

World Journal of *Clinical Cases*

World J Clin Cases 2022 September 16; 10(26): 9180-9549



Contents

Thrice Monthly Volume 10 Number 26 September 16, 2022

REVIEW

- 9180** Assisting individuals with diabetes in the COVID-19 pandemic period: Examining the role of religious factors and faith communities

Eseadi C, Ossai OV, Onyishi CN, Ilechukwu LC

- 9192** Role of octreotide in small bowel bleeding

Khedr A, Mahmoud EE, Attallah N, Mir M, Boike S, Rauf I, Jama AB, Mushtaq H, Surani S, Khan SA

MINIREVIEWS

- 9207** Internet of things-based health monitoring system for early detection of cardiovascular events during COVID-19 pandemic

Dami S

- 9219** Convergence mechanism of mindfulness intervention in treating attention deficit hyperactivity disorder: Clues from current evidence

Xu XP, Wang W, Wan S, Xiao CF

- 9228** Clinical presentation, management, screening and surveillance for colorectal cancer during the COVID-19 pandemic

Akbulut S, Hargura AS, Garzali IU, Aloun A, Colak C

- 9241** Early diagnostic value of liver stiffness measurement in hepatic sinusoidal obstruction syndrome induced by hematopoietic stem cell transplantation

Tan YW, Shi YC

ORIGINAL ARTICLE

Case Control Study

- 9254** Local inflammatory response to gastroesophageal reflux: Association of gene expression of inflammatory cytokines with esophageal multichannel intraluminal impedance-pH data

Morozov S, Sentsova T

Retrospective Study

- 9264** Evaluation of high-risk factors and the diagnostic value of alpha-fetoprotein in the stratification of primary liver cancer

Jiao HB, Wang W, Guo MN, Su YL, Pang DQ, Wang BL, Shi J, Wu JH

- 9276** One-half layer pancreaticojejunostomy with the rear wall of the pancreas reinforced: A valuable anastomosis technique

Wei JP, Tai S, Su ZL

- 9285** Development and validation of an epithelial-mesenchymal transition-related gene signature for predicting prognosis

Zhou DH, Du QC, Fu Z, Wang XY, Zhou L, Wang J, Hu CK, Liu S, Li JM, Ma ML, Yu H

Observational Study

- 9303** Incidence and risk factor analysis for swelling after apical microsurgery

Bi C, Xia SQ, Zhu YC, Lian XZ, Hu LJ, Rao CX, Jin HB, Shang XD, Jin FF, Li JY, Zheng P, Wang SH

CASE REPORT

- 9310** Acute carotid stent thrombosis: A case report and literature review

Zhang JB, Fan XQ, Chen J, Liu P, Ye ZD

- 9318** Congenital ovarian anomaly manifesting as extra tissue connection between the two ovaries: A case report

Choi MG, Kim JW, Kim YH, Kim AM, Kim TY, Ryu HK

- 9323** Cefoperazone-sulbactam and ornidazole for *Gardnerella vaginalis* bloodstream infection after cesarean section: A case report

Mu Y, Li JJ, Wu X, Zhou XF, Tang L, Zhou Q

- 9332** Early-onset ophthalmoplegia, cervical dyskinesia, and lower extremity weakness due to partial deletion of chromosome 16: A case report

Xu M, Jiang J, He Y, Gu WY, Jin B

- 9340** Posterior mediastinal extralobar pulmonary sequestration misdiagnosed as a neurogenic tumor: A case report

Jin HJ, Yu Y, He W, Han Y

- 9348** Unexpected difficult airway due to severe upper tracheal distortion: A case report

Zhou JW, Wang CG, Chen G, Zhou YF, Ding JF, Zhang JW

- 9354** Special epithelioid trophoblastic tumor: A case report

Wang YN, Dong Y, Wang L, Chen YH, Hu HY, Guo J, Sun L

- 9361** Intrahepatic multicystic biliary hamartoma: A case report

Wang CY, Shi FY, Huang WF, Tang Y, Li T, He GL

- 9368** ST-segment elevation myocardial infarction in Kawasaki disease: A case report and review of literature

Lee J, Seo J, Shin YH, Jang AY, Suh SY

- 9378** Bilateral hypocalcaemic cataracts due to idiopathic parathyroid insufficiency: A case report

Li Y

- 9384** Single organ hepatic artery vasculitis as an unusual cause of epigastric pain: A case report

Kaviani R, Farrell J, Dehghan N, Moosavi S

- 9390** Congenital lipoid adrenal hyperplasia with Graves' disease: A case report

Wang YJ, Liu C, Xing C, Zhang L, Xu WF, Wang HY, Wang FT

- 9398** Cytokine release syndrome complicated with rhabdomyolysis after chimeric antigen receptor T-cell therapy: A case report
Zhang L, Chen W, Wang XM, Zhang SQ
- 9404** Antiphospholipid syndrome with renal and splenic infarction after blunt trauma: A case report
Lee NA, Jeong ES, Jang HS, Park YC, Kang JH, Kim JC, Jo YG
- 9411** Uncontrolled high blood pressure under total intravenous anesthesia with propofol and remifentanyl: A case report
Jang MJ, Kim JH, Jeong HJ
- 9417** Noncirrhotic portal hypertension due to peripheral T-cell lymphoma, not otherwise specified: A case report
Wu MM, Fu WJ, Wu J, Zhu LL, Niu T, Yang R, Yao J, Lu Q, Liao XY
- 9428** Resumption of school after lockdown in COVID-19 pandemic: Three case reports
Wang KJ, Cao Y, Gao CY, Song ZQ, Zeng M, Gong HL, Wen J, Xiao S
- 9434** Complete recovery from segmental zoster paresis confirmed by magnetic resonance imaging: A case report
Park J, Lee W, Lim Y
- 9440** Imaging findings of immunoglobulin G4-related hypophysitis: A case report
Lv K, Cao X, Geng DY, Zhang J
- 9447** Systemic lupus erythematosus presenting with progressive massive ascites and CA-125 elevation indicating Tjasma syndrome? A case report
Wang JD, Yang YF, Zhang XF, Huang J
- 9454** Locally advanced cervical rhabdomyosarcoma in adults: A case report
Xu LJ, Cai J, Huang BX, Dong WH
- 9462** Rapid progressive vaccine-induced immune thrombotic thrombocytopenia with cerebral venous thrombosis after ChAdOx1 nCoV-19 (AZD1222) vaccination: A case report
Jiang SK, Chen WL, Chien C, Pan CS, Tsai ST
- 9470** Burkitt-like lymphoma with 11q aberration confirmed by needle biopsy of the liver: A case report
Yang HJ, Wang ZM
- 9478** Common carotid artery thrombosis and malignant middle cerebral artery infarction following ovarian hyperstimulation syndrome: A case report
Xu YT, Yin QQ, Guo ZR
- 9484** Postoperative radiotherapy for thymus salivary gland carcinoma: A case report
Deng R, Li NJ, Bai LL, Nie SH, Sun XW, Wang YS
- 9493** Follicular carcinoma of the thyroid with a single metastatic lesion in the lumbar spine: A case report
Chen YK, Chen YC, Lin WX, Zheng JH, Liu YY, Zou J, Cai JH, Ji ZQ, Chen LZ, Li ZY, Chen YX

- 9502** Guillain-Barré syndrome and hemophagocytic syndrome heralding the diagnosis of diffuse large B cell lymphoma: A case report
Zhou QL, Li ZK, Xu F, Liang XG, Wang XB, Su J, Tang YF
- 9510** Intravitreal injection of conbercept for bullous retinal detachment: A case report
Xiang XL, Cao YH, Jiang TW, Huang ZR
- 9518** Supratentorial hemangioblastoma at the anterior skull base: A case report
Xu ST, Cao X, Yin XY, Zhang JY, Nan J, Zhang J

META-ANALYSIS

- 9524** Certain sulfonylurea drugs increase serum free fatty acid in diabetic patients: A systematic review and meta-analysis
Yu M, Feng XY, Yao S, Wang C, Yang P

LETTER TO THE EDITOR

- 9536** Glucose substrate in the hydrogen breath test for gut microbiota determination: A recommended noninvasive test
Xie QQ, Wang JF, Zhang YF, Xu DH, Zhou B, Li TH, Li ZP
- 9539** A rare cause of acute abdomen after a Good Friday
Pante L, Brito LG, Franciscatto M, Brambilla E, Soldera J
- 9542** Obesity is associated with colitis in women but not necessarily causal relationship
Shen W, He LP, Zhou LL
- 9545** Risk stratification of primary liver cancer
Tan YW

ABOUT COVER

Editorial Board Member of *World Journal of Clinical Cases*, Youngmin Oh, MD, PhD, Associate Professor, Neurosurgeon, Department of Neurosurgery, Jeonbuk National University Medical School/Hospital, Jeonju 54907, Jeollabukdo, South Korea. timoh@jbnu.ac.kr

AIMS AND SCOPE

The primary aim of *World Journal of Clinical Cases* (WJCC, *World J Clin Cases*) is to provide scholars and readers from various fields of clinical medicine with a platform to publish high-quality clinical research articles and communicate their research findings online.

WJCC mainly publishes articles reporting research results and findings obtained in the field of clinical medicine and covering a wide range of topics, including case control studies, retrospective cohort studies, retrospective studies, clinical trials studies, observational studies, prospective studies, randomized controlled trials, randomized clinical trials, systematic reviews, meta-analysis, and case reports.

INDEXING/ABSTRACTING

The WJCC is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Journal Citation Reports/Science Edition, Current Contents®/Clinical Medicine, PubMed, PubMed Central, Scopus, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2022 Edition of Journal Citation Reports® cites the 2021 impact factor (IF) for WJCC as 1.534; IF without journal self cites: 1.491; 5-year IF: 1.599; Journal Citation Indicator: 0.28; Ranking: 135 among 172 journals in medicine, general and internal; and Quartile category: Q4. The WJCC's CiteScore for 2021 is 1.2 and Scopus CiteScore rank 2021: General Medicine is 443/826.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: Hua-Ge Yin; Production Department Director: Xu Guo; Editorial Office Director: Jin-Lei Wang.

NAME OF JOURNAL

World Journal of Clinical Cases

ISSN

ISSN 2307-8960 (online)

LAUNCH DATE

April 16, 2013

FREQUENCY

Thrice Monthly

EDITORS-IN-CHIEF

Bao-Gan Peng, Jerzy Tadeusz Chudek, George Kontogeorgos, Maurizio Serati, Ja Hyeon Ku

EDITORIAL BOARD MEMBERS

<https://www.wjgnet.com/2307-8960/editorialboard.htm>

PUBLICATION DATE

September 16, 2022

COPYRIGHT

© 2022 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/GerInfo/287>

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

PUBLICATION ETHICS

<https://www.wjgnet.com/bpg/GerInfo/288>

PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

ARTICLE PROCESSING CHARGE

<https://www.wjgnet.com/bpg/gerinfo/242>

STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/GerInfo/239>

ONLINE SUBMISSION

<https://www.f6publishing.com>



Retrospective Study

Development and validation of an epithelial–mesenchymal transition-related gene signature for predicting prognosis

De-Hua Zhou, Qian-Cheng Du, Zheng Fu, Xin-Yu Wang, Ling Zhou, Jian Wang, Cheng-Kai Hu, Shun Liu, Jun-Min Li, Meng-Li Ma, Hua Yu

Specialty type: Respiratory system

Provenance and peer review:

Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): A

Grade B (Very good): B

Grade C (Good): 0

Grade D (Fair): 0

Grade E (Poor): 0

P-Reviewer: Qin Y, China; Wang J, China

Received: April 2, 2022

Peer-review started: April 2, 2022

First decision: June 16, 2022

Revised: June 30, 2022

Accepted: July 20, 2022

Article in press: July 20, 2022

Published online: September 16, 2022



De-Hua Zhou, Xin-Yu Wang, Ling Zhou, Hua Yu, Department of General Surgery, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai 200434, China

Qian-Cheng Du, Zheng Fu, Jian Wang, Cheng-Kai Hu, Shun Liu, Department of Thoracic surgery, Shanghai Xuhui Central Hospital, Shanghai 200031, China

Jun-Min Li, Meng-Li Ma, Surgical Intensive Care Unit, Shanghai Xuhui Central Hospital, Shanghai 200031, China

Corresponding author: Hua Yu, MM, Associate Chief Physician, Department of General Surgery, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, No. 1279 Sanmen Road, Shanghai 200434, China. luckyhuhua@163.com

Abstract

BACKGROUND

Currently, there are many therapeutic methods for lung adenocarcinoma (LUAD), but the 5-year survival rate is still only 15% at later stages. Epithelial–mesenchymal transition (EMT) has been shown to be closely associated with local dissemination and subsequent metastasis of solid tumors. However, the role of EMT in the occurrence and development of LUAD remains unclear.

AIM

To further elucidate the value of EMT-related genes in LUAD prognosis.

METHODS

Univariate, least absolute shrinkage and selection operator, and multivariate Cox regression analyses were applied to establish and validate a new EMT-related gene signature for predicting LUAD prognosis. The risk model was evaluated by Kaplan–Meier survival analysis, principal component analysis, and functional enrichment analysis and was used for nomogram construction. The potential structures of drugs to which LUAD is sensitive were discussed with respect to EMT-related genes in this model.

RESULTS

Thirty-three differentially expressed genes related to EMT were found to be highly associated with overall survival (OS) by using univariate Cox regression analysis ($\log_2\text{FC} \geq 1$, false discovery rate < 0.001). A prognostic signature of 7

EMT-associated genes was developed to divide patients into two risk groups by high or low risk scores. Kaplan–Meier survival analysis showed that the OS of patients in the high-risk group was significantly poorer than that of patients in the low-risk group ($P < 0.05$). Multivariate Cox regression analysis showed that the risk score was an independent risk factor for OS ($HR > 1$, $P < 0.05$). The results of receiver operator characteristic curve analysis suggested that the 7-gene signature had a perfect ability to predict prognosis (all area under the curves > 0.5).

CONCLUSION

The EMT-associated gene signature classifier could be used as a feasible indicator for predicting OS.

Key Words: Lung adenocarcinoma; Epithelial–mesenchymal transition; Gene signature; Overall survival

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Lung cancer is one of the major causes of death associated with malignancy, and lung cancer is associated with approximately 2 million new cases and 1.76 million deaths every year. Although some reports have shown that suppression or knockdown of some genes in lung adenocarcinoma (LUAD) could reverse the epithelial–mesenchymal transition (EMT) process that inhibits tumor occurrence and metastasis, the correlations of these genes with overall survival in patients with LUAD remain largely unclear. A cohort of 884 LUAD patients was used to construct and validate a novel predictive model comprising a 7-EMT-related gene signature, which can be used to predict the prognosis of LUAD.

Citation: Zhou DH, Du QC, Fu Z, Wang XY, Zhou L, Wang J, Hu CK, Liu S, Li JM, Ma ML, Yu H. Development and validation of an epithelial–mesenchymal transition-related gene signature for predicting prognosis. *World J Clin Cases* 2022; 10(26): 9285-9302

URL: <https://www.wjgnet.com/2307-8960/full/v10/i26/9285.htm>

DOI: <https://dx.doi.org/10.12998/wjcc.v10.i26.9285>

INTRODUCTION

Worldwide, lung cancer is one of the major causes of death associated with malignancy, and lung cancer is associated with approximately 2 million new cases and 1.76 million deaths every year[1]. Approximately 80%-85% of lung carcinomas are non-small cell lung cancer, and the two major histopathological types are lung adenocarcinoma (LUAD) and lung squamous cell carcinoma[2]. Among patients with lung cancer, 40% have LUAD, which has a 5-year survival rate of only 15% at later stages[3]. Currently, there are many therapeutic methods for LUAD, including curative resection, radiotherapy, chemotherapy, targeted therapy, and immunotherapy. Although the therapeutic prognosis of lung cancer has improved significantly because of advances in technology and selection strategies, the prognosis for lung cancer patients remains poor, as 70%-80% of patients are diagnosed in advanced stages[4,5]. In modern clinical oncology, histopathology is commonly effective in predicting the prognosis of lung cancer patients; however, its value is limited, as individual distinctions in patients with the identical pathological characteristics result in different conclusions. Therefore, more effective biomarkers for the early diagnosis of LUAD and more accurate prognostic predictions are necessary.

Epithelial–mesenchymal transition (EMT) is a multistep process in which morphological changes occur and is involved in neoplastic invasion, metastasis, embryogenesis and wound healing; during EMT, epithelial cells acquire mesenchymal characteristics and gradually lose their epithelial features[6]. EMT has been shown to be closely associated with local dissemination and subsequent metastasis of solid tumors[7]. Exosomes also frequently facilitate communication between cells and the extracellular matrix, complete the EMT process, and increase the transformation of nontumor stem cells into tumor stem cells, which accelerates the development of cancer[8]. There are many EMT states that can occur in cancer, and these states contribute differently to carcinogenesis, invasion, and metastasis. Mixed EMT states have a higher propensity for metastasis[9]. Pinin was shown by Qiao *et al*[10] to induce EMT by controlling m6A modification, and it might also be a potential anticancer target for hepatocellular carcinoma treatment. Previous studies suggested that highly proliferative non-EMT cells are sensitive to chemotherapy, and recurrent EMT-derived metastases were observed after treatment[11]. However, the underlying mechanism of EMT in the occurrence and development of non-small cell lung cancer, especially LUAD, remains largely unclear. Therefore, understanding the roles of EMT-related genes in LUAD diagnosis and prognostic prediction might be useful for identifying prognostic targets.

Table 1 Gene sets enriched in lung adenocarcinoma

GS: follow link to MSigDB	Size	NES	Nom P value	FDR Q value
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_DN	144	-1.66	0.040	0.040
GOBP_CARDIAC_EPITHELIAL_TO_MESENCHYMAL_TRANSITION	30	-1.98	0.006	0.006
GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION	150	-1.71	0.045	0.045
GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION	16	-2.05	< 0.001	< 0.001
GOBP_POSITIVE_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION	5	-1.65	0.006	0.006
GOBP_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION	6	-1.61	0.029	0.029
HOLLERN_EMT_BREAST_TUMOR_DN	121	1.90	0.020	0.020
JECHLINGER_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_DN	62	-1.80	0.042	0.042
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP	166	2.03	0.008	0.008

GS: Gene set; MSigDB: Molecular Signatures Database; NES: Normalized enrichment score; NOM: Normalized; FDR: False discovery rate

However, studies based on the transcriptome that systematically assess the EMT-related gene signature and its correlation with overall survival (OS) in LUAD patients remains rare. In this study, we explored available gene expression data from LUAD patient samples ($n = 497$) and extracted a total of 54774 candidate genes. In addition, GSEA identified 612 EMT-related functional gene sets that presented differential expression between tumor samples and normal samples. Then, we became the first to develop a novel prognostic model composed of differentially expressed EMT-related genes based on LUAD samples. This model based on external data ($n = 443$) was further validated to be an independent prognostic factor related to OS in LUAD patients.

MATERIALS AND METHODS

Baseline characteristics

The LUAD dataset, including whole-transcriptome RNA sequencing data and the corresponding clinical data of LUAD patients up to July 15, 2021, was downloaded from The Cancer Genome Atlas (<https://cancergenome.nih.gov/>). The clinical data of the LUAD patients included sex, grade, age, stage, and tumor, node and metastasis (TNM) classification. Data normalization was carried out utilizing the limma (Linear Models for Microarray Data) package in the R environment (version 4.1.0, 2021-05-18, R Foundation for statistical computing, Vienna, Austria). Gene expression profiles were obtained using raw data, and log2 normalization was performed for each candidate gene for further statistical analysis. If one gene occupied more than one row, we took the mean of its expression value. All genes with an expression value of 0 were deleted from the sample. RNA-seq data and the corresponding clinical baseline characteristics for 443 LUAD patient samples were collected from a Gene Expression Omnibus dataset (GSE68465, <http://www.ncbi.nlm.nih.gov/geo/>). All 443 candidate patients underwent surgery. Normalized count values were utilized. All the data in this article obtained from The Cancer Genome Atlas and the Gene Expression Omnibus dataset were open access and freely available. Therefore, the study did not need approval from the local ethics committee. All methods were carried out following the related policies, regulations and guidelines provided by The Cancer Genome Atlas, Gene Expression Omnibus and other databases for analysis of their data and presentation of the related findings.

The hallmark gene sets, including 612 specific gene sets, comprised significantly expressed signature genes obtained from Molecular Signatures Database (version 7.4) gene sets to show perfectly defined biological states or processes. Gene set enrichment analysis (GSEA) was carried out using GSEA software (v4.1.0) to establish whether a fixed set of genes exhibited a statistically significant difference between certain phenotypes[12]. We acquired nine EMT-related gene sets: SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_DN, GOBP_CARDIAC_EPITHELIAL_TO_MESENCHYMAL_TRANSITION, GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION, GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION, GOBP_POSITIVE_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION, GOBP_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION, GOBP_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION, GOBP_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION, GOBP_REGULATION_OF_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_INVOLVED_IN_ENDOCARDIAL_CUSHION_FORMATION.

ION_FORMATION, HOLLERN_EMT_BREAST_TUMOR_DN, JECHLINGER_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_DN AND SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP, that were obviously enriched with a normalized P value of < 0.05 (Table 1, Figure 1). Then, 612 EMT-associated genes were identified by filtering according to GSEA.

Construction and validation of a new EMT-associated gene signature

To identify differentially expressed genes (DEGs) between cancer tissues and paracancerous tissues, we utilized the “limma” package in the R environment. The criteria used to identify DEGs in The Cancer Genome Atlas-LUAD cohort were $\log_2FC \geq 1$ and false discovery rate (FDR) < 0.001 . Univariate Cox regression analysis was carried out to evaluate OS-related genes (P values < 0.05). To predict correlations among DEGs, we constructed a PPI network using the online STRING database (version 11.0, <https://string-db.org/>). To reduce the risk of overfitting, we utilized the least absolute shrinkage and selection operator (LASSO) Cox regression model (“glmnet” R package), and the minimum criteria were selected [13,14]. The independent variable in the LASSO regression analysis was the standardized expression matrix of candidate prognostic DEGs associated with EMT, and the dependent variable was the OS (or other survival) status of LUAD patients in The Cancer Genome Atlas cohort. The optimal penalty parameter lambda (λ) and the matching coefficients were confirmed by 10-fold cross-validation according to the minimum criteria in the LASSO Cox regression algorithm. Lambda (λ) was the minimum partial likelihood deviance that was used for feature selection. The linear combination of gene expression values weighted by the corresponding regression coefficients was used to calculate the risk scores of LUAD patients according to the following formula: Risk score = $e^{\sum (\text{expression value of each gene} \times \text{corresponding regression coefficient})}$. LUAD patients were classified into a low-risk group and a high-risk group according to the median risk score. Principal component analysis (PCA) was carried out with the “prcomp” function in the R package “stats”. The t-distributed stochastic neighbor embedding (t-SNE) method in the R package “Rtsne” was performed for cluster visualization. For survival analysis based on each gene, the “surv_cutpoint” function in the “survminer” package was used to identify the best cutoff value of the risk score. The R package “survivalROC” was used to assess the predictive prognostic performance of the EMT-associated gene signature via time-dependent receiver operator characteristic curve (ROC) curve estimation.

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses based on the prognostic DEGs (P value < 0.05 , FDR < 0.05) between the high-risk group and the low-risk group were carried out using the R package “clusterProfiler”. Our GO functional enrichment analysis covered three ontologies: biological process, cellular component and molecular function. KEGG pathway enrichment analysis based on a hypergeometric algorithm was carried out to determine the significant KEGG pathways, and the results were visualized with the “ggplot2” package in the R environment. Adjusted P values were calculated by the Benjamini-Hochberg method.

Construction of a predictive nomogram

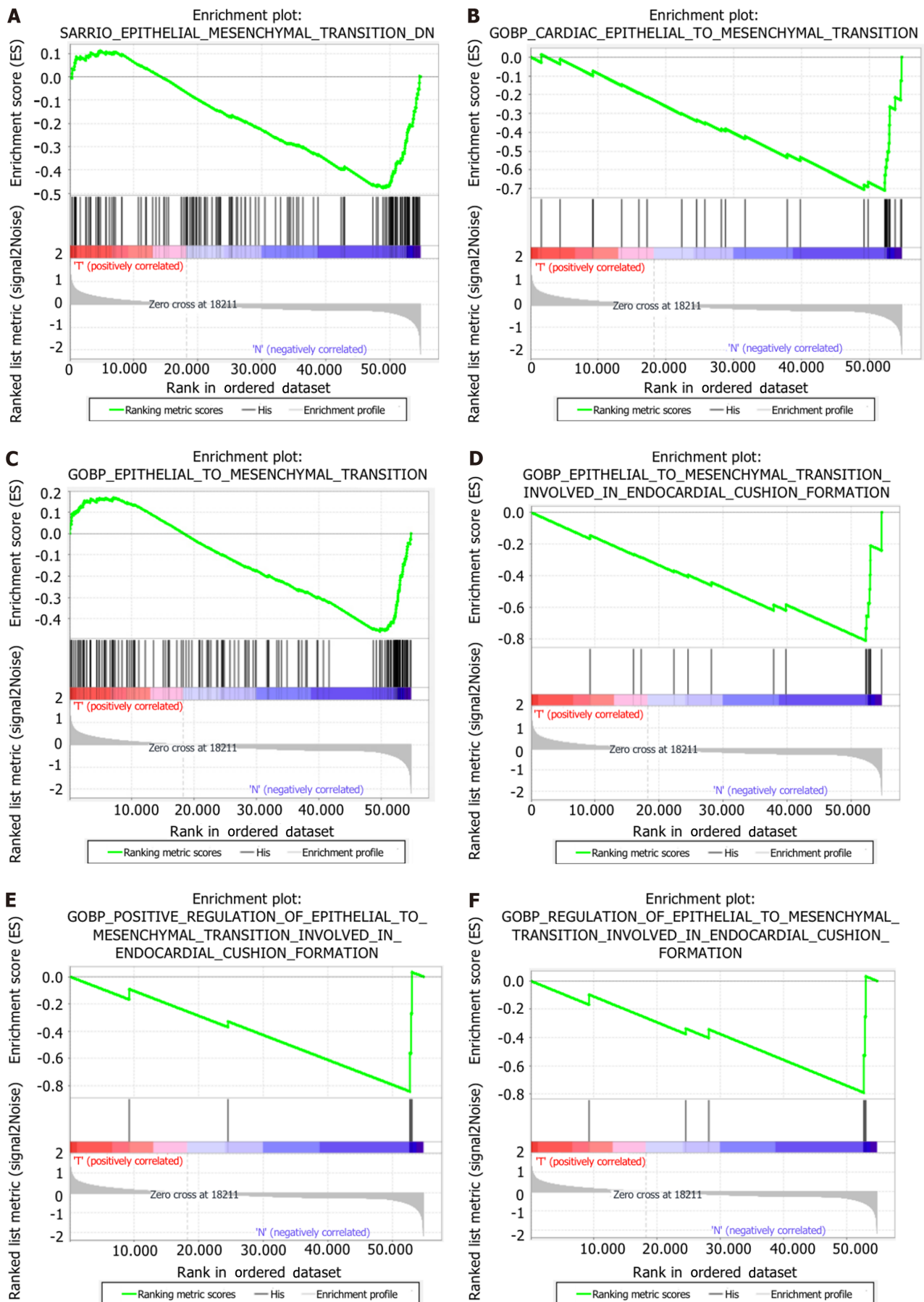
R software was used to develop a novel nomogram model containing predictors (age, risk score, and TNM stage) related to OS at 1, 2 and 3 years. Calibration curves were generated to estimate the agreement of the model between the actual outcomes and predicted outcomes for one-year OS, two-year OS and three-year OS.

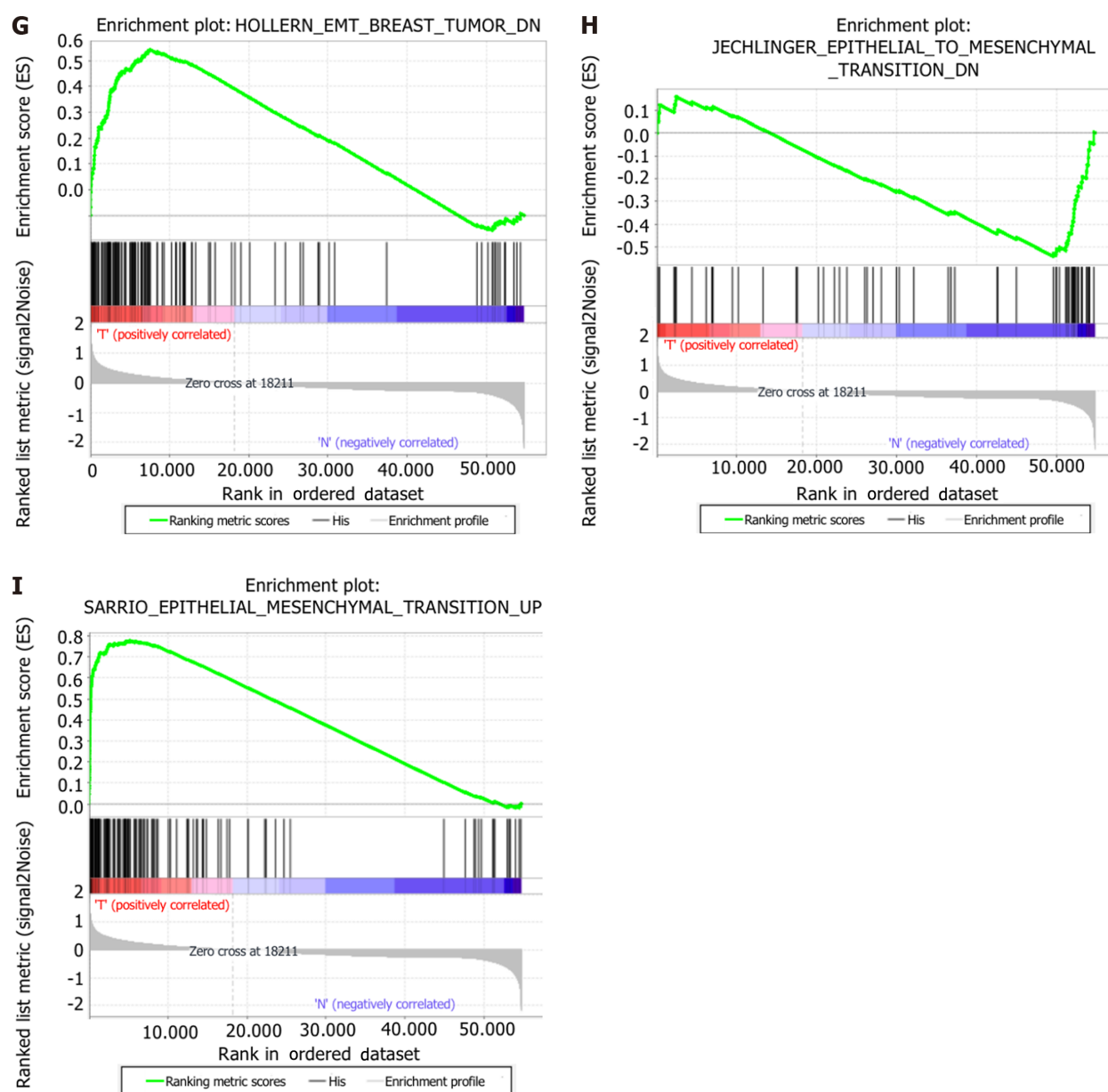
Identifying potential therapeutic targets

We uploaded the prognostic DEGs with a P value of < 0.05 to CMap (<https://portals.broadinstitute.org/cmap>), which is a biological application database of small molecule drugs, gene expression profiles, and diseases. The CMap database was built based on DEGs in human cells treated with small molecule drugs. The enrichment score based on similarity was ultimately calculated. A positive correlation score showed that a small molecule drug increased the risk of death in OS patients. In contrast, a negative correlation score showed that a drug reduced the risk of death. Drugs with negative connectivity scores had a potential therapeutic application. We obtained the three-dimensional structures of the candidate small molecule therapeutic drugs from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>).

Statistical analysis

Student’s t test was performed to identify differences in gene expression between cancer tissues and adjacent paracancerous tissues. Differences between proportions were evaluated by the χ^2 test. Overall survival was evaluated by Kaplan-Meier survival analysis, and the survival of the groups was compared using the log-rank test. Univariate and multivariate Cox regression analyses were carried out to explore independent prognostic factors of OS in LUAD patients. All statistical analyses were performed using SPSS software (version: 23.0, IBM Corporation, 2015, United States) and R software. All P values were two-tailed, and statistical significance was defined as $P < 0.05$.





DOI: 10.12998/wjcc.v10.i26.9285 Copyright ©The Author(s) 2022.

Figure 1 Enrichment plots of the gene set enrichment analysis results of 9 gene sets. GSEA: Gene set enrichment analysis.

RESULTS

The overall workflow of our research in light of model development and subsequent analyses is visualized in Figure 2. All 445 patients with LUAD from The Cancer Genome Atlas cohort and 439 patients from the gene set enrichment analysis cohort (GSE68465) were finally combined. The specific clinical baseline data of these participants are presented in Table 2.

Screening for prognostic EMT-associated DEGs

EMT-associated genes (42/612, 6.86%) were differentially expressed between cancer tissues and adjacent paracancerous tissues, and univariate Cox regression analysis showed that there were 33 genes associated with OS in LUAD (Figure 3A). The heatmap showed that *ANXA8L1* and *INMT* were highly expressed in adjacent paracancerous tissues (Figure 3B). However, the forest plots showed that the expression of *ANXA8L1* was upregulated in cancer tissues, and the results of univariate Cox regression analysis showed that there was a statistically significant association between *ANXA8L1* gene expression and OS (Figure 3C). Therefore, *ANXA8L1* was excluded from further study. In addition, 7 genes (*UHRF1*, *INMT*, *ANLN*, *SKA3*, *SAPCD2*, *GJB2* and *CDCA5*) were not found in the gene expression omnibus cohort. Finally, a total of 25 EMT-related DEGs associated with prognosis were used in subsequent analyses (all *P* values < 0.05, Figure 3B-E). The protein-protein interaction network of these 25 genes showed that *PLK1*, *NDC80* and *BUB1B* were considered the hub genes (Figure 3D). The mutual

Table 2 Clinical baseline data of lung adenocarcinoma patients

Variable	TCGA cohort	GEO cohort
Patients (n)	445	439
Sex		
Female	245 (55.1%)	218 (49.7%)
Male	200 (44.9%)	221 (50.3%)
Age, yr		
< 65	199 (44.7%)	213 (48.5%)
≥ 65	246 (55.3%)	226 (51.5%)
TNM stage		
I	241 (54.2%)	274 (62.4%)
II	107 (24.0%)	95 (21.6%)
III	74 (16.4%)	67 (15.3%)
IV	24 (5.4%)	0
Unknown	0	3 (0.7%)
T stage		
T1	154 (34.6%)	149 (33.9%)
T2	234 (52.6%)	248 (56.5%)
T3	37 (8.3%)	28 (6.4%)
T4	17 (3.8%)	11 (2.5%)
Tx	3 (0.7%)	3 (0.7%)
Survival status		
OS time, median days	653	1418
Censored (%)	170 (38.2%)	233 (53.1%)

OS: Overall survival; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; TNM: Tumor, node and metastasis.

connections among the 25 genes are shown in [Figure 3E](#).

Development of a novel model for predicting prognosis

To assess the risk probability for each patient, LASSO Cox regression analysis was performed to develop a predictive model based on prognostic genes utilizing the expression profiles of the 25 EMT-related genes indicated above. According to the optimum value of lambda (λ), 7 possible factors (*HMMR*, *PLK1*, *ASPM*, *EXO1*, *KRT16*, *SPRR1B* and *EFNA3*) were finally confirmed. Survival analysis according to the optimum threshold expression value of each gene showed that there was a significant correlation between high expression of these genes and poor prognosis (P value < 0.05). The risk score formula was as follows: Risk score = $e^{(0.229 \times \text{expression level of HMMR} + 0.484 \times \text{expression level of PLK1} + 0.018 \times \text{expression level of ASPM} + 0.085 \times \text{expression level of EXO1} + 0.075 \times \text{expression level of KRT16} + 0.178 \times \text{expression level of SPRR1B} + 0.086 \times \text{expression level of EFNA3})}$. The LUAD patients were classified into the high-risk group ($n = 222$) or the low-risk group ($n = 223$) on the basis of the median cutoff value ([Figure 4A](#)). Patients in the high-risk group were significantly more likely to have a more advanced TNM stage ([Table 3](#)). PCA and t-SNE analyses were carried out to ascertain the difference in the distribution between the two risk groups ([Figure 4B and C](#)). The results showed that patients with higher risk scores had shorter survival times than those with lower risk scores ([Figure 4D](#)). Kaplan–Meier analysis showed that the OS of patients in the high-risk group was significantly poorer than that of patients in the low-risk group ($P = 1.335 \times 10^{-6}$, [Figure 4E](#)). Time-dependent ROC curve analysis showed that the model for predicting prognosis had favorable performance in predicting OS [area under the curve (AUC) for one-year OS = 0.638, AUC for two-year OS = 0.653, AUC for three-year OS = 0.688; [Figure 4F](#)].

Validation of the predictive model in an external cohort

Survival analysis based on the 7 EMT-associated genes in the signature proved that these genes were associated with poorer OS. To confirm the robustness of the predictive model, the LUAD patients in the

Table 3 Baseline data of the patients in the different risk groups

Variable	Derivation cohort			Validation cohort		
	High risk	Low risk	P value	High risk	Low risk	P value
Sex			0.235			0.015
Female	116	129		96	122	
Male	106	94		123	98	
Age, yr			0.742			0.416
< 65	101	98		102	111	
≥ 65	121	125		117	109	
TNM stage			0.001			0.017
I + II	159	189		174	195	
III + IV	63	34		44	23	
Unknown				1	2	
T stage			0.002			< 0.001
T1	57	97		49	100	
T2	132	102		142	106	
T3	23	14		21	7	
T4	9	8		6	5	
Tx	1	2		1	2	
Status			< 0.001			0.003
Surviving	116	159		87	119	
Non-surviving	106	64		132	101	

TNM: Tumor, node and metastasis.

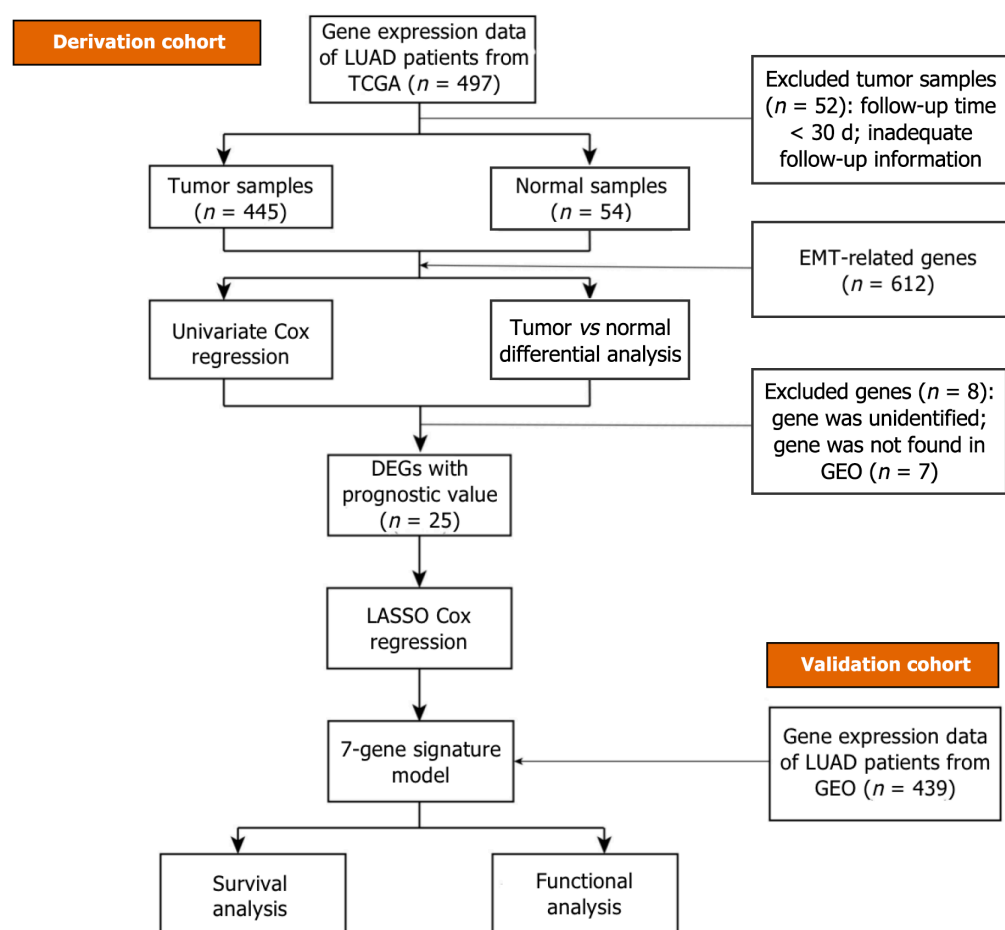
external cohort were also classified into the high-risk group and the low-risk group according to the median risk score calculated by the same formula used in the derivation cohort (Figure 5A). Patients in the high-risk group in this validation cohort were also more likely to have a more advanced TNM stage, and the evidence for this association was entirely consistent (Table 3). Moreover, the results of PCA and t-SNE analysis showed that the LUAD patients in the different risk groups were distributed in two directions (Figures 5B and C). Similarly, we found the number of nonsurviving patients was much higher in the high-risk group, revealing that patients with high risk scores might have poor survival outcomes ($P = 3.656 \times 10^{-3}$, Figure 5D and E). Additionally, the AUCs for survival in the validation model were 0.678 at one year, 0.691 at two years, and 0.651 at three years (Figure 5F).

Assessment of the independent prognostic value of the 7-gene signature

The OS-associated indicators confirmed by univariate Cox regression analysis were subsequently determined utilizing multivariate Cox regression analysis. The results of univariate Cox regression analysis showed that the risk score was considerably associated with OS in both the modeling cohort and the validation cohort (modeling cohort: HR = 3.849, 95%CI: 2.421-6.120, $P < 0.001$; validation cohort: HR = 1.441, 95%CI: 1.198-1.735, $P < 0.001$; Figure 6A and C). After controlling for confounding effects, multivariate Cox regression analysis showed that the risk score was an independent predictor for OS in LUAD patients (derivation cohort: HR = 3.136, 95%CI: 1.976-4.975, $P < 0.001$; validation cohort: HR = 1.330, 95%CI: 1.099-1.609, $P = 0.003$, respectively) (Figure 6B and D).

Functional analyses in the two cohorts

To further elucidate the biological functions and relevant pathways associated with the risk score, GO functional enrichment analysis and KEGG pathway enrichment analysis were carried out using the DEGs between the high-risk group and the low-risk group. The GO molecular function terms enriched with the DEGs in the modeling cohort and validation cohort were mainly associated with the cell cycle and cell division and included nuclear division, organelle fission, and mitotic nuclear division ($P < 0.05$, Figure 7A and C). Surprisingly, the results of GO functional enrichment analysis were completely consistent between the modeling cohort and validation cohort. The results of KEGG pathway



DOI: 10.12998/wjcc.v10.i26.9285 Copyright ©The Author(s) 2022.

Figure 2 A flow chart of the study. TCGA: The Cancer Genome Atlas; LUAD: Lung adenocarcinoma; DEGs: Differentially expressed genes; EMT: Epithelial-mesenchymal transition; GEO: Gene Expression Omnibus.

enrichment analysis showed that the cell cycle pathway was enriched in the two cohorts ($P < 0.05$, Figure 7B and D).

Development and evaluation of the prognostic nomogram

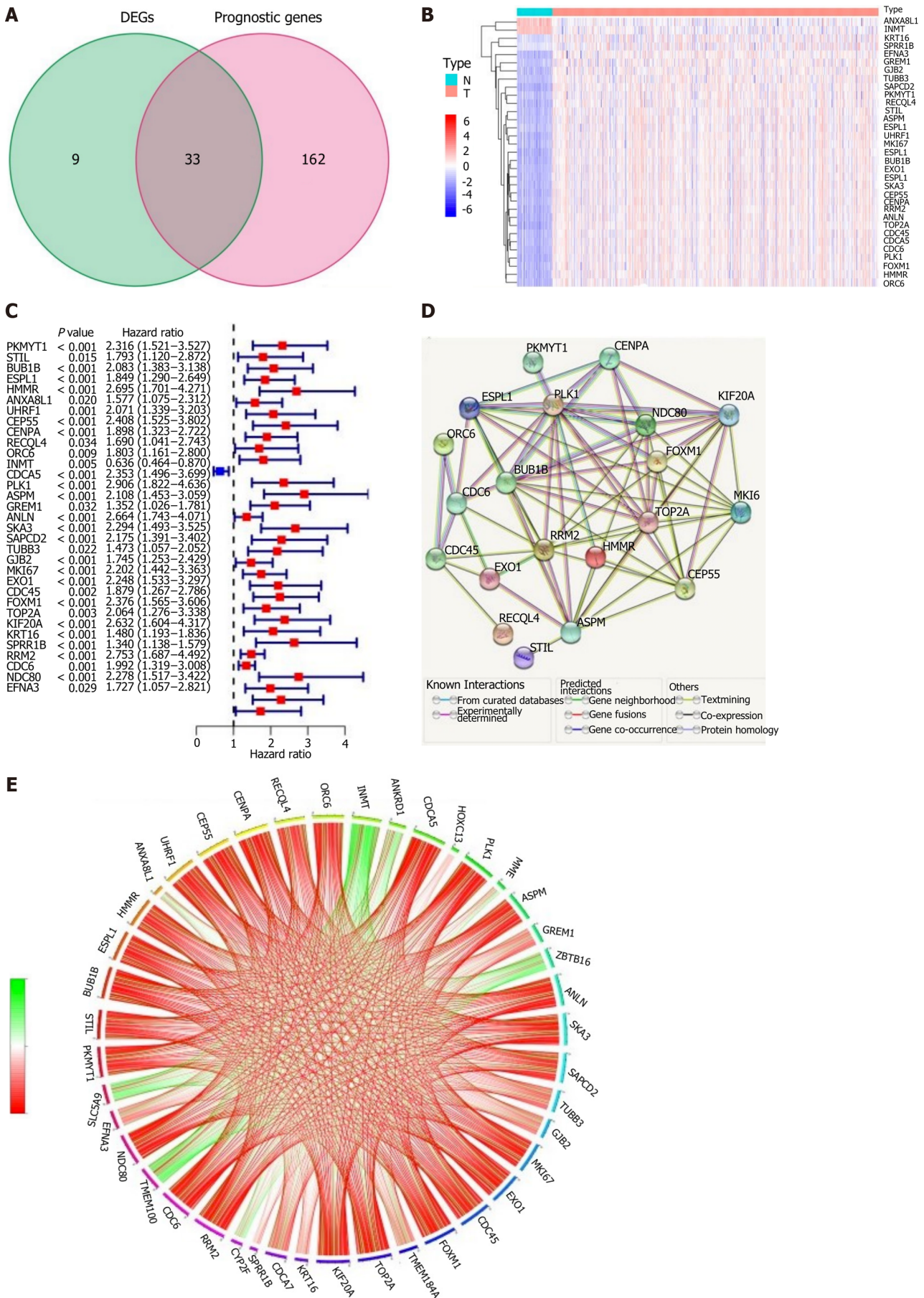
A novel prognostic nomogram including the risk score and clinical risk factors was developed to evaluate one-year, two-year, and three-year OS. Combined with clinical risk factors, the risk level in the model for predicting prognosis had significant predictive ability in the nomogram (Figure 8A). The calibration plots for the novel nomogram for predicting one-year, two-year and three-year OS are shown, and all exhibited satisfactory accuracy (Figure 8B-D).

Identification of promising therapeutic targets

We uploaded 12012 DEGs with P value < 0.05 , namely, 1186 downregulated genes and 4982 upregulated genes, into the CMap network tool. According to the screening criteria of an enrichment score < 0 and a P value < 0.05 , 66 small molecule drugs were found to have significant antitumor effects. Among these highly significantly correlated molecules, GW-8510, menadione, 6-thioguanosine, phenoxybenzamine, medrysone, chlorpromazine, 8-azaguanine, meticrane, morantel and resveratrol were among the 10 drugs with the highest negative correlations with high-risk patients (Figure 9). The three-dimensional structure of 0175029-0000 was not found in the CMap network tool; thus, this drug was not included.

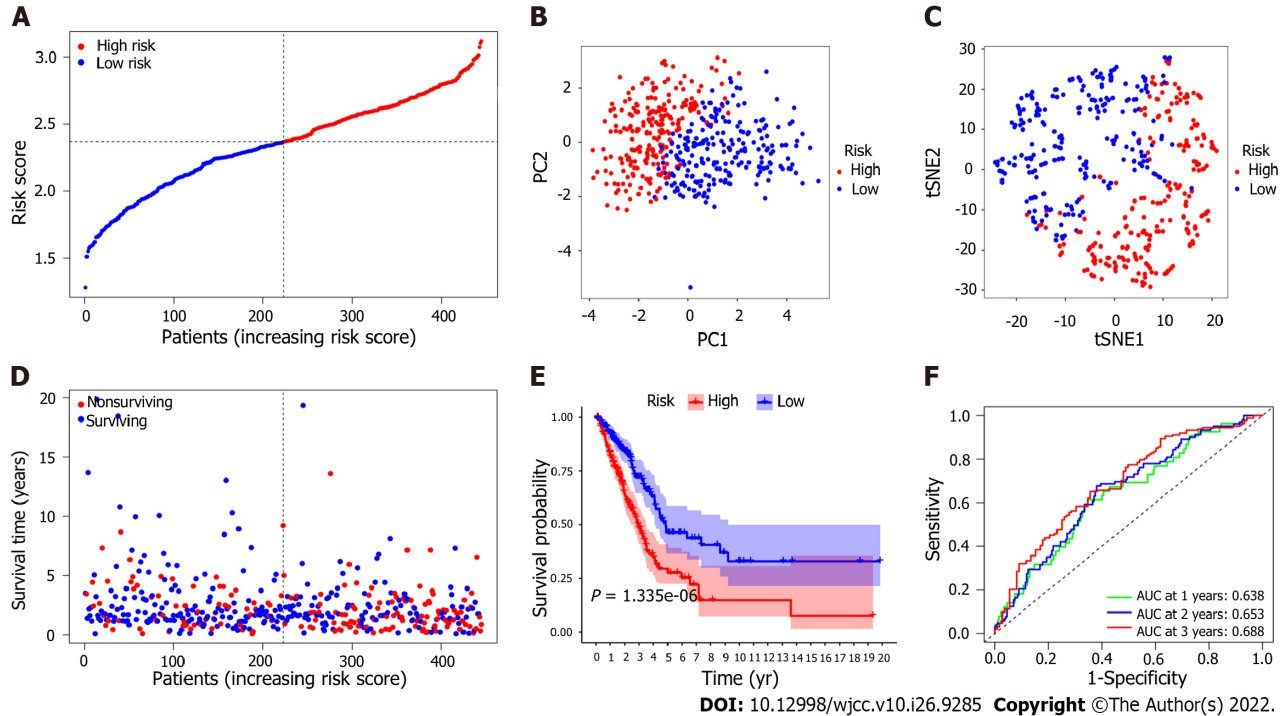
DISCUSSION

LUAD is recognized as the most common and aggressive malignant tumor with variable prognosis and is a threat to health and life[15]. Epithelial cells can lose their epithelial phenotypes, including cell polarity and connection to the basement membrane, and acquire mesenchymal phenotypes with potent migration, invasion, anti-apoptotic and extracellular matrix degradation abilities, thus acquiring



DOI: 10.12998/wjcc.v10.i26.9285 Copyright ©The Author(s) 2022.

Figure 3 Identification of the potential epithelial-mesenchymal transition-associated genes in the derivation cohort. A: Venn diagram to classify differentially expressed genes between tumor and adjacent normal tissue that were related to overall survival (OS); B: Thirty-one of the 33 overlapping genes were upregulated in cancer tissue; C: Forest plots presenting the outcomes of univariate Cox regression analysis between gene expression and OS; D: Protein-protein interaction network downloaded from the STRING database showing the interactions among the potential genes; E: The correlation network of the potential genes: correlation coefficients are shown with different colors; the number of lines indicates the correlation strength.



DOI: 10.12998/wjcc.v10.i26.9285 Copyright ©The Author(s) 2022.

Figure 4 Prognostic analysis of the 7 epithelial-mesenchymal transition-associated gene signature model in the derivation cohort. A: The median value and distribution of the risk scores. B: The results of principal component analysis (PCA) indicated that patients with lung adenocarcinoma (LUAD) were significantly distributed in two regions according to the risk score; C: The results of t-distributed stochastic neighbor embedding (t-SNE) analysis suggested that LUAD patients clustered in two different regions; D: The distributions of the overall survival (OS) status; the results showed that patients with a higher risk score had shorter survival times than those with lower risk scores; E: Kaplan-Meier analysis of OS for patients in the high-risk group and the low-risk group in the derivation cohort. F: The area under the curve (AUC) of the time-dependent receiver operator characteristic curve (ROC) verified the predictive performance of the prognostic risk score in the derivation cohort.

important biological behaviors of migration and invasion[16]. To the best of our knowledge, no prior research has explored the correlation between EMT and LUAD growth. Surprisingly, EMT-related genes were differentially expressed in tumor and normal tissues, and these genes were substantially correlated with OS (33/42, 78.57%). This suggests that EMT might have a role in LUAD prognosis and that this risk score is useful for predicting LUAD survival. Previous studies have also reported that the malignant progression of non-small cell lung cancer is determined by EMT and PD-L1 and is related to poor survival in patients with cancer[17].

GSEA is a gene set enrichment analysis approach that combines information from several sources and levels. Using expression data for 54774 genes from 551 individuals with lung cancer, we attempted to conduct GSEA in this study and discovered significant differences in 9 functions. EMT, with a statistically significant *P* value, was chosen for further study based on the normalized enrichment score, size, and *P* value. Via GSEA, we concentrated on a particular function to select genes for predicting patient survival and thoroughly investigated these genes. A combination of 7 genes with prognostic significance for patients with LUAD, rather than a single gene, was found through univariate and multivariate Cox regression analyses. This selected risk signature might be slightly more focused and have greater predictive power than certain other known prognostic biomarkers, supporting clinical outcomes and serving as a useful categorization tool for LUAD patients. Using the TCGA LUAD dataset, genes associated with EMT were identified, and their expression was compared to that in adjacent noncancerous tissues. High risk scores were linked to worse survival according to Kaplan-Meier analysis. These findings implied that identifying and determining the risk scores for LUAD patients has significant prognostic relevance and might improve the methods currently used to predict patient survival and prognosis. The risk score could also aid physicians in selecting the most appropriate course of action.

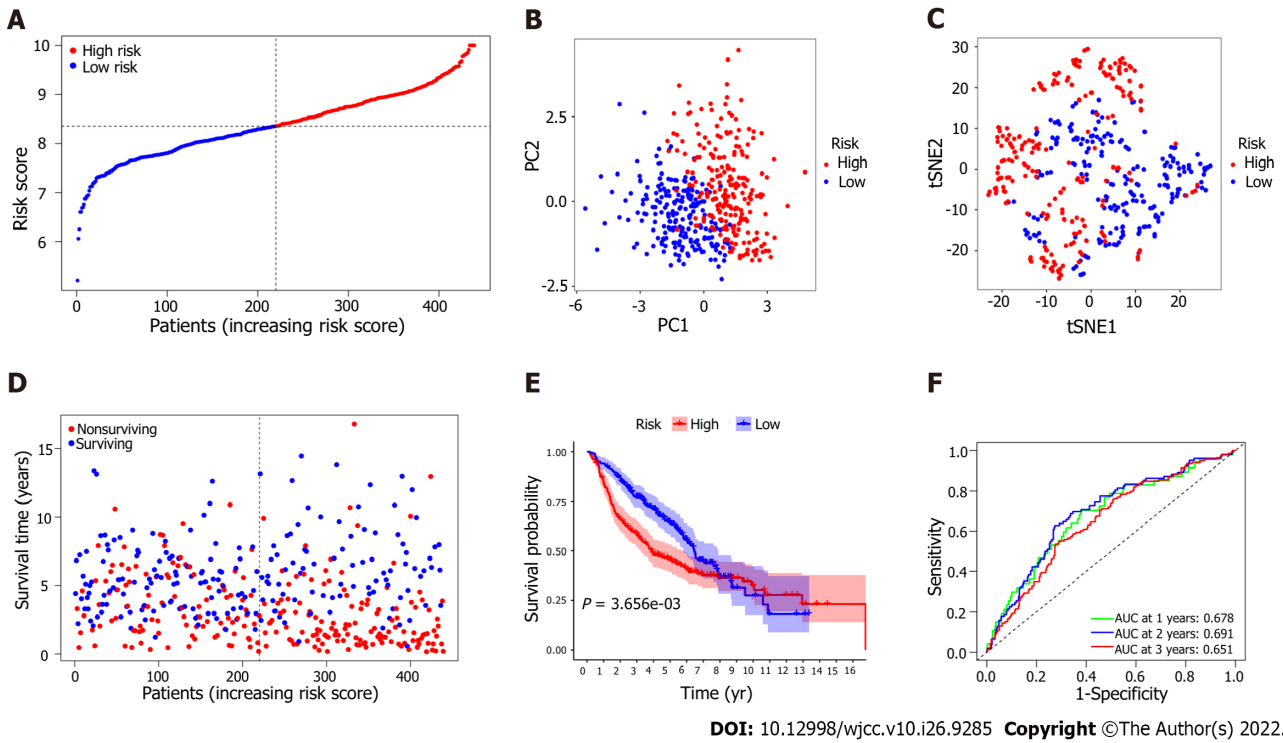


Figure 5 Validation of the 7 epithelial-mesenchymal transition-associated gene signature model in the validation cohort. A: The median value and distribution of the risk scores in the validation cohort; B: Principal component analysis (PCA) plot; C: Results of t-distributed stochastic neighbor embedding (t-SNE) analysis; D: The distributions of overall survival (OS) status for the high-risk group and the low-risk group; E: Kaplan-Meier analysis of OS for lung adenocarcinoma (LUAD) patients in the high-risk group and the low-risk group; F: The area under the curve (AUC) of the time-dependent receiver operator characteristic curve (ROC) validated the predictive performance of the prognostic risk score in the validation cohort.

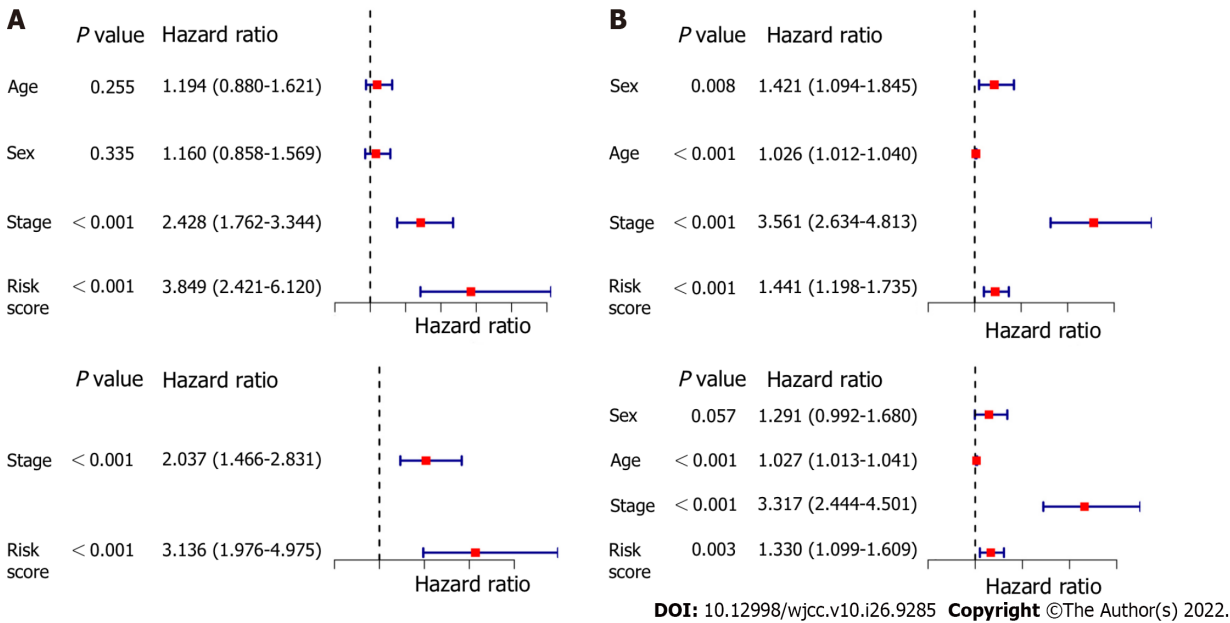
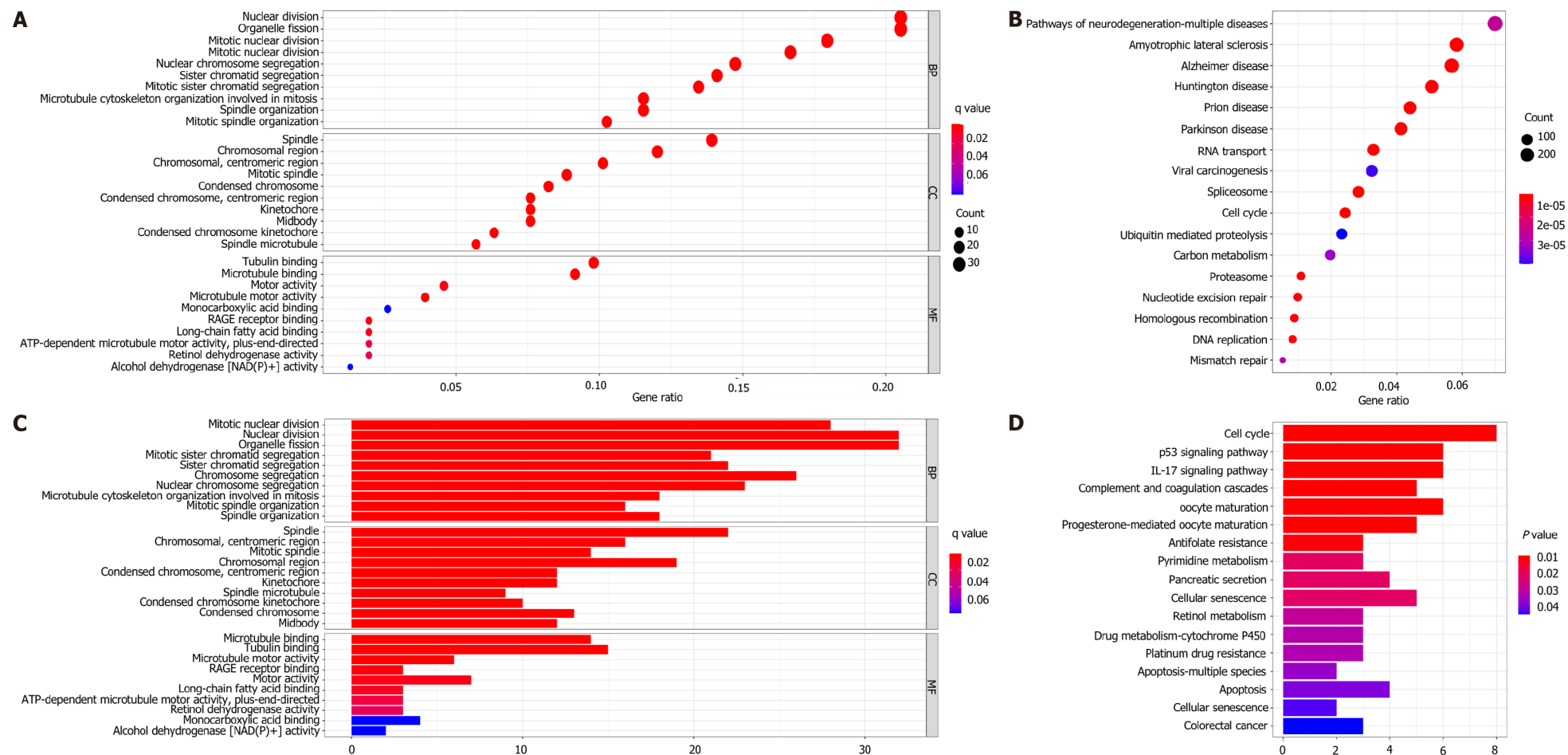


Figure 6 Results of univariate and multivariate Cox regression analyses. A: Derivation cohort; B: Validation cohort.

Stevens *et al*[18] reported that overexpression of the hyaluronic acid receptor HMMR in LUAD was related to molecular characteristics of inflammation and poor prognosis. In LUAD cells, decreasing HMMR expression decreased their ability to induce the development of lung carcinomas and distant metastases. In addition, HMMR is important for the orientation of the mitotic spindle in human mitotic cells and could regulate spindle assembly in mitotic cells and meiotic extracts[19]. Cell cycle analysis showed that in *PLK1*-silenced cells, the cell cycle was arrested at the G2/M phase transition. There was a positive connection between the expression levels of *PLK1* and *KPNB1* in LUAD cell lines. A decrease in



DOI: 10.12998/wjcc.v10.i26.9285 Copyright ©The Author(s) 2022.

Figure 7 Representative results of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. A and C: Most significant or available enriched GO terms in the derivation cohort and validation cohort; C and D: KEGG pathways in the two cohorts.

KPNB1 expression was induced by inhibition of *PLK1* and resulted in apoptosis in LUAD cells[20]. *ASPM* has been reported to be widely expressed in a variety of cancer tissues and is involved in the occurrence and progression of many malignant tumors, including lung tumors[21]. The results of Yuan *et al*[21] showed that *ASPM* depletion significantly suppressed the proliferation of lung squamous cell carcinoma cells, consistent with significant decreases in cyclin D1 and cyclin-dependent kinase 4

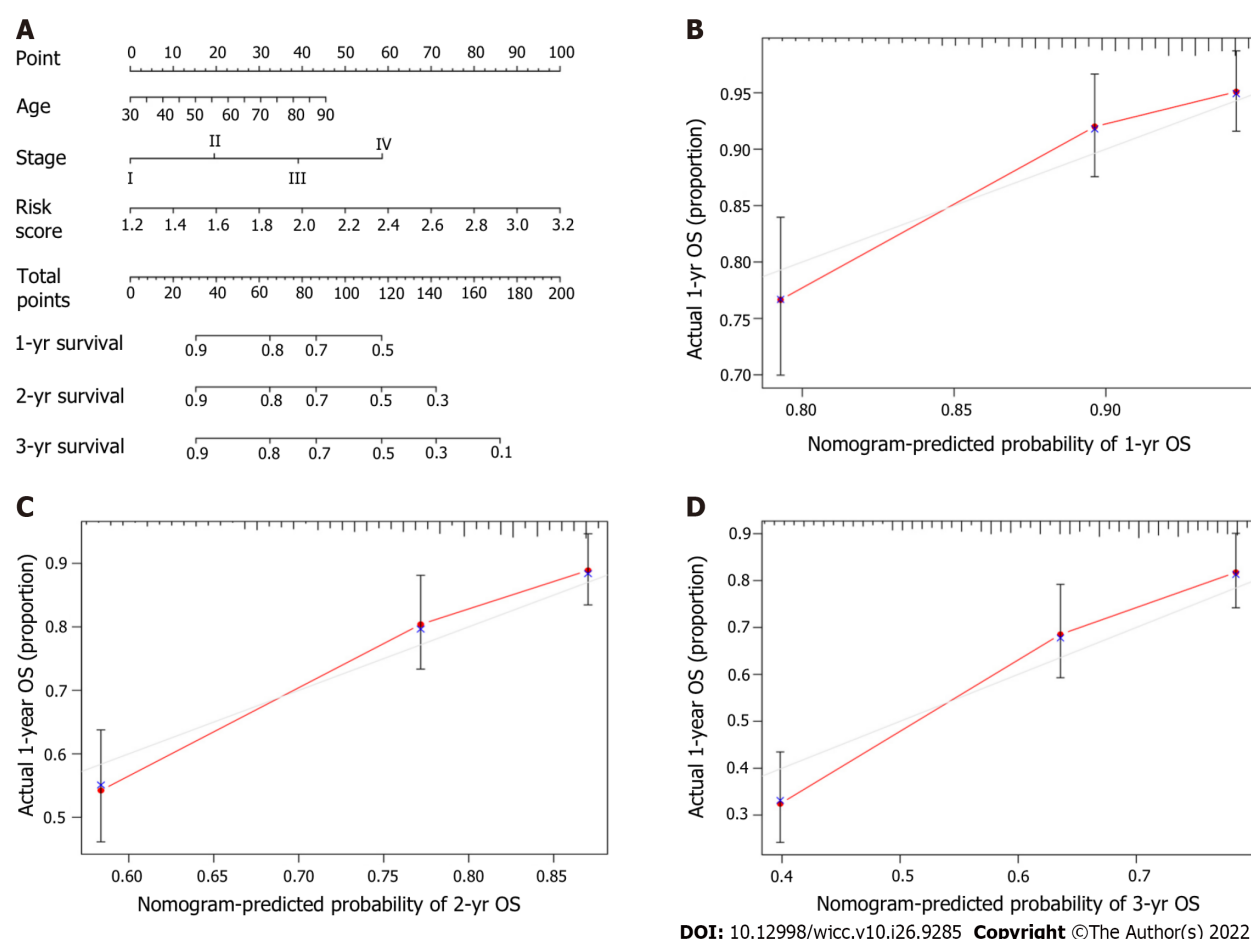


Figure 8 Development and estimation of a prognostic nomogram. A: The nomogram predicted the one-year overall survival (OS), two-year OS, and three-year OS probabilities; B: Calibration plot of the nomogram predicting the one-year OS probability; C: Calibration plot of the nomogram predicting the two-year OS probability; D: Calibration plot of the nomogram predicting the three-year OS probability.

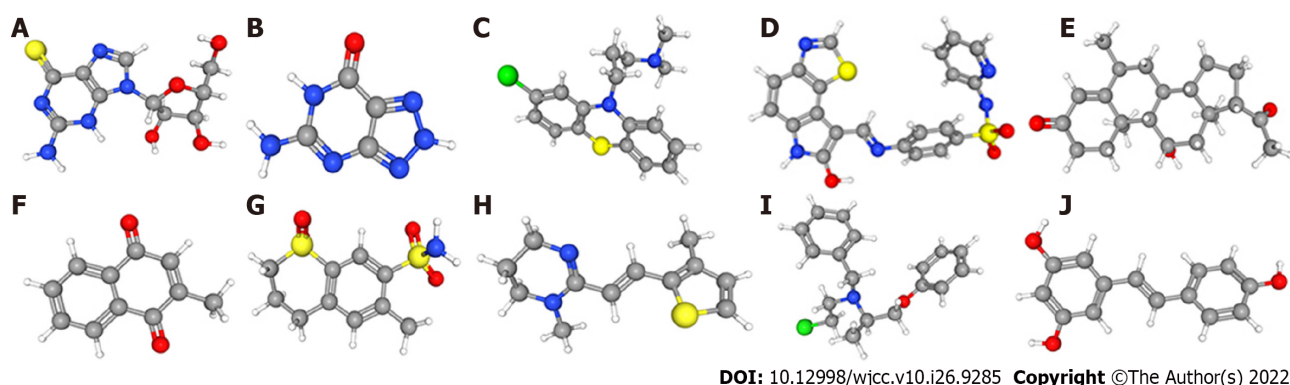


Figure 9 Three-dimensional structures of the ten most significant drugs. A: 6-Thioguanosine; B: 8-Azaguanine; C: Chlorpromazine; D: GW-8510; E: Medrysone; F: Menadione; G: Meticrane; H: Morantel; I: Phenoxybenzamine; J: Resveratrol.

expression. Liu *et al*[22] reported that *ASPM* was positively related to progression and poor prognosis in prostate and hepatocellular carcinoma, but the relationship between *ASPM* and prognosis in LUAD has not been reported. Previous reports suggested that functional polymorphisms of *EXO1* might be related to the occurrence of lung cancer and that *EXO1* might be a novel biomarker for the diagnosis and treatment of lung carcinoma[23]. The KRT family is closely associated with tumor development and metastasis in various cancers, including lung, breast, and colon cancers. Han *et al*[24] suggested that *KRT16* was overexpressed in primary melanoma and metastatic melanoma and indicated significant differences in *KRT16* expression in diverse T stages and postoperative pathological stages[25]. Downregulation of *SPRR1B* suppressed the proliferation, metastasis and invasion of LUAD cells and induced G2/M arrest *in vitro*[26]. Studies have shown that hypoxia has a significant inhibitory effect on

the increase in EFNA3 expression, but there are few studies on EFNA3 in tumor cells[27]. However, hypoxia is clearly associated with prognosis and metastasis in many cancers, including LUAD[28]. All 7 genes were upregulated in LUAD tissues and were related to poor prognosis in this study.

We performed GO functional enrichment analysis based on the differentially expressed genes between the different risk groups and found that these genes were unexpectedly enriched in many processes and cytoskeletal features involved in the cell division cycle. It was rational to hypothesize that EMT might be closely related to the cell division cycle. Interestingly, there were significant differences in the enrichment of mitotic nuclear division processes between the low-risk group and the high-risk group in our research. Wang *et al*[29] showed that Axl inhibition was particularly synergistic with antimitotic drugs in eliminating tumor cells that had undergone EMT. In particular, tumor cells stopped dividing when they underwent EMT; thus, EMT acted as a suppressor of cancer cell division and then ended proliferation[30]. Therefore, how EMT influences tumor metastasis by affecting cell cycle processes, including mitotic nuclear division, requires further study.

Although many attempts have been made to improve the prognosis of OS, the results of these changes have not been obvious. Here, we identified 10 small molecules (GW-8510, menadione, 6-thioguanosine, phenoxybenzamine, medrysone, chlorpromazine, 8-azaguanine, meticrane, morantel and resveratrol) that have potential therapeutic benefits in OS. GW-8510 is a CDK2 inhibitor that usually inhibits the activity of cyclin/CDK complexes and negatively regulates cell cycle progression[31]. Menadione, also named VK3, is a redox cycling agent that has an antitumor effect against prostate, hepatic, lung, and breast cancer[32,33]. A previous study indicated that 6-thioguanosine, which inhibits PTHrP transcription, was used in the treatment of leukemia[34]. Phenoxybenzamine, a nonselective first-generation α 1- and α 2-AR antagonist, has been administered to treat lower urinary tract symptoms [35]. Medrysone is a corticosteroid usually administered to cure eye inflammation[36]. Chlorpromazine is a traditional antipsychotic or neuroleptic that is used for the treatment of schizophrenia and other psychiatric disorders[37]. Liberante *et al*[38] showed that the concurrent targeting of RNA metabolism and splicing by 8-azaguanine offered curative prospects for *SF3B1* mutant myelodysplastic syndromes. A study found that meticrane can target the PPAR signaling pathway and inhibit the reabsorption of sodium and chloride ions in distal convoluted tubules[39]. Morantel is the classic competitive antagonist of dihydro- β -erythroidine, which noncompetitively inhibits morantel-evoked currents[40]. Resveratrol has been reported to treat a variety of cancers, including papillary thyroid cancer and liver cancer[41, 42]. Overall, GW-8510, menadione, 6-thioguanosine, phenoxybenzamine, medrysone, chlorpromazine, 8-azaguanine, meticrane, morantel and resveratrol exhibit high clinical potential and warrant further study, particularly for the improvement of OS through mechanisms affecting the cell cycle.

Undoubtedly, the current research has some limitations that should be addressed. First, this was a retrospective study for the establishment of gene signatures based on public databases. Next, we must use our own esophageal cancer specimens to confirm the expression of the seven EMT-related genes in tumor tissues and adjacent tissues, as well as their correlations with clinicopathological parameters and prognosis. Second, a well-defined EMT mechanism is lacking in lung adenocarcinoma research. The biological functions of the predicted genes were estimated by computational methods, and further study is needed to reveal their mechanisms of action in tumorigenesis. In the future, EMT-related genes could be overexpressed or silenced in animal models to study tumor growth and metastasis. Additionally, we investigated how different EMT states influence tissue canceration, invasion, and metastasis in lung adenocarcinoma. Exosomes need to be studied for their role in lung adenocarcinoma initiation, development, and EMTs. Therefore, other large external multicenter experimental studies are needed to further confirm our results.

CONCLUSION

In conclusion, we constructed and validated a 7-gene prognostic risk prediction signature based on EMT-related genes in LUAD. The risk score determined by the prognostic model of EMT-related genes was an independent risk factor for OS in LUAD. We developed a nomogram for clinicians to help predict OS in LUAD patients. According to the results of functional enrichment analysis, the pathways involved in the cell cycle and cell division were enriched in the high-risk group. In addition, we analyzed the molecular structure of potential therapeutic agents and obtained sufficient evidence indicating that EMT-related genes could be used to predict biomarkers and targeted therapies for LUAD.

ARTICLE HIGHLIGHTS

Research background

The transformation between epithelial and stromal cells during cancer progression is important in cancer progression. Understanding the relationship between epithelial-mesenchymal transition (EMT)-

related genes and lung adenocarcinoma is valuable for improving its prognosis.

Research motivation

To develop an EMT-related gene signature to predict the prognosis of lung adenocarcinoma.

Research objectives

To construct and validate prognostic polygenic signatures of differentially expressed genes associated with EMT.

Research methods

We constructed a prognostic model based on differentially expressed EMT-related genes from 445 lung adenocarcinoma samples and validated its feasibility with another 439 Lung adenocarcinoma samples.

Research results

Seven EMT-related prognostic gene signatures were developed and validated. Kaplan–Meier survival analysis showed that the overall survival of patients in the high-risk group was statistically significantly poorer than that of patients in the low-risk group. The risk score based on EMT-associated genes was an independent risk factor for overall survival.

Research conclusions

The EMT-related gene signature could be used as a feasible indicator to predict the overall survival of lung adenocarcinoma patients. The molecular structures of potential therapeutic agents associated with EMT genes that could be used to treat lung adenocarcinoma were identified.

Research perspectives

More accurate genetic markers and models are needed to predict prognosis because of the low survival rate of lung adenocarcinoma.

ACKNOWLEDGEMENTS

We thank The Cancer Genome Atlas database (<https://cancergenome.nih.gov/>) and the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) for providing us with analytical data.

FOOTNOTES

Author contributions: Zhou DH and Du QC analyzed the data and wrote the manuscript, and both contributed equally to this work; Yu H designed the study; Wang XY, Fu Z, Zhou L, Wang J, Hu CK, Liu S, Li JM and Ma ML collected the data and revised the paper. All authors have read and approved the final manuscript.

Institutional review board statement: The data for the study came from public databases and did not involve blood or tissue samples from humans or animals. Therefore, there were no ethical issues involved in this study.

Conflict-of-interest statement: The authors declare no conflicts of interest.

Data sharing statement: No additional data are available.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: China

ORCID number: De-Hua Zhou 0000-0003-2877-7746; Qian-Cheng Du 0000-0002-0154-2210; Zheng Fu 0000-0002-6373-4998; Xin-Yu Wang 0000-0001-8488-7910; Ling Zhou 0000-0003-2290-3756; Jian Wang 0000-0002-7244-4976; Cheng-Kai Hu 0000-0002-7670-5707; Shun Liu 0000-0001-7607-1567; Jun-Min Li 0000-0002-8369-7821; Meng-Li Ma 0000-0001-7120-8165; Hua Yu 0000-0002-1599-6192.

S-Editor: Zhang H

L-Editor: A

P-Editor: ChenYX

REFERENCES

- 1 **Sung H**, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; **71**: 209-249 [PMID: [33538338](#) DOI: [10.3322/caac.21660](#)]
- 2 **Thai AA**, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung cancer. *Lancet* 2021; **398**: 535-554 [PMID: [34273294](#) DOI: [10.1016/S0140-6736\(21\)00312-3](#)]
- 3 **Ferlay J**, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, Bray F. Cancer statistics for the year 2020: An overview. *Int J Cancer* 2021 [PMID: [33818764](#) DOI: [10.1002/ijc.33588](#)]
- 4 **Ramalingam SS**, Vansteenkiste J, Planchard D, Cho BC, Gray JE, Ohe Y, Zhou C, Reungwetwattana T, Cheng Y, Chewaskulyong B, Shah R, Cobo M, Lee KH, Cheema P, Tiseo M, John T, Lin MC, Imamura F, Kurata T, Todd A, Hodge R, Saggese M, Rukazenzov Y, Soria JC; FLAURA Investigators. Overall Survival with Osimertinib in Untreated, *EGFR*-Mutated Advanced NSCLC. *N Engl J Med* 2020; **382**: 41-50 [PMID: [31751012](#) DOI: [10.1056/NEJMoa1913662](#)]
- 5 **Shen J**, Zhuang W, Xu C, Jin K, Chen B, Tian D, Hiley C, Onishi H, Zhu C, Qiao G. Surgery or Non-surgical Treatment of ≤ 8 mm Non-small Cell Lung Cancer: A Population-Based Study. *Front Surg* 2021; **8**: 632561 [PMID: [34124131](#) DOI: [10.3389/fsurg.2021.632561](#)]
- 6 **Giannopoulou AF**, Velentzas AD, Anagnostopoulos AK, Agalou A, Papandreou NC, Katarachia SA, Koumoundourou DG, Konstantakou EG, Pantazopoulou VI, Delis A, Michailidi MT, Valakos D, Chatzopoulos D, Syntichaki P, Iconomidou VA, Tsitsilonis OE, Papassideri IS, Voutsinas GE, Hatzopoulos P, Thanos D, Beis D, Anastasiadou E, Tsangaris GT, Stravopodis DJ. From Proteomic Mapping to Invasion-Metastasis-Cascade Systemic Biomarkering and Targeted Drugging of Mutant BRAF-Dependent Human Cutaneous Melanomagenesis. *Cancers (Basel)* 2021; **13** [PMID: [33922182](#) DOI: [10.3390/cancers13092024](#)]
- 7 **Kryczka J**, Papiewska-Pajak I, Kowalska MA, Boncela J. Cathepsin B Is Upregulated and Mediates ECM Degradation in Colon Adenocarcinoma HT29 Cells Overexpressing Snail. *Cells* 2019; **8** [PMID: [30818851](#) DOI: [10.3390/cells8030203](#)]
- 8 **Jiang J**, Li J, Zhou X, Zhao X, Huang B, Qin Y. Exosomes Regulate the Epithelial-Mesenchymal Transition in Cancer. *Front Oncol* 2022; **12**: 864980 [PMID: [35359397](#) DOI: [10.3389/fonc.2022.864980](#)]
- 9 **Pastushenko I**, Blanpain C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol* 2019; **29**: 212-226 [PMID: [30594349](#) DOI: [10.1016/j.tcb.2018.12.001](#)]
- 10 **Qiao K**, Chen C, Liu H, Qin Y. Pinin Induces Epithelial-to-Mesenchymal Transition in Hepatocellular Carcinoma by Regulating m6A Modification. *J Oncol* 2021; **2021**: 7529164 [PMID: [34917148](#) DOI: [10.1155/2021/7529164](#)]
- 11 **Chaudhary KR**, Kinslow CJ, Cheng H, Silva JM, Yu J, Wang TJ, Hei TK, Halmos B, Cheng SK. Smurf2 inhibition enhances chemotherapy and radiation sensitivity in non-small-cell lung cancer. *Sci Rep* 2022; **12**: 10140 [PMID: [35710591](#) DOI: [10.1038/s41598-022-14448-8](#)]
- 12 **Tan Z**, Chen K, Wu W, Zhou Y, Zhu J, Wu G, Cao L, Zhang X, Guan H, Yang Y, Zhang W, Li J. Overexpression of HOXC10 promotes angiogenesis in human glioma *via* interaction with PRMT5 and upregulation of VEGFA expression. *Theranostics* 2018; **8**: 5143-5158 [PMID: [30429891](#) DOI: [10.7150/thno.27310](#)]
- 13 **Zhou ZR**, Wang WW, Li Y, Jin KR, Wang XY, Wang ZW, Chen YS, Wang SJ, Hu J, Zhang HN, Huang P, Zhao GZ, Chen XX, Li B, Zhang TS. In-depth mining of clinical data: the construction of clinical prediction model with R. *Ann Transl Med* 2019; **7**: 796 [PMID: [32042812](#) DOI: [10.21037/atm.2019.08.63](#)]
- 14 **Mathew B**, Hauptmann A, Léon J, Sillanpää MJ. NeuralLasso: Neural Networks Meet Lasso in Genomic Prediction. *Front Plant Sci* 2022; **13**: 800161 [PMID: [35574107](#) DOI: [10.3389/fpls.2022.800161](#)]
- 15 **Rong H**, Chen B, Wei X, Peng J, Ma K, Duan S, He J. Long non-coding RNA XIST expedites lung adenocarcinoma progression through upregulating MDM2 expression *via* binding to miR-363-3p. *Thorac Cancer* 2020; **11**: 659-671 [PMID: [31968395](#) DOI: [10.1111/1759-7714.13310](#)]
- 16 **Jasemi S**, Emaneini M, Fazeli MS, Ahmadinejad Z, Nomanpour B, Sadeghpour Heravi F, Sechi LA, Feizabadi MM. Toxigenic and non-toxigenic patterns I, II and III and biofilm-forming ability in *Bacteroides fragilis* strains isolated from patients diagnosed with colorectal cancer. *Gut Pathog* 2020; **12**: 28 [PMID: [32518594](#) DOI: [10.1186/s13099-020-00366-5](#)]
- 17 **Manjunath Y**, Upparahalli SV, Avella DM, Deroche CB, Kimchi ET, Staveley-O'Carroll KF, Smith CJ, Li G, Kaifi JT. PD-L1 Expression with Epithelial Mesenchymal Transition of Circulating Tumor Cells Is Associated with Poor Survival in Curatively Resected Non-Small Cell Lung Cancer. *Cancers (Basel)* 2019; **11** [PMID: [31212653](#) DOI: [10.3390/cancers11060806](#)]
- 18 **Stevens LE**, Cheung WKC, Adua SJ, Arnal-Estapé A, Zhao M, Liu Z, Brewer K, Herbst RS, Nguyen DX. Extracellular Matrix Receptor Expression in Subtypes of Lung Adenocarcinoma Potentiates Outgrowth of Micrometastases. *Cancer Res* 2017; **77**: 1905-1917 [PMID: [28196904](#) DOI: [10.1158/0008-5472.CAN-16-1978](#)]
- 19 **He Z**, Mei L, Connell M, Maxwell CA. Hyaluronan Mediated Motility Receptor (HMMR) Encodes an Evolutionarily Conserved Homeostasis, Mitosis, and Meiosis Regulator Rather than a Hyaluronan Receptor. *Cells* 2020; **9** [PMID: [32231069](#) DOI: [10.3390/cells9040819](#)]
- 20 **Sekimoto N**, Suzuki Y, Sugano S. Decreased KPNB1 Expression is Induced by PLK1 Inhibition and Leads to Apoptosis in Lung Adenocarcinoma. *J Cancer* 2017; **8**: 4125-4140 [PMID: [29187890](#) DOI: [10.7150/jca.21802](#)]
- 21 **Yuan YJ**, Sun Y, Gao R, Yin ZZ, Yuan ZY, Xu LM. Abnormal spindle-like microcephaly-associated protein (ASPM) contributes to the progression of Lung Squamous Cell Carcinoma (LSCC) by regulating CDK4. *J Cancer* 2020; **11**: 5413-5423 [PMID: [32742488](#) DOI: [10.7150/jca.39760](#)]
- 22 **Liu J**, Zhou S, Li S, Jiang Y, Wan Y, Ma X, Cheng W. Eleven genes associated with progression and prognosis of endometrial cancer (EC) identified by comprehensive bioinformatics analysis. *Cancer Cell Int* 2019; **19**: 136 [PMID: [31212653](#) DOI: [10.3390/cancers11060806](#)]

- 31139013 DOI: [10.1186/s12935-019-0859-1](https://doi.org/10.1186/s12935-019-0859-1)]
- 23 **Zhang M**, Zhao D, Yan C, Zhang L, Liang C. Associations between Nine Polymorphisms in EXO1 and Cancer Susceptibility: A Systematic Review and Meta-Analysis of 39 Case-control Studies. *Sci Rep* 2016; **6**: 29270 [PMID: [27387683](https://pubmed.ncbi.nlm.nih.gov/27387683/) DOI: [10.1038/srep29270](https://doi.org/10.1038/srep29270)]
- 24 **Han W**, Hu C, Fan ZJ, Shen GL. Transcript levels of keratin 1/5/6/14/15/16/17 as potential prognostic indicators in melanoma patients. *Sci Rep* 2021; **11**: 1023 [PMID: [33441834](https://pubmed.ncbi.nlm.nih.gov/33441834/) DOI: [10.1038/s41598-020-80336-8](https://doi.org/10.1038/s41598-020-80336-8)]
- 25 **Zhou CS**, Feng MT, Chen X, Gao Y, Chen L, Li LD, Li DH, Cao YQ. Exonuclease 1 (EXO1) is a Potential Prognostic Biomarker and Correlates with Immune Infiltrates in Lung Adenocarcinoma. *Onco Targets Ther* 2021; **14**: 1033-1048 [PMID: [33623391](https://pubmed.ncbi.nlm.nih.gov/33623391/) DOI: [10.2147/OTT.S286274](https://doi.org/10.2147/OTT.S286274)]
- 26 **Zhang Z**, Shi R, Xu S, Li Y, Zhang H, Liu M, Zhu G, Chen C, Pan Z, Liu H, Chen J. Identification of small proline-rich protein 1B (SPRR1B) as a prognostically predictive biomarker for lung adenocarcinoma by integrative bioinformatic analysis. *Thorac Cancer* 2021; **12**: 796-806 [PMID: [33501784](https://pubmed.ncbi.nlm.nih.gov/33501784/) DOI: [10.1111/1759-7714.13836](https://doi.org/10.1111/1759-7714.13836)]
- 27 **Zajkowska M**, Mroczko B. Angiopoietin-like Proteins in Colorectal Cancer-A Literature Review. *Int J Mol Sci* 2021; **22** [PMID: [34445141](https://pubmed.ncbi.nlm.nih.gov/34445141/) DOI: [10.3390/ijms22168439](https://doi.org/10.3390/ijms22168439)]
- 28 **Koren A**, Rijavec M, Krumpstar T, Kern I, Sadikov A, Čufer T, Korošec P. Gene Expression Levels of the Prolyl Hydroxylase Domain Proteins *PHD1* and *PHD2* but Not *PHD3* Are Decreased in Primary Tumours and Correlate with Poor Prognosis of Patients with Surgically Resected Non-Small-Cell Lung Cancer. *Cancers (Basel)* 2021; **13** [PMID: [34065840](https://pubmed.ncbi.nlm.nih.gov/34065840/) DOI: [10.3390/cancers13102309](https://doi.org/10.3390/cancers13102309)]
- 29 **Wang C**, Jin H, Wang N, Fan S, Wang Y, Zhang Y, Wei L, Tao X, Gu D, Zhao F, Fang J, Yao M, Qin W. Gas6/Axl Axis Contributes to Chemoresistance and Metastasis in Breast Cancer through Akt/GSK-3 β / β -catenin Signaling. *Theranostics* 2016; **6**: 1205-1219 [PMID: [27279912](https://pubmed.ncbi.nlm.nih.gov/27279912/) DOI: [10.7150/thno.15083](https://doi.org/10.7150/thno.15083)]
- 30 **Zheng X**, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, Wu CC, LeBleu VS, Kalluri R. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015; **527**: 525-530 [PMID: [26560028](https://pubmed.ncbi.nlm.nih.gov/26560028/) DOI: [10.1038/nature16064](https://doi.org/10.1038/nature16064)]
- 31 **Kuang Y**, Kang J, Li H, Liu B, Zhao X, Li L, Jin X, Li Q. Multiple functions of p21 in cancer radiotherapy. *J Cancer Res Clin Oncol* 2021; **147**: 987-1006 [PMID: [33547489](https://pubmed.ncbi.nlm.nih.gov/33547489/) DOI: [10.1007/s00432-021-03529-2](https://doi.org/10.1007/s00432-021-03529-2)]
- 32 **Ren X**, Santhosh SM, Coppo L, Ogata FT, Lu J, Holmgren A. The combination of ascorbate and menadione causes cancer cell death by oxidative stress and replicative stress. *Free Radic Biol Med* 2019; **134**: 350-358 [PMID: [30703479](https://pubmed.ncbi.nlm.nih.gov/30703479/) DOI: [10.1016/j.freeradbiomed.2019.01.037](https://doi.org/10.1016/j.freeradbiomed.2019.01.037)]
- 33 **Ivanova D**, Zhelev Z, Getsov P, Nikolova B, Aoki I, Higashi T, Bakalova R. Vitamin K: Redox-modulation, prevention of mitochondrial dysfunction and anticancer effect. *Redox Biol* 2018; **16**: 352-358 [PMID: [29597144](https://pubmed.ncbi.nlm.nih.gov/29597144/) DOI: [10.1016/j.redox.2018.03.013](https://doi.org/10.1016/j.redox.2018.03.013)]
- 34 **Gładysz M**, Andrałojć W, Czapik T, Gdaniec Z, Kierzek R. Thermodynamic and structural contributions of the 6-thioguanosine residue to helical properties of RNA. *Sci Rep* 2019; **9**: 4385 [PMID: [30867505](https://pubmed.ncbi.nlm.nih.gov/30867505/) DOI: [10.1038/s41598-019-40715-2](https://doi.org/10.1038/s41598-019-40715-2)]
- 35 **Shah PA**, Shrivastav PS. Novel LC-MS/MS method for the quantification of silodosin and its metabolites in human plasma. *Bioanalysis* 2018; **10**: 7-10 [PMID: [29236515](https://pubmed.ncbi.nlm.nih.gov/29236515/) DOI: [10.4155/bio-2017-0242](https://doi.org/10.4155/bio-2017-0242)]
- 36 **Lo YC**, Senese S, France B, Gholkar AA, Damoiseaux R, Torres JZ. Computational Cell Cycle Profiling of Cancer Cells for Prioritizing FDA-Approved Drugs with Repurposing Potential. *Sci Rep* 2017; **7**: 11261 [PMID: [28900159](https://pubmed.ncbi.nlm.nih.gov/28900159/) DOI: [10.1038/s41598-017-11508-2](https://doi.org/10.1038/s41598-017-11508-2)]
- 37 **Nikvarz N**, Vahedian M, Khalili N. Chlorpromazine vs penfluridol for schizophrenia. *Cochrane Database Syst Rev* 2017; **9**: CD011831 [PMID: [28940256](https://pubmed.ncbi.nlm.nih.gov/28940256/) DOI: [10.1002/14651858.CD011831.pub2](https://doi.org/10.1002/14651858.CD011831.pub2)]
- 38 **Liberante FG**, Lappin K, Barros EM, Vohhodina J, Grebien F, Savage KI, Mills KI. Altered splicing and cytoplasmic levels of tRNA synthetases in SF3B1-mutant myelodysplastic syndromes as a therapeutic vulnerability. *Sci Rep* 2019; **9**: 2678 [PMID: [30804405](https://pubmed.ncbi.nlm.nih.gov/30804405/) DOI: [10.1038/s41598-019-39591-7](https://doi.org/10.1038/s41598-019-39591-7)]
- 39 **Tragni V**, Cotugno P, De Grassi A, Cavalluzzi MM, Mincuzzi A, Lentini G, Sanzani SM, Ippolito A, Pierri CL. Targeting *Penicillium expansum* GMC Oxidoreductase with High Affinity Small Molecules for Reducing Patulin Production. *Biology (Basel)* 2020; **10** [PMID: [33396459](https://pubmed.ncbi.nlm.nih.gov/33396459/) DOI: [10.3390/biology10010021](https://doi.org/10.3390/biology10010021)]
- 40 **Keiser J**, Häberli C. Assessment of FDA-approved drugs against Strongyloides ratti *in vitro* and *in vivo* to identify potentially active drugs against strongyloidiasis. *Parasit Vectors* 2021; **14**: 615 [PMID: [34949209](https://pubmed.ncbi.nlm.nih.gov/34949209/) DOI: [10.1186/s13071-021-05117-2](https://doi.org/10.1186/s13071-021-05117-2)]
- 41 **Bian P**, Hu W, Liu C, Li L. Resveratrol potentiates the anti-tumor effects of rapamycin in papillary thyroid cancer: PI3K/AKT/mTOR pathway involved. *Arch Biochem Biophys* 2020; **689**: 108461 [PMID: [32531316](https://pubmed.ncbi.nlm.nih.gov/32531316/) DOI: [10.1016/j.abb.2020.108461](https://doi.org/10.1016/j.abb.2020.108461)]
- 42 **Rahman M**, Almalki WH, Afzal O, Alfawaz Altamimi AS, Kazmi I, Al-Abbasi FA, Choudhry H, Alenezi SK, Barkat MA, Beg S, Kumar V, Alhalmi A. Cationic Solid Lipid Nanoparticles of Resveratrol for Hepatocellular Carcinoma Treatment: Systematic Optimization, *in vitro* Characterization and Preclinical Investigation. *Int J Nanomedicine* 2020; **15**: 9283-9299 [PMID: [33262588](https://pubmed.ncbi.nlm.nih.gov/33262588/) DOI: [10.2147/IJN.S277545](https://doi.org/10.2147/IJN.S277545)]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

Telephone: +1-925-3991568

E-mail: bpgoffice@wjgnet.com

Help Desk: <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

