Identification and Validation of a Ferroptosis-Related Gene to Predict Survival Outcomes and the Immune Microenvironment in Lung Adenocarcinoma

Biao Deng
Guangdong Medical University
Jing Xiang
Guangdong Medical University
Zhu Liang (liangzhuwsh@163.com)
Guangdong Medical University
Lianxiang Luo
Guangdong Medical University

Research Article

Keywords: Ferroptosis, Lung adenocarcinoma, Prognosis, Immunocytes, Tumor Microenvironment, Immunotherapy

Posted Date: April 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1564484/v1

Abstract

Background: Lung adenocarcinoma (LUAD) is one of the leading causes of cancer-related death worldwide. Ferroptosis, a form of cell death characterized by iron-dependent lipid peroxidation, has been shown to be involved in the regulation of the tumor immune microenvironment by triggering inflammation-related immunosuppression. However, the involvement of ferroptosis in the regulation of immune cell infiltration and immunotherapeutic efficacy in LUAD remains unclear.

Methods: To identify a ferroptosis-related gene to predict survival outcomes and the immune microenvironment in LUAD, the Cancer Genome Atlas (TCGA) LUAD cohort was used to assess the survival prognosis of Ferroptosis-related genes and to construct a 7-genes risk signature. Correlation tests, difference tests and cluster analysis were performed to explore the role of FRGs in the immune microenvironment and immunotherapy efficacy in LUAD. The effects of FRGs on LUAD cells were assessed by methods such as Western Blot, iron assay and lipid peroxidation assay.

Results: The 7-gene risk signature of LUAD patients was established and validated. FRGs clustering based on 70 differentially expressed FRGs are associated with the immune microenvironment, a potential immune subtype of LUAD. 7-gene risk signature is an independent prognostic factor for LUAD and classified LUAD cohort into high- risk and low-risk groups. Immunocytes infiltration levels, immune checkpoints and immunotherapy response rates are significantly different between the two groups. Patients with high-risk scores have lower overall levels of immunocytes infiltration but higher immunotherapy response rates. The key gene RRM2 could influence LUAD prognosis by regulating the level of Mast cells activated and T cells CD4 memory activated infiltration. In addition, RRM2 was involved in the ferroptosis process, and the expression of RRM2 was up-regulated in Lung cancer tissues and LUAD cell H1975. Silencing RRM2 could promote the occurrence of ferroptosis in H1975 cell.

Conclusion: Our results identified RRM2 as promising biomarkers and therapeutic targets associated with tumor immune infiltration in LUAD patients.

1. Introduction

Pulmonary carcinoma is the most serious malignancy and has been highlighted as the leading cause of cancer deaths occurring in developed and developing countries worldwide[1]. Lung adenocarcinoma (LUAD) is the most common histological type of non-small cell pulmonary carcinoma[2], and compared to other lung cancer subtypes, LUAD is closely associated with genomic changes and is highly heterogeneous[3]. Although mortality from pulmonary carcinoma has decreased in recent years, it still causes more deaths than breast, prostate, colorectal, and intracranial tumors combined[4]. More than half of patients with lung adenocarcinoma are diagnosed at advanced stages or have metastasized[5], and the patients have a poor prognosis with a low 5-year survival rate [6].

Ferroptosis, a novel form of programmed cell death characterized by iron-dependent lipid peroxidation, has been widely demonstrated to be associated with a variety of diseases, particularly cancer, and

modulating the occurrence of ferroptosis in cancer cells could be a potential strategy for cancer therapy[7]. In lung adenocarcinoma, mitochondrial-induced cysteine starvation, endoplasmic reticulum-related oxidative stress, lysosome dysfunction and lipid peroxidation related to Golgi stress have been proved to contribute to the induction of ferroptosis[8]. And a growing number of genes have also been found to be involved in the organelle changes that induce ferroptosis. The regulatory mechanisms of FRGs in the occurrence and development of pulmonary carcinoma have attracted increasing attention[9]. Bufu Tang et al. found that FRGs affect pulmonary carcinoma progression and tumor immunocytes infiltration. RRM2 inhibitors effectively promoted M1 macrophage polarization and inhibited M2 macrophage polarization. And Ferrostatin-1 treatment effectively rebalanced macrophage polarization mediated by RRM2 inhibitors[10]. And Huang et al. reported that AKR1C1 induces ferroptosis through multiple pathways and is associated with various cancer-infiltrating immune cells[11]. Studies by Min Wang et al. showed that lncRNA and ceRNA networks also played an important role in tumorigenesis and ferroptosis. Endogenous microRNA 6852 (MIR6852) inhibits cell growth by promoting As MIR6852 sponge, LINC00336 can regulate the expression of cystathionine-β-synthase (CBS), thereby promoting the occurrence of ferroptosis[12]. In addition to chromosome genes, ferroptosis-related lncRNAs in cytoplasm can also inhibit cancer by activating p53 pathway. For example, studies by Chao Mao et al. showed that lncRNA P53RRA binds to protein binding protein 1(G3BP1) to activating G3BP1. P53RRA-G3BP1 interaction can replace p53 from G3BP1 complex, resulting in more p53 retained in the nucleus, leading to cell cycle arrest, apoptosis and ferroptosis[13]. Luo et al. summarized the regulatory mechanism of ferroptosis, the interaction of ferroptosis on tumor cell metabolism and anti-tumor immunity[14], however, the significance of FRGs in the immune microenvironment and immunotherapy of LUAD remains elusive, and FRGs require further investigation.

2. Materials And Methods

2.1. Obtained FRG and gene expression data of LUAD

Download the list of genes related to ferroptosis from the public database FerrDb(http://www.zhounan.org/ferrdb/), delete the non-coding genes and get 382 FRGs. Transcriptome data in FPKM format of LUAD were downloaded online from TCGA(https://ancergenome.nih.gov/) database (a total of 56530 genes from 594 samples, including 535 tumor samples and 59 paracancerous samples; 492 of which have complete clinical information such as survival time, survival status, age, gender, stage and TNM state). The chip data GSE310219 (including 54675 genes from 307 early lung adenocarcinoma samples, 292 of which have complete clinical information) was downloaded online from GEO database(https: //www.ncbi.nlm.nih.gov/geo/).The clinical data of LUAD patients included in this study are shown in Table 1.

Table 1
Clinical characteristics of lung adenocarcinoma patients included in this study.

Characteristic	TCGA-LUAD(n = 492)	GSE310219(n = 292)
age (%)	235(47.8)	179(61.3)
≤ 65	257(52.2)	113(38.7)
> 65		
gender (%)		
male	228 (46.3)	249 (85.3)
female	264 (53.7)	43 (14.7)
stage (%)		
I + II	386 (78.5)	-
III + IV	106 (21.5)	-
T(%)	427(86.8)	235(80.5)
T1 + T2	62(12.6)	52(17.8)
T3 + T4	3(0.6)	5(1.7)
Tx	317(64.4)	198(67.8)
N(%)	165(33.5)	93(31.8)
N0	10(2.0)	1(0.3)
N1 + N2 + N3	328(66.7)	281(96.2)
Nx	25(5.1)	8(2.7)
M(%)	139(28.2)	3(1.1)
M0		
M1		
Mx		

2.2. Differential expression genes (DEG) between LUAD tissues and adjacent tissues

The mRNA data of TCGA-LUAD and GSE310219 cohorts were standardized, and then the transcription sequences were compared with the FRGs, which were analyzed by LIMMA (R software package). When the log2 absolute value (FC) > 1, and adjusted P-value < 0.05, the differential expression gene (DEG) was selected.

2.3. Establishment and validation of a risk signature

Univariate Cox regression analysis was used to screen genes related to overall survival (OS), and multivariate Cox regression analysis was used to identify genes with independent prognostic effect. P-value < 0.05was considered statistically significant. The genes screened by multivariate Cox regression were used to construct risk signature, and then, we established a predictive risk scoring model combining regression coefficient (β) with gene expression level. RiskScore = (β1 * expression level of ALOX12B) + (β2 * expression level of DDIT4) + (β3 * expression level of SLC7A5) + (β4 * expression level of TRIB3) + (β5 * expression level of IL33) + (β6 * expression level of RRM2) + (β7 * expression level of CAV1). The establishment of risk signature needs software packages such as 'survival' and 'survminer' in R. Firstly, the univariate COX and multivariate Cox regression were used to analyze whether the prognostic

2.4. Prognostic potential analysis of risk signature

Indicators were independent of other traditional clinical features, such as age, gender and TNM state. The hazard ratio (HR) of each variable was measured with 95% confidence interval (CI), P-value < 0.05 was considered statistically significant. Then, the LUAD patients in TCGA and GEO cohorts were divided into high-risk group and low-risk group with the median of risk score as the boundary. The difference in the overall survival rate between the high-risk group and the low-risk group was analyzed by Kaplan-Meier (K-M) curve. The prognostic potential and significance of the risk model were tested by ROC curve, and the area under the curve (AUC) was used to evaluate the prediction specificity and sensitivity of the model. This analysis will use the 'Survive' package of R, and 'ggDCA' package will be used to draw the decision curve (DCA) for analyze the net benefit of risk signature. P < 0.05 represents statistical significance.

2.5. Enrichment analysis

GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis were performed using R package 'clusterProfiler', 'org.Hs.eg.db', and 'enrichedplot' to explore the differences of various molecular mechanisms between high and low risk patients. GO analysis includes biological processes (BP), cell components (CC) and molecular functions (MF).

2.6. Construction of predictive nomogram

Using R package 'survival' and 'regplot', all independent predictors of LUAD were identified, and then combined with the survival information of LUAD patients, the nomogram were drawn to predict 1,3 and 5-year survival rates.

2.7. Analysis of tumor microenvironment

Firstly, the StromalScore, ImmuneScore and EstimateScore of each LUAD sample are calculated by the estimation algorithm, and the differences between high and low risk groups are analyzed. Then, transcriptome data were quantitatively converted into specific types of cell absolute abundance, including immune cells and stromal cells, through seven algorithms: CiberSort, TIMER, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, XCELL and EPIC. Thirdly, immunocytes with independent predictive

significance for the prognosis of LUAD were identified by univariate COX and multivariate COX regression analysis, and the correlation between the expression of each risk gene and the level of immunocytes infiltration was analyzed by correlation test. Fourthly, the TCGA-LUAD cohort was clustered according to the level of immunocyte infiltration, and the LUAD samples were divided into high-infiltration group and low-infiltration group. The differences in clinical characteristics between the two groups were compared by difference test.Finally, the scores of 29 immune-related functions or pathways were calculated by ssGSEA analysis, and the differences in scores between high and low risk patients were compared (all tests were considered statistically significant with $P < 0.05$).

2.8 Correlation analysis of immunotherapy

Primarily, TIDE, MSI, Dysfunction and Exclusion scores of TCGA-LUAD cohort were

downloaded from TIDE database (http://tide.dfci.harvard.edu/), and TMB of LUAD patients were downloaded using R packages 'TCGAbiolinks' and 'maftools', then wilcoxon rank sum test was used to compare whether there were significant differences in the scores of five kinds of immunotherapy between high and low risk patients. Moreover, LUAD cohort (GSE126044) with anti-PDL1 treatment history was downloaded from GEO database, and the differences in response of patients with high and low risk to anti-PDL1 immunotherapy were compared. Additionally, the expression matrix of common immune check-points were extracted from TCGA-LUAD queue, and their expression differences between LUAD patients in high and low risk groups were compared. Ultimately, the correlation between first-line targeted therapy driving genes and risk genes was analyzed by correlation test. $P < 0.05$ was considered statistically significant.

2.9. Drug screening

The risk genes were compared with the reference data set of CMap database(https://portals.broadinstitute.org/cmap/), and the drugs were ranked according to the enrichment of risk genes in the reference gene expression profile. The drugs that significantly inhibited the expression of high-risk genes were screened with $P < 0.05$ and the enrichment score < 0.The chemical structures of the selected drugs were searched in the pubchem database (https://pubchem.ncbi.nlm.nih.gov/).Using 'pRRophetic' algorithm, ridge regression model was constructed to predict drug IC50 according to GDSC (https://www.cancerrxgene.org/) cell line expression profile and TCGA gene expression profile, and antitumor drugs with significantly lower IC50 for high-risk LUAD samples were screened.

2.10. Cell Culture and Transfection

A normal lung cell BEAS-2B and seven lung cancer cell lines (H1299, A549, H460, H23, H838, PC-9, H1975) was purchased from American Type Culture Collection (ATCC) and maintained in 1640 medium containing 10% fetal bovine serum (Gibco, Gland Island, USA) and 1% penicillin-streptomycin (Gibco, Gland Island, USA) and cultured at 37°C with 5% CO2. RRM2-siRNA from Sangon Biotech (Shanghai, China) was used to silence the expression of RRM2. In this study, the sequence of RRM2-siRNA is: sense:

5'- GCGAUUUAGCCAAGAAGUUTT-3'. H1975 cell were seeded in 12-well plates at a density of 6 × 104 cells/well, and transfected 24 hours later. RRM2-siRNA or NC-siRNA was transfected to a final concentration of 20 nM using siRNA transfection reagent (Polyplus, France). Finally, the transfected cells were detected after48hours.

2.11. Western Blot

Proteins were extracted from cells and cell lysates were prepared using RIPA lysate with PMSF (Solarbio, Beijing, China), and then protein quantification was performed using a BCA protein assay kit (Sangon Biotech, Shanghai, China). Proteins were then separated by 12% SDS-PAGE and transferred to nitrocellulose membranes. The membranes were blocked with 5% bovine serum albumin (BSA) for 1h and then incubated with primary antibody overnight at 4°C. The next day, after washing the membrane three times with TBST, the membrane was incubated for 1 h at room temperature with horseradish peroxidase-labelled secondary antibody (1:4000), followed by three washes with TBST. Finally, the colors were developed using BeyoECL Moon (Beyotime Biotechnology, Shanghai, China).

2.12. Detection of lipid peroxidation and apoptosis

C11-BODIPY 581 / 591 (10μM; aBclone, Wuhan, China) were added to the H1975 cell and incubated at 37 ° C and 5% CO2 for 1 h. Then, the cells were washed twice with PBS. After trypsin digestion, the cells were resuspended with PBS containing 5% FBS, and finally analyzed by flow cytometry. Apoptosis detection kit (BD Biosciences) was used to prepare fluorescent dyes containing PI and FITC according to the instructions. The H1975 cell were incubated in dark at room temperature for 15 min, and finally analyzed by flow cytometry.

2.13. Iron assay

FerroOrange (Dojindo, China) was used to detect the ferrous level of cells. According to the instructions, cells were incubated with FerroOrange for 0.5 hours. The fluorescence intensity was evaluated using a rotating disk super-resolution laser confocal microscope.

2.14. Immunohistochemical (IHC) staining

From January 2020 to January 2021, 12 cases of lung adenocarcinoma tissues and normal tissues in Department of Cardiothoracic Surgery, Affiliated Hospital of Guangdong Medical University were selected. The obtained tissue was transferred to-80°C refrigerator in an ice box for subsequent IHC experiments. The included research subjects were all determined by experts in the pathology department of our hospital, and none of the patients studied had received radiotherapy, chemotherapy, targeted therapy and immunotherapy before surgery. All patients signed informed consent before surgery. The tissue was fixed with 4% paraformaldehyde and embedded in paraffin to prepare a slice with a thickness of 5 μM. The slices were then dewaxed with xylene and dehydrated in gradient concentrations of alcohol. Subsequently, 0.01M sodium citrate (pH: 6.0) was used for antigen repair, and endogenous peroxidase was blocked by adding 0.3% hydrogen peroxide (H2O2) and incubation in 10% goat serum albumin for 30

min. The slices were then incubated overnight with RRM2 primary antibody at 4°C.The next day, the samples were incubated with 3,3 ' -diaminobenzidine (DAB) after incubated with HRP-conjugated anti-rabbit secondary antibody for 1 hour. The sections were stained with Mayer hematoxylin, dehydrated, removed with xylene, sealed with neutral resin, and finally detected by multi-functional microporous microscope (BIOTEK).

2.15. Statistical analysis

All data were generated, processed and analyzed by R (version 4.1.1), R STUDIO and strawberry-perl (5.32.1.1). Student t test was used to determine the difference between the two groups, and Wilcoxon rank sum test was used to identify the difference for data from non-paired groups. $P < 0.05$ indicated that the difference was statistically significant. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

3. Results

3.1 Obtaining differential expression genes (DEG) of LUAD

382 FRGs were downloaded from the public data center FerrDb, and non-coding genes were deleted from the data set. A total of 70 FRGs are differentially expressed between tumors and normal tissues, and the characteristics of these FRGs in TCGA-LUAD cohort are shown in Fig. 1A. In addition, all LUAD samples are clustered by K-means clustering based on 70 differentially expressed FRGs, and the best clustering effectis achieved when K = 4 (**Supplementary Fig. 1A-D**). And there are significant differences in survival among the four clusters (**Supplementary Fig. 1E**).

3.2. Construction and validation of prognostic model

As shown in the forest map, 15 DEGs associated with LUAD prognosis were identified by R software package 'survival' from the TCGA cohort (Fig. 1B). Then, 7 FRGs with independent prognostic significance for LUAD patients were obtained by multivariate COX regression analysis on the 15 appealed DEGs, and the location of these FRGs in the differential volcanic map is shown in Fig. 1C. Kaplan-Meier curve showed the survival differences of seven FRGs at different levels (**Supplementary Fig. 1F-L**). Ultimately, we established a 7-gene risk signature by multivariate Cox regression coefficient. The risk score of each patient = (0.262297470303534 * expression level of ALOX12B) + (0.0055627086031226 * expression level of DDIT4) + (0.00646442374221167 * expression level of SLC7A5) + (-0.0131510931059613 * expression level of TRIB3) + (-0.0300191156801198 * expression level of IL33) + (0.015384279700975 * expression level of RRM2) + (0.0046423481143834 * expression level of CAV1). According to the risk index ranking of 504 TCGA-LUAD patients, all patients were divided into high-risk group and low-risk group according to the median of risk index. The independent prognostic analysis of the clinical characteristics of LUAD is shown in Fig. 1D, E after screening out the clinical information which including age, gender, staging and risk score. Risk score and clinical stage are independent factors affecting the prognosis of LUAD. The clinical correlation chart in **Supplementary Fig. 4A** shows the proportion of patients with high and low risk score in LUAD patients at each stage. Further analysis showed that the OS

rate of patients in the high-risk group in the TCGA-LUAD cohort was significantly lower than that in the low-risk group ($p < 0.001$) (Fig. 2A). Figure 2B showed that the survival rate difference between low-risk group and high-risk group also existed in GSE30219 cohort, the OS rate of high-risk group was significantly lower than that of low-risk group ($p < 0.001$). Next, the ROC curve (Fig. 2C) shows the prominent sensitivity and specificity of our risk signature in predicting the OS rate of LUAD patients. The AUC were 0.707, 0.685 and 0.673, respectively corresponding to 1, 2 and 3 years. Then we added clinical features including TIDE, TIS, Age, Stage to compare the accuracy of OS prediction. As shown in Fig. 2D, the risk index is the best predictor of OS. It can be seen from the DCA (Fig. 2E) that the net benefit of the risk index is the largest. Finally, a nomogram was constructed based on two independent prognostic factors to quantify the survival probability of LUAD patients in 1, 3 and 5 years (Fig. 2F).

3.3 Enrichment analysis

We conducted enrichment analysis to explore the potential biological characteristics of 70 DEGs. 25 pathways shown in **Supplementary Fig. 2A** were enriched by the KEGG enrichment analysis. Red bubbles represent more prominent pathways, such as fluid shear stress and atherosclerosis, HIF-1 signaling pathways, lipid and atherosclerosis, and ferroptosis. GO enrichment analysis showed that the biological processes including the response to oxidative stress and metal ions, and the cell components including apical plasma membrane, cell apex, and oxidoreductase complex were significantly enriched (**Supplementary Fig. 2C**). Furthermore, GSEA enrichment analysis showed that the high and low risk groups were enriched to completely different functions or pathways. Among high risk patients, small cell pulmonary carcinoma, melanoma, cancer pathway, P53 signaling pathway, pathogenic Escherichia coli infection were significantly enriched (**Supplementary Fig. 2B**).

3.4 We calculate the StromalScore, ImmuneScore and EstimateScore for each LUAD sample by the estimate algorithm. Significant differences in StromalScore, ImmuneScore and EstimateScore between high and low risk groups by T-test were shown in Fig. 3A-C. In order to explore the influence of different levels of immune-related scores on the prognosis of patients, the StromalScore, ImmuneScore and EstimateScore of all LUAD patients were divided into high and low groups by average. The OS rates of the two groups were compared by Kaplan-Meier method, the Fig. 4B showed that higher levels of ESTIMATEScore represented better prognosis. The same is true for ImmuneScore and StromalScore (**Supplementary Fig. 5A, B**). In addition, the T-test showed that there were significant differences in these immune-related scores among the four clusters obtained by K-means clustering analysis (Fig. 3D-F), suggesting that the four clusters obtained by 70 differentially expressed FRGs clustering may serve as promising immune subtypes.

3.5 Further analysis of immune cell infiltration in LUAD TME

CIBERSORT algorithm was used to quantify the levels of 22 immune cells in each LUAD sample, and then T-test was used to compare the levels of immune cells between high-risk and low-risk patients. As shown in Fig. 4A, there were significant differences in the levels of B cell memory, T cell CD4 memory, T cell CD4 memory activated, T cell regulatory (Tregs), NK cells resting, Monocytes, Macrophages M0, Dendritic cells

resting, Mast cells resting and Mast cells activated between high-risk and low-risk patients. The relative proportion of 22 immune cells in all LUAD samples was shown by barplot (**Supplementary Fig. 4E**). When Kaplan-Meier curves were used to compare the effects of immunocytes with different infiltration levels on the OS rates, we found that different infiltration levels of 12 kinds of cells such as T cell CD4 memory activated and Mast cells resting represented different OS rates (**Supplementary Fig. 3A-L**). To learn more about the difference of immune cells between high and low risk patients, we analyzed seven algorithms in CIBERSORT, TIMER, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, XCELL, EPIC, and Fig. 4A showed the differential infiltration of immunocytes in each database.

3.6 Relationship between immunocytes infiltration and clinical characteristics

In order to understand the correlation between immune cell infiltration and clinical characteristics, we calculated the immune cell infiltration level of all LUAD samples in TCGA database by ssGSEA algorithm, and then downloaded the clinical information of all LUAD samples from TCGA database by R package 'TCGAbiolinks'. The 580 LUAD samples with clinical information were divided into two groups by clustering according to the level of immune cell infiltration, and visualized by heat map (Fig. 5A). The left part dominated by blue represents the samples with less immunocytes infiltration, while the right part dominated by red represents the samples with higher immunocytes infiltration. The clinical information such as survival status, stage, gender, tumor location, and related driving gene mutation state are corresponding to the heat map below. The correlation heat map shown in Fig. 4C reveals the relationship between the immune cells, in which blue modules are the majority, suggesting that the immune cell infiltration in the LUAD microenvironment is negatively correlated. In addition, in order to better understand the immune differences between high-risk and low-risk groups, we quantified 29 immune-related functions by ssGSEA algorithm. As shown in Fig. 5C, 16 immune-related functions or pathways such as Adcs and CCR showed significant differences. When Kaplan-Meier curve was used to compare the effects of different levels of immune-related functions on OS rates, we found that various levels of multiple immune-related functions or pathways predicted dissimilar OS rates. Figure 6D showed that discrepant levels of Mast _ cells-related immune functions or pathways corresponded to different prognosis of LUAD patients. After univariate Cox regression analysis, we identified Mast cells activated and T cells CD4 memory activated as prognostic-related immune cells of LUAD, and these two cells with high level of infiltration are the adverse prognostic factors of LUAD (Fig. 6A). Further analysis confirmed that the expression levels of seven risk genes significantly affected the infiltration level of immune cells immunocytes in the LUAD microenvironment, and the infiltration abundance of Mast cells activated, which was most related to the prognosis of LUAD, was affected by the expression levels of RRM2 and IL33 (Fig. 6B). The expression of RRM2 not only regulates the infiltration of Mast cells activated, but also is related to the infiltration of 19 immunocytes such as T cells CD4 memory, dendritic cells and Macrophages (Fig. 6C).

3.7 LUAD immune subtype based on FRGs clustering

In order to explore the relationship between the four clusters obtained by K-means clustering based on 70 differentially expressed FRGs and the two groups clustered according to the level of immunocytes infiltration, we drew a clinical correlation heat map as shown in Fig. 5B. There were significant differences in immunocytes infiltration among the four clusters. For example, the samples with high immunocytes infiltration in the composition of C3 were significantly more than those with low immunocytes infiltration, while the samples with low immunocytes infiltration in the C2 and C4 clusters were more. There are also significant differences in risk scores among the four clusters (Fig. 3G). The mulberry diagram in **Supplementary Fig. 4C** shows the corresponding relationship between immune infiltration level, FRGCluster, risk level and other clinical features. The above results show that the four clusters based on 70 FRGs clustering may be the potential immune subtypes of LUAD.

3.8 Risk signature and immunotherapy

To explore the role of risk signature in guiding immunotherapy, we found 16 pulmonary carcinoma patients receiving anti-PDL1 immunotherapy in the GSE126044 cohort. **Supplementary Fig. 5G** shows the difference of immunocytes infiltration and response to anti-PDL1 immunotherapy between high-risk and low-risk patients. High-risk patients have lower immunocytes infiltration levels, but the response rate to anti-PDL1 immunotherapy is higher. In addition, we found that the expression of RRM2 was higher in patients who had no response to anti-PDL1 treatment(P = 0.052) (**Supplementary Fig. 5F**). Then we downloaded TIDE, MSI, Dysfunction and Exclusion scores of LUAD cohort from TIDE database, and downloaded TMB data of LUAD patients using R packages 'TCGAbiolinks' and 'maftools'. Wilcoxon rank sum test found that there were significant differences in the above five immune therapy-related scores between high-risk and low-risk patients (Fig. 3H, Fig. 5D-G). The LUAD patients with high TMB had worse OS (**Supplementary Fig. 5D**). Subsequent correlation analysis showed that the risk score was positively correlated with TMB (R = 0.23, P = 0.000018) (**Supplementary Fig. 5C**). The correlation heat map in **Supplementary Fig. 4B** showed the distribution of clinical features between high-risk and low-risk groups. The distribution of TMB, Gender, Stage, T and N among high-risk and low-risk groups was significantly different.

In order to study the relationship between the seven genes of risk signature and immunotherapy, we counted the mutations of seven key genes in LUAD patients. As shown in **Supplementary Fig. 5E** waterfall diagram, the mutation probability of ALOX12B was the highest, reaching 18%, followed by RRM2 and IL33, and misscnsc mutation was the most common form. When further exploring the relationship between the expression of 7 risk-genes and immunocytes infiltration, we conducted the correlation analysis as shown in Fig. 7A, B. The expression level of RRM2 positively regulated the infiltration abundance of Mast cells activated (R = 0.12) and T cells CD4 memory activated (R = 0.41), and the infiltration level of Mast cells resting was significantly negatively correlated with the expression level of ALOX12B (R = − 0.13) (**Supplementary Fig. 4F**). When ssGSEA algorithm was used to analyzes the scores of immune-related functions or pathways, it is found that the scores of check point between high and low risk LUAD samples are significantly different. Check_ point is divided into high and low groups by means of average, and Kaplan-Meier curve shows significant differences in survival. Next, we analyzed the

expression differences of 49 common immune checkpoints between high-risk and low-risk patients, as shown in the box plot shown in Fig. 7D. The expression levels of ALK, ROS1, CD44 and other 27 checkpoints were different. The changes of ALK and ROS1 expression with the risk score were shown in Fig. 7E, F, and the expression levels of ALK and ROS1 decreased with the increase of risk score. Furthermore, correlation analysis showed that ALK and ROS1 were significantly associated with key risk genes. The expression of ROS1 was positively correlated with IL33, and negatively correlated with SLC7A5, TRIB3 and RRM2 (Fig. 7H). The expression of ALK was positively correlated with IL33 and CAV1, and negatively correlated with the expression of ALOX12B, SLC7A5 and TRIB3 (Fig. 7I). The expression of two additional checkpoints PD-1 and PD-L1 were also significantly correlated with multiple prognostic FRGs (**Supplementary Fig. 5H, I**).

3.9 Overexpression of RRM2 in tumor tissues

In order to detect the expression of RRM2 in normal lung tissues and lung cancer tissues, we performed immunohistochemical experiments for comparison. The results showed that the expression of RRM2 in tumor tissues was higher than that of normal lung tissues (Fig. 8).

3.10 Silencing RRM2 induced the increase of ferrous ion level and lipid peroxidation accumulation

Ferroptosis is a new form of cell death characterized by lipid peroxidation accumulation and iron dependence. Further study on the specific mechanism of ferroptosis is expected to bring new prospects for cancer treatment. To determine the clinical relevance of RRM2 expression, we detected RRM2 expression in a normal lung cell line (BEAS-2B) and seven lung cancer cell lines (H1299, A549, H460, H23, H838, PC-9, H1975), Western blotting showed that RRM2 protein was highly expressed in H1975 cell. At the same time, we also detected the expression of 4HNE and ACSL4 in different lung cancer cell lines. The results showed that 4HNE and ACSL4 were highly expressed in H1975 cell (Fig. 9A), indicating that H1975 cell may be more sensitive to ferroptosis, Therefore, we used H1975 cell for subsequent experiments. After H1975 cell were transfected with RRM2-siRNA for 48 h, the cell death was detected by flow cytometry. The results showed that silencing RRM2 could induce cell death (Fig. 9B). In order to further study the effect of silencing RRM2 on iron death, we continued to detect the levels of ferrous ion and lipid peroxidation in lung cancer cells after silencing RRM2. The results showed that silencing RRM2 could induce the increase of ferrous ion level (Fig. 9C) and lipid peroxidation accumulation (Fig. 9D) in H1975 cell.

3.11 Screening drugs for high-risk FRGs

In order to predict effective therapeutic drugs for patients with high-risk LUAD, we constructed ridge regression model to predict drug IC_{50} based on GDSC cell line expression profile and TCGA gene expression profile, and screened six antitumor drugs with significantly lower IC50 in high-risk LUAD, including Bosutinib, Dasatinib, Gefitinib, Tipifarnib, Docetaxel and JNK. Inhibitor. VIII (**Supplementary Fig. 6A-F**). These six antitumor drugs can kill tumor cells in high-risk LUAD patients. In addition, we

compared the risk genes with the reference data set of CMap database, and obtained the correlation of 1309 drugs according to the enrichment of risk genes in the reference gene expression profile. With $P < 0.05$ as the standard, 158 drugs were found to be significantly enriched in high-risk genes, of which 86 drugs could promote the expression of high-risk genes, and 72 drugs could inhibit the expression of high-risk genes. Medrysone, phenoxybenzamine, Vorinostat, Thioguanosine, apigenin and chrysin were selected as the six drug candidates enriched with more prominent inhibitors of high-risk gene expression, and their chemical structures were found in the pubchem database (**Supplementary Fig. 6G-L**). The detailed information of these 6 drugs is shown in Table 2.

Table 2
6small molecule drugs for high-risk LUAD patients slected from CMAP database.

Rank	CAMP name	Enrichment	P-value
5	phenoxybenzamine	-0.924	0.00004
8	thioguanosine	-0.914	0.0001
9	apigenin	-0.906	0.00014
19	medrysone	-0.73	0.00073
22	chrysin	-0.913	0.00116
23	vorinostat	-0.51	0.00192

4. Discussion

Recent studies have shown that ferroptosis is closely related to many diseases, such as cancer, blood diseases, neurological diseases, kidney diseases and local ischemia-reperfusion injury[15]. A growing number of studies investigating ferroptosis in cancer have revealed its potential as a immunotherapy strategy [16]. Lung adenocarcinoma is the most common pathological type of non-small cell lung cancer. Lung adenocarcinoma is the most common pathological type of non-small cell lung cancer, and most patients are diagnosed at advanced stages, losing their opportunities for surgical treatment. Chemotherapy, radiotherapy and traditional Chinese medicine treatment are common choices for patients with advanced lung adenocarcinoma. However, the 5-year survival rate is only 15%[17]. The benefits of

targeted therapy and immunotherapy in the treatment of advanced cancer patients have begun to change the treatment of cancer, providing new treatment methods for patients with advanced tumors. The benefits of targeted therapy and immunotherapy in the treatment of advanced cancer patients have begun to change the treatment of cancer, providing new treatment methods for patients with advanced tumors. In this study, the 7-gene risk signature we established was closely related to the immune microenvironment of LUAD, and the immunocytes infiltration and immune function between high- and low-risk groups were significantly different. Moreover, the two groups of LUAD patients with high and low infiltration obtained by immunocytes infiltration clustering had different ferroptosis gene clusters and risk scores, which suggests that our 7-gene signature may provide guidance for the immunotherapy of LUAD. The analysis in immunotherapy showed that patients with higher risk score were more likely to benefit from immunotherapy. However, previous results showed that the overall infiltration level of immunocytes in high-risk patients was lower, perhaps because the insufficient infiltration levels of immunocytes related to good prognosis and increased infiltration levels of Mast cells activated and T cells CD4 memory activated led to worse prognosis in LUAD patients. Ultimately, LUAD patients with lower infiltration level were identified as high-risk by 7 gene signatures. Immunotherapy drugs can also activate their immune system to a greater extent, mobilize more immunocytes related to good prognosis to infiltrate and inhibit the activation of Mast cells, T cells CD4 memory and dendritic cells, and improve the response rate to immunotherapy eventually. These results show that FRGs regulate the TME of LUAD and influencing the efficacy of immunotherapy in LUAD. Gene mutation analysis in LUAD samples showed that seven FRGs had a certain degree of mutations, with RRM2 and ALOX12B mutations being more significant. Further analysis also found that the expression level of RRM2 was strongly correlated with the level of immunocytes in LUAD. In addition, in the present study, we found that silencing RRM2 promoted elevated ferrous ion levels, lipid peroxidation accumulation and induced ferroptosis in LUAD. We therefore suggest that RRM2 may play an important role in LUAD(Fig. 10).

Inhibition of ribonucleotide reductase subunit M2 (RRM2) reduces the level of deoxyribonucleotide triphosphate (dNTP), which is the basis for stable senescence-related cell growth arrest induced by oncogene. And ribonucleotide reductase subunit M2 is a rate-limiting protein for dNTP synthesis[18].Knocking down of RRM2 significantly decreased proliferation during S phase of cell cycle[19, 20]. Rhhman's study found that RRM2 consumption significantly reduced the expression of Bcl-2 protein (Bcl-2 is a key determinant of cell apoptosis)[21]. In contrast, Jin et al. found that by overexpressing RRM2, the activity of the Bcl-2 signaling pathway was increased and the activity of the p53 signaling pathway was decreased[22].Therefore, RRM2 plays an important role in S phase during DNA replication, which may have potential therapeutic significance. Zhong and Li studied the potential of RRM2 as a therapeutic target for gastric cancer and glioblastoma, respectively [23, 24]. Human papillomavirus E7 oncoprotein increases the expression of RRM2 to promote angiogenesis in cervical cancer, and inhibiting RRM2 activity may be a new therapeutic strategy for human cervical cancer[25]. Even Shao developed RRM2 inhibitor to treat breast cancer[26, 27]. Using gambogic acid to reduce the expression of RRM2 can improve the efficacy of gemcitabine in pancreatic cancer, gambogic acid makes

pancreatic cancer cells sensitive to gemcitabine in vitro and in vivo by inhibiting the activation of ERK/E2F1/RRM2 signaling pathway[28].

There are still few studies on RRM2 in lung cancer. RRM2, highly expressed in lung cancer and is related to poor prognosis, regulates the immune microenvironment of LUAD, which is consistent with our results[10]. Cai et al. detected significant changes in RRM2 methylation in NSCLC patients by bioinformatics, and found a significant correlation between RRM2 expression and multiple immunocytes infiltrations[29]. Knockdown of RRM2 suppressed tumor growth in the xenograft tumor models. RRM2 deficiency increased CD8 + T cells in the tumor tissues and spleens[30].Bufu Tang et al. found that the expression level of RRM2 was positively correlated with neutrophil and macrophage infiltration in LUAD tissue, suggesting that RRM2 promoted lung cancer progression and affected macrophage infiltration, stimulated M1 phenotype polarization and inhibited M2 phenotype[10].

However, our research has some limitations. Firstly, our Risk signature was built and validated based on TCGA and GEO databases, with no further external validation was done through real-world prospective LUAD RNAseq cohorts. Secondly, when analyzing the efficacy of immunotherapy, the LUAD sample size of GSE126044 cohort was relatively small. Although the positive results were obtained, the accuracy remains to be discussed. We are also trying to find or establish a better and larger immunotherapy cohort for LUAD patients. Thirdly, although we identified the key role of RRM2 in LUAD by bioinformatics technology, we have not yet verified how RRM2 affects the development of LUAD by regulating Mast cells and T cells CD4 memory or other mechanisms. Fourthly, we screened drugs for high-risk LUAD patients through CMAP and GDSC databases, but the specific efficacy and mechanism need to be further explored by cellular and animal experiments, and even clinical trials. Overall, we need more clinical samples and design prospective experiments to study the influence mechanism of RRM2 on LUAD immune microenvironment and immunotherapy.

5 Conclusions

We established a risk prediction signature based on FRGs. Initially, the signature was closely associated with the immune microenvironment of LUAD. The key risk gene RRM2 can influence the prognosis of LUAD by regulating the infiltration of Mast cells activated and T cells CD4 memory activated. What's more, FRGs-based LUAD clustering was associated with risk signature and immune cell infiltration. In other words, the FRG clustering we provided may be a potential LUAD immune subtype. Furthermore, correlation analysis proves the relationship between the signature and LUAD immunotherapy. In addition, our results also provide a new perspective that FRG such as RRM2 influence anti-PDL1 immunotherapeutic responses by regulating immune checkpoint expression and infiltration of immune cells in TME. Last but not least, RRM2 is an important biomarker affecting the long-term survival of patients with LUAD, and silencing RRM2 can promote the occurrence of ferroptosis in LUAD. Taken together, our study may offer novel insights into the research and treatment of LUAD.

Declarations

Acknowledgements

We thank the Public Service Platform of South China Sea for R&D Marine Biomedicine Resources for support.

Author Contributions

Lianxiang Luo conceived and designed the study; Biao Deng analyzed and interpreted the data; Jing Xiang carried out the experiment; Biao Deng, Jing Xiang and Lianxiang Luo wrote the manuscript; Lianxiang Luo and Zhu Liang reviewed the paper and provided comments, and all the authors reviewed the manuscript. All authors read and approved the final manuscript.

Funding

This project was supported by Basic and Applied Basic Research Program of Guangdong Province (2019A1515110201); The Science and technology program of Guangdong Province (2019B090905011); Discipline Construction Project of Guangdong Medical University (4SG21004G). Innovation and Entrepreneurship Team Leads the Pilot Program of Zhanjiang (2020LHJH005).

Data Availability Statement

The data used to support the findings of this study are included within the article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Conflicts of Interest

The authors declare that they have no competing interests.

References

1. Mattiuzzi C, Lippi G. Cancer statistics: a comparison between World Health Organization (WHO) and Global Burden of Disease (GBD). Eur J Public Health. 2020;30(5):1026–7.
2. Couraud S, Zalcman G, Milleron B, et al. Lung cancer in never smokers–a review. Eur J Cancer. 2012;48(9):1299–311.
3. Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. Lancet Oncol. 2015;16(7):e342-51.
4. Siegel R L, Miller K D, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7–30.

5. Lennon F E, Cianci G C, Cipriani N A, et al. Lung cancer-a fractal viewpoint. Nat Rev Clin Oncol. 2015;12(11):664–75.
6. Siegel R L, Miller K D, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018;68(1):7–30.
7. Torti S V, Torti F M. Iron and Cancer: 2020 Vision. Cancer Res. 2020;80(24):5435-48.
8. Wu Y, Zhang S, Gong X, et al. The epigenetic regulators and metabolic changes in ferroptosis-associated cancer progression. Mol Cancer. 2020;19(1):39.
9. Chen P, Wu Q, Feng J, et al. Erianin, a novel dibenzyl compound in Dendrobium extract, inhibits lung cancer cell growth and migration via calcium/calmodulin-dependent ferroptosis. Signal Transduct Target Ther. 2020;5(1):51.
10. Tang B, Xu W, Wang Y, et al. Identification of critical ferroptosis regulators in lung adenocarcinoma that RRM2 facilitates tumor immune infiltration by inhibiting ferroptotic death. Clin Immunol. 2021;232:108872.
11. Huang F, Zheng Y, Li X, et al. Ferroptosis-related gene AKR1C1 predicts the prognosis of non-small cell lung cancer. Cancer Cell Int. 2021;21(1):567.
12. Wang M, Mao C, Ouyang L, et al. Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. Cell Death Differ. 2019;26(11):2329–43.
13. Mao C, Wang X, Liu Y, et al. A G3BP1-Interacting lncRNA Promotes Ferroptosis and Apoptosis in Cancer via Nuclear Sequestration of p53. Cancer Res. 2018;78(13):3484–96.
14. Luo L, Wang H, Tian W, et al. Targeting ferroptosis for cancer therapy: iron metabolism and anticancer immunity. Am J Cancer Res. 2021;11(11):5508–25.
15. Ding L, Getz G, Wheeler D A, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008;455(7216):1069–75.
16. Mou Y, Wang J, Wu J, et al. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. J Hematol Oncol. 2019;12(1):34.
17. Hirsch F R, Scagliotti G V, Mulshine J L, et al. Lung cancer: current therapies and new targeted treatments. Lancet. 2017;389(10066):299–311.
18. Aird K M, Zhang G, Li H, et al. Suppression of nucleotide metabolism underlies the establishment and maintenance of oncogene-induced senescence. Cell Rep. 2013;3(4):1252–65.
19. Aird K M, Li H, Xin F, et al. Identification of ribonucleotide reductase M2 as a potential target for pro-senescence therapy in epithelial ovarian cancer. Cell Cycle. 2014;13(2):199–207.
20. Nordlund P, Reichard P. Ribonucleotide reductases. Annu Rev Biochem. 2006;75:681–706.
21. Rahman M A, Amin A R, Wang D, et al. RRM2 regulates Bcl-2 in head and neck and lung cancers: a potential target for cancer therapy. Clin Cancer Res. 2013;19(13):3416–28.
22. Jin C Y, Du L, Nuerlan A H, et al. High expression of RRM2 as an independent predictive factor of poor prognosis in patients with lung adenocarcinoma. Aging (Albany NY). 2020;13(3):3518–35.
23. Altinkilic E M, Isbir S, Gormus U, et al. RRM1, RRM2 and ERCC2 Gene Polymorphisms in Coronary Artery Disease. In Vivo. 2016;30(5):611–5.

24. Castelblanco E, Zafon C, Maravall J, et al. APLP2, RRM2, and PRC1: New Putative Markers for the Differential Diagnosis of Thyroid Follicular Lesions. Thyroid. 2017;27(1):59–66.
25. Wang N, Zhan T, Ke T, et al. Increased expression of RRM2 by human papillomavirus E7 oncoprotein promotes angiogenesis in cervical cancer. Br J Cancer. 2014;110(4):1034–44.
26. Shao J, Zhou B, Chu B, Yen Y. Ribonucleotide reductase inhibitors and future drug design. Curr Cancer Drug Targets. 2006;6(5):409–31.
27. Wadler S, Makower D, Clairmont C, et al. Phase I and pharmacokinetic study of the ribonucleotide reductase inhibitor, 3-aminopyridine-2-carboxaldehyde thiosemicarbazone, administered by 96-hour intravenous continuous infusion. J Clin Oncol. 2004;22(9):1553–63.
28. Xia G, Wang H, Song Z, et al. Gambogic acid sensitizes gemcitabine efficacy in pancreatic cancer by reducing the expression of ribonucleotide reductase subunit-M2 (RRM2). J Exp Clin Cancer Res. 2017;36(1):107.
29. Cai X, Lin L, Zhang Q, et al. Bioinformatics analysis of the circRNA-miRNA-mRNA network for non-small cell lung cancer. J Int Med Res. 2020;48(6):300060520929167.
30. Jiang X, Li Y, Zhang N, et al. RRM2 silencing suppresses malignant phenotype and enhances radiosensitivity via activating cGAS/STING signaling pathway in lung adenocarcinoma. Cell Biosci. 2021;11(1):74.

Figures

Figure 1

Construction of 7 - gene Risk Signature. (A)The heatmap of FRDEGs between normal and tumor samples. (B)The forest plot of univariate Cox regression. (C) Differential expressed volcanic maps. The positions of 7 key FRGs in the volcanic map. (D, E) Forest plots show independent prognostic roles of RiskScore and other clinical features.

Figure 2

Verification of 7-gene Risk Signature and Construction of Nomo Graph. (A,B) Survival analysis between the high- and low-risk groups based on OS. (C) ROC curve of the 7-gene signature model. (D) 1-year survival rate predicted by 7-gene signature and other clinical features was compared in ROC curve;(E) Decision curve. (F)Nomogram for predicting 1,3,5-year survival rates in LUAD patients.

Figure 3

Difference of immune related indexes between LUAD patients. (A,B,C)Violin Chart. Differences in StromalScore, ImmuneScore and EstimateScore between high- and low-risk patients. (D, E, F)Differences in StromalScore, ImmuneScore and EstimateScore among the four FRGs clusters. (G)The proportion of risk scores of four FRGs clusters. (H)Difference inTMB among high- and low-risk groups.

Figure 4

Difference and correlation of immunocyte infiltration in LUAD patients.(A)Immune cell infiltration heatmap, 7 algorithms to assess immune cell infiltration level among the high-and low-risk groups. (B)Different OS based on different levels of ESTIMATEScore. (C)Correlation heatmap, correlation of infiltration levels of immune cells in LUAD.

Figure 5

Analysis of Immunocyte Clustering and Immunotherapy Related Indexes in LUAD Patients.(A)Clustering of immunocytes infiltration. TCGA-LUAD cohort was divided into two groups with high-and low-infiltration levels according to 24 immunocytes;(B) Proportion of high- and low-infiltration of immunocytes in four FRGs clusters; (C)The difference of immune-related function scores between high- and low-risk groups; (D, E, F, G)Violin Chart. Differences in Four Immunotherapy Related scores between high- and low-risk Groups.

Figure 6

Seven FRGs regulate the infiltration level of prognosis-related immune cell.(A)Immunocytes which play an independent prognostic role in LUAD patients. (B) Correlation between the expression levels of 7 key FRGs and the infiltration levels of immunocytes. (C)Immunocytes whose infiltration level correlated with RRM2 expression. (D) Different infiltration levels of Mast-cell represent different LUAD OS.

Figure 7

SevenFRGs regulate the expression level of immune checkpoint in LUAD patients. (A,B)The change trend of the infiltration level of Mast-cells and T cells CD4 memory activated with the expression of RRM2. (C)Different Check-point levels represent different LUAD OS. (D)27 common checkpoint genes are

differentially expressed between high and low risk groups. (E, F) The expression of ROS1 and ALK change with risk score. (G, H) Correlation of the expression of ROS1 and ALK with 7 key FRGs.

Figure 8

The expression of RRM2 in normal lung tissue and lung cancer tissue was detected by immunohistochemistry. 674, 677, 712, 728, 733 and 734 are patient numbers. Expression of

RRM2 in LUAD tissues was higher than that in normal tissues.

Figure 9

Effects of RRM2 silencing on ferroptosis. (A)Expression level of RRM2,4HNE and ACSL4 in different lung cancer cell lines by Western blotting. (B)The results of cell death level after silencing RRM2 by flow cytometry. (C)The fluorescence intensity of cells after silencing RRM2 was evaluated using a rotating disk super-resolution laser confocal microscope. (D)Detection of cell lipid peroxidation level after silencing RRM2 by flow cytometry.

Figure 10

The flowchart of this study.

Supplementary Files

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

- SupplementaryFigure1.jpg
- SupplementaryFigure2.jpg
- SupplementaryFigure3.jpg
- SupplementaryFigure4.jpg
- SupplementaryFigure5.jpg
- SupplementaryFigure6.jpg