

RESEARCH

Open Access



Tumor microenvironment-associated lactate metabolism regulates the prognosis and precise checkpoint immunotherapy outcomes of patients with lung adenocarcinoma

Song Qiu^{1†}, Ying Wang^{1†}, Hui Rao^{1†}, Qiuyang Que¹, Yanyang Wu¹, Rui Zhu¹, Xiaofei Feng¹, Jun Chi¹, Weiling Lai¹, Yihang Sun², Qi Xiao³, Huaqiu Shi^{1*} and Yi Xiang^{1*}

Abstract

Background: Despite the wide clinical application of checkpoint inhibitor immunotherapy in lung adenocarcinoma, its limited benefit to patients remains puzzling to researchers. One of the mechanisms of immunotherapy resistance may be the dysregulation of lactate metabolism in the immunosuppressive tumor microenvironment (TME), which can inhibit dendritic cell maturation and prevent T-cell invasion into tumors. However, the key genes related to lactate metabolism and their influence on the immunotherapeutic effects in lung adenocarcinoma have not yet been investigated in depth.

Methods: In this study, we first surveyed the dysregulated expression of genes related to lactate metabolism in lung adenocarcinoma and then characterized their biological functions. Using machine learning methods, we constructed a lactate-associated gene signature in The Cancer Genome Atlas cohort and validated its effectiveness in predicting the prognosis and immunotherapy outcomes of patients in the Gene Expression Omnibus cohorts.

Results: A 7-gene signature based on the metabolomics related to lactate metabolism was found to be associated with multiple important clinical features of cancer and was an independent prognostic factor.

Conclusions: These results suggest that rather than being simply a metabolic byproduct of glycolysis, lactate in the TME can affect immunotherapy outcomes. Therefore, the mechanism underlying this effect of lactate is worthy of further study.

Keywords: Lactate metabolism, Gene signature, Risk score, Prognosis, Immunotherapy benefit, Lung adenocarcinoma

Introduction

An important physiological function of glucose is to provide energy for cell survival. In the human body, glucose is metabolized through three main routes: anaerobic glycolysis, the pentose phosphate pathway, and mitochondrial oxidative phosphorylation [1, 2]. However, metabolic reprogramming is common in tumorigenesis and cancer development, endowing the tumor cells with multiple competitive advantages for survival and

[†]Song Qiu, Ying Wang and Hui Rao have contributed equally to this work

*Correspondence: shq3677274@163.com; xiangyi_xiangyi@126.com

¹ Department of Oncology, The First Affiliated Hospital, Gannan Medical University, No 23, Qingnian Road, Ganzhou, China
Full list of author information is available at the end of the article



progression [3–5]. One such metabolic change is the increase in lactate production due to the Warburg effect of aerobic glycolysis [6–8]. Lactate metabolism in tumor cells is significantly different from that in normal cells [9]. Unlike normal cells that gain energy by metabolizing glucose to pyruvate, which is then transported into the mitochondria for full oxidative phosphorylation, tumor cells rely mainly on aerobic glycolysis for energy and producing precursors for protein, lipid, and nucleotide synthesis [4, 10, 11]. In tumor cells, glucose is first catalyzed to pyruvate by enzymes such as hexokinase, 6-phosphofructokinase-1, and pyruvate kinase [12]. Finally, the high expression of lactate dehydrogenase (LDH) and pyruvate dehydrogenase kinase in tumor tissue promotes the reduction of pyruvate to lactate [13].

This highly efficient aerobic glycolysis can produce a large amount of lactate, which can then be secreted into the extracellular space to acidify the tumor microenvironment (TME) [14, 15]. In the past, lactate was considered to be only metabolic waste as a byproduct of glycolysis. However, an increasing number of studies have found that lactate can activate many important signaling pathways in tumor cells to promote the survival, invasion, immune escape, metastasis, and angiogenesis of multiple types of cancer [16–20]. Baumann et al. [21] demonstrated that lactate could induce the upregulated expression of transforming growth factor-beta 2 (TGFβ2) to stimulate tumor invasion and metastasis. Other studies have shown that lactate acts as an immunosuppressive molecule by promoting the survival of regulatory T lymphocytes (Tregs) [19, 22, 23], suppressing the production of interferon-gamma (IFN-γ) in cytotoxic T cells [24, 25], inhibiting the toxicity and cytolytic functions of natural killer cells [26, 27], and facilitating M2 polarization of tumor-associated macrophages via the ERK/STAT3 signaling pathway [16, 28]. Additionally, the abnormal metabolism and excretion of lactate can suppress the degradation of hypoxia-inducible factor 1-alpha (HIF-1α) [29] and increase the production of vascular endothelial growth factor [30] and fibroblast growth factor [31] in endothelial cells, thereby promoting angiogenesis in the TME [20, 32, 33]. Therefore, more detailed studies on lactate metabolism and the key enzymes involved in the process may provide new insights into tumor pathogenesis as well as potential therapeutic targets.

To date, the effects of the dysregulation of lactate metabolism and related genes in the TME on the therapeutic outcomes and prognosis of cancer patients have not been studied. Moreover, there is a lack of research on the prognostic value of lactate metabolism in lung adenocarcinoma. Therefore, the objectives of this study were to identify the genes associated with lactate metabolism in lung adenocarcinoma and determine their expression

characteristics and biological significance. To this end, data were downloaded from The Cancer Gene Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases to identify the lactate metabolism-related genes in lung adenocarcinoma. Then, using machine learning tools, we constructed a prognostic risk score model based on lactate metabolism-associated genes screened from the TCGA dataset and evaluated its predictive effectiveness on the GEO patient cohorts. The 7-gene signature was found to be associated with multiple important clinical features of cancer. These results suggest that rather than being simply a metabolic byproduct of glycolysis, lactate in the TME can affect immunotherapy outcomes. Further elucidation of the mechanism underlying this lactate effect could provide a practical reference for the formulation of individualized treatment strategies.

Materials and methods

Data acquisition and pretreatment

The lung adenocarcinoma samples of The Cancer Genome Atlas (TCGA) database [34] downloaded from Xena (<https://tcga-xena-hub.s3.us-east-1.amazonaws.com/download/TCGA.LUAD.sampleMap%2FHiSeqV2.gz>) were used as the training set. GSE26939, a lung adenocarcinoma chipset [35], was downloaded from GEO database (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26939>) for verification (shown in Table 1). We converted the gene expression data using log₂-transformed quantile-normalized signal intensity. When there are multiple probes corresponding to the same gene, we take the median as the expression value. And then each probe was converted to corresponding gene symbol according to the labeling information.

Identification of lactate metabolism genes

Lactic acid, and lactate were used as the keyword to search for genes related to lactate metabolism in the GSEA database (<http://www.gsea-msigdb.org/gsea/login.jsp>). Meanwhile, lactate was used as the keyword

Table 1 The expression profile dataset in lung adenocarcinoma downloaded from online databases

Dataset ID	Platform	Tumor	Normal
TCGA-LUAD	Illumina	517	59
GSE26939	GPL9053	116	0
GSE190266	GPL24676	70	0
GSE135222	GPL16791	27	0
GSE136961	GPL24014	21	0
GSE81089	GPL16791	198	22
GSE101929	GPL570	33	33

on Genecard database (<https://www.genecards.org/>), and genes with score >20 were selected as supplementary gene sets. By taking the intersection of different gene sets, we finally obtained 26 candidate genes related to lactate metabolism (the gene symbols and NCBI ID are listed in Additional file 1: Table S1): GOBP_LACTATE_METABOLIC_PROCESS; GOBP_LACTATE_TRANSMEMBRANE_TRANSPORT; GOMF_LACTATE_DEHYDROGENASE_ACTIVITY; GOMF_LACTATE_TRANSMEMBRANE_TRANSPORTER_ACTIVITY.

Differential expression analysis

We then screened the differentially expressed genes (DEGs) of lung adenocarcinoma samples in the TCGA cohorts using R-package DESeq2. By using the Benjamini–Hochberg procedure to control the false discovery rate (FDR), 5530 DEGs were identified with a strict cut-off of $P < 0.01$ and an FDR of less than 0.05. And then 9 genes were screened out after the intersection of the 5530 DEGs with the candidate genes related to lactate metabolism.

Machine learning of prognostic gene signatures based on lactate metabolism

We used R-package *GLMNET* [36] to perform Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis [37] based on the expression matrix of lactate metabolism-related genes from the above univariate Cox regression, and found that the model has the highest accuracy when the degree of freedom (gene number) was 11 (Fig. 4A, B). Finally, the signature of 7 lactate metabolism-related genes was selected to assemble a prognostic risk score model (Fig. 4C). The lactate-related risk score of each patient was calculated by the formula: $\text{riskScore} = (-0.06275159 * \text{ACTN3 expression value} - 0.01796317 * \text{LDHD expression value} + 0.01581625 * \text{LDHA expression value} + 0.01821132 * \text{SLC16A3 expression value} + 0.04737492 * \text{SLC16A1 expression value} + 0.05394477 * \text{HAGH expression value} + 0.05710535 * \text{DARS2 expression value})$.

Immunotherapy cohorts and therapeutic benefit evaluation

In order to investigate the correlation between the 7-gene signature and clinical response, the sequencing and clinical data of patients with lung adenocarcinoma who received anti-PD-(L)1 immunotherapy in six different datasets [GSE190266 ($n=70$) [38], GSE135222 ($n=27$) [39], GSE136961 ($n=21$) [40], GSE81089 ($n=198$) [41], GSE101929 ($n=33$) [42] and the MSS Mixed Solid Tumors cohort, from Dana-Farber Cancer Institute, Boston, USA ($n=47$) [43]] were downloaded from GEO

database and cBioPortal database [44]. We next explored the potential effect of lactate metabolism on anti-PD-(L)1 immunotherapy in patients based on the short-term and long-term analysis. The short-term analysis focused on the tumor objective response, while the long-term analysis focused on the progression-free survival (PFS) and overall survival (OS) of patients after receiving treatments.

Statistical analysis

SPSS version 23.0 software (SPSS, Chicago, IL, USA) or R software (version 4.0.3) was used for statistical analysis. Chi-squared test was used for the clinicopathological parameters among different treatment groups. The log-rank test was used for the Kaplan–Meier survival of each group. Univariate and multivariate analysis were presented through Cox regressions. The reported results covered hazard ratios (HR) and 95% confidence intervals (CI). Curve analysis of the receiver operating characteristic (ROC) was applied for evaluating the predictive performance.

Results

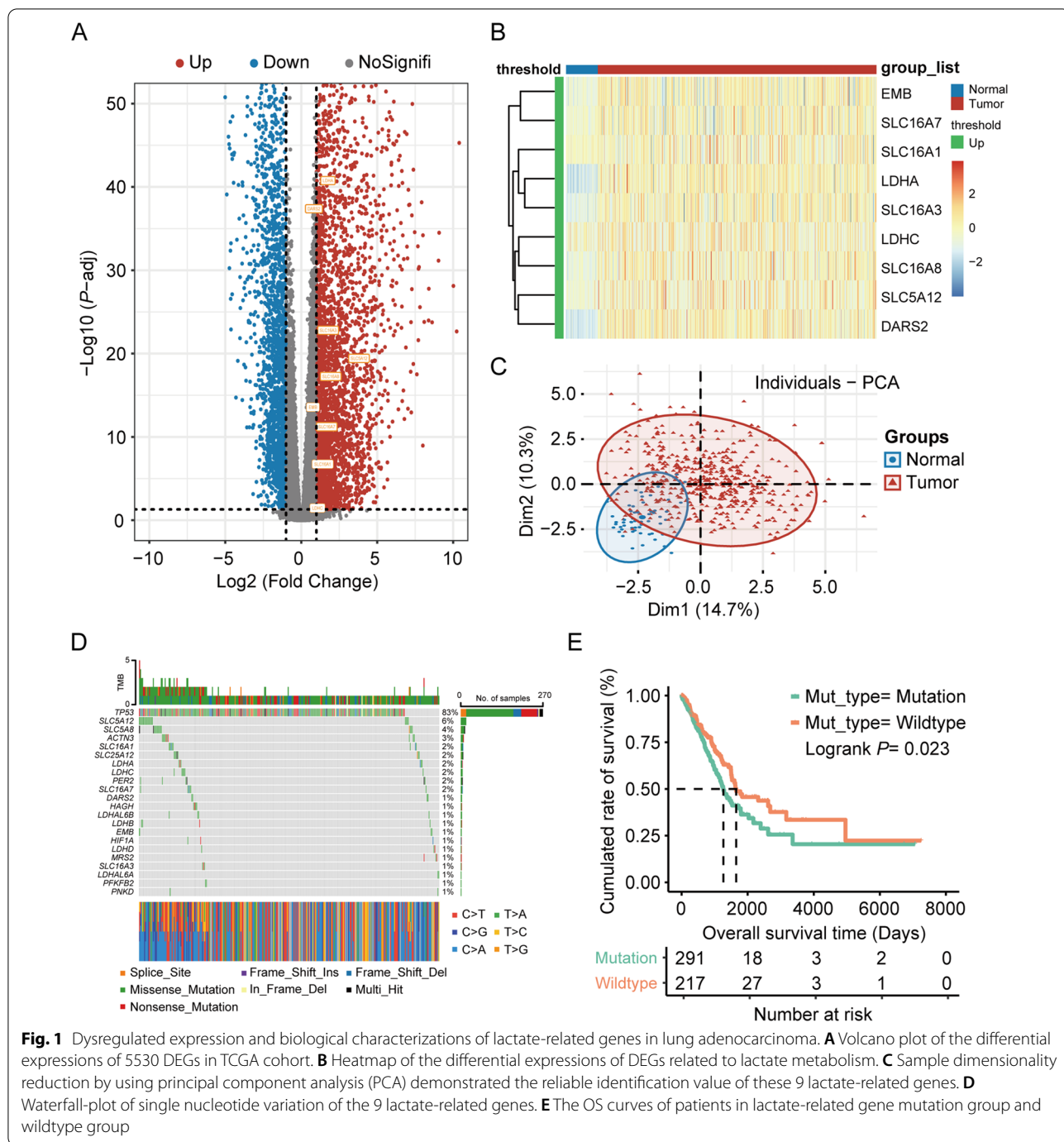
Dysregulated expression and biological characterization of genes related to lactate metabolism in lung adenocarcinoma

First, 5530 DEGs between the 517 cancer samples and 59 normal tissue samples from the TCGA database were extracted using the DESeq2 package. After intersection of these 5530 DEGs with the genes related to lactate metabolism (identified from the GSEA and GeneCards platforms), nine genes were screened for further study (Fig. 1A). The heatmap in Fig. 1B shows the differential expression of these nine genes in the TCGA cohort. To determine whether these lactate metabolism-related genes can distinguish cancer patients from healthy individuals, the sample dimensions of the genes were reduced using principal component analysis. As shown in Fig. 1C, these nine genes had reliable identification value.

Next, on the basis of the mutant MAF file from the TCGA database, a waterfall plot of single nucleotide variations of these nine genes was drawn using maftools in the R package (Fig. 1D) [45]. Subsequently, these genes were divided into mutation and wild-type groups for survival analysis. The results obtained using survminer in the R package demonstrated that the OS of the wild-type group was considerably longer than that of the mutation group (Fig. 1E).

Copy number variations of genes related to lactate metabolism in lung adenocarcinoma

We extracted the CNVs of all lactate metabolism-related genes to evaluate their copy status (Fig. 2A) and



then selected the top six genes with the most CNVs. As illustrated by the box plots (Fig. 2B–G), the following six genes were significantly differentially expressed between the normal and cancerous tissues: aspartyl-tRNA synthetase, mitochondrial (*DARS2*), embigin (*EMB*), lactate dehydrogenase A like 6A (*LDHAL6A*), 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase

2 (*PFKFB2*), solute carrier family 16 member 8 (*SLC16A8*), and tumor protein P53 (*TP53*).

Prognostic performance of the lactate metabolism-related genes in lung adenocarcinoma

Using the forestplot package, univariate Cox regression analysis was performed to evaluate the prognostic

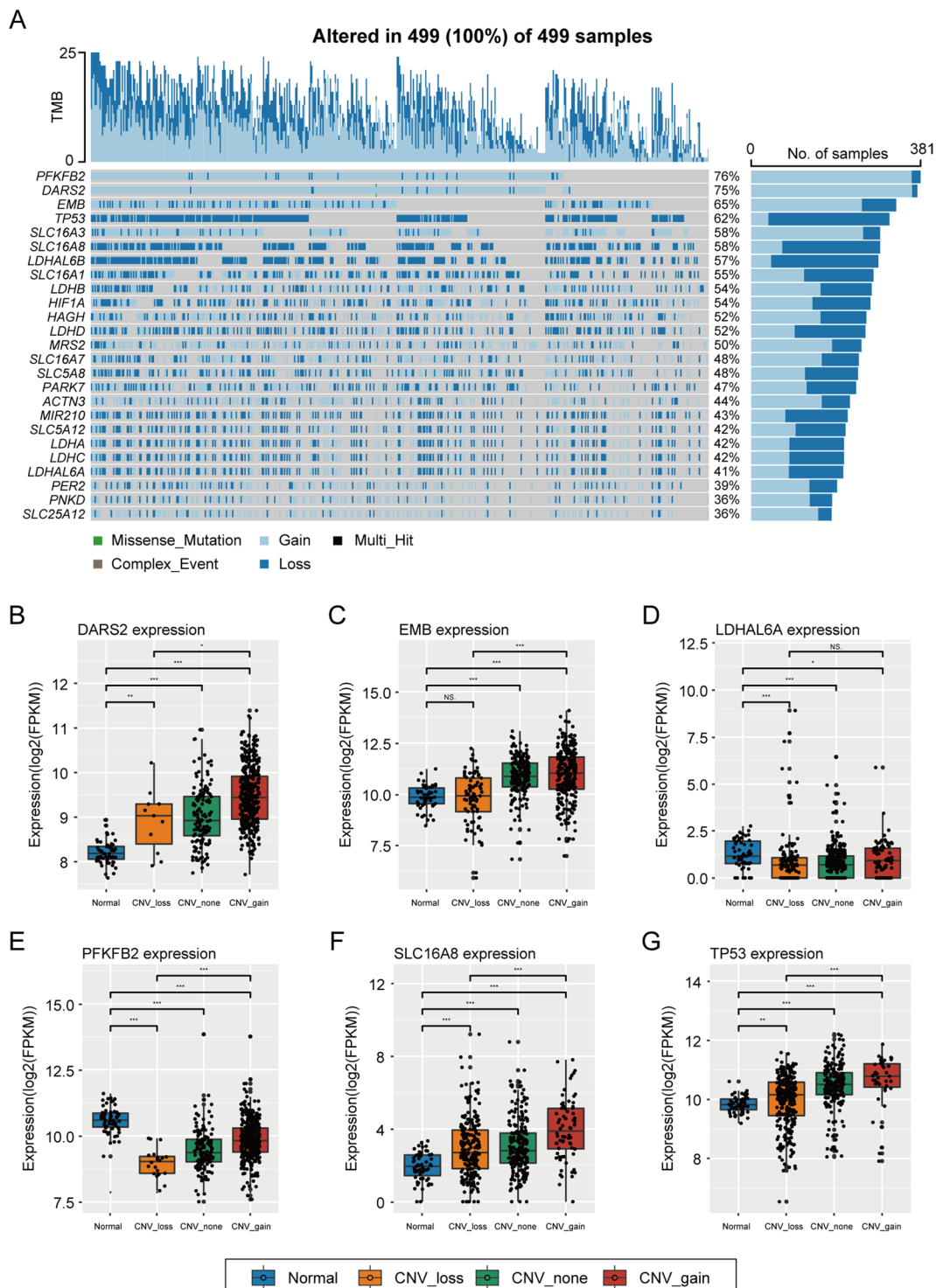


Fig. 2 Copy number variations (CNVs) of lactate-related genes in lung adenocarcinoma. **A** Waterfall-plot of the CNVs of these lactate-related genes. **B–G** Box plots showed the differential expression levels of the top 6 lactate-related genes with the most CNVs. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

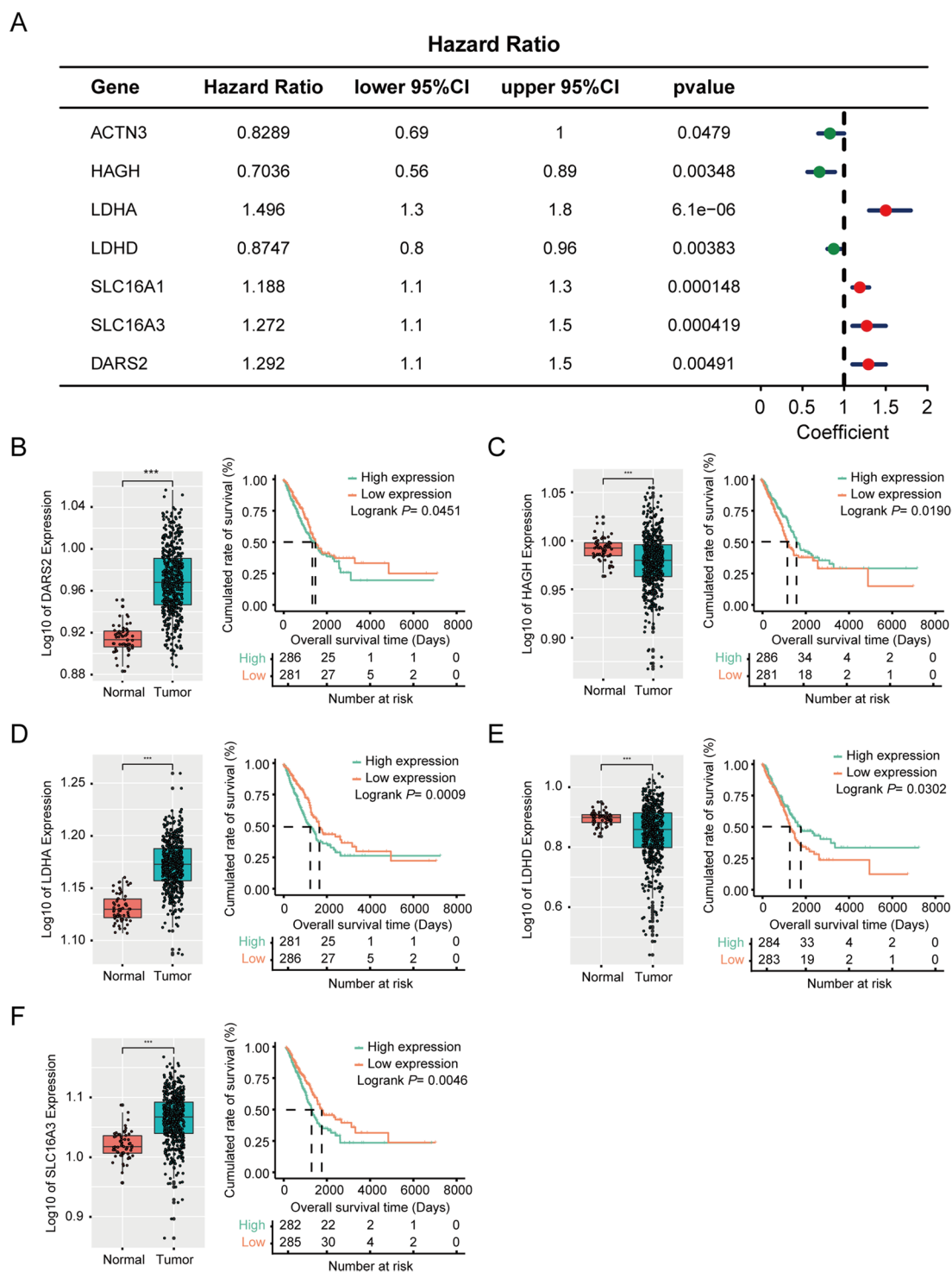
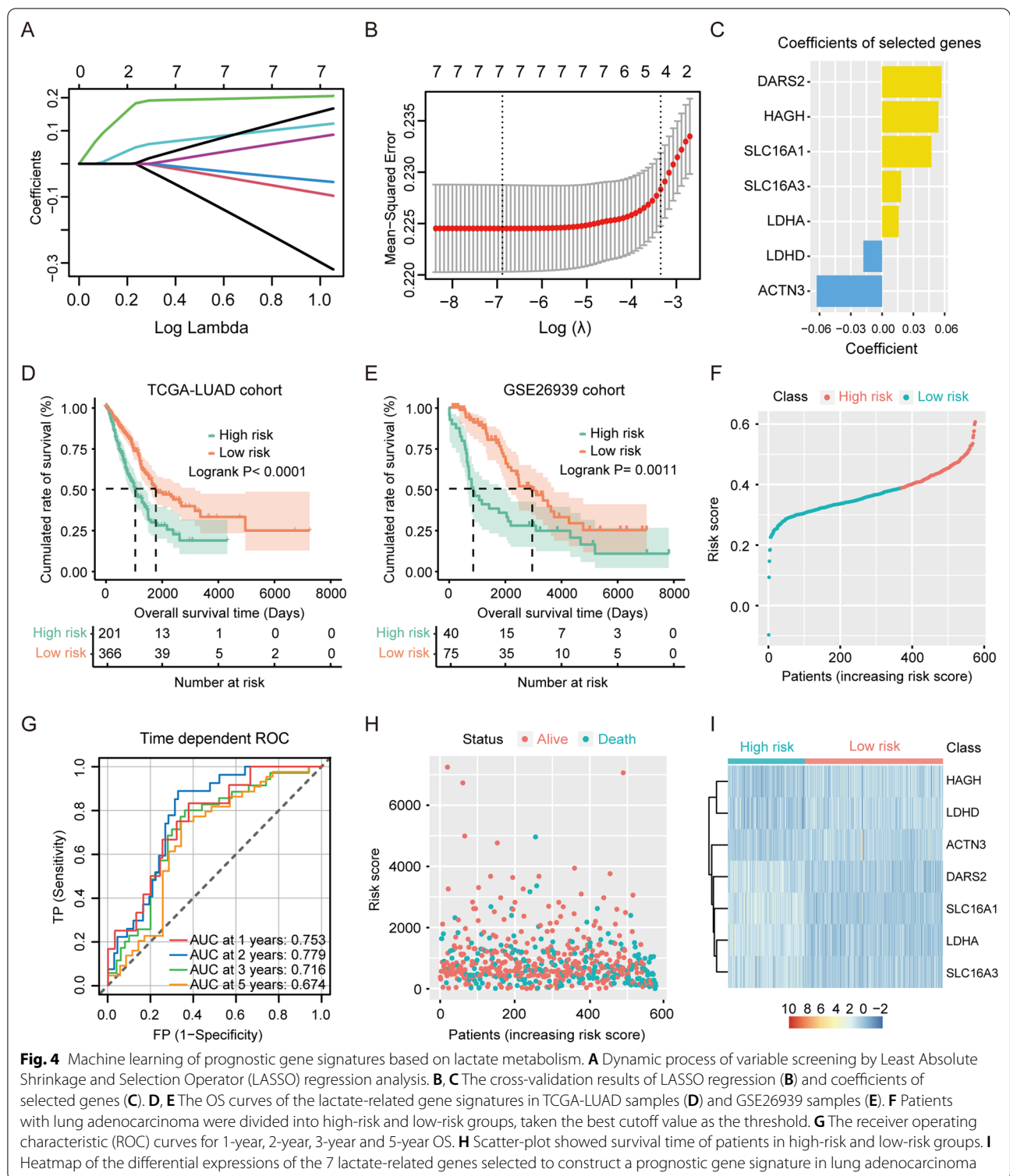


Fig. 3 Prognostic performance of lactate-related genes in lung adenocarcinoma. **A** Forest-plot showed the lactate-related genes associated with the prognosis of patients with lung adenocarcinoma analyzed by univariate Cox regression. **B–F** Box plot and Kaplan–Meier curve exhibited the differential expressions and prognostic performance of the top 5 significant lactate-related genes. * $P < 0.05$; ** $P < 0.01$, *** $P < 0.001$



performance of the lactate metabolism-related genes. The forest plot in Fig. 3A shows the genes associated with patient prognosis. The samples were divided into high and low expression groups according to the

median expression level. The box plot and Kaplan–Meier curve showed the following top five significant prognostic genes to be closely related to patient survival: *DARS2*, hydroxyacyl glutathione hydrolase (*HAGH*),

lactate dehydrogenase A (*LDHA*), lactate dehydrogenase D (*LDHD*), and *SLC16A3* (Fig. 3B–F).

Prognostic signature based on lactate metabolism-related genes screened from the TCGA cohort

By constructing a gene signature of seven lactate metabolism-related genes that were screened using machine learning tools and LASSO regression analysis [37] (Fig. 4A–C), we were able to explore the effect of dysregulated lactate metabolism on patient survival. Patients with lung adenocarcinoma were divided into high- and low-risk groups (with the best cutoff value taken as the threshold) to test the prognostic performance of the constructed gene signature on TCGA-LUAD (training cohort) and GSE26939 (validation cohort) samples (Fig. 4D–F). The areas under the ROC curves (AUCs) for 1-, 2-, 3-, and 5-year OS were 0.753, 0.779, 0.716, and 0.674, respectively (Fig. 4G). The scatter plot in Fig. 4H shows the survival time of patients in the high- and low-risk groups, whereas the heatmap in Fig. 4I displays the differential expression levels of the seven genes selected to construct the gene signature. Additionally, a multivariate Cox analysis (Table 2), performed using the `coxph()` function of survival in the R package, revealed the 7-gene signature to be an independent prognostic factor. These results demonstrated that this lactate metabolism-related gene signature could effectively distinguish the prognosis of patients with lung adenocarcinoma.

Performance of the lactate metabolism-related gene signature in predicting PD-1 immunotherapy efficacy in patients with lung adenocarcinoma

Based on our short- and long-term analyses of six datasets of lung adenocarcinoma immunotherapies downloaded from the GEO database, we found that lactate metabolism may play a potential role in promoting the

therapeutic effect of anti-PD-1 monoclonal antibodies. As shown in Fig. 5A, D, the PFS and OS of patients with high risk scores were significantly shorter than those of patients with low risk scores ($P=0.012$ and $P=0.003$, respectively). The overall response rates (ORRs) of the high- and low-risk groups were 22.5% and 40.9%, respectively (Fig. 5B). Patients who achieved an objective response also had a lower risk score than those who did not respond to treatment ($P=0.028$, Fig. 5C). Ultimately, the ROC curves and the calculated AUCs (Fig. 5E, F) indicate that the constructed gene signature is a promising tool for predicting mortality and recurrence in lung adenocarcinoma patients who received anti-PD-1 immunotherapy.

Correlation of the gene signature with multiple important clinical features of cancer

Figure 6A shows the correlations between the lactate metabolism-related gene signature and common clinical characteristics of patients with lung adenocarcinoma, such as age, sex, TNM stage, and prior radiation therapy. Furthermore, we used the RMSE function in the R package to construct nomograms for guiding clinical practice [46] according to these clinical characteristics (Fig. 6B). The prediction outcomes for 1-, 2-, 3-, and 5-year survival are presented in Fig. 6C–F.

Relationship between lactate metabolism and the tumor immune microenvironment

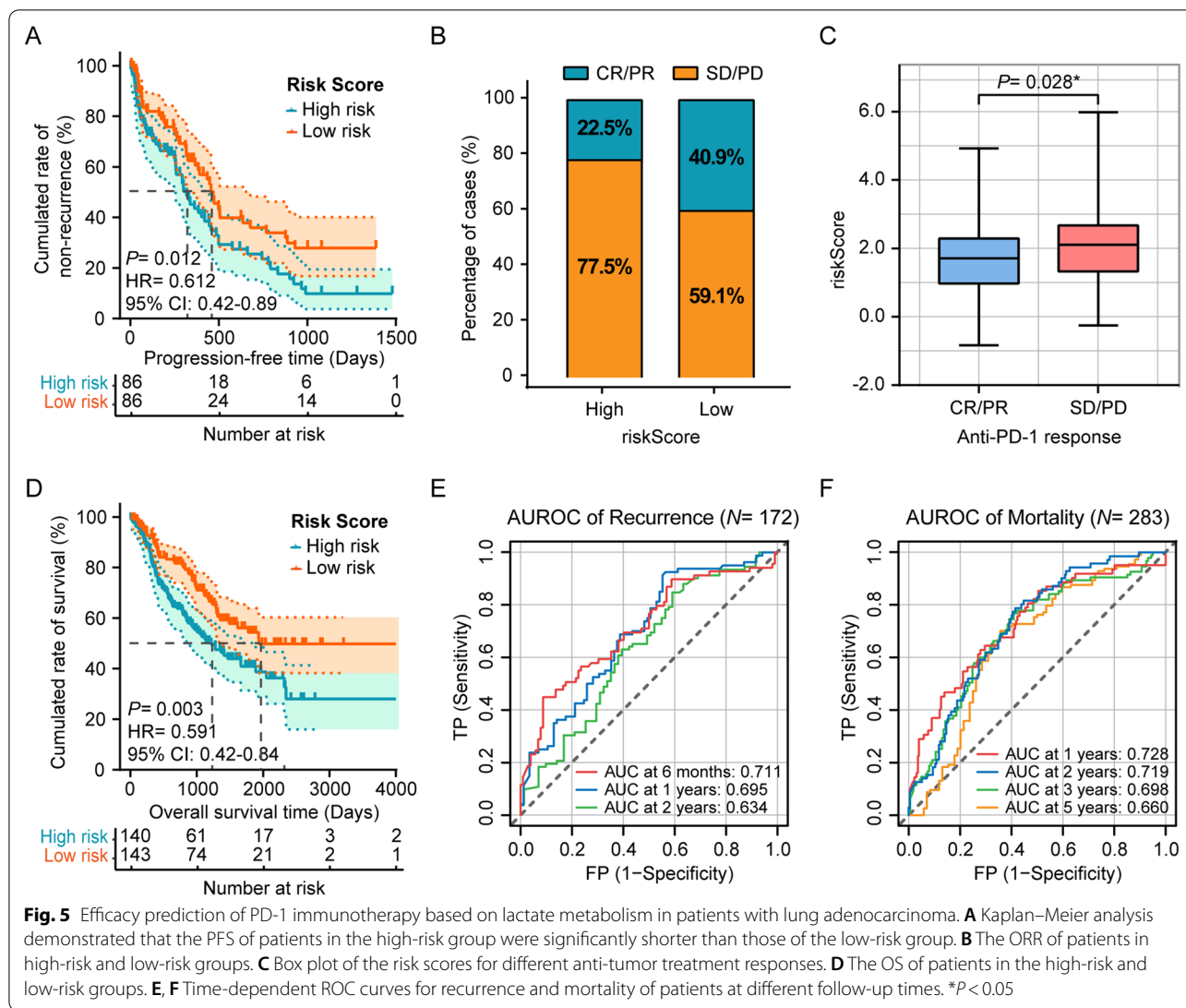
We used CIBERSORT (<https://cibersortx.stanford.edu/runcibersortx.php>) [47] to calculate the proportion of each type of immune cell in the samples to assess the association between lactate metabolism and tumor immune microenvironment. The box plot in Fig. 7 shows the proportions of various infiltrating immune cells between the high and low risk-score groups. The

Table 2 Multivariate Cox regression analysis for survival in patients with lung adenocarcinoma

Variables	Multivariate analysis			P-value
	Hazard ratio	Lower 95%CI	Upper 95%CI	
Lactate-related risk score	340.00	40.00	2900.00	< 0.0001***
Age (≥ 60 vs. < 60)	1.20	0.86	1.60	0.30
Gender (male vs. female)	0.98	0.74	1.30	0.87
Pathological stage (III–IV vs. I–II)	2.10	1.00	4.10	0.036*
Distant metastasis (M1 vs. M0)	0.82	0.37	1.80	0.62
Lymph node metastasis (N2–3 vs. N0–1)	0.92	0.47	1.80	0.80
Tumor invasion (T1–2 vs. T3–4)	0.72	0.11	4.60	0.73
Previous radiation therapy (yes vs. no)	1.50	1.10	2.20	0.02

CI confidence interval

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$



differential expression of several immune checkpoint genes in each group was analyzed using ggplot2 in the R package (Fig. 8A). And the correlation coefficients between genes in our signature and immune checkpoint genes are exhibited in Fig. 8B. Finally, the gene sets of 50 cancer hallmark pathways were used to calculate the GSVA scores for the expression matrix. Spearman’s correlation analysis was performed to determine the relationships between the GSVA scores and the risk score of these samples calculated with the constructed gene signature. The correlations between the gene signature and the 50 cancer hallmark pathways are presented in Additional file 2: Fig. S1.

Discussion

Increasing lines of evidence have shown that the occurrence, development, and prognosis of multiple cancer types are closely related to the TME [3, 48, 49]. As an important stress metabolite, lactate is produced through aerobic glycolysis in tumor cells, where its accumulated level could affect the expression of the cancerous cells as well as T cells, dendritic cells, and macrophages in the TME, thereby triggering intracellular signals [22, 23, 29, 50]. Lactate also has different effects on the biological processes of tumors, including their proliferation and migration, immune escape, and prognosis, as well as on the curative effects of various treatments. It was shown in

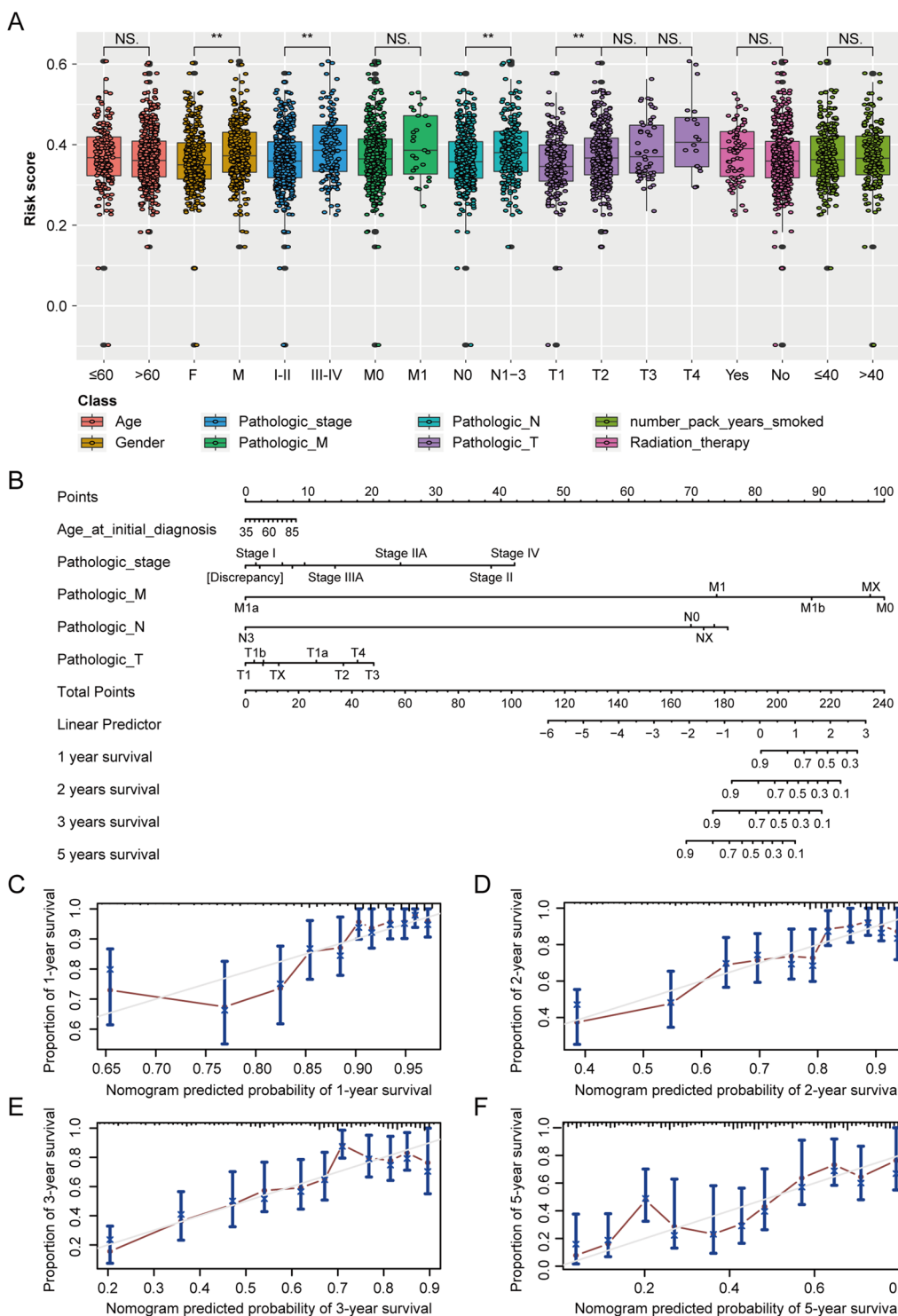
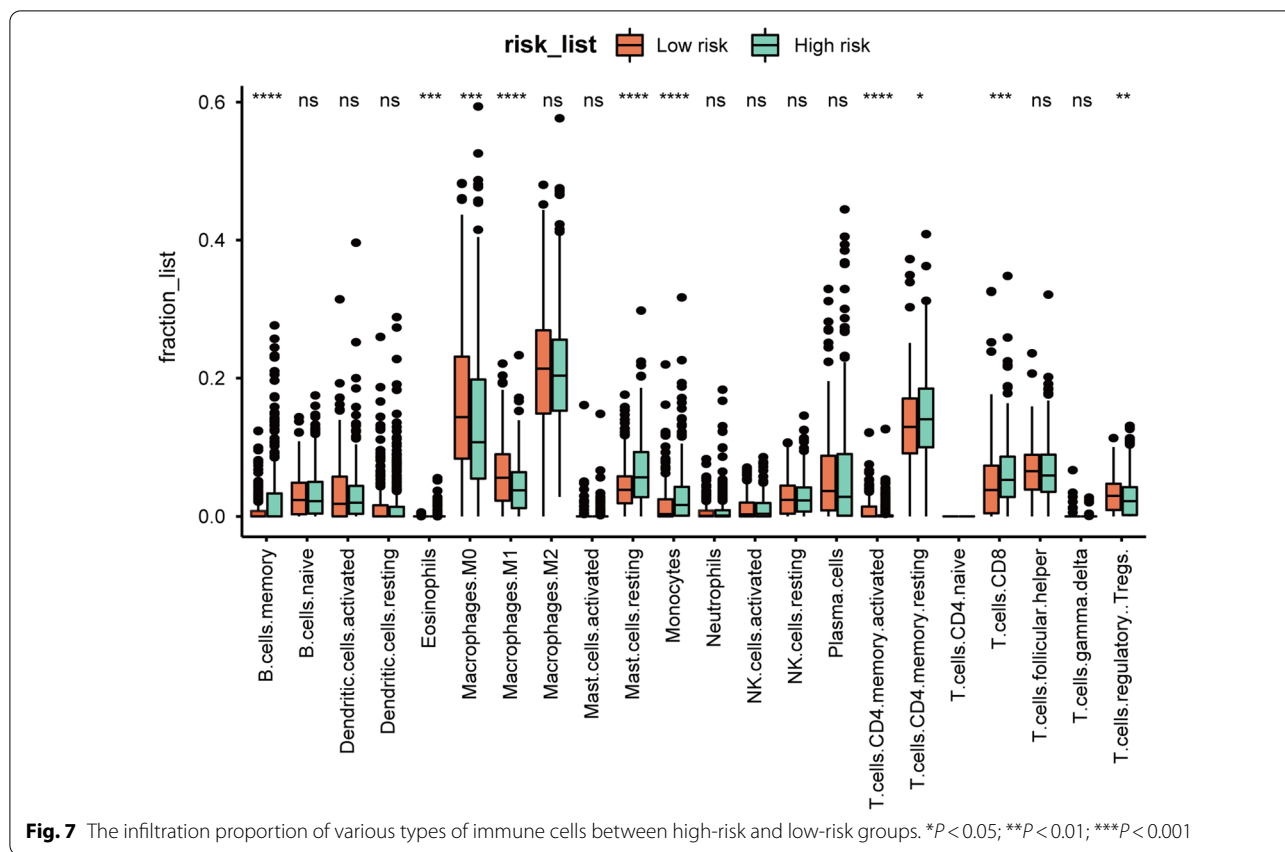


Fig. 6 The lactate-related gene signature is associated with multiple important clinical features of cancer. **A** Box plots exhibited the correlations between the lactate-related gene signatures and various clinical features in lung adenocarcinoma. **B** The nomograms constructed based on the characteristics for guiding clinical practice. **C–F** The prediction outcomes of nomograms for 1-year, 2-year, 3-year and 5-year survival



a previous study that an overflow of proton-coupled lactate in cancerous or stromal cells could promote tumor progression by regulating various processes in the TME (including cell invasion, angiogenesis, survival signaling, metastasis development, and immune surveillance avoidance) [51]. Simultaneously, the acidic microenvironment formed by a large amount of lactate accumulation is conducive to tumor cell metastasis and angiogenesis [32, 52].

Many studies have pointed out that the lactate content in the TME is negatively correlated with the tumor survival rate [53, 54]. Thus, research on abnormal lactate metabolism has long-term prospects. Most of earlier studies about lactate metabolism genes either focus on the overall survival of patients with various cancer types, or explore their potential relationship with tumor immune microenvironment through evaluating the correlation with the expression of common immune biomarkers. However, the identification and performance of lactate metabolism-related gene signatures in predicting the clinical response and long-term benefits for patients receiving checkpoint immunotherapy is still not studied in depth. In this study, patients were stratified into two clusters with the median risk score as the cut-off point. We identified and validated the gene signature based on lactate metabolism in predicting immunotherapeutic

response in lung adenocarcinoma. These results suggest that the lactate in TME can affect immunotherapy outcomes, rather than simply being a metabolic byproduct of glycolysis. The underlying mechanism needs further research and might provide future directions for formulating individualized treatment strategy (Additional file 3).

Machine learning algorithm is an intelligent tool that explores and simulates the law of human intelligent activities based on computer technology. Recently, as high-throughput sequencing has evolved, machine learning algorithms are gaining traction in medical research [55]. Stratifying tumor patients by bioinformatics and machine learning approaches to explore new biomarkers has proven to be reliable and useful. At present, there are few reports on the influence of dysregulated lactate metabolism on the immunotherapy response in lung cancer using these algorithms. LASSO analysis involved in our study is a powerful regression machine learning algorithm for processing high-dimensional data and feature selection. And the prognostic gene signature screened and developed by LASSO regression had been further evaluated by ROC analysis. In general, the combination of machine algorithm and clinical medicine may become a trend in future clinical research. With its high accuracy

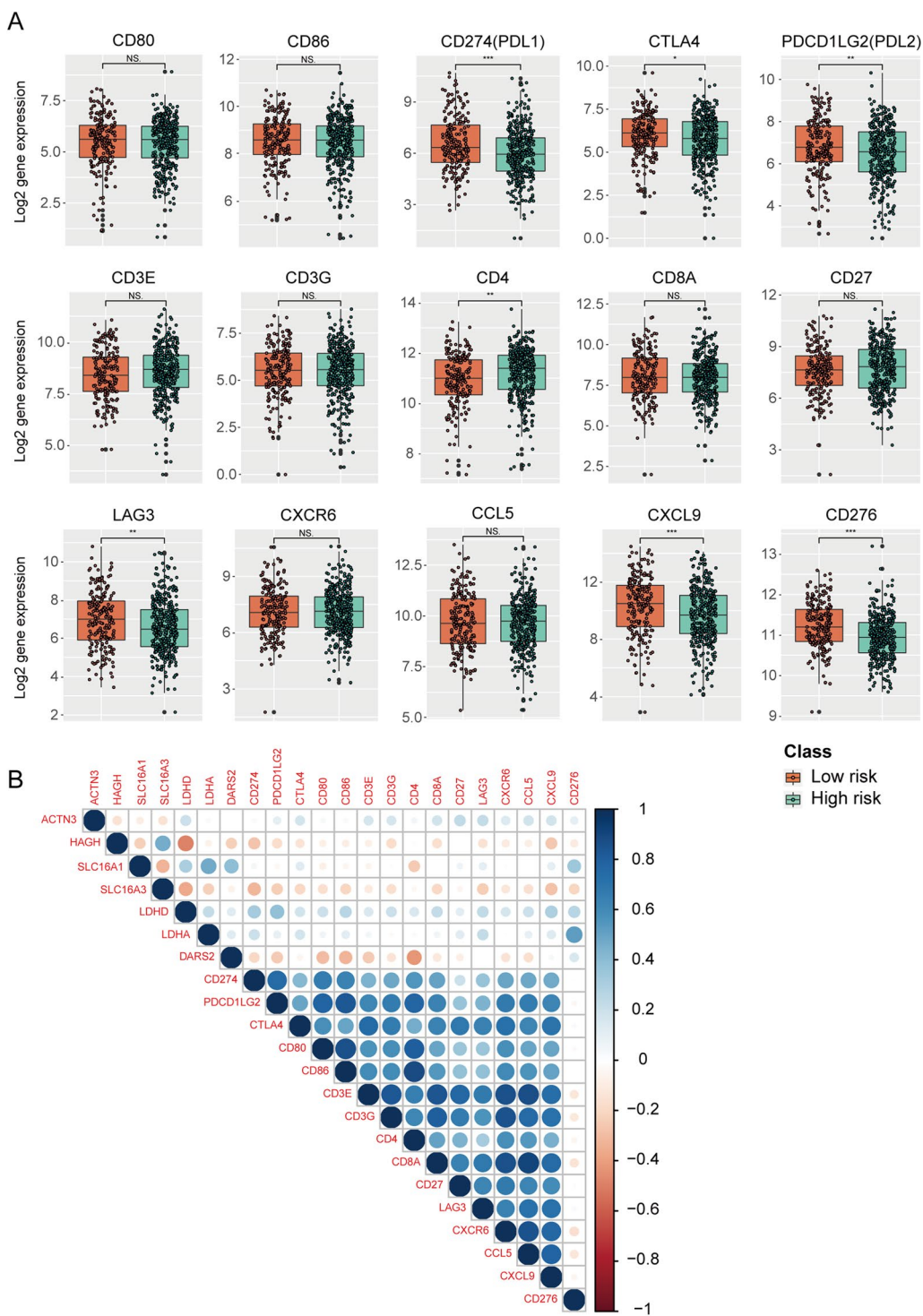


Fig. 8 The relationship between lactate metabolism-related genes and immune checkpoint genes. **A** Box plots of the differential expression levels of several representative immune checkpoint genes between high-risk and low-risk groups. **B** Correlation analysis between genes in the signature and immune checkpoint genes (color of dots represents correlation coefficient). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

and operability, these algorithms can greatly improve the efficiency of clinical practice, and provide a new perspective for diseases research.

The gene signature constructed in this study was based on seven differentially expressed lactate metabolism-related genes screened from TCGA cancer samples: actinin alpha 3 (*ACTN3*), *SLC16A1*, *SLC16A3*, *LDHA*, *LDHD*, *HAGH*, and *DARS2*. *ACTN3*, a structural protein of the Z-line of fast skeletal muscle fibers [56], maintains the orderly arrangement and normal contraction of muscle fibers by cross-linking with fine muscle filaments. *SLC16A1* and *SLC16A3* belong to the solute carrier protein 16A family, also known as the monocarboxylic acid transporter family [57]. They mediate the transmembrane transport of monocarboxylic acids dominated by lactate and short-chain fatty acids. Tumor cells can promote lactate production through the enzyme *LDHA*, thereby destroying tumor-infiltrating T cells as well as IFN- γ in natural killer cells and other cytokines to promote epithelial–mesenchymal transformation, angiogenesis, and invasion [26]. With regard to the human *HAGH* gene, its extended transcripts have been shown to encode both cytosolic and mitochondrial isoforms of glyoxalase II [58]. As for the *DARS2* gene, its abnormal expression is reportedly involved in the occurrence and development of primary liver carcinoma, with the expression level being closely related to the tumor size, stage, and prognosis [59].

Notably, the present study has certain limitations. First, the specific molecular functions of the genes involved in the constructed lactate metabolism model were not clear. Further research is needed to clarify their expression levels and roles in lung adenocarcinoma. Additionally, more clinical practice data are needed to verify the predictive value of this risk score model in the real-world setting.

Conclusion

In conclusion, our novel gene signature based on genes related to lactate metabolism has been proven effective in predicting the prognosis and immunotherapy outcomes of patients with lung adenocarcinoma. Our model provides a practical reference for improving the prediction of the efficacy of individualized immunotherapies and guiding clinical practice.

Abbreviations

TME: Tumor microenvironment; TCGA: The cancer genome atlas; GEO: Gene expression omnibus; LDH: Lactate dehydrogenase; PDH: Pyruvate dehydrogenase kinase; HAGH: Hydroxyacyl glutathione hydrolase; Tregs: T lymphocytes; IFN- γ : Interferon γ ; TAMs: Tumor-associated macrophages; VEGF: Vascular endothelial growth factor; FGF: Fibroblast growth factor; DEGs: Differentially expressed genes; FDR: False discovery rate; GSVA: Gene set variation analysis; LASSO: Least absolute shrinkage and selection operator; HR: Hazard ratios; CI:

Confidence interval; CMVs: Copy number variations; CR: Complete response; PR: Partial response; SD: Stable disease; PD: Progressive disease; ORR: Overall response rate; PFS: Progression-free survival; OS: Overall survival; ROC: Receiver operating characteristic; AUC: Area under the curve; RMSE: Root mean square error.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40001-022-00895-6>.

Additional file 1: Table S1. 26_lactate_metabolism_related_genes.

Additional file 2: Fig. S1. The correlation between the lactate metabolism-related gene signatures and GSVA scores of 50 cancer hallmark pathways.

Additional file 3: Raw documents and tables in the process of data analysis.

Acknowledgements

We would like to thank American Journal Experts (<https://www.aje.com/go/springernature/>) for English language editing.

Author contributions

YX designed the study and take responsibility for the integrity of the data and the accuracy of the data analysis. SQ and YW performed the database analysis from TCGA and GEO. HR performed the clinical data analysis from NSCLC patients. Cox and LASSO regression were carried out by QX and ZL. YW, HR and QQ did the construction and validation of prognostic models. YX, YS and RZ wrote the manuscript, and XF, WL and JC revised the manuscript. All these works were planned and supervised by HS and YX. All authors read and approved the final manuscript.

Funding

This study was supported by the Scientific Research Foundation of Gannan Medical University (2021DQ095 to Yi Xiang), and Beijing Medical and Health Foundation (B183017).

Availability of data and materials

The bioinformatics analysis data utilized in this study are publicly available and listed below: the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>); TCGA (<https://cancergenome.nih.gov>). And the clinical and prognostic data generated and analyzed during the current study are available in supplementary materials or from the corresponding author on reasonable request.

Code availability

All code supporting the current study is deposited in GitHub (<https://github.com/xyouli/DNA-based-predictors-of-non-genetic-cancer-phenotypes>). All computational analyses are done using public R packages.

Declarations

Ethics approval and consent to participate

The collection of human tissue samples or clinical data were not involved in this study. All data was downloaded from network open databases.

Competing interests

The authors have declared that no competing interest exists.

Author details

¹Department of Oncology, The First Affiliated Hospital, Gannan Medical University, No 23, Qingnian Road, Ganzhou, China. ²School of Pharmaceutical Sciences, Southern Medical University, Guangzhou, China. ³Jiangkou Town Central Health Center, Ganxian District, Ganzhou, China.

Received: 18 August 2022 Accepted: 9 November 2022
Published online: 21 November 2022

References

- Bui T, Thompson C. Cancer's sweet tooth. *Cancer Cell*. 2006;9(6):419–20.
- Fantin V, St-Pierre J, Leder P. Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. *Cancer Cell*. 2006;9(6):425–34.
- Dey P, Kimmelman A, DePinho R. Metabolic codependencies in the tumor microenvironment. *Cancer Discov*. 2021;11(5):1067–81.
- Raggi C, Taddei M, Rae C, Braconi C, Marra F. Metabolic reprogramming in cholangiocarcinoma. *J Hepatol*. 2022;77(3):849–64.
- Wang C, Dong Z, Hao Y, et al. Coordination polymer-coated CaCO reinforces radiotherapy by reprogramming the immunosuppressive metabolic microenvironment. *Adv Mater*. 2022;34(3): e2106520.
- Bononi G, Masoni S, Di Bussolo V, Tuccinardi T, Granchi C, Minutolo F. Historical perspective of tumor glycolysis: a century with Otto Warburg. *Semin Cancer Biol*. 2022;86:325–33.
- Icard P, Shulman S, Farhat D, Steyaert J, Alifano M, Lincet H. How the Warburg effect supports aggressiveness and drug resistance of cancer cells? *Drug Resist Updates*. 2018;38:1–11.
- Poff A, Koutnik A, Egan K, Sahebjam S, D'Agostino D, Kumar N. Targeting the Warburg effect for cancer treatment: Ketogenic diets for management of glioma. *Semin Cancer Biol*. 2019;56:135–48.
- Faubert B, Li K, Cai L, et al. Lactate metabolism in human lung tumors. *Cell*. 2017;171(2):358–71.e9.
- Cao L, Wu J, Qu X, et al. Glycometabolic rearrangements— aerobic glycolysis in pancreatic cancer: causes, characteristics and clinical applications. *J Exp Clin Cancer Res*. 2020;39(1):267.
- Reinfeld B, Rathmell W, Kim T, Rathmell J. The therapeutic implications of immunosuppressive tumor aerobic glycolysis. *Cell Mol Immunol*. 2022;19(1):46–58.
- Ganapathy-Kanniappan S. Molecular intricacies of aerobic glycolysis in cancer: current insights into the classic metabolic phenotype. *Crit Rev Biochem Mol Biol*. 2018;53(6):667–82.
- Brooks G. The science and translation of lactate shuttle theory. *Cell Metab*. 2018;27(4):757–85.
- Becker L, O'Connell J, Vo A, et al. Epigenetic reprogramming of cancer-associated fibroblasts deregulates glucose metabolism and facilitates progression of breast cancer. *Cell Rep*. 2020;31(9): 107701.
- Lee S, McIntyre D, Honess D, et al. Carbonic anhydrase IX is a pH-stat that sets an acidic tumour extracellular pH in vivo. *Br J Cancer*. 2018;119(5):622–30.
- Colegio O, Chu N, Szabo A, et al. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature*. 2014;513(7519):559–63.
- Gao F, Tang Y, Liu W, et al. Intra/extracellular lactic acid exhaustion for synergistic metabolic therapy and immunotherapy of tumors. *Adv Mater*. 2019;31(51):e1904639.
- Hui S, Ghergurovich J, Morscher R, et al. Glucose feeds the TCA cycle via circulating lactate. *Nature*. 2017;551(7678):115–8.
- Watson M, Vignali P, Mullett S, et al. Metabolic support of tumour-infiltrating regulatory T cells by lactic acid. *Nature*. 2021;591(7851):645–51.
- Doherty J, Cleveland J. Targeting lactate metabolism for cancer therapeutics. *J Clin Investig*. 2013;123(9):3685–92.
- Baumann F, Leukel P, Doerfelt A, et al. Lactate promotes glioma migration by TGF-beta2-dependent regulation of matrix metalloproteinase-2. *Neuro Oncol*. 2009;11(4):368–80.
- Kumagai S, Koyama S, Itahashi K, et al. Lactic acid promotes PD-1 expression in regulatory T cells in highly glycolytic tumor microenvironments. *Cancer Cell*. 2022;40(2):201–18.e9.
- Multhoff G, Vaupel P. Lactate-avid regulatory T cells: metabolic plasticity controls immunosuppression in tumour microenvironment. *Signal Transduct Target Ther*. 2021;6(1):171.
- Decking S, Bruss C, Babl N, et al. LDHB overexpression can partially overcome T cell inhibition by lactic acid. *Int J Mol Sci*. 2022;23(11):5970.
- Mendler A, Hu B, Prinz P, Kreutz M, Gottfried E, Noessner E. Tumor lactic acidosis suppresses CTL function by inhibition of p38 and JNK/c-Jun activation. *Int J Cancer*. 2012;131(3):633–40.
- Brand A, Singer K, Koehl G, et al. LDHA-associated lactic acid production blunts tumor immunosurveillance by T and NK cells. *Cell Metab*. 2016;24(5):657–71.
- Scott K, Cleveland J. Lactate wreaks havoc on tumor-infiltrating T and NK cells. *Cell Metab*. 2016;24(5):649–50.
- Zhou H, Yan X-Y, Yu W, et al. Lactic acid in macrophage polarization: the significant role in inflammation and cancer. *Int Rev Immunol*. 2022;41(1):4–18.
- Zhang L, Li S. Lactic acid promotes macrophage polarization through MCT-HIF1 α signaling in gastric cancer. *Exp Cell Res*. 2020;388(2): 111846.
- Song J, Lee K, Park S, et al. Lactic acid upregulates VEGF expression in macrophages and facilitates choroidal neovascularization. *Invest Ophthalmol Vis Sci*. 2018;59(8):3747–54.
- Alber J, Föllner M. Lactic acid induces fibroblast growth factor 23 (FGF23) production in UMR106 osteoblast-like cells. *Mol Cell Biochem*. 2022;477(2):363–70.
- Brown T, Ganapathy V. Lactate/GPR81 signaling and proton motive force in cancer: role in angiogenesis, immune escape, nutrition, and Warburg phenomenon. *Pharmacol Ther*. 2020;206: 107451.
- Gao Y, Zhou H, Liu G, Wu J, Yuan Y, Shang A. Tumor microenvironment: lactic acid promotes tumor development. *J Immunol Res*. 2022;2022:3119375.
- Chang J, Lee Y, Huang R. The impact of the Cancer Genome Atlas on lung cancer. *Transl Res*. 2015;166(6):568–85.
- Wilkerson MD, Yin X, Walter V, et al. Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS ONE*. 2012;7(5):e36530.
- Tibshirani R, Bien J, Friedman J, et al. Strong rules for discarding predictors in lasso-type problems. *J R Stat Soc Ser B Stat Methodol*. 2012;74(2):245–66.
- McEligot A, Poynor V, Sharma R, Panangadan A. Logistic LASSO regression for dietary intakes and breast cancer. *Nutrients*. 2020;12(9):2652.
- Limagne E, Nuttin L, Thibaudin M, et al. MEK inhibition overcomes chemioimmunotherapy resistance by inducing CXCL10 in cancer cells. *Cancer Cell*. 2022;40(2):136–52.e12.
- Jung H, Kim HS, Kim JY, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun*. 2019;10(1):4278.
- Hwang S, Kwon AY, Jeong JY, et al. Immune gene signatures for predicting durable clinical benefit of anti-PD-1 immunotherapy in patients with non-small cell lung cancer. *Sci Rep*. 2020;10(1):643.
- Goldmann T, Marwitz S, Nitschkowski D, et al. PD-L1 amplification is associated with an immune cell rich phenotype in squamous cell cancer of the lung. *Cancer Immunol Immunother*. 2021;70(9):2577–87.
- Mitchell KA, Zingone A, Toulabi L, Boeckelman J, Ryan BM. Comparative transcriptome profiling reveals coding and noncoding RNA differences in NSCLC from African Americans and European Americans. *Clin Cancer Res*. 2017;23(23):7412–25.
- Miao D, Margolis CA, Vokes NI, Liu D, Taylor-Weiner A, Wankowicz SM, Adeegbe D, Keliher D, Schilling B, Tracy A, Manos M, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet*. 2018;50(9):1271–81.
- Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269):pl1.
- Mayakonda A, Lin D, Assenov Y, Plass C, Koeffler H. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747–56.
- Wu J, Zhang H, Li L, et al. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: a population-based analysis. *Cancer Commun (Lond, Engl)*. 2020;40(7):301–12.
- Kawada J, Takeuchi S, Imai H, et al. Immune cell infiltration landscapes in pediatric acute myocarditis analyzed by CIBERSORT. *J Cardiol*. 2021;77(2):174–8.
- Bejarano L, Jordão M, Joyce J. Therapeutic targeting of the tumor microenvironment. *Cancer Discov*. 2021;11(4):933–59.
- Meurette O, Mehlen P. Notch signaling in the tumor microenvironment. *Cancer Cell*. 2018;34(4):536–48.
- Marin E, Bouchet-Delbos L, Renoult O, et al. Human tolerogenic dendritic cells regulate immune responses through lactate synthesis. *Cell Metab*. 2019;30(6):1075–90.e8.

51. San-Millan I, Sparagna GC, Chapman HL, et al. Chronic lactate exposure decreases mitochondrial function by inhibition of fatty acid uptake and cardiopilin alterations in neonatal rat cardiomyocytes. *Front Nutr.* 2022;9: 809485.
52. Morland C, Andersson K, Haugen Ø, et al. Exercise induces cerebral VEGF and angiogenesis via the lactate receptor HCAR1. *Nat Commun.* 2017;8:15557.
53. Pelizzari G, Basile D, Zago S, et al. Lactate dehydrogenase (LDH) response to first-line treatment predicts survival in metastatic breast cancer: first clues for a cost-effective and dynamic biomarker. *Cancers.* 2019;11(9):1243.
54. Wulaningsih W, Holmberg L, Garmo H, et al. Serum lactate dehydrogenase and survival following cancer diagnosis. *Br J Cancer.* 2015;113(9):1389–96.
55. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2016;316(22):2402–10.
56. Agarwal S, Simonds W, Marx S. The parafibromin tumor suppressor protein interacts with actin-binding proteins actinin-2 and actinin-3. *Mol Cancer.* 2008;7:65.
57. Sivaprakasam S, Bhutia Y, Yang S, Ganapathy V. Short-chain fatty acid transporters: role in colonic homeostasis. *Compr Physiol.* 2017;8(1):299–314.
58. Cordell PA, Futers TS, Grant PJ, Pease RJ. The Human hydroxyacylglutathione hydrolase (HAGH) gene encodes both cytosolic and mitochondrial forms of glyoxalase II. *J Biol Chem.* 2004;279(27):28653–61.
59. Qin X, Li C, Guo T, et al. Upregulation of DARS2 by HBV promotes hepatocarcinogenesis through the miR-30e-5p/MAPK/NFAT5 pathway. *J Exp Clin Cancer Res.* 2017;36(1):148.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

