A Distinct Glucose Metabolism Signature of Lung Adenocarcinoma With Prognostic Value

Ding Li†, Jiaming Liang‡, Wenzhou Zhang†, Xuan Wu‡ and Jie Fan§

1Department of Pharmacy, The Affiliated Cancer Hospital of Zhengzhou University and Henan Cancer Hospital, Zhengzhou, China, 2Department of Internal Medicine, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, 3Department of Respiratory and Critical Care Medicine, Zhengzhou University People’s Hospital, Zhengzhou, China, 4Academy of Medical Science, Zhengzhou University, Zhengzhou, China, 5Department of Head Neck and Thyroid Surgery, The Affiliated Cancer Hospital of Zhengzhou University and Henan Cancer Hospital, Zhengzhou, China

Background: Lung adenocarcinoma (LUAD) remains the most common type of lung cancer and is the main cause of cancer-related death worldwide. Reprogramming of glucose metabolism plays a crucial role in tumorigenesis and progression. However, the regulation of glucose metabolism is still being explored in LUAD. Determining the underlying clinical value of glucose metabolism will contribute in increasing clinical interventions. Our study aimed to conduct a comprehensive analysis of the landscape of glucose metabolism-related genes in LUAD and develop a prognostic risk signature.

Methods: We extracted the RNA-seq data and relevant clinical variants from The Cancer Genome Atlas (TCGA) database and identified glucose metabolism-related genes associated with the outcome by correlation analysis. To generate a prognostic signature, least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed.

Results: Finally, ten genes with expression status were identified to generate the risk signature, including FBP2, ADH6, DHDH, PRKCB, INPP5J, ABAT, HK2, GNPNAT1, PLCB3, and ACAT2. Survival analysis indicated that the patients in the high-risk group had a worse survival than those in the low-risk group, which is consistent with the results in validated cohorts. And receiver operating characteristic (ROC) curve analysis further validated the prognostic value and predictive performance of the signature. In addition, the two risk groups had significantly different clinicopathological characteristics and immune cell infiltration status. Notably, the low-risk group is more likely to respond to immunotherapy.

Conclusion: Overall, this study systematically explored the prognostic value of glucose metabolism and generated a prognostic risk signature with favorable efficacy and accuracy, which help select candidate patients and explore potential therapeutic approaches targeting the reprogrammed glucose metabolism in LUAD.

Keywords: lung adenocarcinoma, glucose metabolism, prognosis, risk signature, biomarker
INTRODUCTION

As the most common type of lung cancer, lung adenocarcinoma (LUAD) is often diagnosed at an advanced stage with distant metastatic disease (Denisenko et al., 2018). Owing to the substantial advances in the understanding of disease biology, application of predictive biomarkers, refinements in treatment, and therapeutic strategies for LUAD patients ranged from nonselective cytotoxic chemotherapy to personalized precision medicine. Precision medicine is based on validated biomarkers to better classify patients according to probable disease risk, prognosis, and/or treatment response and assists in improving the outcome by combining biomarker measurements and clinical data to a great extent. Therefore, it is important to identify new specific biomarkers to detect more aggressive disease subgroups with poor prognosis (Yuxia et al., 2012). Although a single biomarker has been identified and progressed into the clinic, molecular biomarker panels are still in the discovery stage (Vallée et al., 2014). Biomarker panels consisting of various molecules are promising while genes, proteins, and metabolites work together to promote the development of cancer hallmarks, which could provide a more accurate prediction than a single biomarker (Luo et al., 2018; Song et al., 2021).

Glucose metabolism is reprogrammed in cancer cells to provide energy, biosynthetic precursors, and intermediates for cancer cells (Allen and Locasale, 2018). In addition, reprogrammed glucose metabolism is closely related to the clinical outcome and drug resistance (Boroughs and DeBerardinis, 2015; Faubert et al., 2017). Here, we extracted the RNA-seq profile and relevant clinical information from The Cancer Genome Atlas (TCGA) to systematically and comprehensively analyze the clinical value of glucose metabolism in LUAD. This study aims to provide a distinct signature to better classify patients with different risk scores, as well as potential biomarkers for the use of glucose metabolism and metabolic pathways as therapeutic targets for LUAD.

MATERIALS AND METHODS

Data Collection

The RNA-seq profiles and relevant clinical information were acquired from the University of California, Santa Cruz (UCSC) Xena Browser (https://xenabrowser.net/) on 23 October 2021. The samples with missing clinical information and overall survival (OS) less than 30 days were excluded, and a total of 492 samples were included in the analysis. The other LUAD cohorts, GSE30219, GSE31210, and GSE50081, were downloaded from Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/).

Estimation of the Glucose Pathways

Fifteen glucose metabolism-related pathways comprising 356 genes were acquired from Molecular Signature Database v7.1 (MSigDB) (http://www.broad.mit.edu/gsea/msigdb/). Single sample gene set enrichment analysis (ssGSEA) from the R package “GSVA” was conducted to determine the activity of each glucose metabolic pathway in LUAD (Yi et al., 2020).

Generation of a Prognostic Risk Signature

The least absolute shrinkage and selection operator (LASSO) removes coefficients that become zero from the signature by adding a penalty equal to the absolute value of some coefficient magnitudes. Thus, a signature with few coefficients could be created. We randomly split the TCGA LUAD cohort (n = 492) into a training (n = 368) and testing dataset (n = 124) in a ratio of 7:3. A survival analysis for the 356 genes was conducted to select the candidate genes to construct the prognosis signature with p < 0.05 based on the log-rank test. Then, LASSO Cox regression analysis was performed with the candidate gene expression profiles from the training dataset to reduce coefficients using the R package “glmnet” (Friedman et al., 2010). Multivariate Cox analysis was followed to identify the most robust markers for the construction of the risk score signature, which included ten genes. The risk score of each sample was calculated as the following formula:

\[ \text{Risk Score} = 0.293387631789007\times \text{GNPNAT1} + 0.270363599212865\times \text{PLCB3} + 0.21736762334296\times \text{ACAT2} + 0.16127906670295\times \text{HK2} + 0.116014046444865\times \text{ADH6} + (-0.234392324167846)\times \text{PRKCB} + (-0.125128962864713)\times \text{ABAT} + (-0.114088018724961)\times \text{DHDH} + (-0.057340883024442)\times \text{FBP2} \]

Prediction of the Immune Response

The response of each sample to anti-PD-1/PD-L1 and anti-CTLA4 immunotherapy was evaluated using the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm according to the gene expression profiles of the LUAD cohort.

Evaluation of Immune Cell Infiltration

Gene Set Variation Analysis (GSVA), as shown by the R package “GSVA,” carried out a non-parametric unsupervised way to evaluate the underlying pathway activity based on gene expression profiles (Hänzelmann et al., 2013). The marker gene set, consisting of 782 genes that represent 28 immune cell types, was used to assess immune cell infiltration in the tumor microenvironment. The ssGSEA algorithm was performed to estimate the infiltration level of each immune cell type based on the expression profiles (Yoshihara et al., 2013).

Construction and Evaluation of Nomogram

We constructed a nomogram based on the clinical stage, T stage, and the signature score using the R package “rms.” To assess the application of the nomogram, the R package “ROCSurvival” was performed to construct ROC curves to predict the 1-, 3-, and 5-year OS by the nomogram. The R package “rms” was used to construct calibration curves to assess the accuracy for the prediction of 1-, 3-, and 5-year OS prediction (Li et al., 2021).

Survival Analysis

The risk score for each sample was used to assess the association between the prognosis of LUAD patients and...
the risk signature. A Kaplan–Meier curve and log-rank test were performed to compare the differences in OS outcomes between the two risk groups. $p < 0.05$ was set as the significance value. The log-rank test was performed using the R package "survival" while "surviminer" was performed to plot Kaplan–Meier curves (Zeng et al., 2019).

### Statistical Analysis
Student’s t-tests were performed to determine statistical significance among variables. $p < 0.05$ was defined as statistical significance. All statistical analysis was performed in the R version 4.0.2.

### RESULTS

#### Construction of Glucose Metabolism-Related Genes’ Prognostic Signature

Through univariate Cox regression analysis, 77 genes significantly associated with prognosis were identified from the 356 glucose metabolism-related genes ($p < 0.05$) (Figure 1A). To eliminate collinearity of the variables and avoid over-fitting of the prognostic model, these 77 genes underwent the LASSO regression analysis in the training dataset. Subsequently, 20 candidate genes were identified for further multivariate Cox regression analysis (Figure 1B). Finally, the risk signature was constructed according to the expression levels of ten genes (FBP2, ADH6, DHDH, PRKCB, INPP5J, ABAT, HK2, GNPNAT1, PLCB3, and ACAT2) (Figure 1C).

The risk score of each sample was calculated with the above formula defined by expression levels of the signature genes and regression coefficients. And, the samples were assigned to high-risk groups and low-risk groups by median risk score both in the training and testing datasets. The scatter plot showed that the high-risk group was associated with a higher mortality rate than the low-risk group (Figures 2A,B). Kaplan–Meier curves indicated that the high-risk group has significantly poor outcomes compared with the low-risk group (Figures 2C,D). To evaluate the predictive performance of the signature, we performed a time-dependent receiver operating characteristic (ROC) curve based on the risk score. The area under the curves (AUCs) of the 1-, 3-, and 5-year OS were 0.751, 0.731, and 0.648 in the training dataset, and 0.739, 0.628, and 0.614 in the testing dataset, respectively (Figures 2E,F). The results showed the signature displayed great specificity and sensitivity in predicting the prognosis of LUAD patients in TCGA cohort.

#### Validation of Glucose Metabolism-Related Genes’ Prognostic Signature Using the GEO Dataset

To validate the predictive reliability of the signature, we calculated the risk scores of samples in the GEO LUAD cohort using the same formula and similarly classified the samples into high-risk and low-risk groups, and the high-risk
group had a higher mortality rate than the low-risk group (Figures 3A–C). As expected, in the GEO database, the high-risk group tended to have a significantly shorter survival time than the low-risk group (Figures 3D–F). Above all, these results showed that the signature had robust and stable predictive power for the LUAD cohort.
Evaluation of the Signature in Different Subgroups of LUAD Patients

Stratified analysis was carried out according to the clinical variables including age (Figures 4A,B), gender (Figures 4C, D), tumor stage (Figures 4E,F), and TNM stage (Figures 4G–L). Kaplan–Meier curve analyses showed that the high-risk group had a worse survival outcome than the low-risk group when stratified by the different clinical features, except for M1, probably because of the small sample size of M1 patients (n = 12).

Construction and Validation of the Nomogram

To explore the potential value of the signature in clinical practice, we constructed a nomogram based on the risk score and clinical variables to predict the 1-, 3-, and 5-year survival rates through univariate and multivariate Cox regression analysis. Univariate Cox regression analysis demonstrated that risk score, tumor stage, and TNM stage were significantly associated with the survival of LUAD patients (Figure 5A). Multivariate Cox regression analysis showed that the risk score was an independent prognostic factor for LUAD patients after adjusting for these clinical parameters, although tumor and T stage were also independent (Figure 5B). Then we constructed the nomogram with the risk score, tumor, and T stage for their independent prognostic ability and clinical accessibility (Figure 5C). Calibration plots revealed that the nomogram showed perfect concordance between the observed and predicted survival rates at 1-, 3-, and 5-years (Figures 5D–F). The time-dependent ROC curves demonstrated that the nomogram had excellent predictive accuracy in predicting the 1-, 3-, and 5-year survival of LUAD patients. The AUCs for 1-, 3-, and 5-year survival was 0.762, 0.752, and 0.669, which indicated that the nomogram has robust and stable ability to predict the survival of LUAD patients (Figure 5G).

Correlation Between Immune Cell Infiltration and Risk Score

To explore the potential correlation of the signature with the immune microenvironment, we performed the CIBERSORT algorithm to evaluate the infiltrating level of immune cells in the tumor microenvironment and made comprehensive comparisons with the risk score. The results showed that the proportions of 28 immune cell types were significantly different between the two risk groups, and the low-risk group tended to have significantly higher infiltrating levels of the most immune cell types than the high-risk group, which may represent an intrinsic feature that can characterize individual differences (Figure 6A).

Furthermore, we also evaluated the difference in the response rate of immunotherapy between the two risk groups. Samples in the low-risk group had a higher response rate to immunotherapy than those in the high-risk group (Figure 6B). The aforementioned results indicated that signature could predict...
the immune cell infiltration level and the response to immunotherapy in LUAD.

**Gene Set Enrichment Analysis**

Given that risk scores were inversely associated with prognosis in patients with LUAD, further functional annotation was performed between the two risk groups using GSEA. The result showed that enriched gene sets of the HALLMARK collection in the high-risk group were mainly involved in tumor-related pathways, including E2F, G2/M checkpoint, glycolysis, mTORC1, MYC, oxidative phosphorylation, and unfolded protein response, which are closely related to the malignant proliferation of tumor cells (Figure 7).

**DISCUSSION**

LUAD is the most common histological subtype of NSCLC, often with the presence of specific genetic mutations for further molecular stratification (Xiong et al., 2020). Since the patients at an early stage could have a favorable prognosis, most patients have developed distant metastasis at the first time of diagnosis, with poor survival (Denisenko et al., 2018). Risk stratification is important to assess the prognosis of patients, which may promote the development of new strategies for LUAD management. Furthermore, prognostic prediction plays an important role in treatment selection and the identification of potential prognostic biomarkers (Wang et al., 2020; Yi et al., 2021).

Tumorigenesis and progression are primely required for metabolic reprogramming in cancer cells (Taubes 2012). Cancer cells could alter their fluxes via various metabolic pathways to meet increased biosynthetic and bioenergetic demands and alleviate oxidative stress required for cancer cell proliferation and survival (Boroughs and DeBerardinis, 2015). In recent years, there has been a growing interest in developing cancer genetic analysis for patient stratification in combination with therapies that target metabolism (Hay 2016). Although it is well known that metabolic reprogramming is a hallmark of cancer, regulation of
glucose metabolism in LUAD is still being explored, and identifying the underlying clinical value of glucose metabolism in LUAD phenotype may contribute to increased clinical interventions. In addition, there is an urgent need to identify new strategies for patient stratification with easier access to gene abnormality detection in cancers, which will promote the efficiency and velocity of translation from basic research to clinical practice (Wettersten et al., 2017; Qin et al., 2020). However, studies regarding transcriptome-wide analysis on the correlation between glucose metabolism and LUAD are limited. We evaluated the correlation between glucose metabolism-related pathways and clinical characteristics as well as the immune phenotype in LUAD. The ssGSEA was conducted to calculate the enrichment score of each gene set regulating glucose metabolism-related pathways, and the results showed that the citrate cycle (TCA cycle) pathway had the highest score, whereas the enrichment score of the ascorbate and aldarate metabolism pathways are the lowest (Supplementary Figure S1). To better understand the clinical significance of the glucose metabolism-related pathways in LUAD, we compared the discrepancies in the pathways between different subgroups of LUAD. The result showed that the samples with the N2-N3 stage had a
significantly higher enrichment score in glyoxylate and dicarboxylate metabolism than that of the ones with the N0-N1 stage ($p < 0.05$), whereas there is no significant difference in the pathways in the subgroups stratified by T and M stage (Supplementary Figures S2–S4). In addition, the citrate cycle (TCA cycle), glyoxylate and dicarboxylate, and pentose phosphate metabolism pathways have a significantly elevated enrichment score in tumor stage III-IV LUAD samples compared with tumor stage I-II LUAD samples ($p < 0.05$) (Supplementary Figure S5). The results demonstrated that the specific glucose metabolism pathway was significantly associated with the specific subgroup of LUAD patients.

Here, we first introduce a glucose metabolism-related prognosis signature for the malignancy of LUAD and the survival of LUAD patients. From 356 glucose metabolism-related genes involved in 15 pathways, we finally included ten genes, of which their expressions were significantly associated with prognosis, to construct a risk signature. The prognostic risk signature showed great predictive ability both in the training and testing datasets and was an independent indicator for the prognosis of LUAD patients.

Furthermore, we also evaluated the distribution trends of glucose metabolism-related pathways between the two risk groups in the
TCGA database (Supplementary Figure S6). It can be seen that among the 15 pathways, the ascorbate and aldarate metabolism pathway, citrate cycle (TCA cycle) pathway, fructose and mannose metabolism pathway, galactose metabolism pathway, glyoxylate and dicarboxylate metabolism pathway, pentose and glucuronate interconversions pathway, pentose phosphate pathway, and starch and sucrose metabolism pathway increased with an increase in the risk score, suggesting that these pathways’ imbalances had a significantly positive correlation with tumor development. The results may provide some insight in the glucose metabolism scope of tumor development.

The most included genes in the risk signature have been reported to play important roles in tumorigenesis and progression in various cancer types, which enhance the predictive performance of the signature.

Among the ten genes, fructose-1,6-biphosphatase 2 (FBP2) has been demonstrated to inhibit glycolysis and growth in gastric cancer cells (Li et al., 2013). Alcohol dehydrogenase (ADH) had shown potential prognostic values in pancreatic adenocarcinoma and hepatocellular carcinoma (Liao et al., 2017; Liu et al., 2020). DHDH had been reported to be included in a metabolism-related prognostic signature for hepatocellular carcinoma (Yang et al., 2021). PRKCB has also been reported to be included in the prognostic signature for adult T-cell leukemia/lymphoma and prostate cancer (Kataoka et al., 2018; Daniuaitė et al., 2021). INPP5J regulates AKT1-dependent breast cancer growth and metastasis and predicts recurrence in lung adenocarcinoma (Ooms et al., 2015; Zhang et al., 2020). ABAT and HK2 have been reported to play crucial roles in cancer metabolism, progression, and therapeutic resistance of cancers (Jansen et al., 2015; Garcia et al., 2019; Shen et al., 2020). GNPAT1 and PLCB3 had shown the independent prognostic potential in NSCLC (Zhang et al., 2019; Zheng et al., 2020). ACAT2 could promote cell proliferation and associated with malignant progression in colorectal cancer (Weng et al., 2020). The evidence mentioned earlier demonstrated that these included signature genes might play vital roles in cancer, and their roles in LUAD should be further explored.

Previous studies have demonstrated that immune cell infiltration and immune checkpoints are correlated with the response rate of immunotherapy in LUAD (Bodor et al., 2020). We assessed the correlations between the risk signature and immune cell infiltration. The proportions of 28 immune cell types in the tumor microenvironment were significantly different between the two risk groups, and the low-risk group tended to have significantly higher infiltrating levels of the most immune cell types than the high-risk group. Notably, the glucose metabolism-related signature was significantly correlated to CD4+ and CD8+ T cells, and the samples in the high-risk group tended to have a lower number of CD8+ T cells and a higher number of CD4+ T cells. The result indicated that patients of higher risk tend to have an unfavorable tumor-infiltrating lymphocyte pattern. Moreover, the signature was also significantly associated with innate immune cell types, including macrophages, monocytes, and NK cells, which is consistent with the results of previous research that showed tryptophan metabolic adaptation in lung cancer was related to evasion of innate immune by cancer cells (Cassetta and Pollard, 2018; Dejima et al., 2021). Since immune checkpoint inhibitors have shown promising anti-tumor effects by reversing the immunosuppressive effects of tumors, the expression of immune checkpoints has attracted widespread attention as a biomarker for identifying patients with LUAD to receive immunotherapy. Immune checkpoints could be used to predict the efficacy of immune checkpoint blockade and have been proven to be a biomarker for identifying patients who can benefit from immunotherapy in several cancer types. In this study, we analyzed the association between the signature genes and immune checkpoints. The expression of the ten signature genes was significantly associated with the expression of the four checkpoint markers, PD-1, PD-L1, PD-L2, and CTLA-4 (Supplementary Figure S7). The findings showed that the risk signature based on glucose metabolism-related genes was involved in the altered immune microenvironment of LUAD.

Overall, we constructed a risk signature based on the glucose metabolism-related genes for the prognosis, malignancy, and immune phenotype of LUAD, which might provide a better understanding of the glucose metabolic role in immune phenotype and carcinogenesis. This study also suggested that glucose metabolism could be a potential target and that the glycolytic inhibitor combined with immunotherapy maybe a novel strategy for LUAD treatment.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AUTHOR CONTRIBUTIONS

XW, DL, WZ, and JF designed the research. JL analyzed the data. DL and XW wrote the paper with contributions from all the authors. All authors have read and approved the manuscript.

FUNDING

This study was supported by the Henan Provincial Science and Technology Research Project (202102310157) and Medical Science and Technology Research Plan (Joint Construction) Project of Henan Province (LHGJ20190676).

ACKNOWLEDGMENTS

We are grateful to the contributors of the public databases used in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.860677/full#supplementary-material

Supplementary Figure 1 | Enrichment score of each gene set regulating glucose metabolism-related pathway.
Yuxia, M., Zhennan, T., and Wei, Z. (2012). Circulating miR-125b Is a Novel Biomarker for Screening Non-Small-Cell Lung Cancer and Predicts Poor Prognosis. J. Cancer Res. Clin. Oncol. 138, 2045–2050. doi:10.1007/s00432-012-1285-0

Zeng, L., Fan, X., Wang, X., Deng, H., Zhang, K., Zhang, X., et al. (2019). Bioinformatics Analysis Based on Multiple Databases Identifies Hub Genes Associated with Hepatocellular Carcinoma. Cg 20, 349–361. doi:10.2174/1389202920666191011092410

Zhang, T., Song, X., Liao, X., Wang, X., Zhu, G., Yang, C., et al. (2019). Distinct Prognostic Values of Phospholipase C Beta Family Members for Non-Small Cell Lung Carcinoma. Biomed. Res. Int. 2019, 4256524. doi:10.1155/2019/4256524

Zhang, Y., Fan, Q., Guo, Y., and Zhu, K. (2020). Eight-Gene Signature Predicts Recurrence in Lung Adenocarcinoma. Cbm 28, 447–457. doi:10.3233/cbm-190329

Zheng, X., Li, Y., Ma, C., Zhang, J., Zhang, Y., Fu, Z., et al. (2020). Independent Prognostic Potential of GNPNAT1 in Lung Adenocarcinoma. Biomed. Res. Int. 2020, 8851437. doi:10.1155/2020/8851437

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, Liang, Zhang, Wu and Fan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.