A Novel Fatty Acid Metabolism Signature Predicts Prognoses, Tumor Immune Microenvironment, and Immunotherapy Response In Lung Adenocarcinoma

Denggang Fu  
Medical University of South Carolina

Biyu Zhang  
Wuhan Institute of Technology

Wenyan Fan  
Jiujiang University

Xin Wang ( wangxin0072@126.com )  
Jiujiang University

Research Article

Keywords: lung adenocarcinoma, fatty acid metabolism, prognosis, immune microenvironment, immunotherapy

Posted Date: January 16th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2468569/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

**Background:** Lung adenocarcinoma (LUAD) is the most common and aggressive subtype of non-small cell lung cancer. Aberrant fatty acid metabolism (FAM) has been demonstrated to play an essential role in the tumorigenesis of human cancers, yet limited studies in LUAD.

**Methods:** The RNA-sequencing dataset of LUAD patients with clinical features from the TCGA database was used as the training set. Six independent LUAD cohorts totaling 1,368 encompassing diverse platforms from the GEO database were employed as validation sets. The prognostic signature was constructed by multivariate Cox regression analysis with the Akaike information criterion. The tumor immune microenvironment (TIME) was analyzed by ESTIMATE and infiltrated immune cell subsets were calculated using multiple deconvolution algorithms. Tumor characteristics such as T cell receptors richness and diversity, and tumor mutation burden (TMB) were assessed. The implication of the signature in predicting immunotherapy response was also investigated.

**Results:** Overall survival (OS) related FAMGs were identified. A robust prognostic signature for OS prediction was developed. Patients were divided into high- and low-risk groups and decreased OS was observed in low-risk patients. Furthermore, the signature could be an independent prognostic indicator after adjusting for clinicopathological features. Receiver operating characteristic curve analysis indicated the validity of the signature. The predictive power was validated using six LUAD validation cohorts. The signature also has strong risk stratification utility for patients’ disease relapse. TIME analysis showed increased immune activity in low-risk patients, which was convinced by higher infiltrated CD8+ T, natural killer, and B cells, as well as lower tumor purity, stemness index, TMB, and cell proliferation. Additionally, elevated activated and less senescence of immune cells were observed in low-risk patients. Differentially expressed pathways that related to resistance to immune checkpoint blockades such as DNA repair, hypoxia, cell cycle, epithelial-mesenchymal-transition, and oxidative phosphorylation were enriched in high-risk patients. T cell receptor richness and diversity were higher in low-risk patients. Responders had lower risk scores in contrast to non-responders for LUAD patients receiving anti-PD-1 treatment.

**Conclusions:** The study was the first time to establish a novel FAMGs-based signature in recognition of the prognosis for LUAD patients and evaluation of the possibility of immunotherapy response in personalized treatment.

Introduction

Lung cancer including non-small cell lung cancer (NSCLC) and small lung cancer (SLC), ranks as the leading cause of cancer-associated deaths worldwide. Lung adenocarcinoma (LUAD) is the most prevalent histological subtype, accounting for 50% – 70% of all NSCLC cases\(^1\). Remarkable advancement in prevention and therapeutics has been made, while most patients are diagnosed at advanced stages, leading to lower 5-year overall survival. Many endeavors have been taken to develop novel therapies including molecular targeted therapy, immunotherapy, and combination therapy apart from surgery,
radio/chemotherapy, and chemotherapy for NSCLC patients\textsuperscript{2;3}. Immune checkpoint inhibitors targeting PD-1/PD-L1, and CTLA-4 have made remarkable forward in various human malignancies, while most patients are reported to obtain a short-term complete or partial remission following treatment, the majority of patients eventually raised drug resistance and succumb to tumor recurrence\textsuperscript{4;5}. Increasing evidence witnessed that the heterogeneity of tumor microenvironment (TME) among patients, especially the variations in the infiltration of immune cells in the niche, and the interactions of tumor and host are the main determinants of responsiveness to treatment\textsuperscript{6;7}. Emerging signatures have been proposed to predict the responsiveness to immunotherapy, such as tumor mutation burden (TMB)\textsuperscript{8;9} and immune checkpoint molecule expression\textsuperscript{10–12}, while these tools are insufficient to characterize the landscape of the TME heterogeneity. This posed an unmet need to identify additional reliable biomarkers for predicting therapeutic efficacy.

Lipid metabolism abnormality is a hallmark in diverse cancers, and perturbed metabolism enabled cancer cells to acquire a rapid proliferative rate by over-activating endogenous lipid synthesis or increasing the uptake of exogenous lipids or lipoproteins\textsuperscript{13–15}. Cellular fatty acids (FAs) are involved in many biological processes including being incorporated into membrane structure, energy storage, signaling macromolecules, and oxidized into carbon dioxide for energy production. Tremendous evidence showed that FAs metabolic reprogramming in cancer has an integral role in tumorigenesis including lung cancer\textsuperscript{16}. Aberrantly activated enzymes involved in the metabolism of FAs such as ATP citrate lyase (ACLY), fatty acid synthase (FASN), and acetyl-CoA carboxylase (ACC) in normal cells accelerated cancerous transformation in lung cancer\textsuperscript{16}. Endogenous FAs metabolisms were found to be reversely correlated with EGFR expression\textsuperscript{17} and promoted epithelial-mesenchymal-transition (EMT) regulation\textsuperscript{18}, which contributes to the invasive and metastatic capacity of lung cancer cells. Multiple molecules mediate the activity of FAs metabolism such as sterol regulatory element binding proteins (SREBPs) and FASN. High FASN expression was closely associated with decreased survival and increased cell proliferation and invasion in lung cancer\textsuperscript{19}. Inhibition of the availability and metabolism of FAs represented a promising therapeutic strategy by blocking FAs synthesis, increasing oxidative degradation, and limiting FAs release from storage\textsuperscript{20}. However, the regulatory mechanism of the FA metabolism process in LUAD has not been well interrogated. Therefore, it will provide new perspectives in understanding tumor heterogeneity and novel targets for patients with LUAD by analyzing FAMGs.

This study elucidated the prognostic utility of FAMGs in LUAD. A robust risk signature based on FAMGs was established in the TCGA cohort and prediction power was validated in six independent LUAD cohorts. Signature-related features including tumor immune microenvironment (TIME) landscape, differentially expressed pathways, and immunotherapy predictive potential were delineated. Our study may shed insight into developing FAMGs-related targeted therapy in the treatment of LUAD in further in-depth work.

**Materials And Methods**

**Lung adenocarcinoma datasets collection and pre-processing**
The processed RNA-seq gene expression profile (FPKM normalized) of 535 lung adenocarcinoma (LUAD) and 59 normal samples were downloaded from the UCSC Xena database (https://xenabrowser.net/datapages) for identifying survival-related FAMGs and constructing prognostic signatures. Six independent LUAD microarrays with available clinical features including GSE30219 (n=85), GSE31210 (n=226), GSE50081 (n=127), GSE68465 (n=442), GSE72094 (n=398), and GSE11969 (n=90), were also obtained from GEO database for external validation. These publicly available datasets had ethical approval in their original studies. The details for these datasets were listed in Table S1.

**Construction and validation of FAMGs-associated prognostic signature**

The FAMGs were retrieved from the MSigDB database. To investigate the prognostic roles of these FAMGs, OS-related FAMGs were identified using univariate Cox regression analysis. GO terms enrichment analysis including biological process (BP), molecular function (MF), and cellular components (CC) of these OS-related FAMGs were conducted using clusterProfiler package\(^\text{21}\). The KEGG pathways were interrogated as well.

These OS-related FAMGs were employed to develop a prognostic signature using multivariate Cox stepwise regression analysis. The minimum number of features that comprised the signature was determined by both directions of stepwise selection and the Akaike information criterion (AIC)\(^\text{22}\). Patients were divided into low- and high-risk groups based on the median risk score. The difference in OS was compared using the log-rank test and visualized Kaplan-Meier curve.

Subset analysis was conducted to calculate the predictive utility of the signature for patients with different clinical features such as age, gender, and tumor stage.

To validate the prognostic capability of the signature, six independent LUAD datasets with overall survival information were used.

The signature predictive capacity was assessed by the receiver operating characteristic (ROC) curve using the survivalROC package.

**Predictive Utility of Signature for Patients’ Relapse**

To further assess the potential of signature in predicting patient’ relapse or disease progression, the available information on disease-free survival (DFS) or first progression (FP) was collected from TCGA (n=858), GSE30219 (n=85), GSE31210 (n=226), and GSE50081 (n=124). Patients were divided into low- and high-risk groups according to the median risk score. The survival of patients in high- and low-risk groups was compared using the log-rank test.

**Tumor Immune Microenvironment Analysis**

To characterize the signature associated TIME, the immune and stromal cell infiltration was assessed using the ESTIMATE algorithm by “estimate” package. Then, the proportions of immune cell infiltration in
high- and low-risk groups were dissected using multiple cell deconvolution algorithms including CIBERSORT\textsuperscript{23}, xCell\textsuperscript{24}, TMIER\textsuperscript{25}, and EPIC\textsuperscript{26}. In addition, the expression levels of stimulatory, inhibitory checkpoint molecules, and immune senescence markers were also compared in patients within high- and low-risk groups.

**Immunotherapy Response Prediction**

To predict the responsiveness to immune checkpoint blockade, we analyzed PD-1/PD-L1 mRNA and protein expression, Tumor mutation burden (TMB) mutant-allele tumor heterogeneity (MTH), and TCR repertoire. Gene mutation profiles of LUAD patients were downloaded from the TCGA database and TMB was calculated using “maftools” package\textsuperscript{27}. TMB was calculated as the number of somatic indels and base substitutions per million bases in the coding region of the detected genome\textsuperscript{28}. The richness and Shannon diversity indexes, which were retrieved from the Pan-Cancer Atlas study\textsuperscript{29}, were used to characterize the diversity of the TCR repertoire. The richness measures the number of unique TCRs in the sample, while the Shannon diversity index reflects the relative abundance of the different TCRs.

Immunophenoscore (IPS) is a measure of the overall immunogenicity of an individual tumor, which is calculated as 0 -10 scale based on the sum of the weighted averaged Z score. The Z score was identified by a random forest based on the expression of major determinants included in MHC molecules, immunomodulators, effector cells, and suppressor cells in the Cancer Immunome Atlas (TCIA, https://tcia.at/). A higher score was indicative of a better prognosis and better response to immunotherapy. To predict patients’ responsiveness to ICI therapy. The IPS of LUAD patients who respond to the anti-PD-1/anti-CTLA-4 treatment were extracted from the TCIA database. The IPS levels between immunity High and Low subtypes were compared.

We applied the signature to NSCLC patients receiving anti-PD-1 therapy (GSE126044) including 5 responders and 11 non-responders. The patient’s risk score was calculated and compared between responders and non-responders using the Wilcox test with p<0.05 as statistically significant.

**Differential expression and Gene Set Enrichment Analysis (GSEA)**

The differentially expressed genes with |Log FC| > 1 and false discovery rate (FDR) < 0.05 were identified between high- and low-risk groups by edgeR package\textsuperscript{30}. GSEA was conducted to identify the differentially expressed pathways between low- and high-risk patients using all the genes ranked by fold change. Hallmark and KEGG pathway gene sets were used as an enriched portal from MSigDB database\textsuperscript{31}. The adjusted $p$-value of less than 0.05 was considered statistically significant.

**Results**

**Identification of Overall Survival-related FAMGs**
To investigate the biological process of the FAMGs, functional enrichment analysis showed that these FAMGs were enriched in the cellular response to stimulus and metabolic-related biological processes such as steroid metabolism, bile acid, and epoxygenase P450 pathway (Figure S1A). KEGG analysis confirmed that cytochrome P450 metabolism, Wnt, steroid, Phenylalanine, and bile acid synthesis ranked as the top pathways (Figure S1B). Then, we explore the interplay of these FAMGs through PPI network analysis, and two modules (named CYP3A4, and KRT19) were identified by the number of nodes (Figure S1C). In module CY3A4, these membrane-associated CYPs are the major enzymes involved in drug metabolism. Genes in module KRT19 are the regulators in Wnt signaling which control cell proliferation and growth. This further promoted us to assess their clinical relevance in LUAD. A total of 22 FAMGs were found to be correlated with patients’ outcomes by univariate Cox regression analysis in the TCGA LUAD dataset (Figure 1A). Among these OS-related FAMGs, decreased survival was observed in patients with high expression of 11 OS-related FAMGs such as TCN1, RHOV, DKK1, and GPR115, while elevated expression of the remaining FAMGs was considered as protective factors.

Construction and Validation of The Prognostic Signature Based on FAMGs

The 22 FAMGs with prognostic significance were delivered for further analysis. Stepwise multivariate Cox proportional hazard regression analysis was employed to determine the minimum set of features that generated the most powerful prediction for patients’ prognoses. This identified 10 genes of significance (GPR115, SOAT2, CDH17, MOGAT2, COL11A1, TCN1, LGR5, SLC34A2, RHOV, and DKK1) which comprise the optimal prognostic signature (Figure 1B). The formulation was listed as follows:

$$FAScore = \left[ \text{Expression level of } GPR115 \times (-0.0573) \right] + \left[ \text{Expression level of } SOAT2 \times (-0.0987) \right] + \left[ \text{Expression level of } CDH17 \times (0.0608) \right] + \left[ \text{Expression level of } MOGAT2 \times (-0.1079) \right] + \left[ \text{Expression level of } SLC34A2 \times (-0.0651) \right] + \left[ \text{Expression level of } COL11A1 \times (0.0522) \right] + \left[ \text{Expression level of } TCN1 \times (0.0502) \right] + \left[ \text{Expression level of } LGR5 \times (-0.0663) \right] + \left[ \text{Expression level of } RHOV \times (0.0917) \right] + \left[ \text{Expression level of } DKK1 \times (0.0771) \right]$$

The FAScore for an individual patient was calculated, and patients were divided into low- and high-risk groups using median FAScore as the cutoff value. We generated a Kaplan-Meier curve and decreased survival was observed in patients within the high-risk group as compared to those patients in the low-risk group (Figure 1C, \(p=1.19e-11\)). The number of deaths was increasing along with elevated FAScore (Figure 1D-E). We noted that six model genes (GPR115, CDH17, MOGAT2, COL11A1, TCN1, and RHOV) were significantly up-regulated in LUAD patients as compared to adjacent normal tissues (Figure S1D), while two genes (SLC34A2 and LGR5) were markedly down-regulated in patients (Figure S1D). Additionally, six model genes were up-regulated in high-risk patients (Figure S1E), while the expression of the remaining four genes (SLC34A2, LGR5, MOGAT2, and SOAT2) was observed to be increased in low-risk patients (Figure S1E). The area under curve (AUC) values of the receiver operating characteristics (ROC) curve for 1-, 3-, and 5-year were 0.80, 0.67, and 0.67 (Figure 1F), indicating that the prognostic signature has a robust capacity for monitoring prognosis.

Clinicopathological features such as clinical stages and age were correlated with disease progression and patient’s prognosis. To further test its predictive independence, univariate Cox regression analysis
was performed in the TCGA LUAD dataset, and we found that FAScore, involved lymph nodes, tumor size, and clinical stages correlated with decreased survival, while patients receiving treatment have favorable survival (Figure 2A). Multivariate Cox regression analysis showed that the prognostic signature could serve as an independent predictor after adjusting for other clinical parameters such as treatment (Figure 2B). Similar observations were found in the GSE68465 LUAD dataset (Figure 2C-D). In the GSE72094 dataset, the FAScore could serve as an independent indicator for OS prediction (Figure 2E-F). In addition, patients with oncogenic driver mutations including EGFR and KRAS were associated with clinical survival, which was consistent with previous reports.

**Validation of the Prognostic Signature in Independent Cohorts**

To verify the reproducibility of the prognostic signature in LUAD, we applied this signature formula in six independent publicly available cohorts (GSE30219 (n=85), GSE31210 (n=226), GSE50081 (n=127), GSE68465 (n=442), GSE72094 (n=398), and GSE11969 (n=90)) totaling 1,368 LUAD patients to assess the predictive capability of the proposed signature for patients’ prognosis. Patients were also stratified into low- and high-risk groups, and we found that patients in the low-risk group in these six cohorts have prolonged OS than those in the high-risk group (Figure 3A-F). The predictive performance of the signature in each validation set was also calculated and showed that the AUCs of 1-, 3-, and 5-year all exceeded 0.6. Specifically, the AUC of 1-year in the GSE31210 validation set was 0.91 (Figure 3B). These data convinced that the signature has robust risk stratification in the microarray-based platform for LUAD patients (Figure 3A-F).

**Differentially expressed and Functional Enrichment Analysis**

The differentially expressed genes (DEGs) were identified using the *edgeR* package. To look at the pathways involved in the difference of the malignant characteristics between low- and high-risk groups, Gene Set Enrichment Analysis (GSEA) was performed in Hallmark and KEGG pathway gene sets. Epithelial-mesenchymal-transition (EMT), DNA repair, oxidative phosphorylation, G2M checkpoint, and hypoxia were found to be enriched in high-risk patients (Figure S2A-D, Table S2), suggesting dysregulated metabolism and cell cycle, and the malignant transition was active in patients within the high-risk group. These pathways were also observed in the KEGG gene set enrichment analysis (Figure S2E-H). Increasing reports indicated that these pathways were correlated with resistance to ICB therapy.

**Predictive Potential of The Signature for Patients’ Relapse**

Relapse has been one of the main challenges for patients receiving various types of treatment such as chemotherapy or radiotherapy. We used four LUAD data sets (TCGA, GSE30219, GSE31210, and GSE50081) with available clinical information on patients’ relapse or progression to investigate the predictive potential of the signature. We found that patients in the low-risk group have significantly favorable progression-free survival as compared to those patients in the high-risk group (Figure 4A). A similar result was also observed in the GSE30219 set (Figure 4B). As for disease relapse, patients with low FAScores in the GSE31210 set showed better survival (Figure 4C), suggesting the signature may be
an indicator for monitoring patients’ relapse or disease progression, which was verified in an independent LUAD set (GSE50081) as well (Figure 4D).

**Tumor Immune Microenvironment Analysis**

Recognition of the dual role of TIME in anti-tumor immunity has led to remarkable leaps forward in tumor immunotherapy. To delineate the TIME landscape in both risk groups, each patient was scored based on profiling of 29 immune signatures using single sample gene set enrichment analysis (ssGSEA) and found that patients in the low-risk group showed enhanced immune activities (Figure 5A). Further ESTIMATE analysis showing higher immune scores and ESTIMATE scores, and decreased tumor purity (Figure 5B-D) in patients within the low-risk group in contrast to the high-risk group demonstrated the notion (Figure 5B-C). Next, we interrogated the infiltrated immune cell subsets by multiple deconvolution algorithms. CIBERSORT analysis showed that CD8+ T cells and B cells increased in patients in the low-risk group, whereas activated CD4+ T cells, neutrophils, and mast cells were elevated (Figure 5E), which was confirmed by TIMER (Figure 5F) and EPIC (Figure 5G) infiltration analysis. In addition, xCell deconvolution also convinced that CD8+ and CD8+ central memory T cells, and microenvironment scores were higher than those of patients in the high-risk group (Figure 5H). Several co-stimulatory molecules were up-regulated in low-risk patients such as CD27/28, and ICOS (Figure 5I), while the co-inhibitory molecules were shown differential expressed in both groups, CD47, IDO2, CTLA4, and PD-1, were increased expression in low-risk patients, but PD-1/L1 and B7-H3 were decreased (Figure 5J), indicating that patients in low-risk groups might benefit from the immunotherapy targeting PD-1/CTLA4. We noted that IL2, CD160, and KLAG1, the positive regulators of immune cell proliferation and activation expressed highly in patients in the low-risk group, and H2AX, a key immune cell senescence marker, expressed higher in high-risk patients, this might be a sign of immune cell exhaustion in this group (Figure 5K).

**The Signature Correlates with Anti-tumor Immunity and Therapy Response**

Given the difference of elevated tumor-infiltrating lymphocytes (TILs) (Figure 5E-H, Figure 6A) such as cytotoxic T cells and NK cells in high- and low-risk groups and their critical roles in predicting the efficacy and treatment response. We investigated the association between the signature and widely used immunotherapy markers PD-L1 expression (Figure 5J) and tumor mutation burden (TMB) in the LUAD cohort. FAScore was positively correlated with the TMB of patients (Figure 6J). Accounting reports have revealed that the repertoire of T cell receptors (TCR), which recognize antigens presented by the major histocompatibility complexes (MHC), could serve as a predictive indicator of responsiveness to immunotherapy. We conducted the repertoire analysis of TCR and found that patients in the low-risk group exhibited higher TCR richness and diversity (Figure 6B). B cell receptor richness in low-risk patients was also increased (Figure 6C), while the BCR diversity was comparable (Figure 6D).

IFN-γ is a pleiotropic cytokine with antitumor or pro-tumorigenic roles, and TGF-β is an important cancer-promoting cytokine that contributes to the suppression of anti-tumor immunity (TGF-β and Cancer Immunotherapy). We subsequently scored the IFN-γ and TGF-β responses and found that both...
cytokine responses were enhanced in high-risk patients (Figure 6E-F). The proliferation index was accordingly markedly higher in patients with high-risk scores (Figure 6G). In addition, we found that FAScore was significantly correlated with TMB (Figure 6H). A lower aneuploidy score has been observed in patients with complete or partial responses to immune checkpoint blockade\textsuperscript{38}, and the same decreased trend was found in low-risk patients as compared to those in high-risk patients (Figure 6I). Mutant-allele tumor heterogeneity (MATH), a hallmark of cancer that is a promising biomarker for clinical outcomes and patients’ response to therapy\textsuperscript{39}, was decreased in low-risk patients in contrast to those high-risk patients (Figure 6J). Furthermore, the number of cancer stem cells (CSCs) was estimated using an mRNA expression-based stemness index and found that CSCs was decreased in low-risk patients (Figure 6K).

Assessment of responsiveness to therapy including immunotherapy is a critical challenge before treatment. These identified features supported that the signature implies predicting responsiveness to therapies. We found that patients with progressive disease have higher FAScores than those patients with responders after receiving chemo/radiotherapy in the TCGA dataset (Figure 6L). IPS analysis indicated that low-risk patients might benefit from the anti-PD-1/CTLA-4-based immunotherapies (Figure 6M). Next, to verify the hypothesis, we applied the signature to the dataset from patients receiving anti-PD-1 treatment and found that responders have significantly lower risk scores as compared to patients in the high-risk group (Figure 6N).

**Discussion**

Lung adenocarcinoma is one of the most frequently histological subtypes of non-small cell lung cancer (NSCLC). Great advancements in high-throughput genomics studies have been made in recent years which accelerate the understanding of tumor heterogeneity [35]. although patients with LUAD benefited from the emerging treatments including molecular targeted therapy, anti-PD-1/CTLA4 immunotherapy, and chimeric antigen receptor (CAR) T cell therapy, the remission duration is limited. Increasing evidence indicated that the tumor microenvironment context matters in anti-cancer immunity, while few biomarkers that predict immune therapy responses and prognoses could delineate the TME in LUAD. Dysregulated fatty acid metabolism in cancer can serve as the essential accomplice in cancer progression and metastasis by reprogramming the TME\textsuperscript{40}. Targeting fatty acid metabolism has become a promising therapeutic strategy in fighting cancer\textsuperscript{41}, whereas the roles of fatty acid metabolism-associated genes in LUAD have not been fully investigated. In this study, we systematically explored the associations between the gene expressions of FAMGs and clinical outcomes in LUAD patients. We identified a robust FAMGs-related risk signature that is tightly correlated with the OS and DFS of patients. This signature was validated in six publicly independent available LUAD cohorts. The utility of the signature was further confirmed as an independent indicator for patients’ prognosis by adjusting for clinical features. The relevant mechanisms of the signature in predicting tumor microenvironment landscape, anti-immunity regarding antigen-specific tumor killing, and responsiveness to immunotherapy were also analyzed. Therefore, the proposed signature may shed a light on monitoring outcomes and understanding personal immunotherapy for patients with LUAD.
The biological role and therapeutic potential of fatty acid metabolism have attracted interest in cancer, and rare comprehensive evaluations of their clinical relevance in LUAD was reported. We found that they have diverse effects on prognosis by Kaplan-Memier curve analysis, indicating that differential expression of these FAMGs has implications for cancer progression. Overexpression of Type XI collagen (COL11A1) promoted cell proliferation, migration, and drug resistance in NSCLC or recurrent NSCLC by in vitro and in vivo functional assays. Patients with high COL11A1 correlated with decreased survival was consistent with the notion that it is a prognostic biomarker for various cancers including NSCLC. Similar functions were observed for IGF2BP1, FAM83B, GPR115, TCN1, and PRAME in NSCLC. It might be a novel druggable target for COL11A1-high cancers. RHOV has been identified as one of the most up-regulated Rho GTPase members in lung adenocarcinoma and was associated with unfavorable survival by bulk and single RNA sequencing. RHOV silencing inhibited proliferation and migration, as well as improved the sensitivity of EGFR-TKI and increased cancer cell apoptosis in gefitinib-resistant PC9 cells. In addition, high Dickkopf-1 (DKK1), an inhibitor of the Wnt/β-catenin signaling pathway, in NSCLC was linked with the proliferative and invasive capacity, and it could be a potential therapeutic target. CDH17 was considered as a tissue-specific diagnostic marker for adenocarcinomas. Thus, further investigation of these tumor-promoting FAMGs in mediating metabolism in the TME might represent promising therapeutics for patients with LUAD. This prompted us to systematically profile these genes in LUAD by mathematical modeling, which has been widely employed to monitor outcomes or predict treatment response in recent years. Many tools have been in the preclinical phase or approved by FDA after multi-centers sets validation, such as Guardant360® CDx, a qualitative next-generation sequencing-based diagnostic device that uses targeted high throughput hybridization-based capture technology for detecting mutations of 55 genes to identify non-small cell lung cancer patients who may benefit from treatment with the targeted therapies. This study confirmed that FAMGs are important predictors of survival among LUAD patients. We constructed a robust signature with great prognostic value for OS prediction utilizing TCGA RNA-sequencing data as the training set. The reliability of the signature was then verified via multiple diverse microarray platforms, suggesting that it has powerful risk discrimination in pooled populations and strong translational potential. The signature still showed superior prognostic utility when integrated clinical characteristics such as tumor stage, grade, and oncogenic drivers’ mutations. Its predictive performance was also confirmed by ROC curve analysis exhibiting moderate to high AUC values.

Secreted FAs accumulation in TME promoted the infiltrated immune cell function and phenotype. Abnormal fatty acid metabolism including fatty acid oxidation and lipid synthesis would nourish cancer cell survival, increase resistance to chemotherapeutic/radiation treatments, and weaken cellular stresses. Tumor immune microenvironment (TIME) analysis indicated elevated immune scores in low-risk patients. Further immune cell subset deconvolution by different algorithms confirmed that higher infiltrated total T cells, central memory CD8+ T cells, B cells, and NK cells in low-risk patients in opposite to that of high-risk patients. These tumor-infiltrating lymphocytes (TILs) are the main killers in anti-tumor immunity. Amounting reports noted that these TILs usually are less located in tumor sites and
growing exhaustion\textsuperscript{57,58}. Many co-stimulatory and inhibitory molecules were found increased expression in low-risk patients, as well as T cell activation markers such as KLRG1, which might imply a hyporesponsive state of T cells in tumor-killing activities that decrease tumor burden to improve survival. In addition, patients with risk scores may benefit from immune checkpoint blockade (ICB). Meanwhile, markers of T cell senescence including CD57 and H2AX were increased in high-risk patients, indicating dysfunctional protective immunity\textsuperscript{59}. Low precursor frequency and T cell receptor (TCR) affinity have been demonstrated in the most tumor-specific T cells, and antigen presentation was also impaired in TME, which result in weak priming and boosting of T cells\textsuperscript{60}. TCR, a unique molecule on the T cell surface, can recognize antigens presented by MHC, and it has been considered as a potent indicator to predict immunotherapy response\textsuperscript{61,62}. The effectiveness of ICB therapy was tightly correlated with the abundance and functionality of infiltrated T cells in the tumoral niche. The richness and diversity of TCR repertoire were assessed and we found that low-risk patients showed higher TCR richness and diversity, which indicated enhanced T cell functionality in recognizing antigens and killing tumor cells. Additionally, biomarkers of ICB response PD-L1 were increased in high-risk patients, while low-risk patients harbored significantly lower TMB, suggesting decreased immunogenicity in low-risk tumors. IPS analysis confirmed that low-risk patients might be sensitive to ICB. Therefore, we validated the predictive value of the signature using an NSLCLC cohort that received anti-PD-1 treatment (GSE126044) and found that responders have lower risk scores in contrast to those of non-responders. This might explain why the functionality of TILs is more important than immunogenicity. Overall, low-risk patients were prone to benefit from ICB. Further validation in more cohorts will convince the signature is a reliable marker for immunotherapy response.

To seek differentially expressed pathways that are related to malignant traits and immunotherapy, GSEA was conducted using the Hallmark and KEGG pathway gene sets. We noted that EMT, unfolded protein response, TNFα-signaling via NF-κB, DNA repair, oxidative phosphorylation, glycolysis, G2M checkpoint, and hypoxia were significantly enriched in high-risk patients, which were also validated in KEGG pathways analysis. These pathways have been associated with responsiveness or resistance to ICB therapy\textsuperscript{32,63,64}, suggesting the relevance of the signature to the biology of tumor progression and immunotherapy.

The robustness of the signature showed robust performance in multiple cohorts and could be an independent predictor for LUAD, while several limitations should be taken into consideration when interpreting the signature. The retrospective nature of the signature requires multi-centered large cohorts’ validation. We found that several FAMGs correlated with clinical outcomes, further work to explore their biological mechanisms in LUAD is imperative. The signature revealed the different TIME phenotypes of LUAD and predicted immunotherapy effectiveness, the correlations of these genes in mediating tumor niche that was related to therapeutics need investigation.

**Conclusions**
In summary, we developed a FAMG-based prognostic signature used for predicting prognoses, reflecting TIME, and stratifying the benefits of immunotherapy. Further verification of these findings may provide a meaningful perspective view on understanding the role of FAMGs in LUAD and enabling personalized immunotherapy.

**Abbreviations**

NSCLC, non-small cell lung carcinoma; LUAD, lung adenocarcinoma; DEGs, differentially expressed genes; FAs, fatty acids; FAMGs, fatty acid metabolism related genes; TME, Tumor microenvironment; CIBERSORT, Cell-type Identification by Estimating Relative Subsets of RNA Transcripts; TILs, Tumor infiltrating lymphocytes; ROC, the receiver operating characteristics curve; GEO, Gene Expression Omnibus; TCGA, The Cancer Genome Atlas. PPI, protein-protein interaction; ICB, immune checkpoint blockade; MATH, mutant-allele tumor heterogeneity; TMB, tumor mutation burden; TCR, T cell receptor; BCR, B cell receptor; TIME, tumor immune microenvironment.

**Declarations**

**Acknowledgments**

We are grateful to the contributors of data to public databases including TCGA and GEO databases.

**Authors’ contributions**

Conceptualization and design: Denggang Fu; Data acquisition: Denggang Fu, Biyu Zhang, Wenyan Fan; Methodology: Denggang Fu, Biyu Zhang, Wenyan Fan, Xin Wang; Data analysis and interpretation: Denggang Fu, Biyu Zhang; Writing (original draft): Denggang Fu, Biyu Zhang; Writing (review and editing): Denggang Fu, Xin Wang; Project administration: Denggang Fu, Xin Wang.

**Funding**

This study was supported by the Natural Science Foundation of Jiangxi Province (20192BAB215001;20224ACB206038).

**Availability of data and materials**

The data analyzed in this study are available in the following repositories:

2. GEO: https://xenabrowser.net/datapages/.
3. xCell: https://xcell.ucsf.edu/.
4. TIMER: http://timer.cistrome.org/.
5. STRING: https://string-db.org/cgi/input.pl.

More data are available from the corresponding author upon reasonable request.

**Ethics approval and consent to participate**

This article does not contain either human or animal experiments and written consent is not required for the current study.

**Consent for publication**

All authors consent for publication.

**Competing interests**

The authors declare that they have no conflicts of interest regarding the publication of this manuscript.

**Author details**

1 College of Medicine, Medical University of South Carolina, Charleston, SC, USA 29425; 2 Key Laboratory of Green Chemical Engineering Process of Ministry of Education, Hubei Key Laboratory of Novel Reactor and Green Chemical Technology, School of Chemical Engineering and Pharmacy, Wuhan Institute of Technology, Wuhan, China 430079; 3 School of Basic Medicine, Jiujiang University, Jiujiang, Jiangxi, 332005, China.

**References**


**Figures**
Figure 1

Construction of fatty acid metabolism associated prognostic signature

(A). Forest plot showing the overall survival-related fatty acid metabolism genes in LUAD. (B). The signature formula. (C). Kaplan-Meier curve showing the survival difference of low- and high-risk patients. (D). Distribution of risk scores among LUAD patients. (E). The number of deaths varies with increasing risk scores. (F). 1-, 3-, and 5-year of AUCs value of the signature.
Figure 2

Univariate and multivariate Cox regression analysis of the signature


Figure 3

Signature validation in six independent LUAD cohorts
Figure 4

Figure 5

Tumor immune microenvironment landscape analysis

(A). single sample gene set enrichment analysis of 29 immune signatures in low- and high-risk patients. (B-C). Immune scores and ESTIMATE scores in low- and high-risk patients analyzed by ESTIMATE algorithm. (D). The tumor purity in low- and high-risk patients. (E). Infiltrated CD8+ T cells, B cells,
activated CD4+ T cells, neutrophils, and activated mast cells in low- and high-risk patients analyzed by CIBERSORT. (G). Infiltrated CD8+ T cells and B cells in low- and high-risk patients analyzed by EPIC (H). Infiltrated CD8+/central memory CD8+ T cells, NK cells, B cells, and microenvironment scores in low- and high-risk patients analyzed by xCell. (I-J). Co-stimulatory and co-inhibitory receptor molecule expression in low- and high-risk patients. (K). The markers of cellular senescence expression in low- and high-risk patients.

Figure 6

Association of the signature with anti-tumor immunity and immunotherapy

(A). Lymphocyte infiltration signature score in low- and high-risk patients. (B). T cell receptor richness and diversity in low- and high-risk patients. (C-D). B cell receptor richness and diversity in low- and high-risk

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.tif
- Fig.S2.tif
- TableS1.xls
- TableS2.xlsx