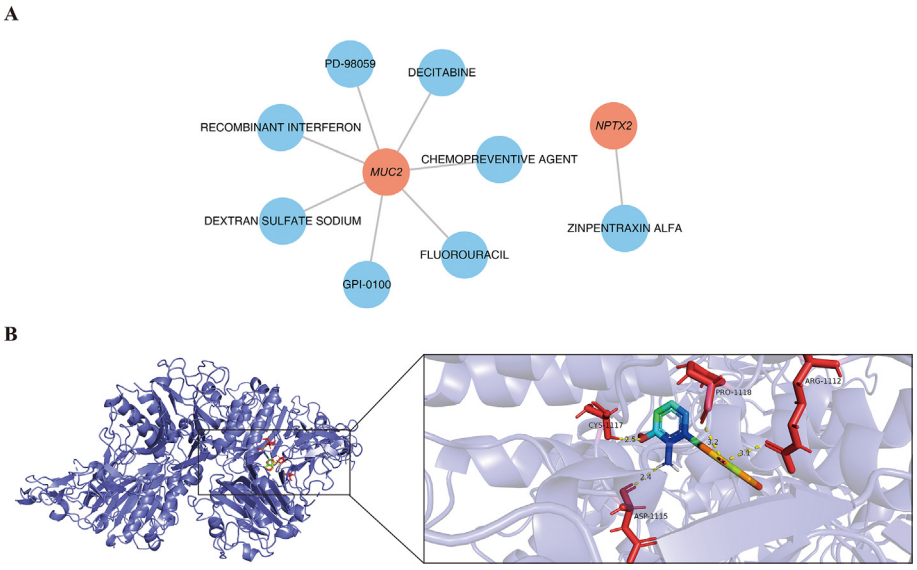


Fig. 8. Biomarker-drug interaction network and molecular docking. (A) The biomarker-drug network. (B) Molecular docking results for MUC2 and three drugs (decitabine, fluorouracil, PD-98059).

Table 1
Binding affinity between biomarker and drug.

Gene	Drug	Score (kcal/mol)
MUC2	DECITABINE	−6.6
MUC2	FLUOROURACIL	−5.0
MUC2	PD-98059	−7.8

tases from mucinous adenocarcinomas arising from other organs [41]. NPTX2 is a gene encoding the pentraxin-related protein NPTX2 and belongs to the pentraxin family. It plays a crucial role in the formation, function, and plasticity of neural synapses [42]. There is increasing evidence that NPTX2 may act as either an oncogene or a tumor suppressor in different tumor types. For example, compared with control tissues, NPTX2 is significantly upregulated in CRC and often downregulated in pancreatic cancer. Higher NPTX2 expression levels accelerate the progression of CRC and are indicative of a poor prognosis for CRC patients [43,44]. PPIP5K1 (inositol-pentakisphosphate 2-kinase 1) is a gene encoding an enzyme that is closely associated with inositol phospholipid metabolism. Many studies have demonstrated that abnormalities in inositol phosphate metabolism in tumor cells are often associated with dysregulation of signaling. Notably, the inositol pyrophosphate 2 kinase PPIP5K represents a potential target for cancer treatment [45,46].

Through the miRNA-lncRNA/circRNA regulatory network analysis, it has been demonstrated that three biomarkers (PPIP5K1, GREB1, and NPTX2) have predicted 7 miRNAs, 16 lncRNAs, and 10 circRNAs. Specifically, 5 miRNAs (mm-miR-122-5p, mmu-miR-450b-3p, mmu-miR-743a-3p, mmu-miR-743b-3p, and mmu-miR-761) target PPIP5K1, 1 miRNA (mmu-miR-345-3p) targets GREB1, and 1 miRNA (mmu-miR-96-5p) targets NPTX2. Additionally, 1 lncRNA (Malat1) simultaneously targets 4 miRNAs (mmu-miR-345-3p, mmu-miR-743a-3p, mmu-miR-743b-3p, and mmu-miR-96-5p). In lung cancer cells, the high expression of Malat1 can sequester more mmu-miR-743b-3p. As a large quantity of mmu-miR-743b-3p is sequestered, its inhibitory effect on PPIP5K1 is attenuated, and the expression of PPIP5K1 protein is upregulated. This activation of related signaling pathways promotes the proliferation of lung cancer cells, inhibits cell apoptosis, enhances the migration and invasion capabilities of lung cancer cells, and further exacerbates the occurrence, development, deterioration, and metastasis of lung cancer [47,48]. The prediction of this regulatory network suggests that Wenyang Hualiu Decoction may disrupt the ceRNA competitive mechanism by alleviating the "sponge" adsorption effect of MALAT1 on mmu-miR-743b-3p, thereby restoring the inhibitory effect of this miRNA on the proto-oncogene PPIP5K1. This ultimately suppresses the proliferation and metastasis of lung cancer cells. It provides a new molecular perspective for understanding the multi-target and network-based regulatory characteristics of traditional Chinese medicine compounds.

Subsequent studies have shown that the biomarker MUC2 has the highest binding energy among the results of molecular docking between the biomarker MUC2 and the drug PD-98059. Research has indicated that PD-98059 is a MEK1/2 inhibitor. The Ras-Raf-MEK-ERK signaling pathway is often aberrantly activated in lung cancer cells. PD-98059 binds to the active site of MEK1/2, inhibits its phosphorylation, blocks the activation of downstream ERK, inactivates this signaling pathway, thereby inhibiting the proliferation of lung cancer cells, arresting them at a specific cell-cycle stage, reducing the migration and invasion capabilities of lung cancer cells, and decreasing the risk of metastasis [49,50,51]. Our prediction links MUC2 to the MEK inhibitor PD-98059, suggesting that MUC2 may be involved in regulating this key pro-survival signaling pathway. This provides a novel intersection point for integrating traditional Chinese medicine efficacy mechanism research with modern targeted drug development.

In addition to elucidating potential mechanisms of action, our integrated analytical pipeline—spanning from biomarker identification to regulatory network construction and drug prediction—exemplifies how novel molecular insights can reshape diagnostic and therapeutic frameworks, as highlighted in systems oncology reviews. The Malat1-miR-743b-3p-PPIP5K1 regulatory axis and the high-affinity interaction between MUC2 and PD-98059 not only provide testable hypotheses for the effects of WYHLT but also offer candidate targets for the future development of RNA-targeted therapies (e.g., targeting specific miRNAs or PPIP5K1) or combinatorial strategies with existing tyrosine kinase inhibitors. This translational perspective underscores the critical value of integrating traditional Chinese medicine research with modern RNA biology to discover actionable targets within complex disease networks.

Combined with the RT-qPCR results, although the expression trends of the four key biomarkers are consistent with previous

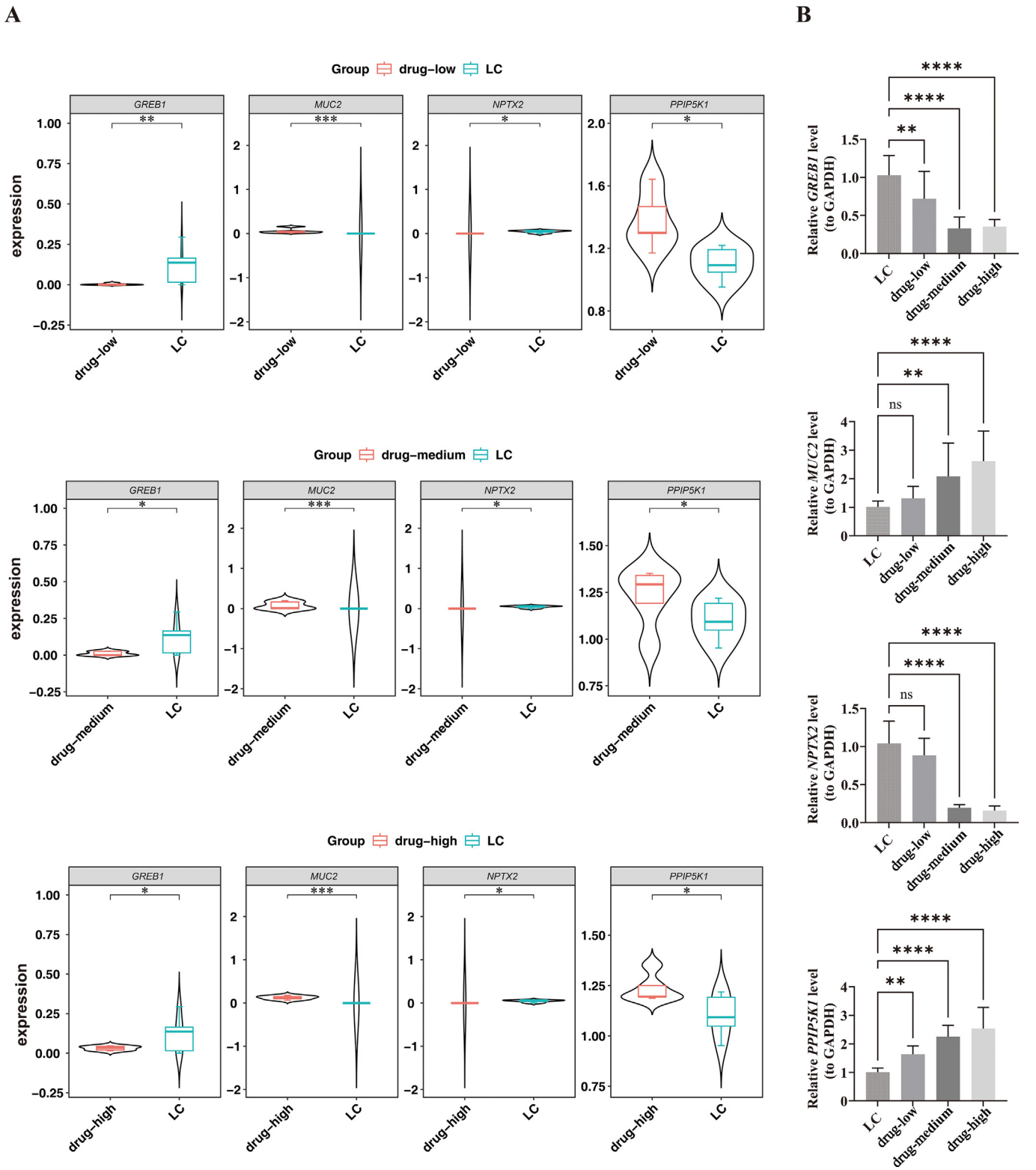


Fig. 9. Expression analysis of biomarkers in drug treatment groups. (A) Expression levels of PPIP5K1, MUC2, GREB1, and NPTX2 in the drug-low, drug-medium, and drug-high intervention groups compared to the LC group, with significant differences ($p < 0.05$). (B) RT-PCR validation of biomarker expression in LC tissue samples from the animal model. The laboratory was requested to perform statistical analysis using Prism GraphPad. All experimental data are expressed as mean ± standard deviation (Mean ± SD). Comparisons of measurement data between groups were conducted using t tests, while comparisons of measurement data among multiple groups were performed using one way analysis of variance (ANOVA). A p value < 0.05 was considered statistically significant, $p < 0.01$ as highly statistically significant, and $p < 0.001$ as extremely statistically significant.

studies, due to the limited sample size of the animals used in the experiment, the quantitative results are less than satisfactory. The main limitations cannot be overlooked, such as potential data noise, poor interpretability caused by algorithm complexity, and the occurrence of false-positive and false-negative results. Moreover, given the complexity and diversity of biological systems,

the results invariably need to be verified through clinical trials to confirm their authenticity and reliability. Overall, this study provides a solid theoretical foundation and experimental guidance for a more profound understanding of the pathogenesis and diagnostic model of LC, and offers crucial reference for future clinical practice and basic research. In the follow-up, continuous efforts will be made to focus on the mechanism of action of these biomarkers in LC, providing valuable support for the development of clinical drugs.

5. Conclusions

In conclusion, this study provides a new theoretical basis for a deeper understanding of the pathogenesis of lung cancer. The four biomarkers identified (GREB1, MUC2, NPTX2, and PPIP5K1) offer potential targets for the precise diagnosis of lung cancer. Moving forward, we plan to validate the expression and clinical significance of these markers in human lung cancer tissue samples and further elucidate the specific molecular pathways through which Wenyang Huayu Tang exerts its antitumor effects by regulating these ARE-related genes, thereby promoting the clinical translation of this research.

CRediT authorship contribution statement

Wen-xin Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Teng-yu Wang:** Supervision, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Hui-ting Wu:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Project administration, Funding acquisition. **Mingfeng Xiong:** Visualization, Validation, Supervision, Software, Resources, Project administration, Funding acquisition, Conceptualization.

Ethical approval (animals)

The study was approved by the Medical Ethics Committee of the First Affiliated Hospital of Nanchang University and was conducted in accordance with the Declaration of Helsinki Ethical approval number: NO. SYXK2021-0003.

Financial support

We would like to express our deep gratitude to the Center for the Treatment of Future Diseases of the First Affiliated Hospital of Nanchang University, as well as the support of the Research Start-up Fund for the Introduction of Full-time Talents (NO. RSC-0015).

Declaration of competing interest

The authors declare that they have no competing interests.

Supplementary material

https://doi.org/10.1016/j.ejbt.2026.100714.

Data availability

No data was used for the research described in the article.

References

[1] Mikaeili Namini A, Jahangir M, Mohseni M, et al. An in silico comparative transcriptome analysis identifying hub lncRNAs and mRNAs in brain metastatic small cell lung cancer (SCLC). Sci Rep 2022;12:18063. https://doi.org/10.1038/s41598-022-22252-7. PMid: 36302939.
[2] Thai AA, Solomon BJ, Sequist LV, et al. Lung cancer. Lancet 2021;398(10299):535–54. https://doi.org/10.1016/s0140-6736(21)00312-3. PMid: 34273294.
[3] Zhang C, Jiang M, Zhou N, et al. Use tumor suppressor genes as biomarkers for diagnosis of non-small cell lung cancer. Sci Rep 2021;11:3596. https://doi.org/10.1038/s41598-020-80735-x. PMid: 33580150.
[4] Siegel RL, Wagle NS, Cercek A, et al. Colorectal cancer statistics, 2023. CA Cancer J Clin 2023;73(3):233–54. https://doi.org/10.3322/caac.21772. PMid: 36856579.
[5] Zheng X, Song X, Zhu G, et al. Nanomedicine combats drug resistance in lung cancer. Adv Mater 2024;36(3):e2308977. https://doi.org/10.1002/adma.202308977. PMid: 37968865.
[6] Wei Z, Chen J, Zuo F, et al. Traditional Chinese Medicine has great potential as candidate drugs for lung cancer: A review. J Ethnopharmacol 2023;300:115748. https://doi.org/10.1016/j.jep.2022.115748. PMid: 36162545.
[7] Dan W, Liu J, Guo X, et al. Study on medication rules of traditional Chinese medicine against antineoplastic drug-induced cardiotoxicity based on network pharmacology and data mining. Evid Based Complement Alternat Med 2020;2020(1):7498525. https://doi.org/10.1155/2020/7498525. PMid: 33281914.
[8] Guo M, Fang W, Hu Z. Traditional Chinese medicine and its components effectively reduce resistance mediated by immune checkpoint inhibitors. Front Immunol 2024;15:1429483. https://doi.org/10.3389/fimmu.2024.1429483. PMid: 39660124.
[9] Chen X, Liu T, Luo J, et al. Data for teenagers' stressor, mental health, coping style, social support, parenting style and self-efficacy in South China. Data Brief 2020;29:105202. https://doi.org/10.1016/j.dib.2020.105202. PMid: 32071981.
[10] Shi YH, Zhang XL, Ying PJ, et al. Neuroprotective effect of Astragaloside IV on cerebral ischemia/reperfusion injury rats through Sirt1/Mapt pathway. Front Pharmacol 2021;12:639898. https://doi.org/10.3389/fphar.2021.639898. PMid: 33841157.
[11] Ye F, Xu R, Ge Y, et al. LINC00963 confers oncogenic properties in glioma by regulating the miR-506/BCAT1 axis. Cancer Manag Res 2020;12:2339–51. https://doi.org/10.2147/cmar.S246332. PMid: 32273770.
[12] Özdemir-Şimşek Ö, Erfidan G, Arslansoyu-Çamlar S, et al. A rare cause of chronic tubulointerstitial nephritis in childhood: Answers. Pediatr Nephrol 2022;37(3):571–5. https://doi.org/10.1007/s00467-021-05326-y. PMid: 34734329.
[13] Roberts TJ, Burns AT, MacIsaac RJ, et al. Sildenafil enhances central hemodynamic responses to exercise, but not $\dot{V}o_{2peak}$, in people with diabetes mellitus. J Appl Physiol 2019;127(1):1–10. https://doi.org/10.1152/japplphysiol.00947.2018. PMid: 31046521.
[14] Dai J, Xu Y, Wang T, et al. Exploring the relationship between socioeconomic deprivation index and Alzheimer's disease using summary-level data: From genetic correlation to causality. Prog Neuropsychopharmacol Biol Psychiatry 2023;123:110700. https://doi.org/10.1016/j.pnpbp.2022.110700. PMid: 36566903.
[15] Rao Z, Wang Z, Deng H, et al. Role of traditional Chinese medicine in lung cancer management: A review. Am J Chin Med 2025;53(1):97–117. https://doi.org/10.1142/s0192415x25500053. PMid: 39880665.
[16] Huiting W, Chunhong W, Liqin W, et al. Research progress of Warming Yang and dispersing nodules treatment in malignant tumors. Chin Med Modern Distance Educ China 2024;9:166–8. https://doi.org/10.3969/j.issn.1672-2779.2024.09.054.
[17] Mingfeng X, Wenxin L, Huiting W, et al. Assessment of the clinical efficacy of Wenyang Hualiu Decoction and chemotherapy in the treatment of non-small cell lung cancer. Chin Sci J Database Med Health 2024:0135–9.
[18] Tian F, Yu CT, Ye WD, et al. Cinnamaldehyde induces cell apoptosis mediated by a novel circular RNA hsa_circ_0043256 in non-small cell lung cancer. Biochem Biophys Res Commun 2017;493(3):1260–6. https://doi.org/10.1016/j.bbrc.2017.09.136. PMid: 28958934.
[19] Xiao Q, Zheng F, Tang Q, et al. Repression of PDK1- and LncRNA HOTAIR-mediated EZH2 gene expression contributes to the enhancement of Atractylenolide 1 and Erlotinib in the inhibition of human lung cancer cells. Cell Physiol Biochem 2018;49(4):1615–32. https://doi.org/10.1159/000493497. PMid: 30223276.
[20] Zhang J, Liang Y, Lin Y, et al. IRE1α-TRAF2-ASK1 pathway is involved in CSTMP-induced apoptosis and ER stress in human non-small cell lung cancer A549 cells. Biomed Pharmacother 2016;82:281–9. https://doi.org/10.1016/j.biopha.2016.04.050. PMid: 27470364.
[21] Bakheet T, Frevel M, Williams BR, et al. ARED: Human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins. Nucl Acids Res 2001;29(1):246–54. https://doi.org/10.1093/nar/29.1.246. PMid: 11125104.
[22] Zubiaga AM, Belasco JG, Greenberg ME. The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. Mol Cel Biol 1995;15(4):2219–30. https://doi.org/10.1128/mcb.15.4.2219. PMid: 7891716.

[23] DeMaria CT, Brewer G. AUF1 Binding affinity to A+U-rich elements correlates with rapid mRNA degradation. J Biol Chem 1996;271(21):12179–84. https://doi.org/10.1074/jbc.271.21.12179. PMid: 8647811.
[24] Khabar KS. Post-transcriptional control during chronic inflammation and cancer: A focus on AU-rich elements. Cell Mol Life Sci 2010;67(17):2937–55. https://doi.org/10.1007/s00018-010-0383-x. PMid: 20495997.
[25] Legrand N, Dixon DA, Sobolewski C. AU-rich element-binding proteins in colorectal cancer. World J Gastroint Oncol 2019;11(2):71–90. https://doi.org/10.4251/wjgo.v11.i2.71. PMid: 30788036.
[26] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8. PMid: 25516281.
[27] Gustavsson EK, Zhang D, Reynolds RH, et al. *ggtranscript*: An R package for the visualization and interpretation of transcript isoforms using *ggplot2*. Bioinformatics 2022;38(15):3844–6. https://doi.org/10.1093/bioinformatics/btac409. PMid: 35751589.
[28] Gu Z, Hübschmann D. Make interactive complex heatmaps in R. Bioinformatics 2022;38(5):1460–2. https://doi.org/10.1093/bioinformatics/btab806. PMid: 34864868.
[29] Chen H, Boutros PC. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinf 2011;12:35. https://doi.org/10.1186/1471-2105-12-35. PMid: 21269502.
[30] Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation 2021;2(3):100141. https://doi.org/10.1016/j.xinn.2021.100141. PMid: 34557778.
[31] Shannon P, Markiel A, Ozier O, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303. PMid: 14597658.
[32] Lateef RA. Machine and deep learning techniques in cancer prediction and risk stratification using bioinformatics in Big Data Era: A review. EDRAAK 2024;2024:118–27. https://doi.org/10.70470/EDRAAK/2024/015.
[33] Orifjon S, Jammatov J, Sousa C, et al. Translation and adaptation of the adult developmental coordination disorder/dyspraxia checklist (ADC) into Asian Uzbekistan. Sports 2023;11(7):135. https://doi.org/10.3390/sports11070135. PMid: 37505622.
[34] Hu Y, Yan C, Hsu CH, et al. OmicCircos: A simple-to-use R package for the circular visualization of multidimensional omics data. Cancer Inf 2014;13:13–20. https://doi.org/10.4137/cin.S13495. PMid: 24526832.
[35] Zhang H, Meltzer P, Davis S. RCircos: An R package for Circos 2D track plots. BMC Bioinf 2013;14:244. https://doi.org/10.1186/1471-2105-14-244. PMid: 23937229.
[36] Ru Y, Kechris KJ, Tabakoff B, et al. The *multiMiR* R package and database: Integration of microRNA-target interactions along with their disease and drug associations. Nucleic Acids Res 2014;42(17):e133. https://doi.org/10.1093/nar/gku631. PMid: 25063298.
[37] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods 2001;25(4):402–8. https://doi.org/10.1006/meth.2001.1262. PMid: 11846609.
[38] Al-Rawi NH, Rizvi Z, Mkadmi S, et al. Differential expression profile of salivary oncomiRNAs among smokeless tobacco users. Eur J Dent 2023;17(4):1215–20. https://doi.org/10.1055/s-0043-1761191. PMid: 36812928.
[39] Li R, Liu J, Fang Z, et al. Identification of mutations related to cisplatin-resistance and prognosis of patients with lung adenocarcinoma. Front Pharmacol 2020;11:572627. https://doi.org/10.3389/fphar.2020.572627. PMid: 33192515.
[40] Cox KE, Liu S, Lwin TM, et al. The mucin family of proteins: Candidates as potential biomarkers for colon cancer. Cancers 2023;15(5):1491. https://doi.org/10.3390/cancers15051491. PMid: 36900282.
[41] Malmros K, Lindholm A, Vidarsdottir H, et al. Diagnostic gastrointestinal markers in primary lung cancer and pulmonary metastases. Virchows Arch 2024;485(2):347–57. https://doi.org/10.1007/s00428-023-03583-w. PMid: 37349623.
[42] Hruska-Plochan M, Wiersma VI, Betz KM, et al. A model of human neural networks reveals NPTX2 pathology in ALS and FTLD. Nature 2024;626:1073–83. https://doi.org/10.1038/s41586-024-07042-7. PMid: 38355792.
[43] García-Ortiz MV, Cano-Ramírez P, Toledano-Fonseca M, et al. Circulating *NPTX2* methylation as a non-invasive biomarker for prognosis and monitoring of metastatic pancreatic cancer. Clin Epigenetics 2023;15(1):118. https://doi.org/10.1186/s13148-023-01535-4. PMid: 37481552.
[44] Xu G, Fan L, Zhao S, et al. Neuronal pentraxin II (NPTX2) hypermethylation promotes cell proliferation but inhibits cell cycle arrest and apoptosis in gastric cancer cells by suppressing the p53 signaling pathway. Bioengineered 2021;12(1):1311–23. https://doi.org/10.1080/21655979.2021.1915658. PMid: 33896384.
[45] Machkalyan G, Trieu P, Pétrin D, et al. PPIP5K1 interacts with the exocyst complex through a C-terminal intrinsically disordered domain and regulates cell motility. Cell Signal 2016;28(5):401–11. https://doi.org/10.1016/j.cellsig.2016.02.002. PMid: 26854614.
[46] Gu C, Nguyen HN, Ganini D, et al. KO of 5-InsP(7) kinase activity transforms the HCT116 colon cancer cell line into a hypermetabolic, growth-inhibited phenotype. Proc Natl Acad Sci USA. 2017;114(45):11968–73. https://doi.org/10.1073/pnas.1702370114. PMid: 29078269.
[47] Guo F, Yu F, Wang J, et al. Expression of *MALAT1* in the peripheral whole blood of patients with lung cancer. Biomed Rep 2015;3(3):309–12. https://doi.org/10.3892/br.2015.422. PMid: 26137228.
[48] Bartel DP. MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 2004;116(2):281–97. https://doi.org/10.1016/s0092-8674(04)00045-5. PMid: 14744438.
[49] Li Z, Liu H, Zhang M, et al. An apoptosis-related specific risk model for breast cancer: From genomic analysis to precision medicine. Front Biosci 2024;29(7):239. https://doi.org/10.31083/j.fbl2907239. PMid: 39082332.
[50] Zhang X, Zeng Y, Qu Q, et al. PD-L1 induced by IFN-γ from tumor-associated macrophages via the JAK/STAT3 and PI3K/AKT signaling pathways promoted progression of lung cancer. Int J Clin Oncol 2017;22(6):1026–33. https://doi.org/10.1007/s10147-017-1161-7. PMid: 28748356.
[51] Ding C, Tang W, Fan X, et al. Overexpression of PEAK1 contributes to epithelial-mesenchymal transition and tumor metastasis in lung cancer through modulating ERK1/2 and JAK2 signaling. Cell Death Dis 2018;9:802. https://doi.org/10.1038/s41419-018-0817-1. PMid: 30038287.