


Research Article

Network Pharmacology Integrated Molecular Docking to Explore the Mechanism of Blister Beetle Therapy for Lung Adenocarcinoma

Shoujun Deng,¹ Ying Mao,² Heng Li,¹ and Gaofeng Li ¹

¹Department of Thoracic Surgery II, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Kunming 650118, China

²Department of Thyroid Breast Surgery, The Second Affiliated Hospital of Kunming Medical University, Kunming 650101, China

Correspondence should be addressed to Gaofeng Li; ligaofeng698@163.com

Received 21 April 2022; Revised 20 May 2022; Accepted 24 May 2022; Published 14 July 2022

Academic Editor: Mohammad Farukh Hashmi

Copyright © 2022 Shoujun Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lung adenocarcinoma (LUAD) is one of the major causes of cancer death in the world. Studies show that the effective anticancer component in blister beetles is cantharidin, which can improve chemotherapy efficacy, median survival, and prognosis of LUAD. However, the antitumor mechanism of blister beetles has not been fully clarified. This study aimed to identify the key targets of the treatment of LUAD by blister beetles based on the principle of network pharmacology. An integrated approach including network pharmacology and a molecular docking technique was conducted, which mainly comprises target prediction, weighted gene correlation network analysis (WGCNA) analysis, network construction, gene ontology, and pathway enrichment analysis. 35 key targets were obtained and significantly associated with response to external stimuli, collagen binding, cyclin binding, organic acid binding, pyruvate metabolism, glycolysis, and amino acid biosynthesis pathways. Both LASSO regression and the RF model had a high predictive ability, and 9 candidate genes were screened, among which BIRC5 and PLK1 were the key targets for the treatment of LUAD by using blister beetles and showed significant survival significance. Cantharidin exerts its antitumor effects through 8 targets in 32 pathways, while BIRC5 and PLK1 have obvious survival significance.

1. Introduction

Lung cancer is the most common malignant tumor in clinical practice, and its mortality and morbidity rank first in malignant tumors [1]. Moreover, the incidence has been increasing in recent years. According to data released by China's National Cancer Center in 2019, there were about 787,000 new cases of lung cancer and 631,000 deaths in China in 2015. Lung cancer is roughly divided into two categories: non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC). Non-small-cell lung cancer (NSCLC) accounts for about 80% of all lung cancers, mainly including large cell carcinoma, squamous cell carcinoma, and adenocarcinoma [2–4]. Chemotherapy is currently one of the main methods for the treatment of advanced lung adenocarcinoma, but the quality of life of patients has not been significantly improved due to the available drug

resistance of tumor cells and severe adverse reactions of chemotherapy drugs [5, 6]. Therefore, it is of great significance to study the molecular mechanism of the occurrence and progression of lung cancer and to find new therapeutic methods and strategies for the clinical diagnosis and treatment of lung cancer.

The blister beetle is a famous Chinese medicine and has been demonstrated in vitro to have significant antitumor effects in cancer cells such as lung, breast, gastric, hepatocellular, and ovarian cancers [7–12]. In China, drugs from blister beetles include Aidi injection (an injection mainly composed of blister beetles), compound cantharidin capsule, sodium cantharidate vitamin B6 injection, and methylcantharidine tablets, which are used in clinical antitumor therapy [13]. Studies show that the effective anticancer component in blister beetles is cantharidin, which has no obvious immunosuppressant effect on the body while

inhibiting tumor cells, including cantharidin and its dimethylcantharidin, sodium cantharidate, and sodium methyl cantharidate [14, 15]. Clinical practice has proved that the chemotherapy efficacy, median survival, and prognosis of lung cancer patients are improved by drugs derived from blister beetles. For example, compared with platinum-based chemotherapy alone, the clinical efficacy of stage IIIB/IV NSCLC patients was improved by Aidi injection combined with platinum-based chemotherapy, including reduced chemotherapeutic toxicity [1]. In vitro studies have shown that cantharidin (CTD) inhibits cell growth and migration and promotes autophagy and apoptosis by inhibiting the PI3K/Akt/mTOR signaling pathway in non-small-cell lung cancer [16]. Although these studies reported part of the molecular mechanism of cantharidin against tumors, the antitumor mechanism of blister beetles has not been fully elucidated due to the multiactive component, multi-target, and multipathway nature of herbal medicine.

Nowadays, network pharmacology integrates network biology and polypharmacology on the basis of existing databases to provide a new approach to explore the mechanisms and synergistic effect of TCM formulas for the treatments of diseases [17–19]. For example, Li et al. [20] combined network pharmacology, machine learning, and molecular dynamics simulation methods to identify two key genes, insulin-like growth factor 1 receptor (IGF1R) and insulin receptor (IR), and three potential longevity-related herbs from established TCM databases [20]. Tang et al. [21] integrated network pharmacology with molecular docking to unravel the active compounds and potential mechanism of the simiao pill in treating rheumatoid arthritis [21]. Many studies attempting to combine network science with ancient Chinese medicine to study multiple molecular mechanisms have been successful [22–24].

In order to systematically explore the targets of blister beetles in inhibiting LUAD, we integrated complex pharmacology and molecular docking experiments based on four databases to obtain the key chemical components and their targets, which not only provided a new direction for the mechanism of action of LUAD inhibition by blister beetles, but also identified clinically relevant biomarkers that could be used to monitor the treatment of blister beetles.

2. Materials and Methods

2.1. Weighted Gene Correlation Network Analysis. TCGA database (<https://portal.gdc.cancer.gov/>) was used to download transcriptomic data of TCGA-LUAD. In this study, WGCNA was used to screen for key modules associated with LUAD [25]. TCGA mRNA expression data profiles were normalized using the “normalizeBetweenArrays” function of the R package. The coexpression network of genes with mean expression values greater than 1 in LUAD samples ($n = 535$) and normal samples ($n = 59$) was used to filter out key modules and genes associated with LUAD. First, the samples were clustered and the outlier samples were deleted to ensure the accuracy of the subsequent analysis. Then, the soft threshold is determined and the coexpression network is constructed by

using the dynamic tree cutting algorithm. The minimum number of genes per gene module was set to 50 and MEDissThres was set to 0.2. The Limma package (version 3.46.0) [26] was used to compare the difference in gene expression levels between the LUAD/normal sample groups ($\log_2FC \geq 1$, $p < 0.05$). Ggplot2 (version 3.3.3) was used to map volcano plots to show differentially expressed genes, The VennDiagram package (Version 1.6.20) was used for differentially expressed genes between LUAD and the control with intersection of key module genes associated with LUAD obtained by WGCNA. The TCMSP (<https://tcmssp.com/tcmssp.php>) database and the DrugBank (<https://go.drugbank.com/>) database were used to obtain the active ingredients of blister beetles and their drug targets; the Genecards (<https://www.genecards.org/>) database was used to predict disease targets in LUAD, retaining only genes with a category of protein coding (category = protein coding and relevance score ≥ 1). The VennDiagram package was used to intersect key module genes obtained from WGCNA with drug targets and LUAD disease targets from the database to screen for key targets.

2.2. Gene Ontology Functional Enrichment Analysis and Kyoto Encyclopedia of Gene and Genomes Analysis. GO stands for Gene Ontology, which is an international standard classification system for gene function. It aims to establish a linguistic vocabulary standard for qualifying and describing gene and protein functions that are applicable to various species. The GO system consists of three components: biological process (BP), molecular functions (MF), and cellular components (CC). Using the GO database, we can obtain what our target genes are mainly related to at the CC, MF, and BP levels. Enrichment analysis of KEGG and GO databases was used to find common functions and related pathways of a large number of genes within the module key genes. In this study, the VennDiagram package was used to obtain key targets by taking intersections of differentially expressed key module genes, drug targets, and LUAD disease targets. The clusterProfiler package [27] was used for GO and KEGG enrichment analysis ($P < 0.05$ and COUNTS ≥ 2). The enrichplot package (Version 1.10.2) was used to draw bar and bubble graphs to show GO enrichment and KEGG pathway enrichment results.

2.3. Network Construction. The key targets of GO_BP, GO_CC, and GO_MF were presented to construct a key target-function regulatory network. The top 50 pathways of KEGG and their corresponding key targets were presented to construct a key target-pathway regulatory network map and were visualized using Cytoscape (version 3.8.2) [28]. To investigate whether there are interactions between key targets, the STRING (<https://string-db.org>) website was used to construct a PPI network map of 35 key targets with confidence = 0.4 and remove discrete proteins.

2.4. LUAD Diagnostic Model Construction. The LASSO regression model and RF algorithm were used to construct a diagnostic model to determine the predictive power of the

model constructed based on candidate genes for LUAD. The LASSO analysis tool is a glmnet package (Version 4.1-1) [29]. All TCGA samples ($n = 594$) were used as the training set, and GSE10072 was used as the verification set ($n = 107$) to construct and verify the diagnostic model. The pROC (Version 1.17.0.1) package is used to draw the receiver operating characteristic curve (ROC) of the training set and verification set, and to calculate the area under curve (AUC) value. The higher the AUC value is, the more accurate the prediction is. When the RF algorithm was used to construct the optimal diagnostic model, the expression values of each sample of 35 key targets in TCGA database were combined with the grouping information of the samples, and the RF model was constructed by using the Caret package (Version 6.0-86). Then, the EXPLAIN function of the DALEX package (Version 2.3.0) was used for explanatory analysis of the RF model, and the plot function was used to visualize the distribution of model representation; then, the cumulative residual distribution and boxplot distribution were drawn.

2.5. Gene Expression Analysis. The VennDiagram package (Version 1.6.20) was used to take the intersection of the genes obtained by LASSO and the RF algorithm. The ggpubr R package (Version 0.4.0) and the ggplot2 package were used for scatter plotting and visualization to study the expression of nine biomarkers in LUAD samples and normal samples. The test was a rank sum test.

2.6. Construction of a Pharmacological Regulatory Network Map and Molecular Docking. The active components corresponding to 9 biomarkers were extracted to construct the drug-active component-key target gene network map, and Cytoscape was used for visual analysis. The protein structure of the key target was obtained from the PDB (<https://www1.rcsb.org/>) database, and the small molecules and water molecules were removed. Molecular docking is used to determine the presence of interactions between biomarkers and molecules, and to predict their binding patterns and affinities. The AutoDock tools (version 1.5.6) were used to complete protein hydrogenation and charge calculation. The active ingredient structures were downloaded from the PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) database. Charge balanced and rotatable bond checking for small molecules was performed using AutoDock tools. Then, AutoDock Vina was used to calculate the receptor-ligand docking according to the range of docking boxes selected by the receptor-active center. The structure with the lowest binding free energy in the output results was selected. Finally, PyMol (Version 2.5) software is used for visualization and beautification.

2.7. Function Prediction, Prognostic Analysis, and Clinical Correlation Analysis of Biomarkers. GO and KEGG were used for functional and pathway analysis of biomarkers. Kaplan–Meier survival analysis of the biomarkers was performed using R language's survival (Version 3.2-3) and survminer (Version 0.4.8) packages, and $P < 0.05$ indicated

that the genes had a significant survival significance. Subsequently, the correlation between biomarkers and clinical factors (age, sex, TNM stage, and stage stage) was studied.

3. Results

3.1. Active Drug Ingredients of Blister Beetles and Corresponding Targets. To obtain key targets for the interaction between blister beetles and LUAD disease, a total of 27 active drug ingredients and 503 drug targets of the blister beetle were obtained from the TCMSP database. Furthermore, 1236 drug targets were obtained from the DrugBank database and 7031 LUAD disease targets were obtained from the GeneCards database. Subsequently, these targets and key module genes were intersected to obtain key targets.

3.2. WGCNA Analysis. We clustered all LUAD samples ($n = 594$) and found that there were no outliers, as shown in Figure 1(a). When the ordinate scale-free R^2 approaches 0.85 (red line), the soft threshold (β) is equal to 5. At this time, the network approaches the scale-free distribution, and mean connectivity also approaches 0. Therefore, we choose the optimal soft threshold (β) as 5, as shown in Figure 1(b). A total of 13 modules were obtained, and 11 modules were obtained after module combination, as shown in Figure 1(c). In the module-trait correlation heat map, the MEbrown module had a significant positive correlation with LUAD ($\text{Cor} = 0.53$, $P = 0.003$), and the absolute value of the correlation coefficient is the largest, which is the key module of lung adenocarcinoma, as shown in Figure 1(d). The correlation coefficient between the MEbrown gene and module traits was 0.67 ($P < 0.05$), as shown in Figure 1(e), and there were 2894 genes in the MEbrown module. A total of 1239 differentially expressed genes were found between LUAD and normal samples, of which 571 genes were upregulated and 668 genes were downregulated in LUAD samples, as shown in Figure 1(f). The intersection of the differentially expressed genes between LUAD, control, and the key module genes related to LUAD obtained by WGCNA was conducted to obtain 396 differentially expressed key module genes, which will be studied in the future, as shown in Figure 1(g).

3.3. Analysis Results of GO and KEGG. A total of 35 key targets were obtained by taking the intersection of differentially expressed key module genes ($n = 396$), drug targets ($n = 1653$), and LUAD disease targets ($n = 7031$), as shown in Figure 2(a). In terms of biological processes, a total of 50 terms were obtained, and key target genes were significantly associated with responses to toxic substances, oxidative stress, and hydrogen peroxide, including response to nutrient levels, L-alpha-amino acid transmembrane transport, oxidative stress, vitamin metabolic process, toxic substance, reactive oxygen species, cellular response to hydrogen peroxide, and hydrogen peroxide. In terms of molecular functions, 12 terms were obtained, and the key target genes were significantly related to collagen binding, cell cycle protein binding, and organic acid binding functions,

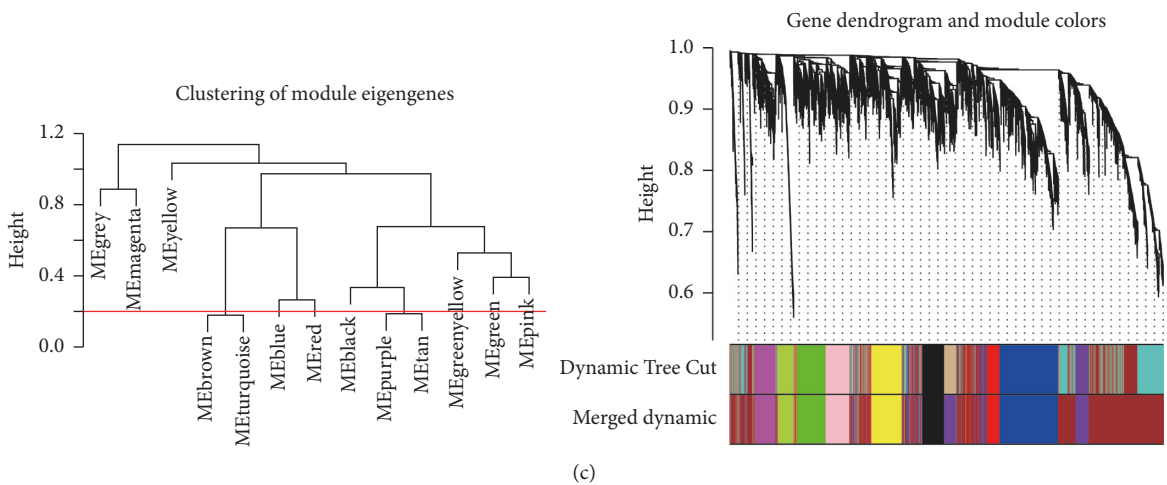
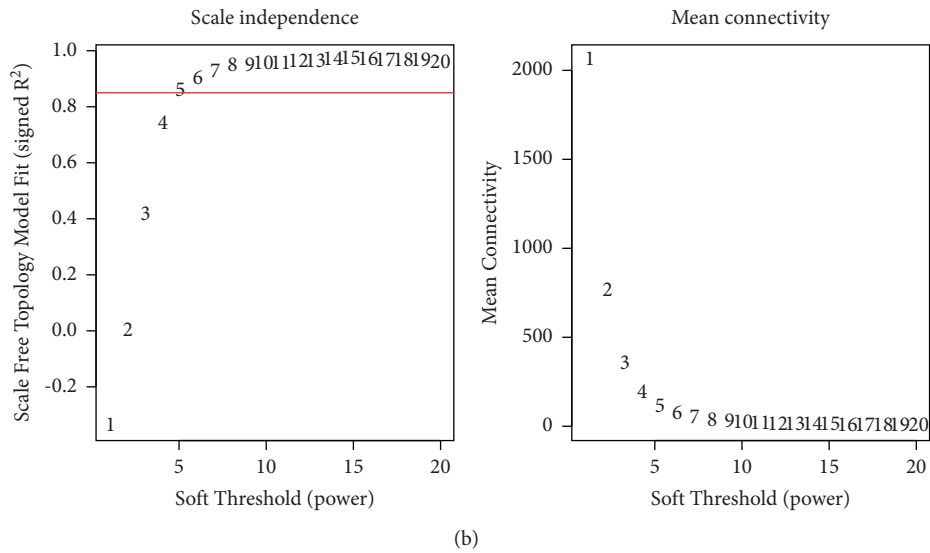
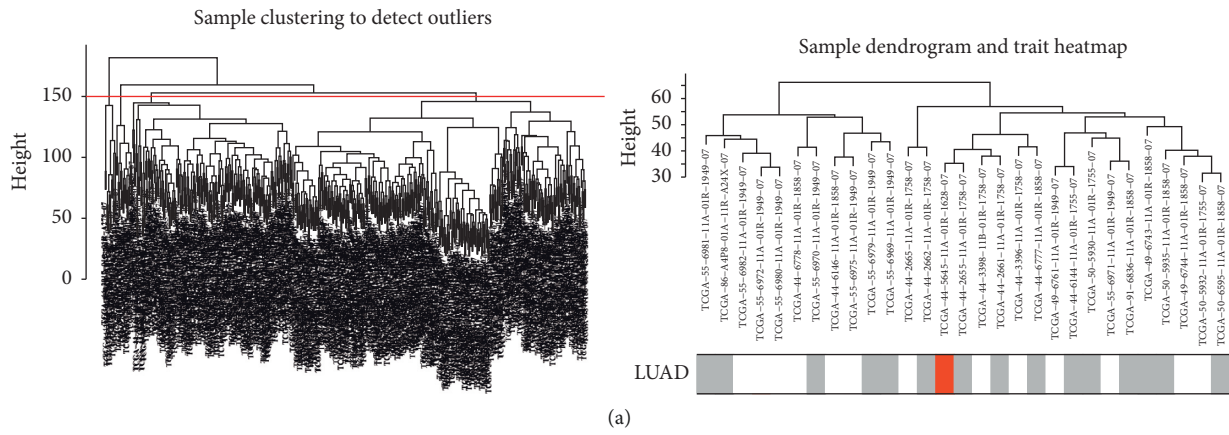


FIGURE 1: Continued.

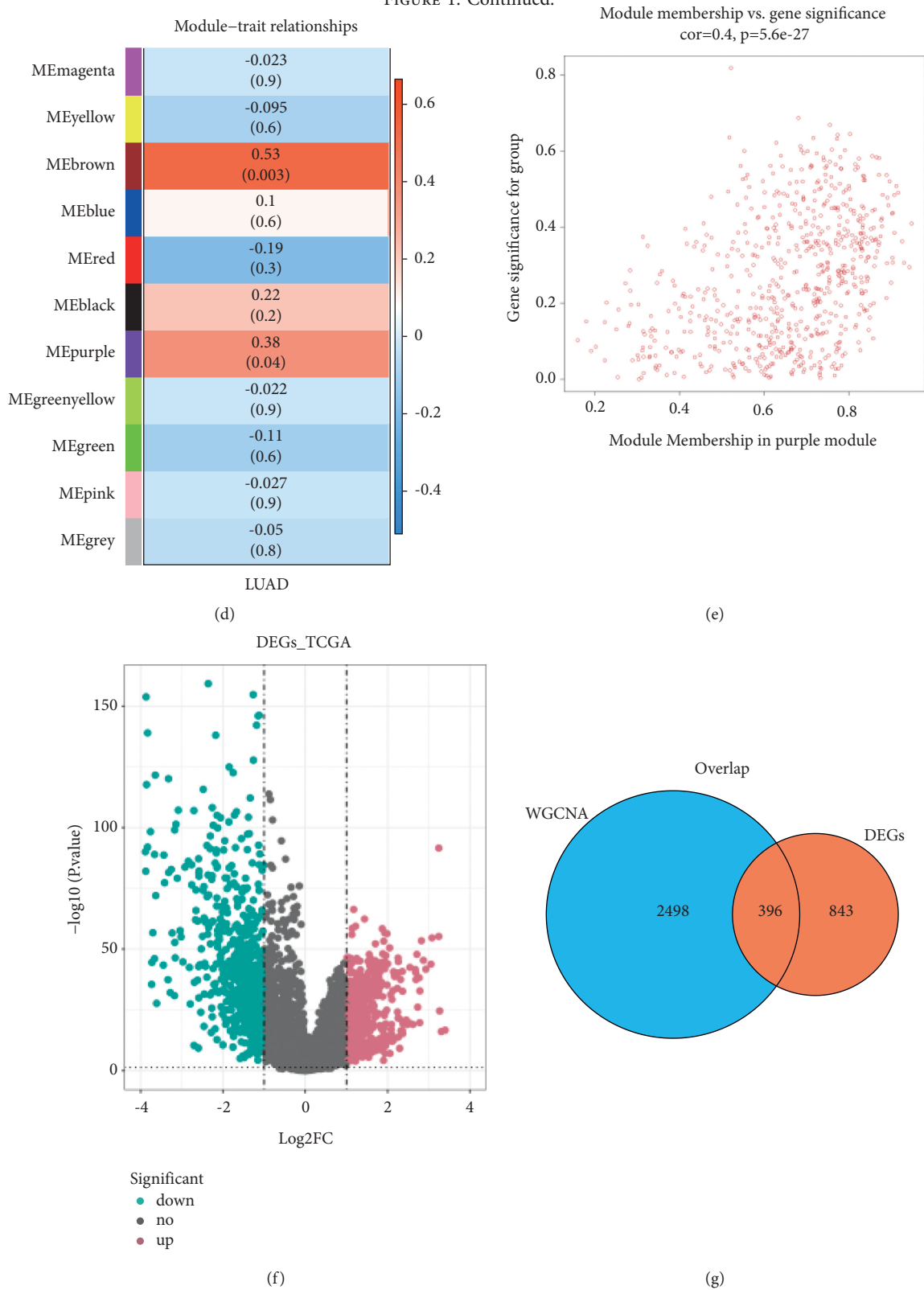


FIGURE 1: WGCNA analysis. (a) Sample clustering and phenotypic heat map. (b) Filtering of soft thresholds. (c) The merge module. (d) Heat map of the correlation between modules and LUAD. (e) Correlation between modular genes and LUAD. (f) A volcano map of differentially expressed genes between LUAD/normal samples. (g) Screening of key genes in modules. Blue represents the differentially expressed genes between LUAD/control and pink represents the key module genes associated with lung adenocarcinoma acquired by WGCNA.

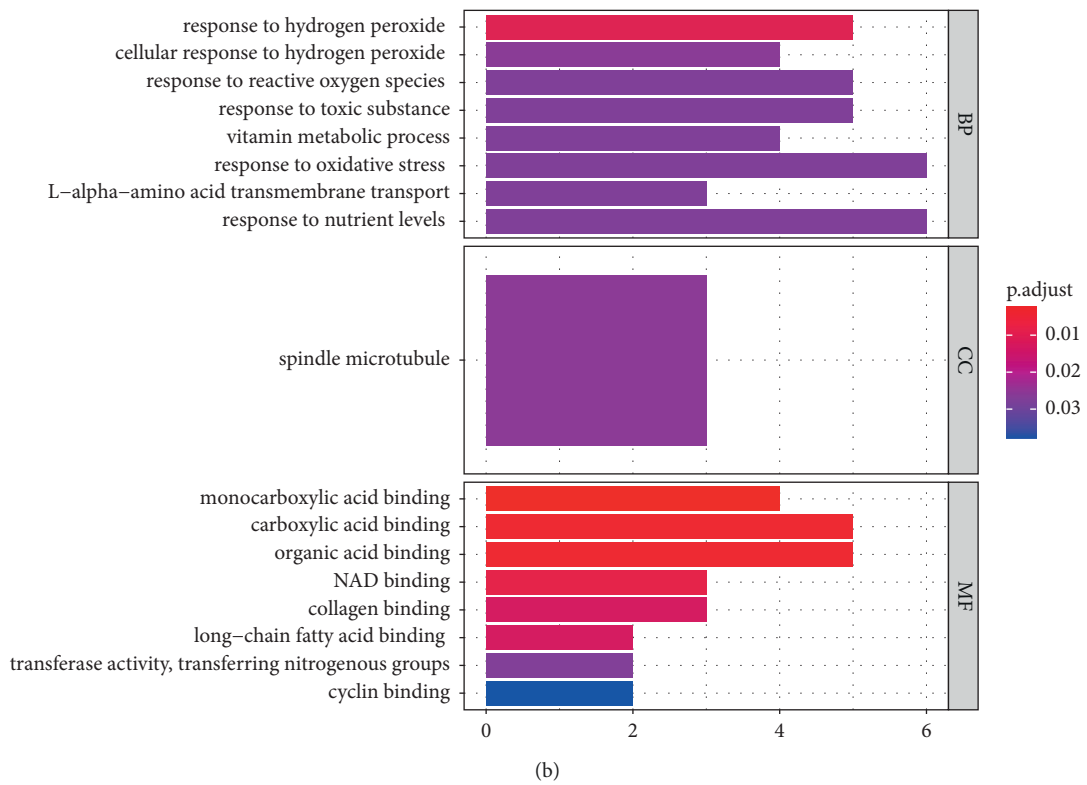
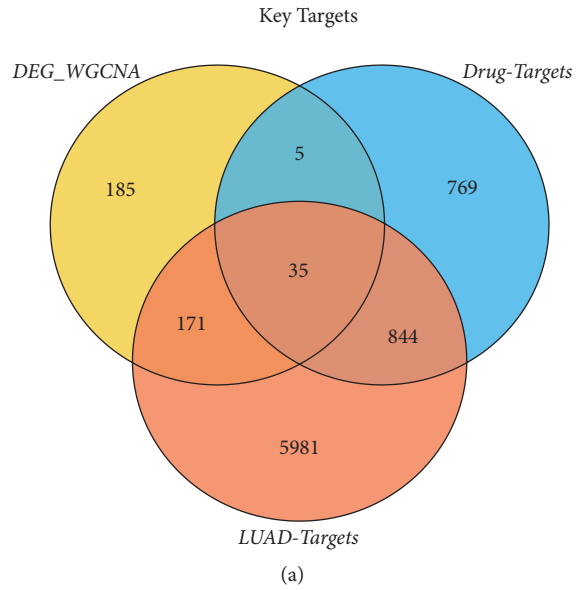


FIGURE 2: Continued.

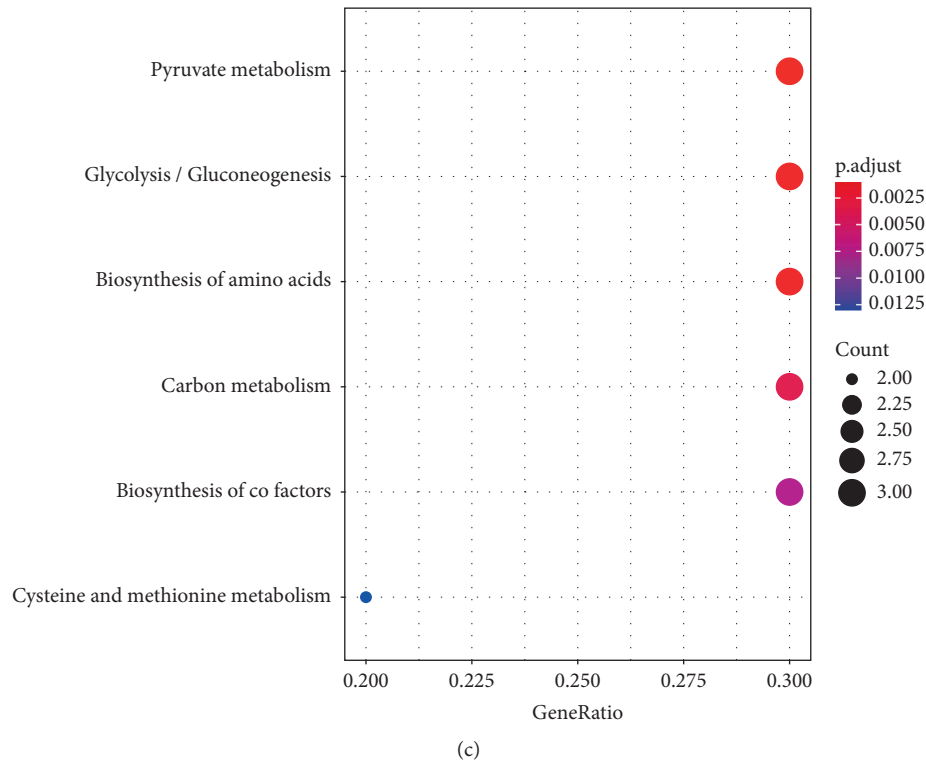


FIGURE 2: Enrichment analysis results of GO and KEGG of target genes. (a) Venn diagram of the intersection of genes, drug targets, and disease targets. (b) GO enrichment bars of key target genes. (c) KEGG pathway enrichment bubble map of key target genes.

including cyclin binding, transferase activity, transferring nitrogenous groups, long-chain fatty acid binding, collagen binding, NAD binding, organic acid binding, carboxylic acid binding, and monocarboxylic acid binding. The results of the study were obtained. In terms of cell composition, a total of 1 term was obtained, and key target genes were significantly associated with spindle microtubules, including spindle microtubules, as shown in Figure 2(b). KEGG functional enrichment analysis of 35 key target genes revealed 6 related pathways, including cysteine and methionine metabolism. Biosynthesis of cofactors, carbon metabolism, biosynthesis of amino acids, glycolysis/gluconeogenesis, and pyruvate metabolism, suggest that the key target genes are related to pyruvate metabolism, glycolysis, and amino acid biosynthesis. Glycolysis and amino acid biosynthesis are shown in Figure 2(c).

3.4. Construction of Regulatory Network Diagram. In the key target-function network diagram, 21 terms, 26 key targets, and 78 relationship pairs were included, as shown in Figure 3(a). In the key target-pathway regulatory network map, 50 KEGG pathways, 100 key targets, and 891 relationship pairs were included, as shown in Figure 3(b). The PPI regulatory network diagram of the key targets showed an interaction network of 35 proteins, including 35 nodes and 90 edges, as shown in Figure 3(c).

3.5. Construction of the LUAD Diagnostic Model. In this study, all TCGA samples ($n = 594$) were used as the training set, and GSE10072 was used as the verification set ($n = 107$)

to construct and verify the diagnostic model. The graph of the gene coefficient and the error graph of cross validation were obtained by LASSO regression analysis, and a total of 11 characteristic genes were screened, as shown in Figure 4(a). The diagnostic model was constructed from these 11 genes, and the AUCs of the training set and verification set were 0.967 and 0.718, respectively, indicating that the diagnostic model had a high predictive ability for LUAD, as shown in Figure 4(b). When the RF algorithm is used to build the optimal diagnostic model, the sample cumulative residual distribution diagram and boxplot distribution diagram are drawn, as shown in Figures 4(c) and 4(d). In the sample cumulative residual distribution diagram, the smaller the curve area is, the smaller the sample cumulative residual value is. In the boxplot distribution, the smaller the sum of squares of residuals, the better the fitting effect. Different variables have different degrees of relative importance to model prediction, as shown in Figure 4(e). In the RF model, IGHG1, LCN2, MMP7, SLC7A11, CYP27B1, PAEP, SLC7A7, MMP13, BIRC5, FEN1, ALOX5AP, AXL, ABCC3, CDK1, EPCAM, GAPDH, CRABP2, PSAT1, DPYSL2, KAT2A, PLK1, and FABP4 were the predicted values of the response variable (score = 0.07946986) (genes above the response variable were selected as characteristic genes), therefore, these 22 genes were used as diagnostic markers for the next step of analysis in this study. In the RF model, the AUCs of the training and validation sets were 0.985 and 0.978, respectively, indicating that the RF model has high a predictive ability for LUAD, as shown in Figure 4(f).

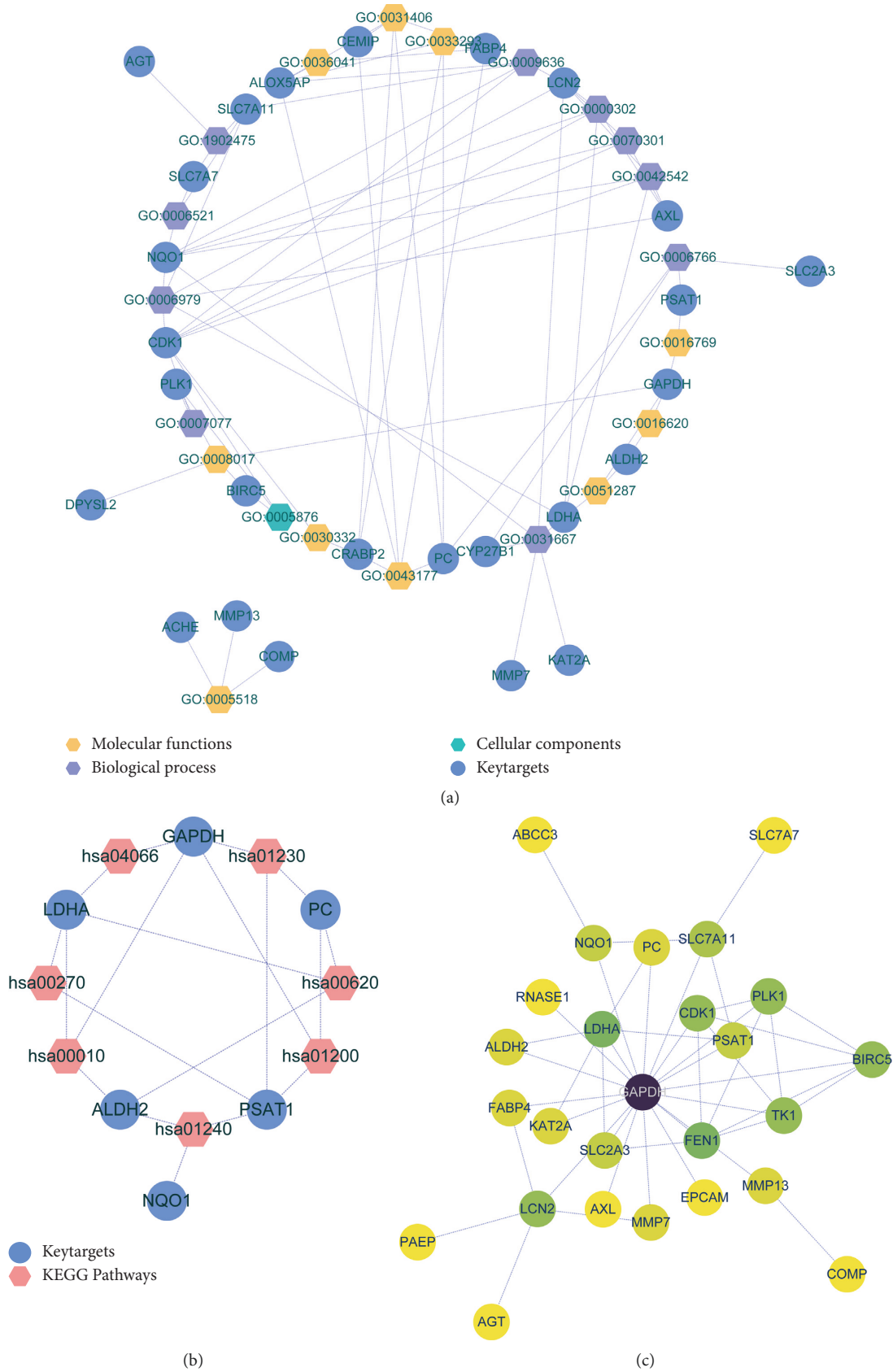


FIGURE 3: Analysis results of the key target-function/-pathway regulatory network diagram and PPI regulatory network diagram. (a) Regulatory network diagram of the key targets and GO function. The yellow hexagon represents GO_MF terms, the purple hexagon represents GO_BP terms, the green hexagon represents GO_CC Terms, and the blue dot represents the key target. (b) Regulatory network diagram of key targets and pathways. The pink hexagons represent KEGG pathways and the blue dots represent key targets. (c) Protein interaction networks of key targets. The lines represent the interaction between them, the color indicates their degree value; the darker the color, the higher the degree value and the higher the core position.

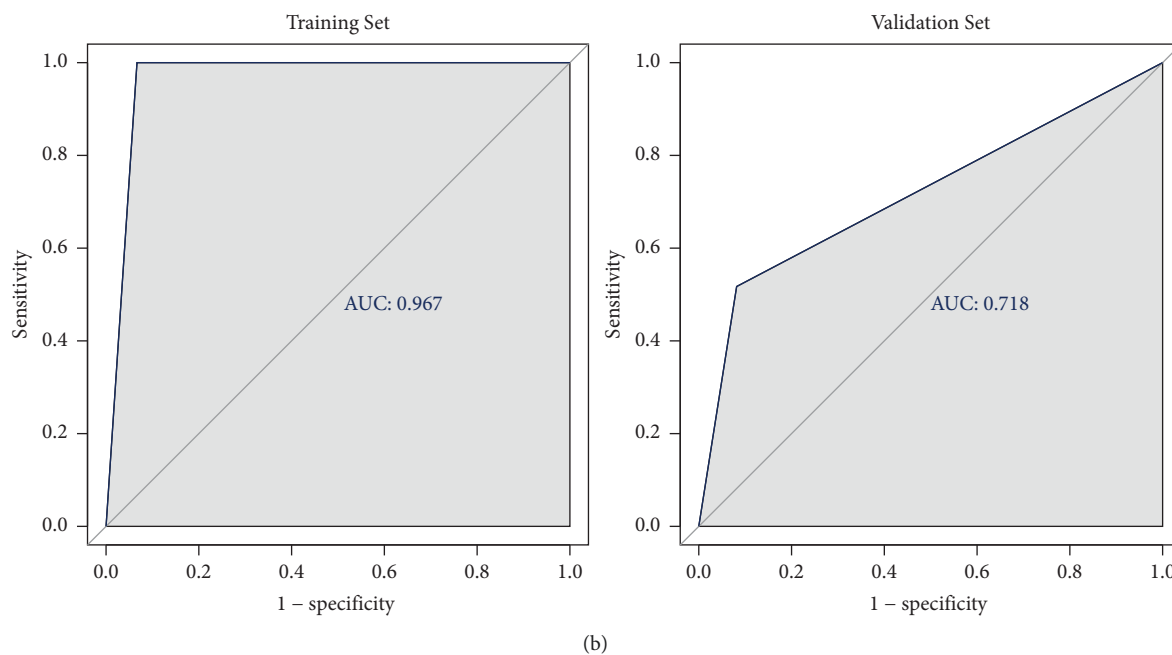
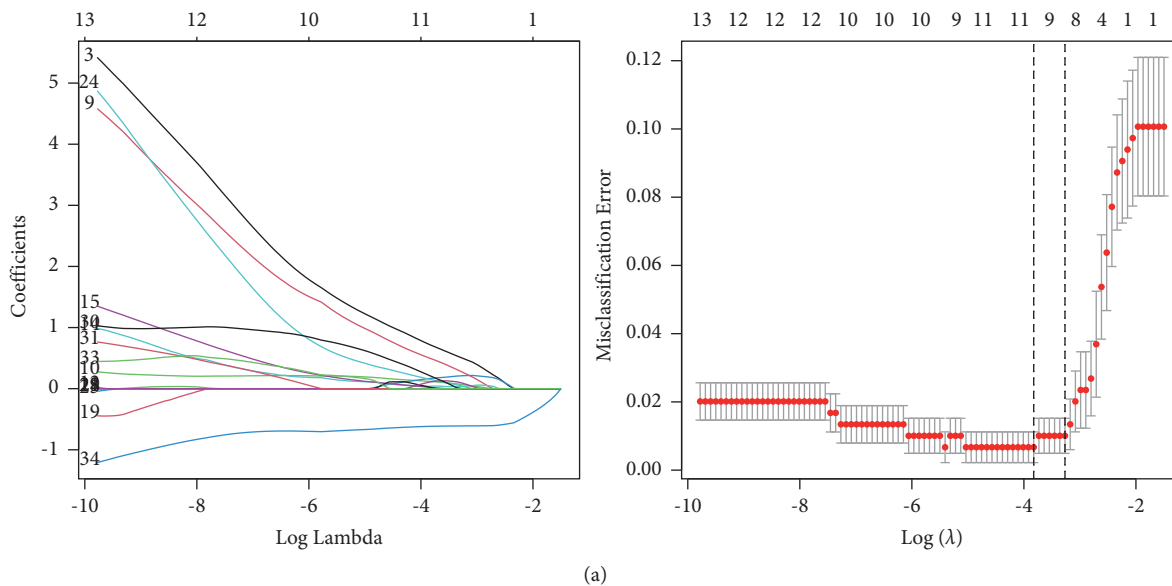
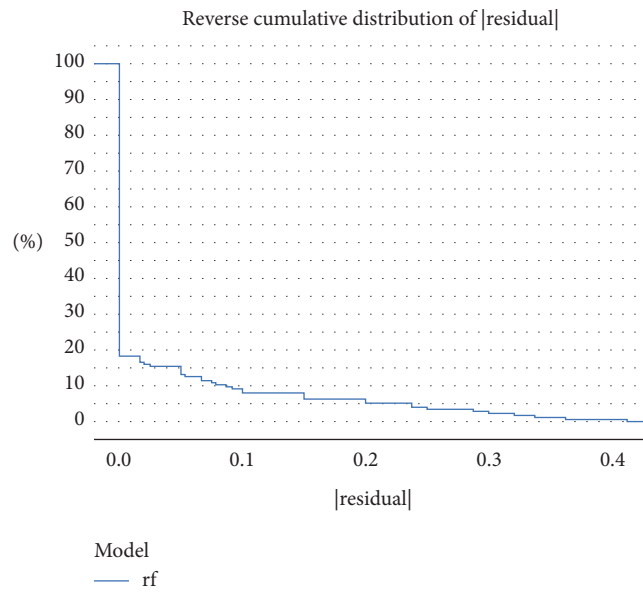
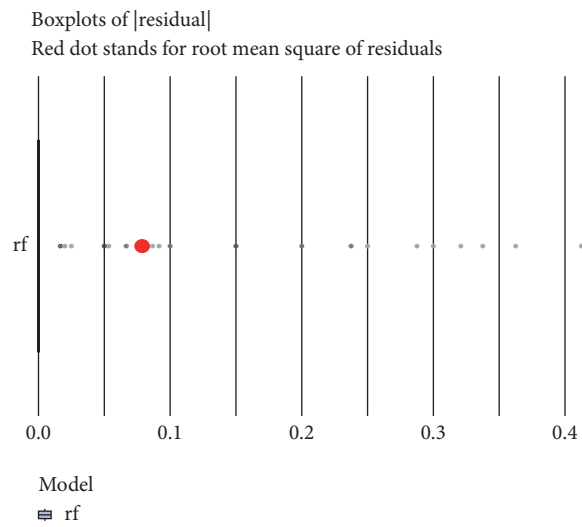


FIGURE 4: Continued.



(c)



(d)

FIGURE 4: Continued.

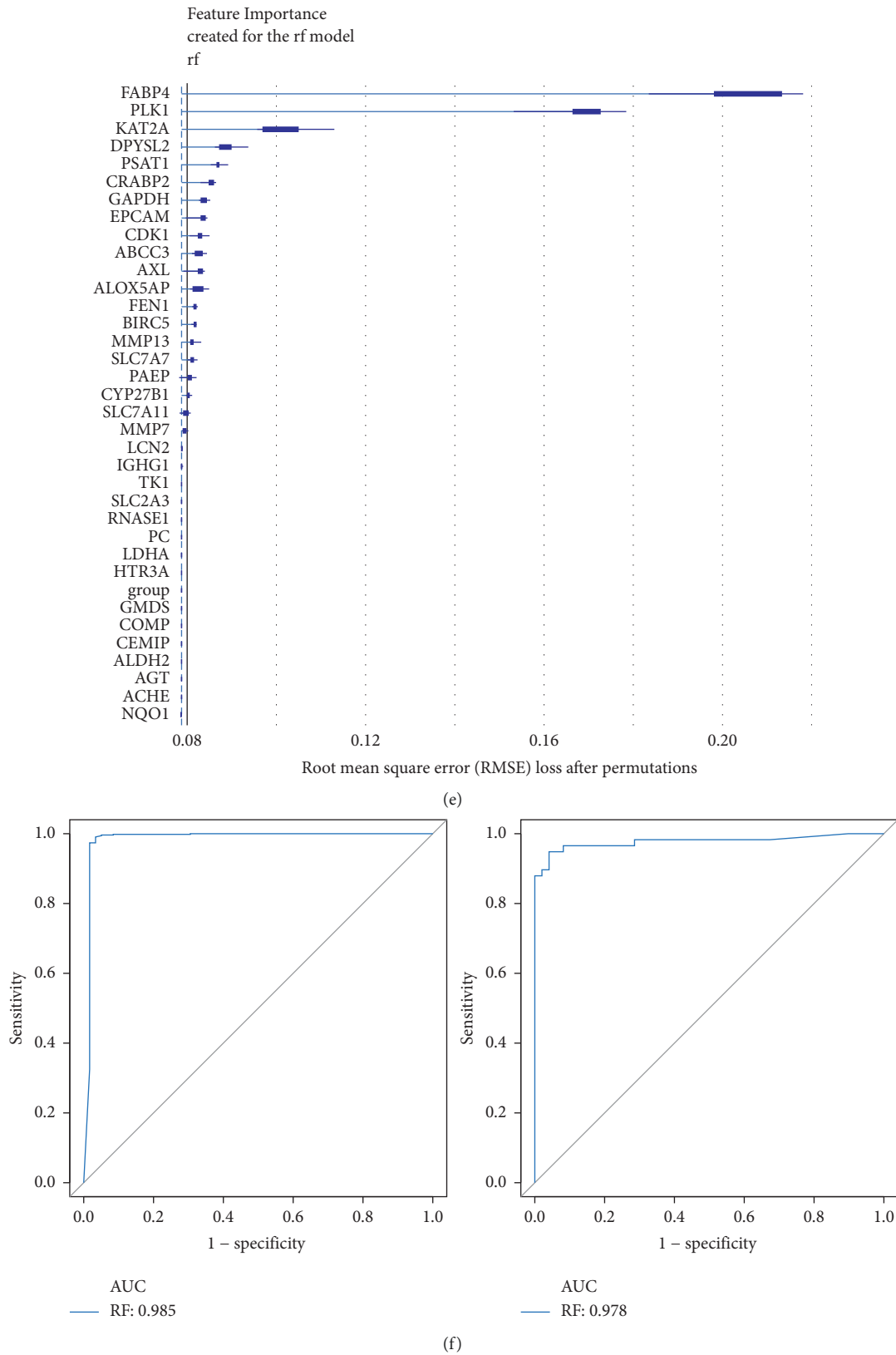


FIGURE 4: Construction of the LUAD diagnostic model and screening of biomarkers. (a) The characteristic genes were screened by LASSO regression analysis. Deviance on the horizontal axis represents the proportion of residual explained by the model, showing the relationship between the number of characteristic genes and the proportion of residual explained by the model (Dev), and the coefficient of genes on the vertical axis (left). The x-coordinate is log (Lambda), and the y-coordinate represents the error of cross-validation (right). (b) Evaluation and validation of the diagnostic model by the ROC curve. (c, d) The cumulative residual distribution diagram of the sample (left) and the boxplot of the sample residual (right). The curve area indicates the cumulative residual value of the whole sample. The smaller the curve area is, the smaller the cumulative residual value of the sample is. The red dot represents the root mean square of the residual. (e) Importance of genetic variables in the RF models. (f) Evaluation of the RF model by the ROC curve.

3.6. Genes Expression Analysis. 8 candidate genes (CRABP2, KAT2A, BIRC5, ABCC3, PLK1, FABP4, IGHG1, and EPCAM) were obtained from the intersection of genes obtained by the LASSO model and RF model, as shown in Figure 5(a). In this study, single-gene ROC curves were plotted for eight genes, and AUC values were calculated and validated at GSE10072. The AUC values of all the intersecting genes in the TCGA dataset and the GSE10072 validation set were above 0.8, as shown in Figure 5(b). CRABP2, KAT2A, BIRC5, ABCC3, IGHG1, EPCAM, and PLK1 were upregulated in the LUAD samples, while FABP4 was downregulated in the LUAD samples, as shown in Figure 5(c), and there was a strong correlation between these biomarkers, as shown in Figure 5(d). In addition, the expression and correlation of biomarkers were validated at GSE10072, and the results were consistent with TCGA, as shown in Figure S1.

3.7. Construction of a Pharmacological Regulatory Network Map and Molecular Docking. The drug-active ingredient-key target gene network diagram contains 1 drug (blister beetles), 7 drug-active ingredients, 8 biomarkers, and 32 relationship pairs, as shown in Figure 6(a). According to the molecular docking results of the 3D conformer structure of cantharidin, as shown in Figure 6(b), and the crystal structure of BIRC5 with PDB ID 3UEH, there are hydrogen bond interactions between residues of SER-81, ASN-111, CYS-60, and cantharidin molecules. The docking affinity between the active molecule and the protein was -5.2 kcal/mol, as shown in Figure 6(c). According to the molecular docking results of the 3D conformer structure of oleic acid, as shown in Figure 6(d), and the crystal structure of KAT2A with PDB ID 5TRM, there are hydrogen bond interactions between residues of TRP-519, ARG-515, and oleic acid molecules. The docking affinity between the active molecule and the protein was -5.9 kcal/mol, as shown in Figure 6(e). According to the molecular docking results of the 3D conformer structure of 3-Phenyl-4-azafluorene, as shown in Figure 6(f), and the crystal structure of PLK1 with PDB ID 1Q40, the TRP-414 residue has hydrogen bonding interactions with the 3-phenyl-4-azafluorene molecule. The docking affinity between the active molecule and protein was -8.4 kcal/mol, as shown in Figure 6(g). According to the molecular docking results of the 3D conformer structure of 3-phenyl-4-azafluorene, as shown in Figure 6(h), and the crystal structure of BIRC5 with PDB ID 3UEH, there are hydrogen bond interactions between residues of ALA-109, LYS-112, PHE-61, ILE-44, and 3-phenyl-4-azafluorene molecules. The docking affinity between the active molecule and the protein was -6.8 kcal/mol, as shown in Figure 6(i). All have a strong affinity.

3.8. Potential Function, Prognostic Significance, and Clinical Relevance of Biomarkers. GO function analysis was performed on the biomarkers, and the biomarkers were mainly enriched into two functional categories of CC and MF ($P < 0.05$ and $\text{count} \geq 2$). In terms of cell composition, biomarkers were associated with the STAGA complex,

dihydro lipid-based dehydrogenase complex, concentrated chromosome, centromere central centromere concentrated chromosome, centromere concentrated spindle microtubules, etc. In terms of molecular function, biomarkers were associated with retinal binding, H4 histone acetyltransferase activity, long-chain fatty acid transporter activity, retinol binding, long-chain fatty acid binding, microtubule binding, organic acid binding, carboxylic acid binding, lipid transporter activity, and monocarboxylic acid binding, as shown in Figure 7(a). Then, KEGG pathway enrichment analysis was performed on the biomarkers ($P < 0.05$ and $\text{count} \geq 1$), and the biomarkers were mainly enriched into CC and MF functional categories. The results showed that biomarkers were associated with PPAR signaling pathway, platinum resistance, notch signaling pathway, lipolysis regulation of adipocytes, ABC transporter, apoptosis, and antifolate resistance, as shown in Figure 7(b). Finally, we explored the prognostic significance of biomarkers. Only BIRC5 and PLK1 were found to have significant survival significance, as shown in Figure 7(c), while the remaining genes were not significant, as shown in Figure S2. We also studied the correlation between biomarkers and clinical factors (age, sex, TNM stage, and stage stage). The results showed that BIRC5 and PLK1 were strongly correlated with clinical factors. In addition, IGHG1 was significantly negatively correlated with *T* staging and stage staging, as shown in Figure 7(d).

4. Discussion

In China, blister beetles and their derivatives are used clinically, especially in antitumor therapy, which has been proven to have good effects. However, blister beetles and their derivatives exhibit cardiac, liver, and urinary system toxicity, which limits their clinical application and dose setting [30]. In addition, the application of raw materials of blister beetles leads to high poisoning, including abuse, unreasonable processing, dosage, and accumulation of toxicity [31]. Therefore, the analysis of pharmacodynamic constituents and target of cantharidin, and the molecular mechanism of blister beetles against tumor are helpful to the development of new antitumor drugs.

This study reveals that the main pharmacodynamic components of blister beetles are related to the 8 targets in lung adenocarcinoma. Among them, CRABP2, KAT2A, BIRC5, ABCC3, PLK1, IGHG1, and EPCAM were upregulated in lung adenocarcinoma samples, while FABP4 was downregulated in lung adenocarcinoma samples. Further survival analysis showed that BIRC5 and PLK1 had significant survival significance. The baculoviral inhibitor of apoptosis protein repeat-containing 5 (BIRC5), also known as survivin, is a member of the inhibitor of apoptosis protein (IAP) family and also a target for cancer therapy [32] which exists at the crossroads of many cancer cell signaling networks. Survivin is not only involved in the progression of NSCLC, but also related to the development of drug resistance. For instance, reduced survivin expression is a key marker for evaluating prognosis and survival in stage III NSCLC patients who receive platinum-based radiotherapy after surgery [33]. Survivin silencing resensitized A549/VCR

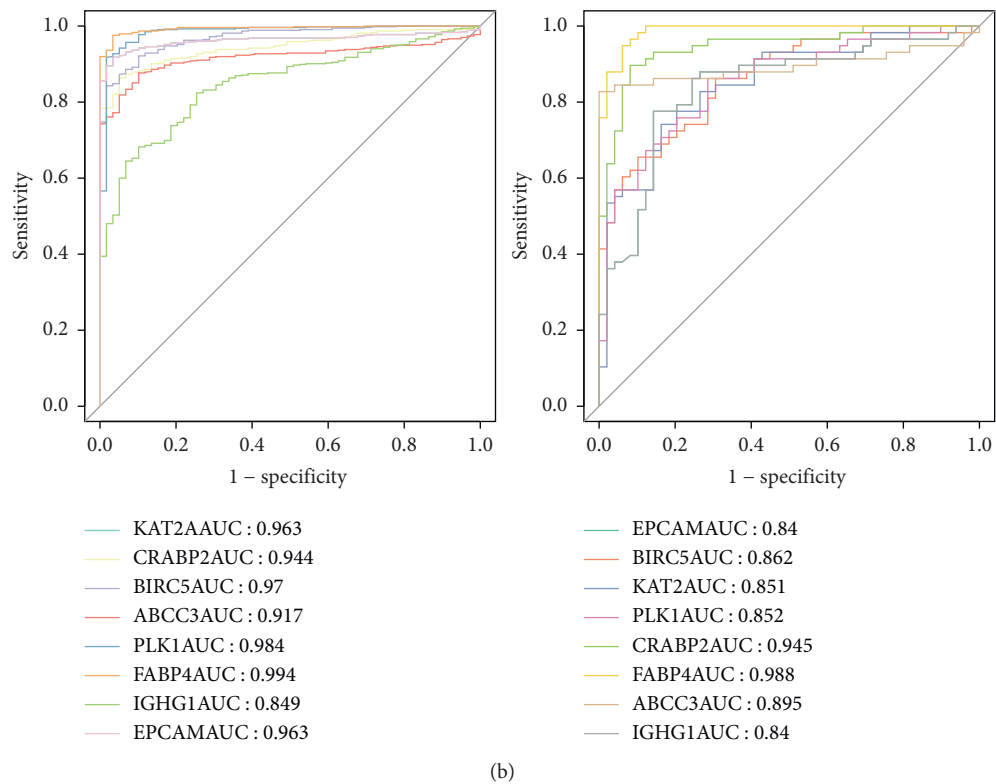
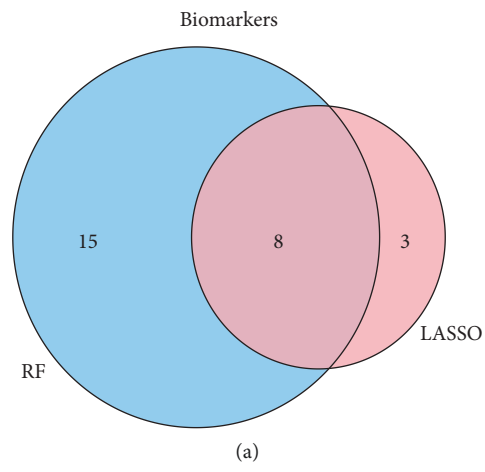
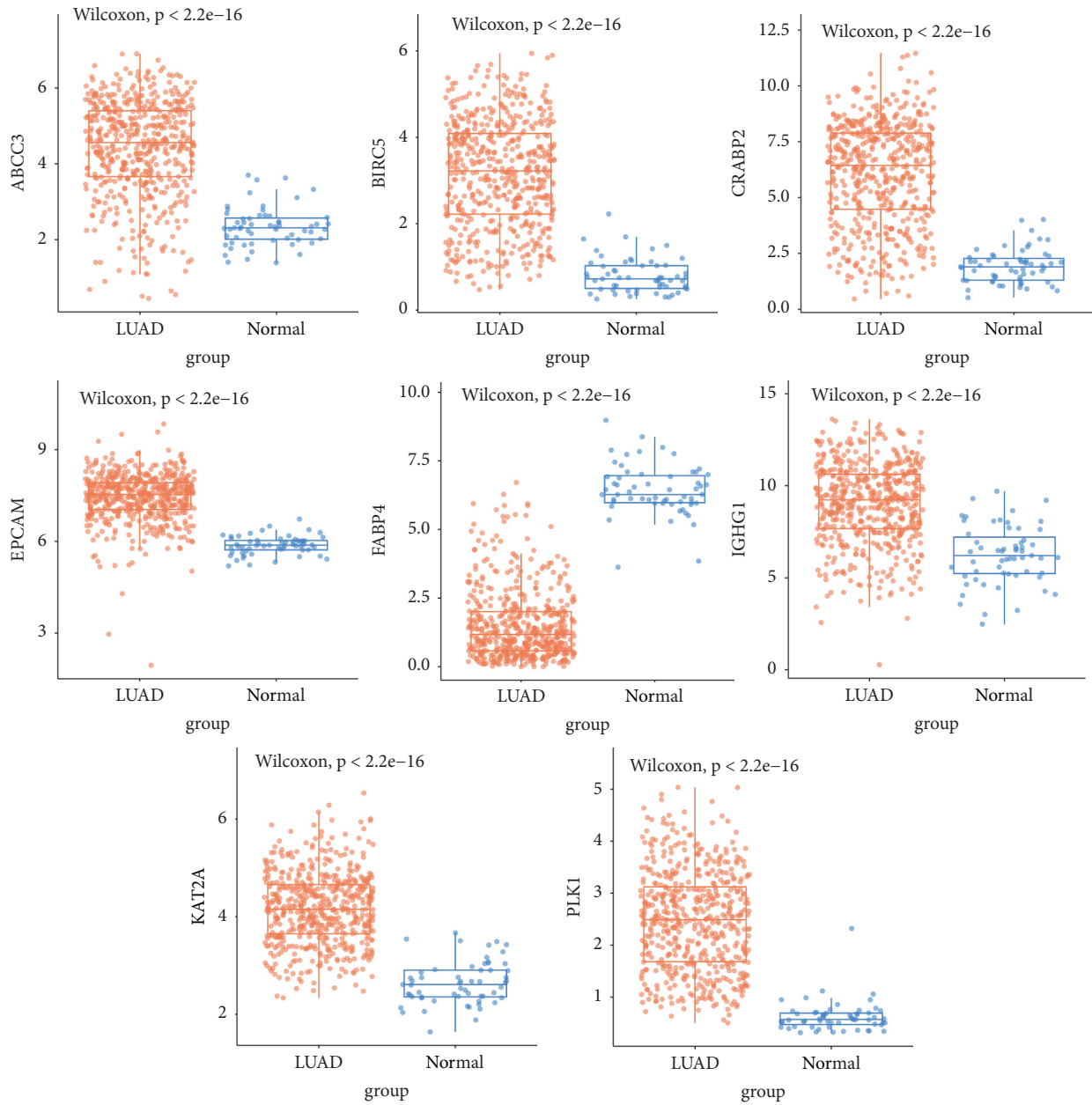


FIGURE 5: Continued.



(c)

FIGURE 5: Continued.

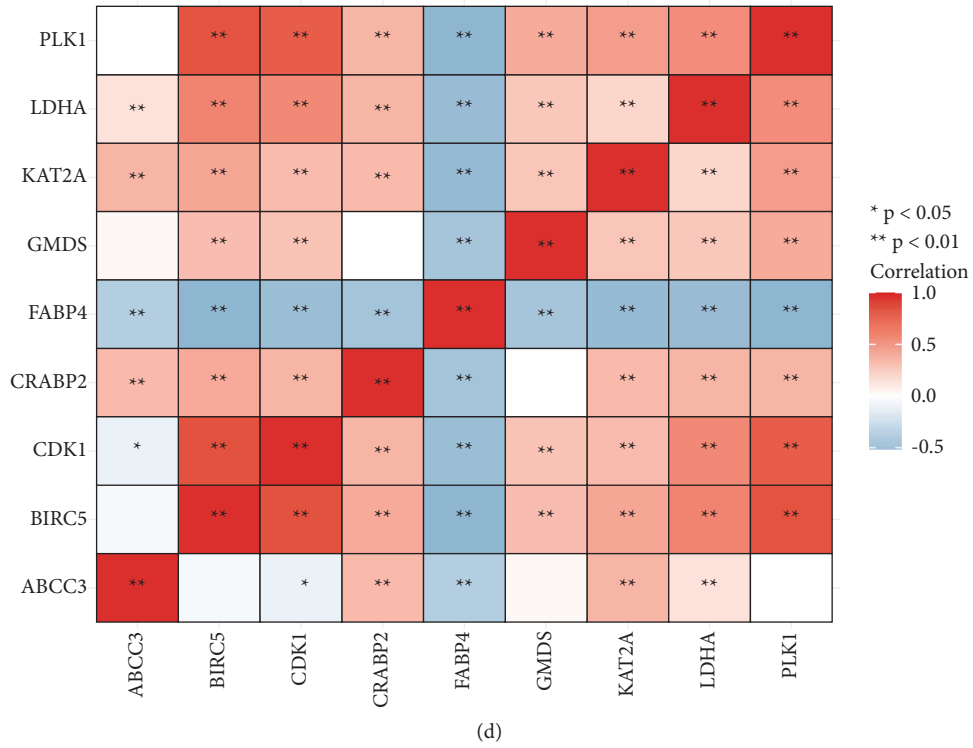


FIGURE 5: Screening and expression analysis of characteristic genes. (a) Intersection of the LASSO diagnostic model gene and RF model gene. Pink represents the LASSO diagnostic model genes and blue represents the RF model genes. (b) ROC curves for candidate biomarkers. (c) Expression of biomarkers. (d) Correlation of biomarkers. Blue is negative and red is positive.

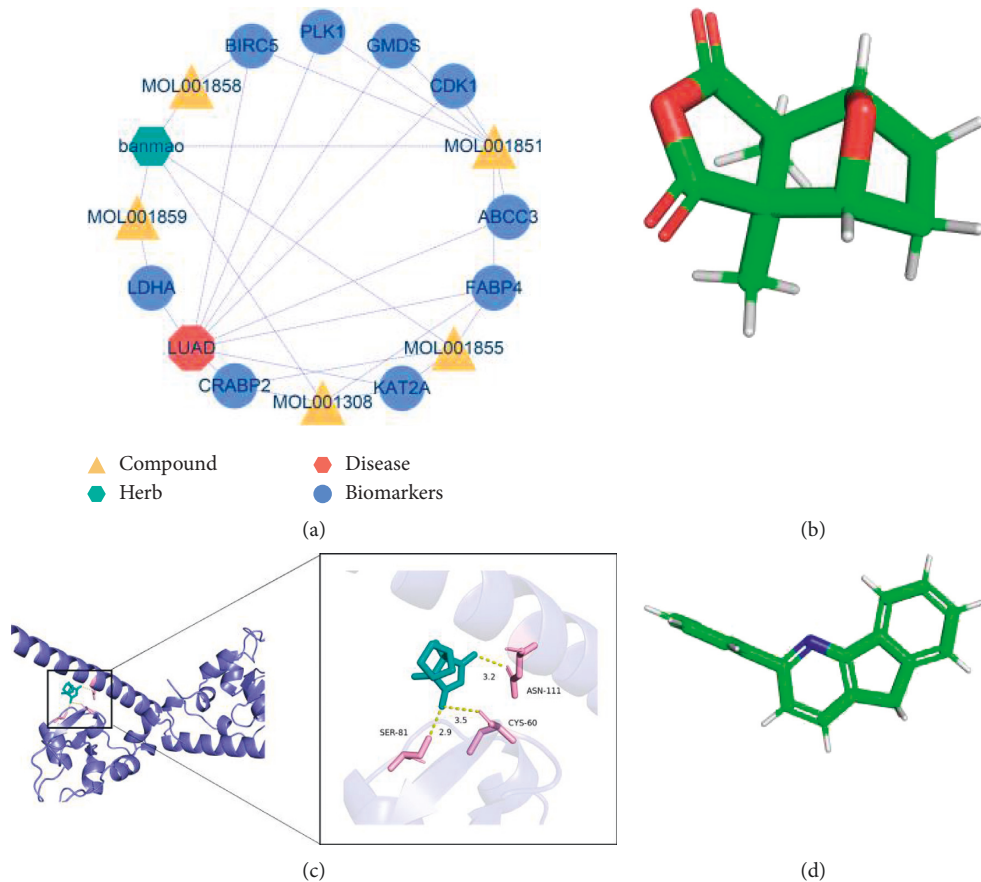


FIGURE 6: Continued.

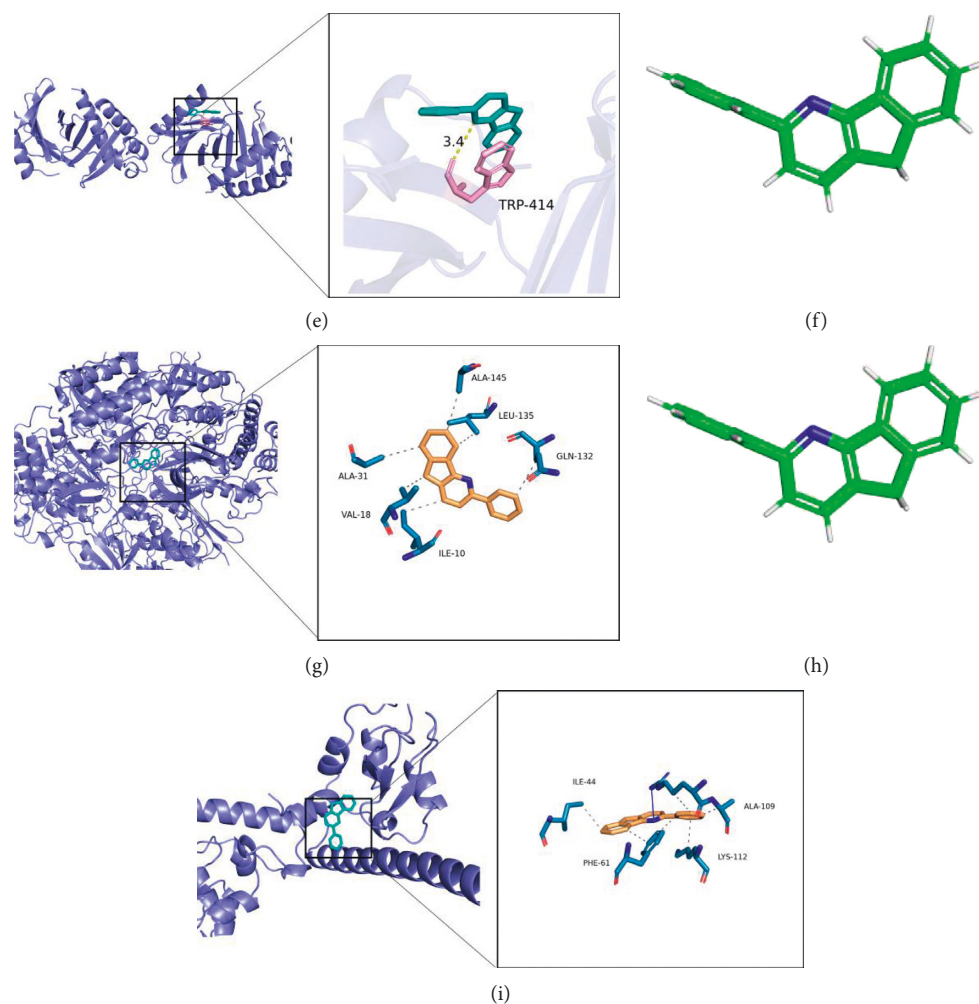


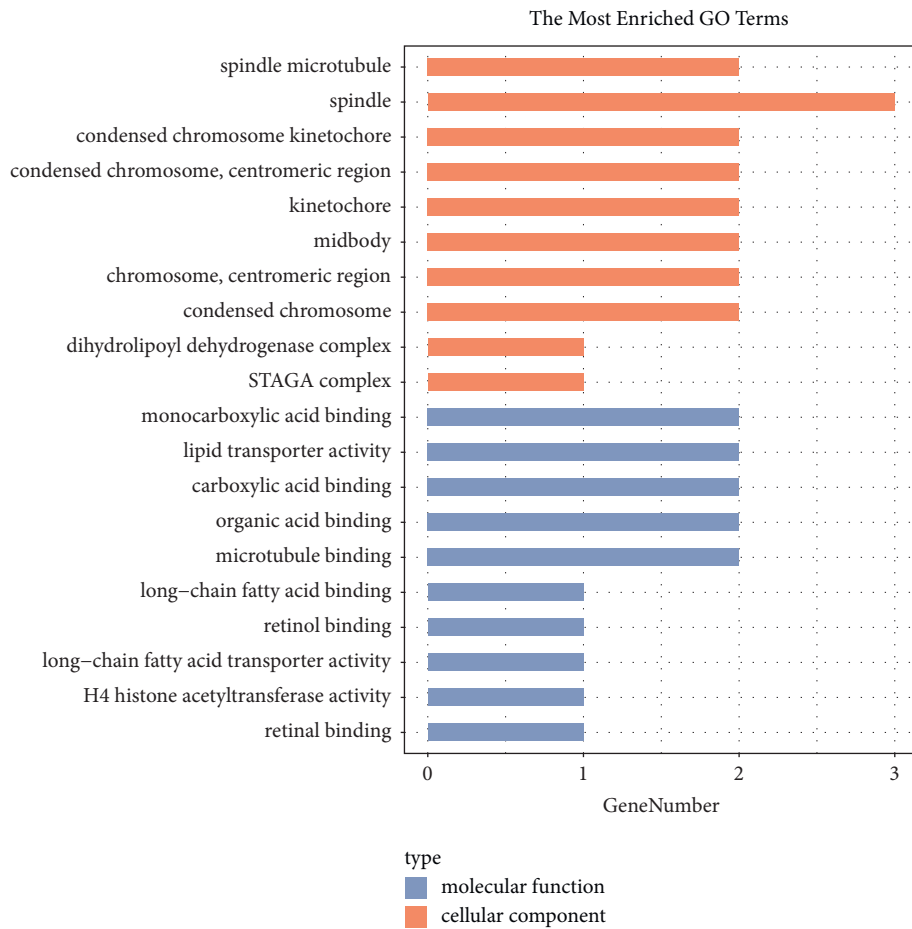
FIGURE 6: Construction of the pharmacological regulatory network map of Chinese medicine with biomarkers and results of molecular docking. (a) Network diagram of pharmacological regulation of biomarkers in traditional Chinese medicine. (b) 3D conformer structure schematic of cantharidin. (c) Docking result diagram of 3UEH and cantharidin. The green double-ring stick model is the active molecule cantharidin. (d) 3D conformer structure schematic of oleic acid. (e) Docking result diagram of 5TRM and oleic acid. The green double-ring stick model is the active molecule oleic acid. (f) 3D conformer structure schematic of 3-phenyl-4-azafluorene. (g) Docking result diagram of 1Q40 and 3-phenyl-4-azafluorene. The green double-ring stick model is the active molecule 3-phenyl-4-azafluorene. (h) 3D conformer structure schematic of 3-phenyl-4-azafluorene. (i) Docking result diagram of 3UEH and 3-phenyl-4-azafluorene. The green double-ring stick model is the active molecule 3-phenyl-4-azafluorene. Dotted gray/yellow lines represent hydrophobic bonds formed between active ingredients and amino acid residues. Each dotted gray/yellow line represents a hydrophobic bond.

cells to vincristine and methotrexate *in vitro* [34]. Inhibition of survivin expression by the short hairpin RNA expression vector significantly reduced the growth of lung cancer cells *in vivo* and *in vitro* [35]. In addition, studies have shown that the polymorphisms of survivin-31 G/C and 9194 A/G are related to the risk of lung cancer. Carriers of CGGC and GCAT haplotypes are more likely to develop lung cancer [36]. These studies suggest that inhibition of survivin is likely to be a major way for cantharidin to inhibit LUAD.

Polo-like kinase 1 (PLK1) is a highly conserved silk/threonine protein kinase that plays an important role in cell cycle regulation and DNA damage repair [37]. Current studies have found that PLK1 is highly expressed in a variety of tumors, including breast cancer, colorectal cancer, endometrial cancer, ovarian cancer, pancreatic cancer, and

non-small-cell lung cancer [38]. The PLK1 expression level is related to the degree of malignancy and prognosis of patients [39–42]. Cao et al. [43] found that the downregulation of Mir-886-3p in small-cell lung cancer is closely related to the shorter survival of patients, and Mir-886-3p can inhibit the expression of target genes PLK1 and TGF- β 1, thus inhibiting the proliferation, migration, and invasion of small-cell lung cancer. Reda et al. [44] found that the application of C-SiPLK1-NP reduced tumor growth and led to prolonged survival. Therefore, PLK1 may be a new prognostic marker and a good target for chemotherapy intervention.

In this study, we obtained the key targets of blister beetles treating LUAD and the biomarkers related to LUAD by using network pharmacology and molecular docking methods based on four databases, indicating that the



(a)
 FIGURE 7: Continued.

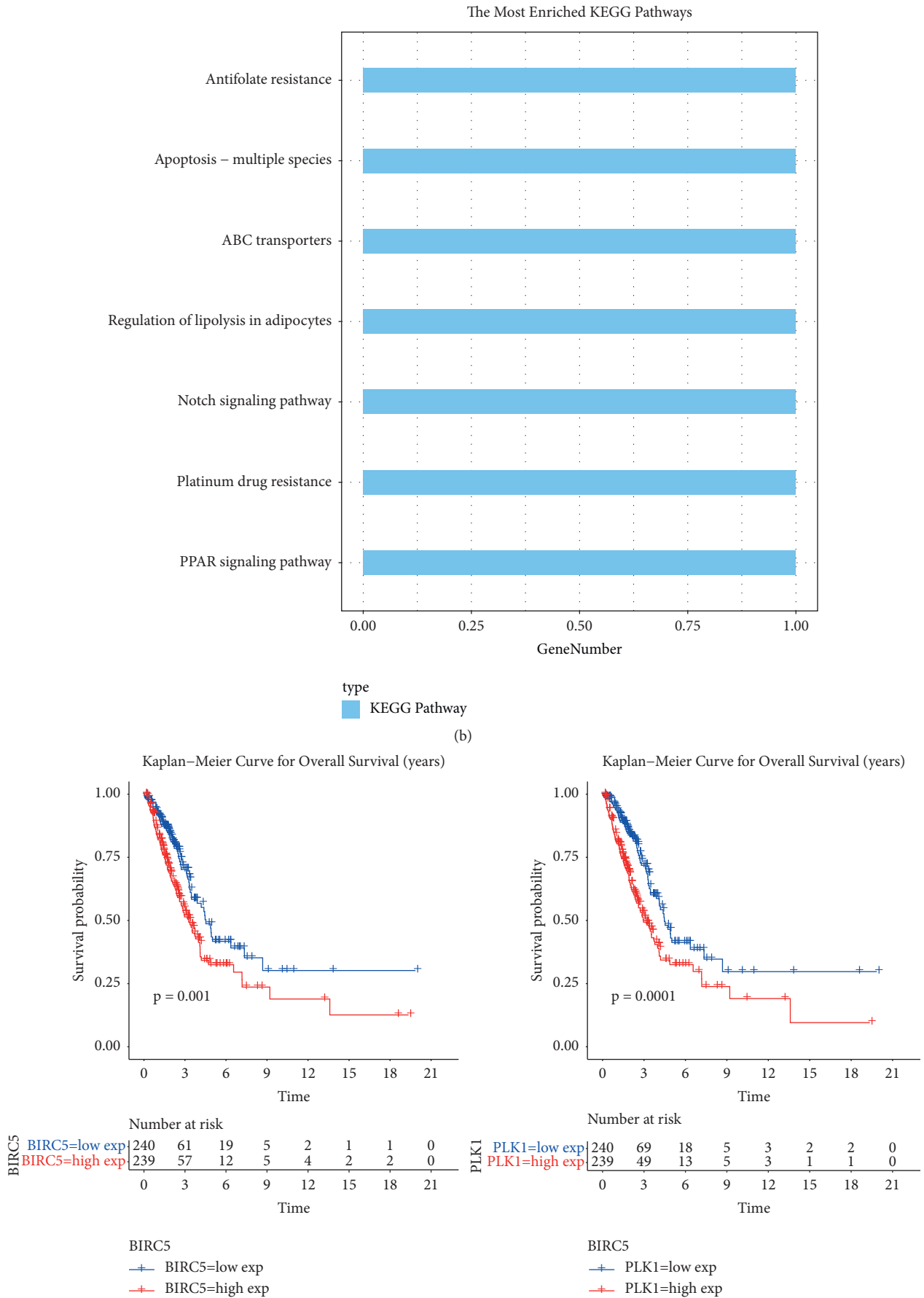


FIGURE 7: Continued.

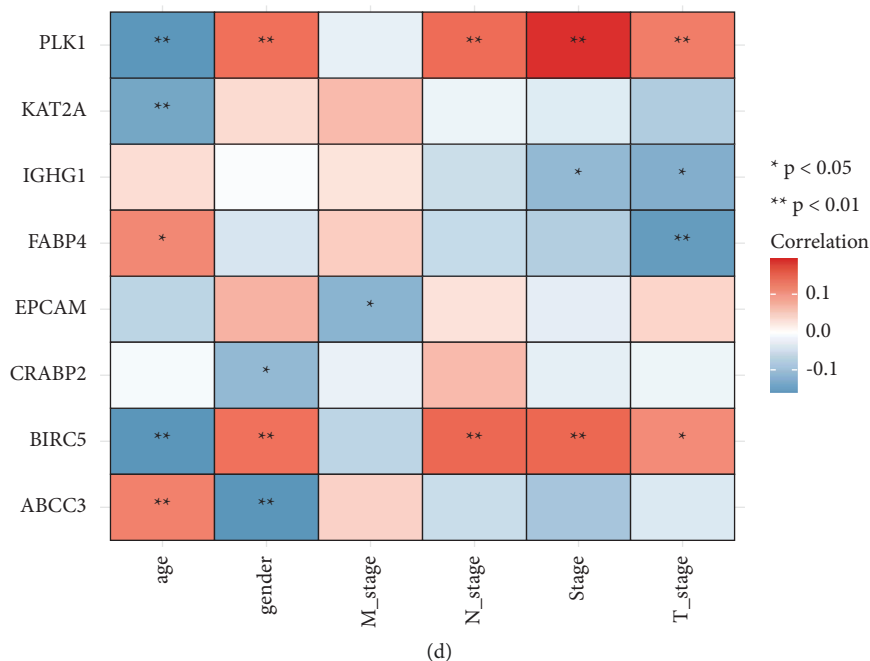


FIGURE 7: Potential function, prognostic significance, and clinical relevance of biomarkers. (a) Bar chart of GO enrichment results for biomarkers. (b) Bar chart of KEGG enrichment results for biomarkers. (c) Kaplan–Meier survival profile of biomarkers. (d) Heat maps of biomarkers associated with clinical factors.

application of network pharmacology is a scientific and feasible method. However, these results need to be verified in more cell experiments to be better applied in clinical treatment and drug development, including the potential pathway of the Blister beetle antilung cancer activity, the mechanism of action in treating lung cancer, and in vivo experiments to confirm its activity are further needed.

5. Conclusions

In conclusion, we screened the key targets of cantharidin for the treatment of LUAD and the biomarkers in LUAD based on network pharmacology and bioinformatics gene analysis. The results revealed that the 7 active compounds exert their antitumor effects via 8 targets in 32 pathways, and BIRC5 and PLK1 had a significant survival significance. This is consistent with the TCM concept of “multiple compounds, multiple targets, and multiple effects.” This provides a systematic view of the potential anticancer mechanism of cantharidin.

Data Availability

The data that support the findings of this study are openly available in TCMSP (<https://tcmssp.com/tcmssp.php>); Drug Bank (<https://go.drugbank.com/>); GeneCards (<https://www.genecards.org/>); TCGA database (<https://portal.gdc.cancer.gov/>); STRING (<https://string-db.org/>); PDB database (<https://www.rcsb.org/>); and PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>).

Consent

Not applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Shoujun Deng contributed to the conception, designed the study, analyzed the data, and wrote the initial draft of the manuscript. Ying Mao and Heng Li completed statistical analysis and the visualization of the results, discussed the results, and revised the manuscript. Gaofeng Li provided financial support for this study.

Acknowledgments

This work was supported in part by the Science and Technology Project of Yunnan Province (202101AY070001-205) and the Department of Education Scientific Research Fund of Yunnan Province (2021J0280).

Supplementary Materials

1. Figure S1: expression of biomarkers in GSE10072 dataset.
2. Figure S2: the prognostic significance of the remaining biomarkers. (*Supplementary Materials*)

References

- [1] J. Wang, G. Li, L. Yu, T. Mo, Q. Wu, and Z. Zhou, “Aidi injection plus platinum-based chemotherapy for stage IIIB/IV non-small cell lung cancer: a meta-analysis of 42 RCTs following the PRISMA guidelines,” *Journal of Ethnopharmacology*, vol. 221, pp. 137–150, 2018.

- [2] Q. Y. Hong, G. M. Wu, and G. S. Qian, "Prevention and management of lung cancer in China," *Cancer*, vol. 121, pp. 3080–3088, 2015.
- [3] R. S. Zheng, K. X. Sun, and S. W. Zhang, "Report of cancer epidemiology in China," *Zhonghua Zhongliu Zazhi*, vol. 41, no. 1, pp. 19–28, 2019.
- [4] M. Zheng, "Classification and pathology of lung cancer," *Surgical Oncology Clinics of North America*, vol. 25, no. 3, pp. 447–468, 2016.
- [5] Z. F. Lim and P. C. Ma, "Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy," *Journal of Hematology & Oncology*, vol. 12, no. 1, p. 134, 2019.
- [6] W. J. Liu, Y. Du, R. Wen, M. Yang, and J. Xu, "Drug resistance to targeted therapeutic strategies in non-small cell lung cancer," *Pharmacology & Therapeutics*, vol. 206, Article ID 107438, 2020.
- [7] G. Delannoy and C. Vallet, "Blister beetle," *Revue du Praticien*, vol. 67, no. 10, p. 1094, 2020.
- [8] Y. M. Kim, M. J. Ku, Y. J. Son, J. M. Yun, S. H. Kim, and S. Y. Lee, "Anti-metastatic effect of cantharidin in A549 human lung cancer cells," *Archives of Pharmacal Research*, vol. 36, no. 4, pp. 479–484, 2013.
- [9] X. Yang, Y. Tian, and H. Liu, "Heavy metal pollution analysis and Health risk assessment of two medicinal insects of mylabris," *Biological Trace Element Research*, vol. 200, no. 4, pp. 1892–1901, 2022.
- [10] H. C. Li, Z. H. Xia, and Y. F. Chen, "Cantharidin inhibits the growth of triple-negative breast cancer cells by suppressing autophagy and inducing apoptosis in vitro and in vivo," *Cellular Physiology and Biochemistry*, vol. 43, no. 5, pp. 1829–1840, 2017.
- [11] L. C. Zheng, M. D. Yang, and C. L. Kuo, "Norcantharidin-induced apoptosis of AGS human gastric cancer cells through reactive oxygen species production, and caspase- and mitochondria-dependent signaling pathways," *Anticancer Research*, vol. 36, no. 11, pp. 6031–6042, 2016.
- [12] Z. Jiang, J. Chi, B. Han, and W. Liu, "Preparation and pharmacological evaluation of norcantharidin-conjugated carboxymethyl chitosan in mice bearing hepatocellular carcinoma," *Carbohydrate Polymers*, vol. 174, pp. 282–290, 2017.
- [13] G. S. Wang, "Medical uses of mylabris in ancient China and recent studies," *Journal of Ethnopharmacology*, vol. 26, no. 2, pp. 147–162, 1989.
- [14] D. W. Whitman, M. F. Andrés, R. A. Martínez-Díaz, A. Ibáñez-Escribano, A. S. Olmeda, and A. González-Coloma, "Antiparasitic properties of cantharidin and the blister beetle *berberomeloe majalis* (Coleoptera: meloidae)," *Toxins*, vol. 11, pp. 234–4, 2019.
- [15] C. L. Maynard-Burrows, J. V. McLennan, T. C. Nicholson-Roberts, and H. Takano, "Blister beetle dermatitis caused by cantharidin in South Sudan in Op TRENTON: a case series," *BMJ Mil Health*, vol. 167, no. 5, pp. 362–364, 2019.
- [16] J. Chun, M. K. Park, H. Ko, K. Lee, and Y. S. Kim, "Bioassay-guided isolation of cantharidin from blister beetles and its anticancer activity through inhibition of epidermal growth factor receptor-mediated STAT3 and Akt pathways," *Journal of Natural Medicines*, vol. 72, no. 4, pp. 937–945, 2018.
- [17] S. Li and B. Zhang, "Traditional Chinese medicine network pharmacology: theory, methodology and application," *Chinese Journal of Natural Medicines*, vol. 11, no. 2, pp. 110–120, 2013.
- [18] A. L. Hopkins, "Network pharmacology: the next paradigm in drug discovery," *Nature Chemical Biology*, vol. 4, no. 11, pp. 682–690, 2008.
- [19] T. T. Luo, Y. Lu, S. K. Yan, X. Xiao, X. L. Rong, and J. Guo, "Network pharmacology in Research of Chinese medicine formula: methodology, application and prospective," *Chinese Journal of Integrative Medicine*, vol. 26, no. 1, pp. 72–80, 2020.
- [20] J. Y. Li, H. Y. Chen, W. J. Dai, Q. J. Lv, and C. Y. Chen, "Artificial intelligence approach to investigate the longevity drug," *Journal of Physical Chemistry Letters*, vol. 10, no. 17, pp. 4947–4961, 2019.
- [21] M. Tang, X. Xie, and P. Yi, "Integrating network pharmacology with molecular docking to unravel the active compounds and potential mechanism of simiao pill treating rheumatoid arthritis," *Evid Based Complement Alternat Med*, vol. 2020, Article ID 5786053, 16 pages, 2020.
- [22] Q. Guo, K. Zheng, and D. Fan, "Wu-tou decoction in rheumatoid arthritis: integrating network pharmacology and in vivo pharmacological evaluation," *Frontiers in Pharmacology*, vol. 8, p. 230, 2017.
- [23] G. Xie, W. Peng, and P. Li, "A network pharmacology analysis to explore the effect of astragalus radix-radix angelica sinensis on traumatic brain injury," *BioMed Research International*, vol. 2018, p. 3951783, 2018.
- [24] P. Li, J. Chen, W. Zhang, H. Li, W. Wang, and J. Chen, "Network pharmacology based investigation of the effects of herbal ingredients on the immune dysfunction in heart disease," *Pharmacological Research*, vol. 141, pp. 104–113, 2019.
- [25] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC Bioinformatics*, vol. 9, p. 559, 2008.
- [26] M. E. Ritchie, B. Phipson, and D. Wu, "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, p. e47, 2015.
- [27] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.
- [28] P. Shannon, A. Markiel, and O. Ozier, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [29] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.
- [30] N. A. Al-Dawsari and K. S. Masterpol, "Cantharidin in dermatology," *Skinmed*, vol. 14, no. 2, pp. 111–114, 2016.
- [31] F. Polticelli, L. Leboffe, and V. Tortosa, "Cantharidin inhibits competitively heme-Fe(III) binding to the FA1 site of human serum albumin," *Journal of Molecular Recognition*, vol. 3011 pages, 2017.
- [32] F. Li, I. Aljahdali, and X. Ling, "Cancer therapeutics using survivin BIRC5 as a target: what can we do after over two decades of study?" *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 368, 2019.
- [33] H. J. Cho, H. R. Kim, Y. S. Park, Y. H. Kim, D. K. Kim, and S. I. Park, "Prognostic value of survivin expression in stage III non-small cell lung cancer patients treated with platinum-based therapy," *Surg Oncol*, vol. 24, no. 4, pp. 329–334, 2015.
- [34] C. Zhou, Y. Zhu, B. Lu, W. Zhao, and X. Zhao, "Survivin expression modulates the sensitivity of A549 lung cancer cells resistance to vincristine," *Oncology Letters*, vol. 16, no. 4, pp. 5466–5472, 2018.

- [35] Q. Huang, Y. Zeng, H. Lin, H. Zhang, and D. Yang, "Transfection with Livin and Survivin shRNA inhibits the growth and proliferation of non-small cell lung cancer cells," *Molecular Medicine Reports*, vol. 16, no. 5, pp. 7086–7091, 2017.
- [36] G. Guo, Q. Zhang, and Z. Yu, "Correlation between survivin genetic polymorphisms and lung cancer susceptibility," *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 6, pp. 7426–7430, 2015.
- [37] S. Zitouni, C. Nabais, S. C. Jana, A. Guerrero, and M. Bettencourt-Dias, "Polo-like kinases: structural variations lead to multiple functions," *Nature Reviews Molecular Cell Biology*, vol. 15, no. 7, pp. 433–452, 2014.
- [38] Y. Degenhardt, J. Greshock, and S. Laquerre, "Sensitivity of cancer cells to Plk1 inhibitor GSK461364A is associated with loss of p53 function and chromosome instability," *Molecular Cancer Therapeutics*, vol. 9, no. 7, pp. 2079–2089, 2010.
- [39] S. Kumar, A. R. Sharma, G. Sharma, C. Chakraborty, and J. Kim, "PLK-1: angel or devil for cell cycle progression," *Biochimica et Biophysica Acta*, vol. 1865, no. 2, pp. 190–203, 2016.
- [40] G. Wolf, R. Elez, and A. Doermer, "Prognostic significance of polo-like kinase (PLK) expression in non-small cell lung cancer," *Oncogene*, vol. 14, no. 5, pp. 543–549, 1997.
- [41] R. Zhang, H. Shi, and F. Ren, "Misregulation of polo-like protein kinase 1, P53 and P21WAF1 in epithelial ovarian cancer suggests poor prognosis," *Oncology Reports*, vol. 33, no. 3, pp. 1235–1242, 2015.
- [42] P. Ramani, R. Nash, E. Sowa-Avugrah, and C. Rogers, "High levels of polo-like kinase 1 and phosphorylated translationally controlled tumor protein indicate poor prognosis in neuroblastomas," *Journal of Neuro-Oncology*, vol. 125, no. 1, pp. 103–111, 2015.
- [43] J. Cao, Y. Song, and N. Bi, "DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer," *Cancer Research*, vol. 73, no. 11, pp. 3326–3335, 2013.
- [44] M. Reda, W. Ngamcherdtrakul, and S. Gu, "PLK1 and EGFR targeted nanoparticle as a radiation sensitizer for non-small cell lung cancer," *Cancer Letters*, vol. 467, pp. 9–18, 2019.