

## LETTER TO EDITOR

# The prognostic significance of CLEC-2-related gene signal in lung adenocarcinoma: A multicenter development and validation cohort

Dear Editor,

Non-small cell lung cancer (NSCLC) is one of the most common carcinomas, in terms of both incidence and mortality. The standard treatment for patients with early-stage NSCLC is radical resection, as tumor metastasis or relapse can lead to treatment failure and postoperative mortality.<sup>1</sup> CLEC-2 can promote cancer metastasis and cancer-related thrombosis.<sup>2</sup> Anti-CLEC-2 drugs are expected to prevent tumor metastasis and alleviate tumor-related symptoms, thus extending survival in cancer patients.<sup>3</sup> A model that integrates multiple CLEC-2-related cancer genes can provide better prognosis prediction accuracy than one that includes only one gene. Therefore, we used a Cox regression analysis to identify genes related to the prognosis from nine CLEC-2-related genes. We constructed a risk mode. Patients were divided into high- or low-risk group by median risk score.

We searched for CLEC-2-related genes in the GENE ONTOLOGY RESOURCE database (<http://geneontology.org/>). Six hundred and eighteen differential gene expression profiles in normal and tumor tissues in the TCGA-LUAD cohort are shown in the heat maps and volcanoes (Figures 1A and 1B). Twenty-four genes had statistically significant correlations with the TCGA-LUAD's overall survival (OS) (Figure 1C). We calculated their regression coefficients with multiple COX regression analysis. Since the UCK2 gene was not available in the GSE68465 data set, only the remaining nine genes were included in the model. Among the nine genes, seven genes have area under the curve (AUC) values greater than 0.75 for diagnosing LUAD, of which PKP2 and FGF2 have the best efficacy, greater than 0.9 (Figure 1D).

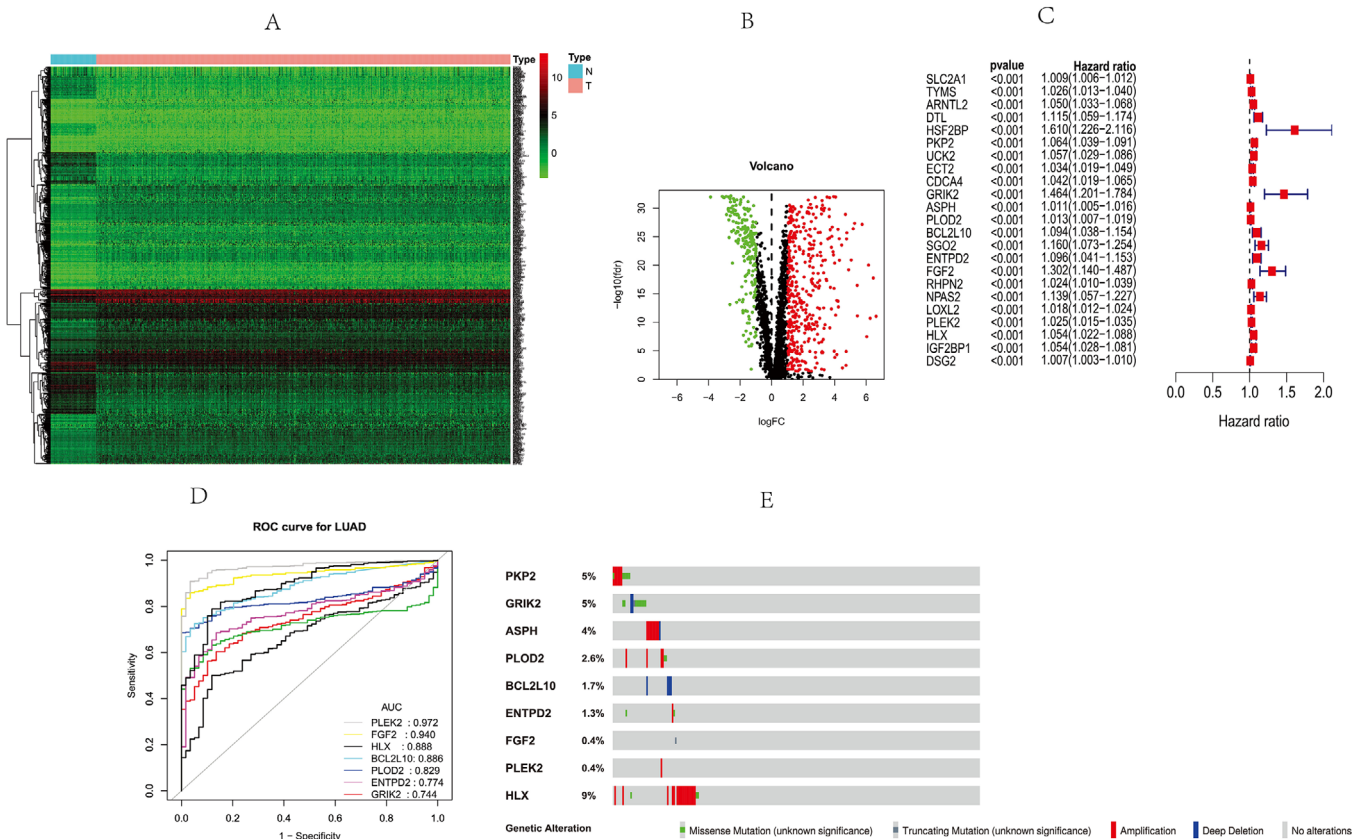
Among the 230 patients who had taken the questionnaire, the number of patients with these nine gene mutations is 54.54 (23%) (Figure 1E). Risk score =  $\text{PKP2} \times 0.039302135 + \text{GRIK2} \times 0.171236654 + \text{ASPH} \times 0.00661721 + \text{PLOD2} \times 0.00684938 + \text{BCL2L10} \times 0.0771317$

$+ \text{ENTPD2} \times 0.056883973 + \text{FGF2} \times 0.1691995447 + \text{PLEK} \times 0.011420101 + \text{HLX} \times 0.058816735$

In the TCGA-LUAD cohort, Figure 2A shows the relationship between risk scores and survival times. Univariate analysis showed a statistically significant correlation between risk scores and OS ( $P = 1.776e-05$ ). The risk score was an independent prognostic factor (HR = 2.602,  $P < .001$ ). The highest AUC for TCGA-LUAD OS was a stage (AUC = 0.721) and risk score, (AUC = 0.718), respectively. AUC values for TCGA-LUAD OS in these two queues still achieved good results (Figures 2B and 2C). We generated a nomogram to predict the likelihood of OS at the first, third, and fifth year by integrating nine CLEC-2 gene signatures, age, sex, T, N, M, and TNM stages. The calibration curve showed that the actual and predicted survival rates matched, especially those of the 3-year survival period. Moreover, we provided a free online web prediction tool that could provide personalized patient information, thus accurately predicting the individual survival prognosis (<https://zcmnewnomogram.shinyapps.io/zcm-riskscore/>). We analyzed whether our nine CLEC-2 gene markers were related to immune infiltration in LUAD, such as B cells and neutrophils. The CLEC-2-related gene features are negatively correlated with the b cells ( $r = -0.154$ ,  $P = 7.554e-04$ ); the CLEC-2-related gene features are positively correlated with the neutrophils ( $r = 0.124$ ,  $P = .007$ ). The CLEC-2-related gene features are positively correlated with the B7-H3 ( $r = 0.235$ ,  $P = 2.16e-07$ ) and PD-L1 ( $r = 0.252$ ,  $P = 2.36e-08$ ); the CLEC-2-related gene features are positively correlated with the neutrophils ( $r = 0.124$ ,  $P = .007$ ). The CLEC-2-related gene features are negatively correlated with the VEGFB ( $r = -0.151$ ,  $P = 9.239e-04$ ) and positively correlated with the MKI67 ( $r = 0.343$ ,  $P = 1.334e-14$ ). There were more TP53 gene mutations in the high-risk group. We also conducted a subgroup analysis in an independent queue (GSE 72094). It showed that high risk was meaningless in the TP53

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics



**Figure 1** Differentially expressed CLEC-2-related genes and identification of hub CLEC-2-related genes. A, Heat map of differentially expressed CLEC-2-related genes in the TCGA. B, Volcano plot of differentially expressed CLEC-2-related genes in the TCGA. C, Single-factor analysis for differentially expressed CLEC-2-related genes. D, ROC results of CLEC-2-related genes in diagnosis of lung adenocarcinoma. E, Mutation of CLEC-2-related genes in cBioportal database

mutation subgroup. However, the risk score can effectively distinguish the OS rate in their wild group. There were differential copy numbers in nine CLEC-2 gene markers with high and low risk, which get copy number in chromosomes 2, 3, 4, 7, 10, 11, 12, 15, 16, 17, 19, 20, and 22, and mitigated in 4, 5, 6, 10, 11, 12, 19, and 20. In the Human Protein Atlas (HPA) proteomics database, the proteins of the genes constituting the risk score were also different. GSEA showed the potential mechanism of the high-risk group may be mainly related to the P53 signaling pathway and cell cycle. Besides, among the other four cancers (brain lower grade glioma, bladder urothelial carcinoma, skin cutaneous melanoma, and thymoma), the prognosis of the two groups is also different according to the risk score ( $P < .05$ ).

In recent years, reliable prognostic biomarkers have emerged, and these markers are changing the prognosis. With these biomarkers, high-risk patient groups that may benefit from personalized treatment can be identified. However, the accuracy of the prognosis prediction remains unclear.<sup>4,5</sup> Now many studies have shown that CLEC-2-related genes may be associated with cancer.<sup>6-10</sup>

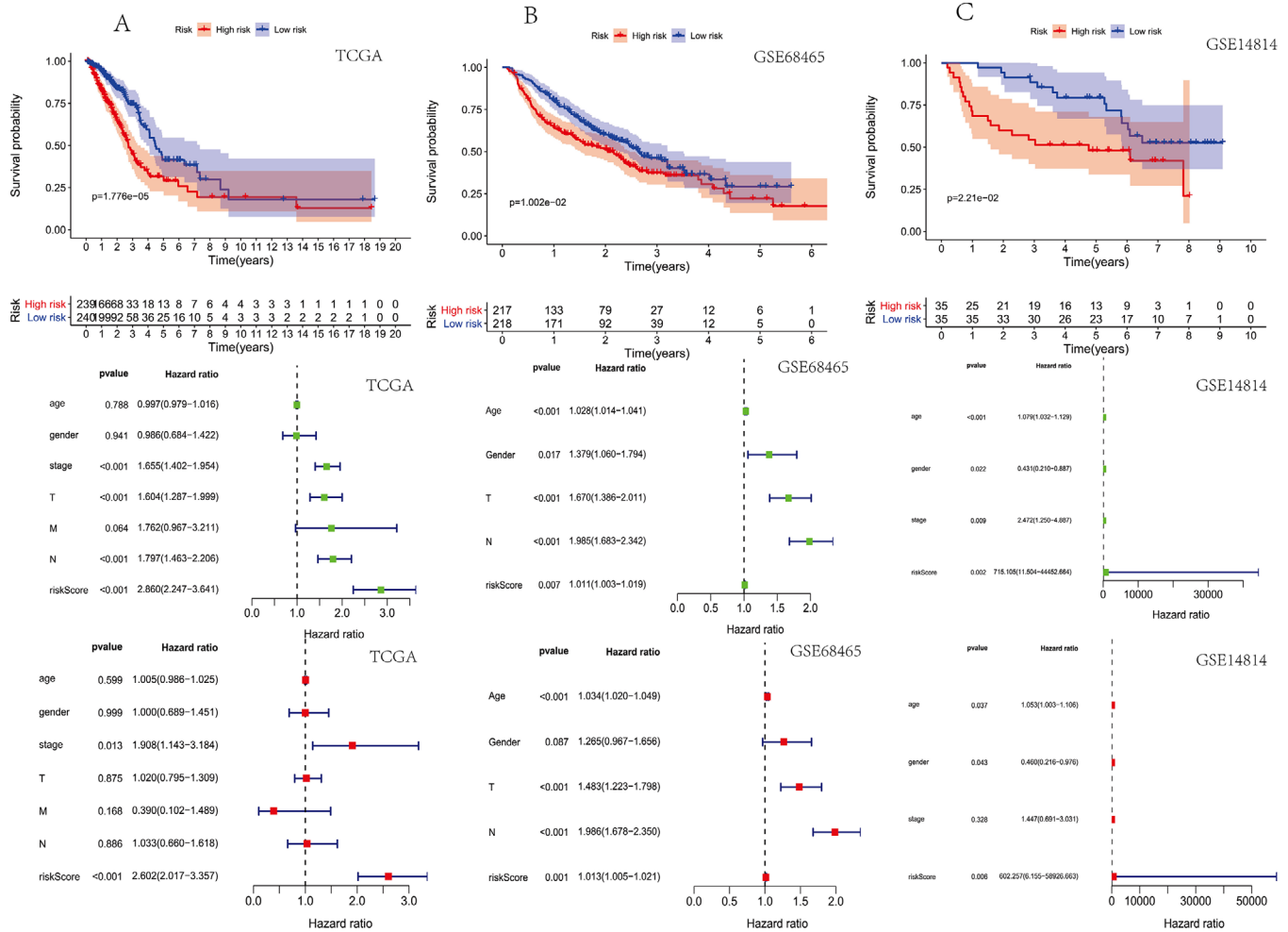
We constructed nine CLEC-2 gene signatures based on the TCGA and GEO lung cancer cohorts (GSE 68465, GSE 14814, and GSE 72094). This genetic marker is an independent prognosis predictor. The mechanism may be related to the P53 signaling pathway and cell cycle. The nomogram, which is based on genetic and clinicopathological characteristics, can accurately predict lung cancer patients' 1-, 3-, and 5-year survival rates. At the same time, we also provided an online web prediction tool for nomograms (<https://zcmnewnomogram.shinyapps.io/zcm-riskscore/>). Our findings indicate that nine CLEC-2 gene signatures may improve personalized treatment in clinical settings.

#### ACKNOWLEDGMENTS

The authors would like to thank the Geo and TCGA databases for the availability of the data.

#### AUTHORS' CONTRIBUTIONS

All authors searched the literature, designed the study, interpreted the findings, and revised the manuscript. ZCM, LL, YJJ, and XHP carried out data management and sta-



**Figure 2** The prognostic value of nine CLEC-2-related risk scores. A, Kaplan-Meier curves of overall survival based on the risk score and Single-factor and multifactor analysis in the TCGA. B, Kaplan-Meier curves of overall survival based on the risk score and Single-factor and multifactor analysis in the GSE68465 cohort. C, Kaplan-Meier curves of overall survival based on the risk score and Single-factor and multifactor analysis in the GSE14814 cohort

tistical analysis and drafted the manuscript. ZCM, LL, JMH, YJJ, and XHP helped with cohort identification and data management. ZCM, YJJ, and XHP performed project administration.

## FUNDING

This study was supported by grants from the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX19\_0113) and Jiangsu Province. Innovative and Entrepreneurial Team Grant and Southeast University-Nanjing Medical University Cooperative Research Project.

Chengmao Zhou<sup>1</sup>  
Lei Lei<sup>2</sup>  
Mu-Huo Ji<sup>2</sup>  
Jian-Jun Yang<sup>1,2</sup>  
Hongping Xia<sup>1,3</sup>

<sup>1</sup> School of Medicine, Southeast University, Nanjing, China

<sup>2</sup> Department of Anesthesiology and Perioperative Medicine, First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

<sup>3</sup> School of Basic Medical Sciences & Sir Run Run Hospital & Key Laboratory of Antibody Technique of National Health Commission, Nanjing Medical University, Nanjing, China

## Correspondence

Jian-Jun Yang, School of Medicine, Southeast University, Nanjing 210009, China.

Email: [yjyangjj@126.com](mailto:yjyangjj@126.com)

Hongping Xia, School of Medicine, Southeast University, Nanjing 210009, China.

Email: [xiahongping@njmu.edu.cn](mailto:xiahongping@njmu.edu.cn)

## REFERENCES

1. Burt ME. Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *J Thorac Cardiovasc Surg.* 1995;109:120-129.
2. Suzuki-Inoue K. Platelets and cancer-associated thrombosis: focusing on the platelet activation receptor CLEC-2 and podoplanin. *Blood.* 2019;134:1912-1918.
3. Katsue SI. Roles of the CLEC-2–podoplanin interaction in tumor progression. *Platelets.* 2018;29:786-792.
4. Shukla S, Evans JR, Malik R, Feng FY, Dhanasekaran SM, Cao X, et al. Development of a RNA-seq based prognostic signature in lung adenocarcinoma. *J Natl Cancer Inst.* 2017; 109:1-9.
5. Wistuba II, Behrens C, Lombardi F, et al. Validation of a proliferation-based expression signature as prognostic marker in early stage lung adenocarcinoma. *Clin Cancer Res.* 2013;19:6261-6271.
6. Hao XL, Tian Z, Han F, Chen JP, Gao LY, Liu JY. Plakophilin-2 accelerates cell proliferation and migration through activating EGFR signaling in lung adenocarcinoma. *Pathol Res Pract.* 2019;215:152438.
7. Ogawa K, Lin Q, Li L, et al. Prometastatic secretome trafficking via exosomes initiates pancreatic cancer pulmonary metastasis. *Cancer Lett.* 2020;481:63-75.
8. Sheng X, Li Y, Li Y, et al. PLOD2 contributes to drug resistance in laryngeal cancer by promoting cancer stem cell-like characteristics. *BMC Cancer.* 2019;19:840.
9. Du H, Chen Y, Hou X, et al. PLOD2 regulated by transcription factor FOXA1 promotes metastasis in NSCLC. *Cell Death Dis.* 2017;8:e3143.
10. Okumura Y, Noda T, Eguchi H, et al. Hypoxia-induced PLOD2 is a key regulator in epithelial-mesenchymal transition and chemoresistance in biliary tract cancer. *Ann Surg Oncol.* 2018;25:3728-3737.