

## Research Paper

# Prognostic implication of glycolysis related gene signature in non-small cell lung cancer

Jie Yao<sup>1</sup>, Rui Li<sup>1</sup>, Xiao Liu<sup>1</sup>, Xijia Zhou<sup>1</sup>, Jianping Li<sup>1</sup>, Tingting Liu<sup>1</sup>, Chen Huo<sup>1</sup>, Yiqing Qu<sup>2</sup>✉

1. Department of Pulmonary and Critical Care Medicine, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan 250012, China.
2. Department of Pulmonary and Critical Care Medicine, Qilu Hospital of Shandong University, Jinan 250012, China.

✉ Corresponding author: Dr. Yiqing Qu, Department of Pulmonary and Critical Care Medicine, Qilu Hospital of Shandong University, Jinan 250012, China. E-mail: quyiqing@sdu.edu.cn; Tel.: +86 531 8216 9335.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2020.07.03; Accepted: 2020.10.24; Published: 2021.01.01

## Abstract

Abnormal glycolysis is one of the hallmarks of cancer and plays an important role in its development. This study was devoted to identify glycolysis related genes as prognostic biomarkers for non-small cell lung cancer (NSCLC). The mRNA expression profile and clinical follow-up data were obtained using The Cancer Genome Atlas (TCGA) database. The validation set was obtained by bootstrap method of random repeated sampling. A total of 200 glycolysis-related genes were obtained from Gene Set Enrichment Analysis (GSEA) and 46 genes were significantly associated with overall survival (OS). Five genes (PKP2, LDHA, HMMR, COL5A1 and B3GNT3) were eventually identified to calculate risk score of NSCLC patients.

The univariate and multivariate Cox regression analysis indicated that the risk score was an independent prognostic factor (training set: HR=2.126, 95% CI [1.605, 2.815],  $p < 0.001$ ; validation set: HR=2.298, 95%CI [1.450, 3.640],  $p < 0.001$ ). Patients assigned to the high-risk group were associated with poor OS compared with patients in the low-risk group (training set:  $P = 7.946e-06$ ; validation set:  $P = 6.368e-07$ ). Receiver operating characteristic (ROC) curve and stratification analysis also demonstrated the potential prognostic performance. In conclusion, we constructed a novel glycolysis related risk signature which might contribute to predicting the prognosis of NSCLC.

Key words: non-small cell lung cancer; glycolysis; TCGA; risk score; prognosis

## Introduction

Lung cancer is one of the leading causes of cancer-related death [1], especially in developed countries, where it has ascended to number one in cancer deaths [2]. NSCLC accounts for approximately 80% of all lung cancer cases [3]. Despite progress in treatment, the prognosis of NSCLC remains poor, due to the lack of early identification markers [4]. To date, the prognosis mainly depends on histopathologic diagnosis and tumor staging [5]. However, this has limitations as patients with the same degree of progression will show distinct outcomes due to individual differences [6]. Some evidence has shown that the discovery and application of molecular biomarkers can provide prognostic value [7-10]. Therefore, new diagnosis markers must be urgently identified for assessing prognosis of NSCLC.

Glycolysis is one of the earliest evidence of

metabolic changes in tumor [11,12]. Even in oxygenated circumstances, the glycolysis activity of tumor cells is also active. This metabolic characteristic is known as Warburg effect, which shows high glucose uptake, active glycolysis and large amounts of lactic acid production [13,14]. Increased glycolysis satisfies the increasing proliferation of cancer cells [15]. In addition, studies have shown that enhanced glucose metabolism could alter apoptosis pathway in cancer cells [16]. The oncogenic regulation of glycolysis provides us with new explanation for tumor progression [17]. Many studies have found that glycolysis changes take an important role in cancer progression. For example, tumor glycolysis promotes immune evasion by participating in immunosuppressive networks [18]. The tumor suppressor protein p53 has a serious effect on glucose metabolic

reprogramming [19]. Glucose transporter GLUT1 is involved in the uptake of glucose and has significantly lower expression in highly differentiated endometrial carcinoma, than in poorly differentiated tumors [20]. However, the concrete role of glycolysis in tumorigenesis of NSCLC is poorly known. Ma et al. found that NSCLC harboring EML4-ALK rearrangements displayed higher glucose metabolism [21]. Li et al. reported that the anaerobic metabolism offers new targets and orientation for tumors, especially lung cancer [22]. Further work is needed to demonstrate the significance of glycolysis gene expression signatures in NSCLC.

In the present study, transcriptome profiling was downloaded from TCGA database and 200 mRNAs were identified that significantly relate to glycolysis. Furthermore, a risk predictive signature including five genes was conducted to calculate the risk of NSCLC patients and further successfully validated in the validation set. Univariate and multivariate Cox regression analysis revealed the signature was an independent predictive factor for OS. The risk score also performed better than other clinical parameters in prognosis. This study might present a new thinking for assessing the prognosis of NSCLC.

## Materials and Methods

### Data collection and processing

Detailed information of NSCLC patients' containing transcriptome profiling and clinical follow-up data, were downloaded from TCGA database on April 2, 2020. Patients meeting the following criteria were excluded: 1) patient had no survival time or survival status, 2) patient had clinical information but no gene expression data. There were 935 patients included in this study. The detailed clinical characteristics are summarized in Table 1.

### Gene set enrichment analysis

GSEA analysis was used to explore whether the selected gene sets were associated with downloaded transcriptome data [23]. We determined gene sets for further investigation at normalized  $p$  values ( $p < 0.05$ ) and normalized enrich score ( $|NES| \geq 1$ ).

### Establishment and validation of glycolysis related gene signature

We intersected glycolysis related genes and their expression in TCGA for matching. Differentially expressed glycolysis related genes was obtained using limma R package with  $|\log_2 FC| \geq 1.5$  and false discovery rate (FDR)  $< 0.05$  [24]. Then differentially expressed genes combined with TCGA clinical data for screening prognostic genes. We identified differentially glycolysis related genes whose

expression was significantly associated with OS by univariate Cox regression analysis ( $P < 0.05$ ). Signature genes were ultimately determined by multivariate Cox regression analysis. The prognostic signature for NSCLC patients was established. Risk score =  $([\text{Coefficient mRNA}_1] \times [\text{Expression of mRNA}_1]) + ([\text{Coefficient mRNA}_2] \times [\text{Expression of mRNA}_2]) + \dots + ([\text{Coefficient mRNA}_n] \times [\text{Expression of mRNA}_n])$ . Independent analyses were applied to verify the independence of the signature. Then NSCLC patients were divided into two groups based on a median risk score. The Kaplan-Meier curve was visualized using the survival R package and survminer R package.  $P < 0.05$  was regarded as statistically significant. Survival ROC R package was used to compare the ROC areas. The validation set was acquired by bootstrap method [25]. Furthermore, stratification analysis also verified the prognostic implication of the signature.

**Table 1.** Clinical characteristics of included patients with non-small cell lung cancer (NSCLC)

Clinical characteristic	N	Percentage (%)
<b>Age (years)</b>		
≤65	407	43.5
>65	528	56.5
<b>Gender</b>		
Male	561	60.0
Female	374	40.0
<b>TNM stage</b>		
I-II	746	79.8
III-IV	189	20.2
<b>T stage</b>		
T1-2	785	84.0
T3-4	150	16.0
<b>N stage</b>		
N0	599	64.1
N1-3	336	35.9
<b>M stage</b>		
M0	697	74.5
M1-3	238	25.5

### Prognostic genes expression and alteration were shown by online databases

In our study, five prognostic genes involved in the signature were all upregulated in NSCLC. The gene expression was further validated in the Oncomine database and Tumor Immune Estimation Resource (TIMER) database [26,27]. The alternation of the signature genes was shown by cBioportal for Cancer Genomics [28,29].

## Results

### Establishment a prognostic signature of five glycolysis related genes in NSCLC

Clinical information and corresponding gene expression data from 935 NSCLC patients were obtained from TCGA. GSEA results showed gene sets

which were significantly enriched, with normalized  $p$ -values  $<0.05$  and NES  $\geq 1$ , including glycolysis, oxidative phosphorylation, DNA repair, inflammatory response, mitotic spindle, bile acid metabolism, E2F targets and G2M checkpoint (Table 2, Figure 1). We selected GLYCOLYSIS ( $P = 0.000$ , NES =2.46) for further analysis. Total 200 glycolysis related genes were obtained from GSEA, of which 46 genes were considered statistically significant (FDR  $< 0.05$ ,  $|\log_2FC| \geq 1.5$ ). The heatmap (Figure 2a) and boxplot (Figure 2b) showed the expression of these genes. A total of 10 genes expression pattern had a significant association with OS ( $P < 0.05$ ) using univariate COX regression analysis. Subsequently, 5 of 10 genes were identified as independent prognosis genes to establish model using multivariate COX regression analysis. Finally, five prognosis genes and the expression coefficient of each gene were obtained, and a gene-based prognostic signature was constructed for calculating risk score of each NSCLC patient. The five genes were LDHA, HMMR, B3GNT3, PKP2, and COL5A1 (Table 3). The risk score =  $(0.2170 \times \text{Expression of LDHA}) + (0.1200 \times \text{expression of HMMR}) + (0.1088 \times \text{expression of B3GNT3}) + (0.0960 \times \text{expression of PKP2}) + (0.1344 \times \text{expression of COL5A1})$ . The heatmap displayed the expression of five signature genes (Figure 3a). The five genes were all significantly up-regulated (Figure 3b-3f). Based on the median risk score, the NSCLC patients were divided into high- and low-risk groups. Patients of high-risk group tended to have a shorter survival time than those in low-risk group ( $P=7.946e-06$ ) (Figure 3g). The risk score and survival state distribution were also visualized (Figure 3h, 3i). Furthermore, we used cox regression analysis to identify the independence of the risk score. Univariate independent prognostic analysis illustrated that TNM stage, T stage, N stage and risk score have statistical significance with OS ( $P<0.001$ ) (Figure 4a). Multivariate independent prognostic analysis illustrated that risk score was an independent prognostic factor ( $P<0.001$ ) (Figure 4b). Thus, our results confirmed that the glycolysis-related gene signature could be used as an independent prognostic factor in clinical practice (Table 4).

**Table 3.** The information of five mRNAs involved in the signature

mRNA	Ensemble ID (ENSG0000)	Location	coef	HR	HR.95L	HR.95H	$p$ -value
LDHA	0134333	Chr11: 18, 394, 389-18, 408, 425	0.21695	1.24228	1.05311	1.46544	0.01006
B3GNT3	0179913	Chr19: 17, 794, 828-17, 813, 576	0.10884	1.11498	1.03914	1.19636	0.00246
PKP2	0057294	Chr12: 32, 790, 745-32, 896, 846	0.09602	1.10067	0.98912	1.22481	0.07853
COL5A1	0130635	Chr9: 134, 641, 774-134, 844, 843	0.13439	1.14384	1.05674	1.23811	0.00088
HMMR	0072571	Chr5: 163, 460, 203-163, 491, 945	0.12004	1.12754	0.96947	1.31138	0.11932

## Clinical correlation analysis

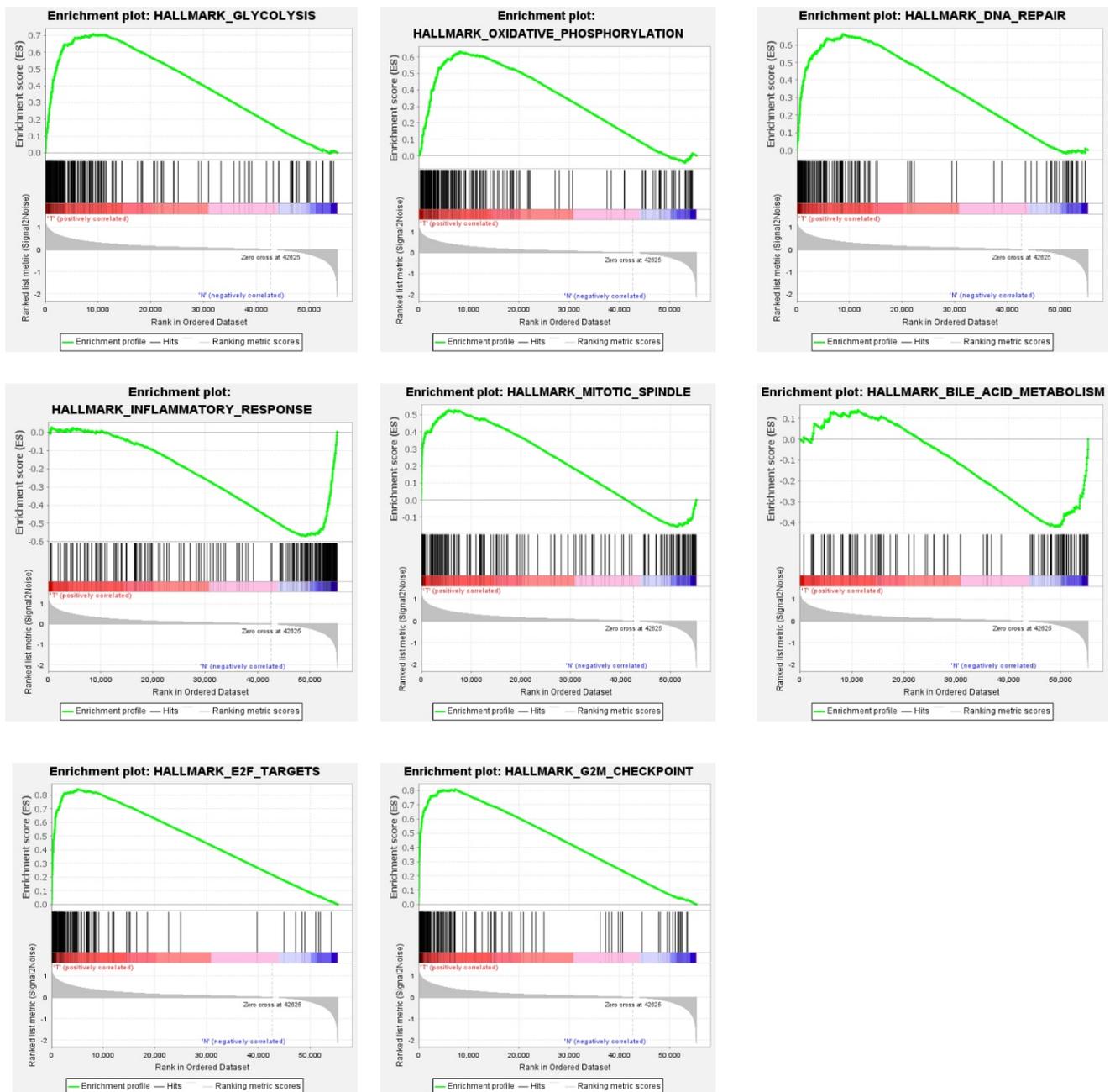
ROC curves of OS helped to assess the prognosis significance of gene signature with AUC of 0.649, whereas lower scores were shown in other clinical parameters, such as age (AUC = 0.547), gender (AUC = 0.551), TNM stage (AUC = 0.634), T stage (AUC = 0.629), N stage (AUC = 0.579) and M stage (AUC = 0.501) (Figure 5a). The connection between signature and other clinical pathological factors was also used to demonstrate prognostic performance. Patients grouped by age ( $\leq 65$ ,  $>65$ ) or gender (male, female) or TNM stage (I-II, III-IV). Risk score was found to be significantly associated with age ( $p = 0.027$ ) and TNM stage ( $p < 0.001$ ), but not with gender (Figure 5b-d). Additionally, these five genes in signature with showed substantial difference among some clinical parameters. Differential expression of HMMR showed across different gender, N stage and TNM stage (Figure 5e-g). The differential expression of PKP2 exhibited across age and gender (Figure 5h, 5i). B3GNT3 was expressed differently across different gender and TNM stage (Figure 5j, 5k). The differential expression of LDHA was found in N stage and TNM stage (Figure 5l, 5m).

## Stratification analysis

Patients were grouped in the same manner as the previous step for stratification analysis where high-risk patients with poorer OS were grouped by age, no matter if they were older or younger than 65 (Figure 6a, 6b). The same was true of TNM stage (Figure 6e, 6f). Therefore, the potential associations between risk score with patients' age and TNM stage were statistically significant. However, there was no significant difference in gender stratification (Figure 6c, 6d).

**Table 2.** Gene sets enriched in non-small cell lung cancer (NSCLC)

GS follow link to MSigDB	SIZE	NES	NOM $p$ -value	Rank at MAX
Glycolysis	200	2.46	0	9022
Oxidative phosphorylation	200	1.84	0.034	8207
DNA repair	150	2.16	0	8719
Inflammatory response	200	-1.88	0.024	6162
Mitotic spindle	198	1.81	0.014	5517
Bile acid metabolism	112	-1.68	0.015	6354
E2F targets	198	2.18	0	4960
G2M checkpoint	196	2.18	0	7277



**Figure 1. Gene sets significantly enriched in the NSCLC using GSEA.** This includes the following: glycolysis, oxidative phosphorylation, DNA repair, inflammatory response, mitotic spindle, bile acid metabolism, E2F targets, G2M checkpoint.

### Prognostic significance in validation set

Internal validation was conducted using the bootstrapping method. The original dataset was acted as training set. The validation set was performed as described above, including survival analysis, risk analysis, independent prognosis analysis and ROC curve drawing. The AUC of ROC was 0.631, which showed that this model has superior accuracy (Figure 7a). The OS was notably shorter in high-risk group than in the low-risk group ( $P$ -value=6.368e-07; Figure 7b). Besides, univariate and multivariate regression analysis were performed on the validation set to

demonstrate the independent prognostic significance of the signature (Figure 4c, 4d). The heatmap, risk score curve and survival status data in validation set were shown in Figure 7c-e. The prognostic significance of our signature was further verified in the validation set.

### External validation using online database

We analyzed the alterations of the five genes in 1144 NSCLC samples in the cBioPortal for Cancer Genomics. The results showed that the genes in signature were mutated in 191 (17%) of the sequenced cases. As shown in Figure 8b. The five genes in the

signature were all up-regulated in NSCLC. This is also confirmed in the Oncomine database (Figure 8a) and the TIMER database (Figure 9).

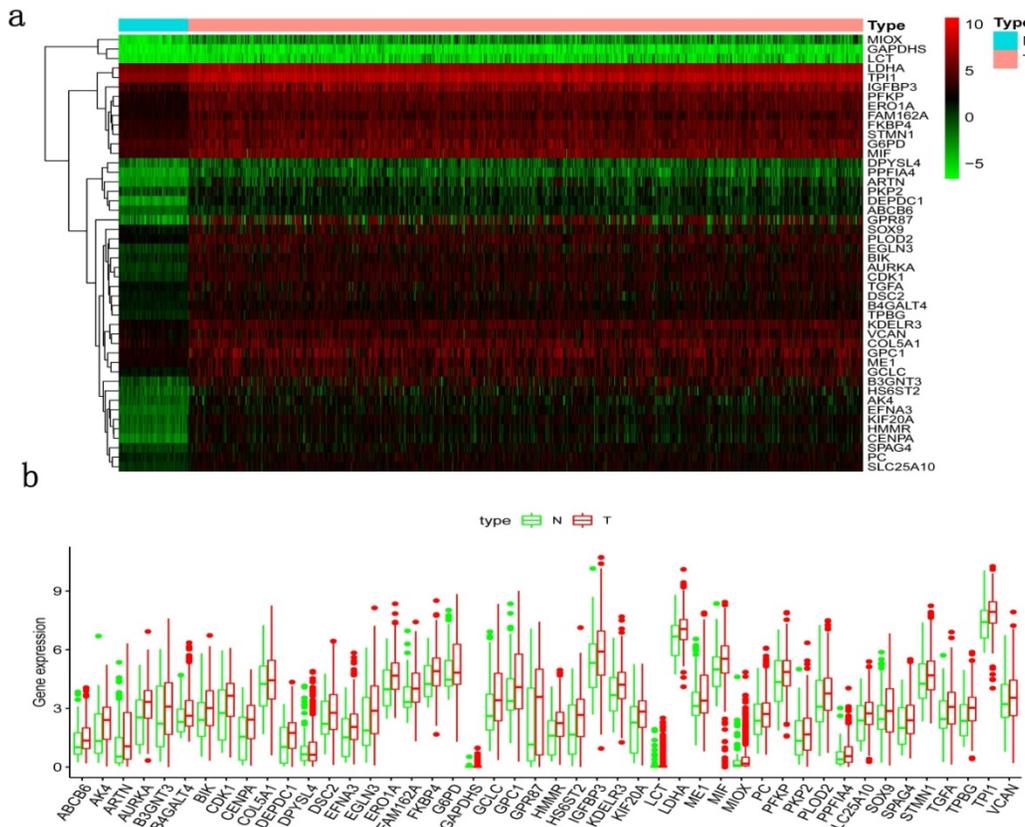
### Discussion

NSCLC remains a major challenge for global public health, with incidence and mortality still increasing in many countries. Although great progress has been made, the underlying molecular pathogenesis of NSCLC is still unclear. Considering the poor prognosis of NSCLC due to its later diagnosis, identification reliable prognostic markers

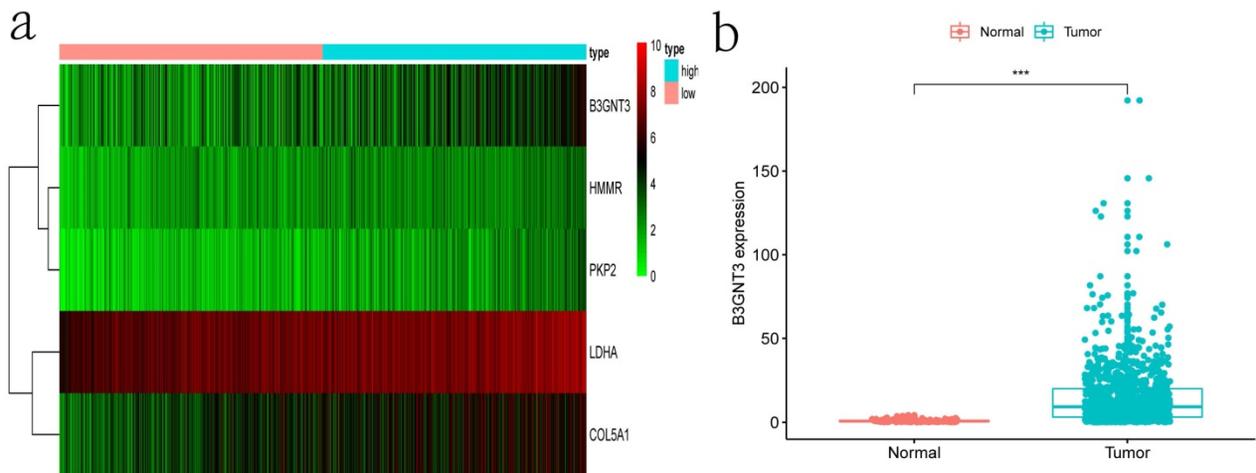
and establishment of more accurate prognostic models are urgently needed.

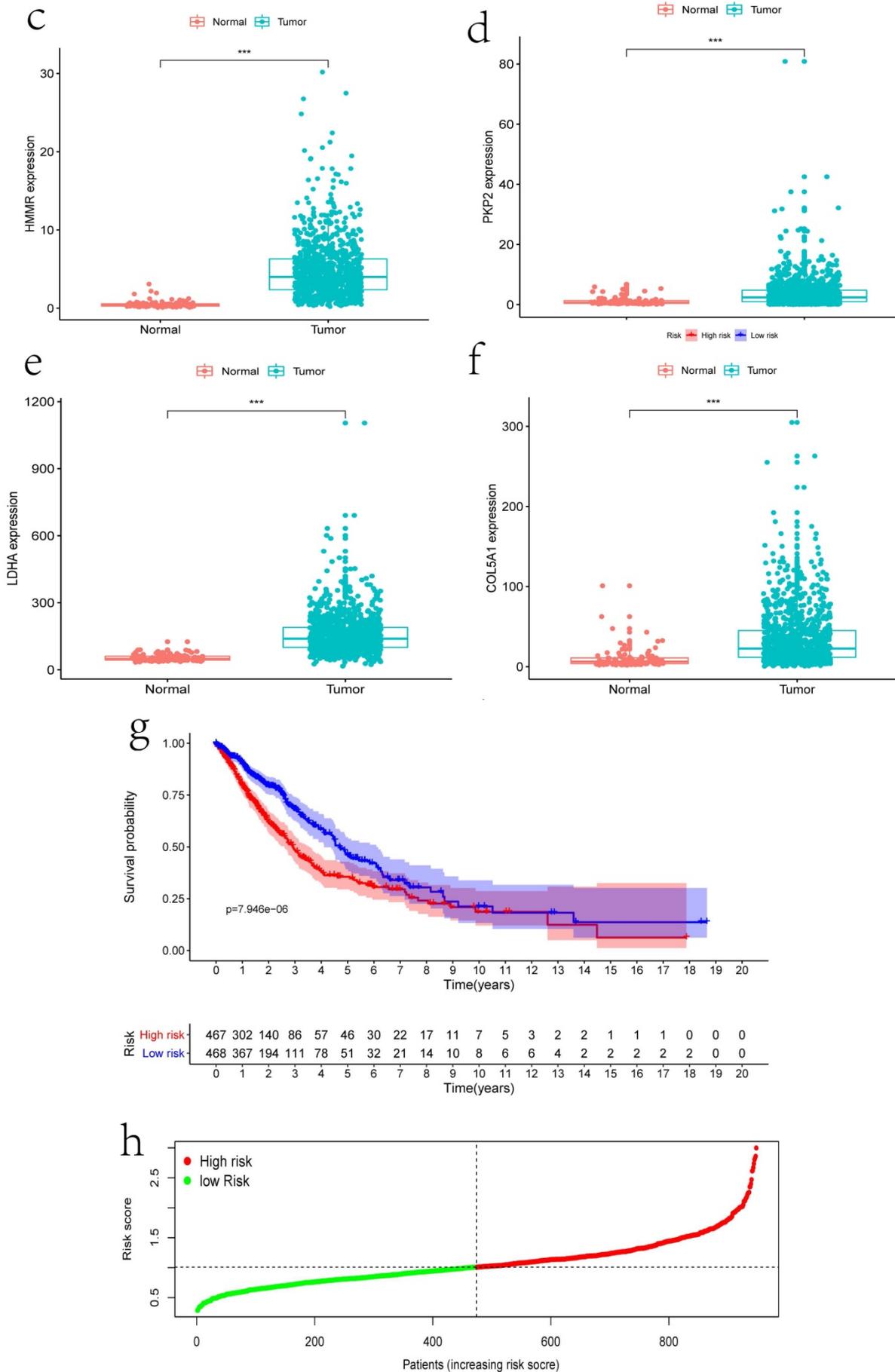
**Table 4.** Univariate and multivariate analysis for each clinical feature

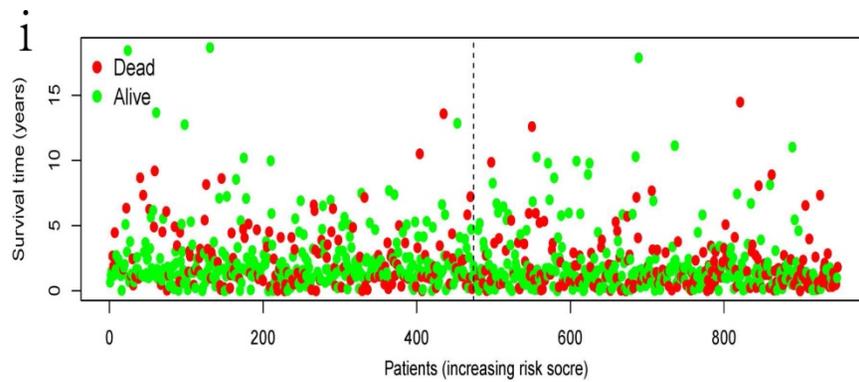
Clinical feature	Univariate analysis			Multivariate analysis		
	HR	95%CI	P	HR	95%CI	P
Age	1.012	0.998-1.026	0.083	1.017	1.003-1.032	0.015
Gender	1.11	0.862-1.43	0.417	1.108	0.859-1.428	0.430
Stage	1.452	1.282-1.644	<0.001	1.253	0.93-1.686	0.138
T	1.411	1.216-1.636	<0.001	1.138	0.94-1.378	0.184
M	1.835	1.088-3.095	0.023	1.115	0.505-2.46	0.787
N	1.438	1.239-1.669	<0.001	1.127	0.869-1.463	0.368
Risk score	2.383	1.810-3.137	<0.001	2.126	1.605-2.815	<0.001



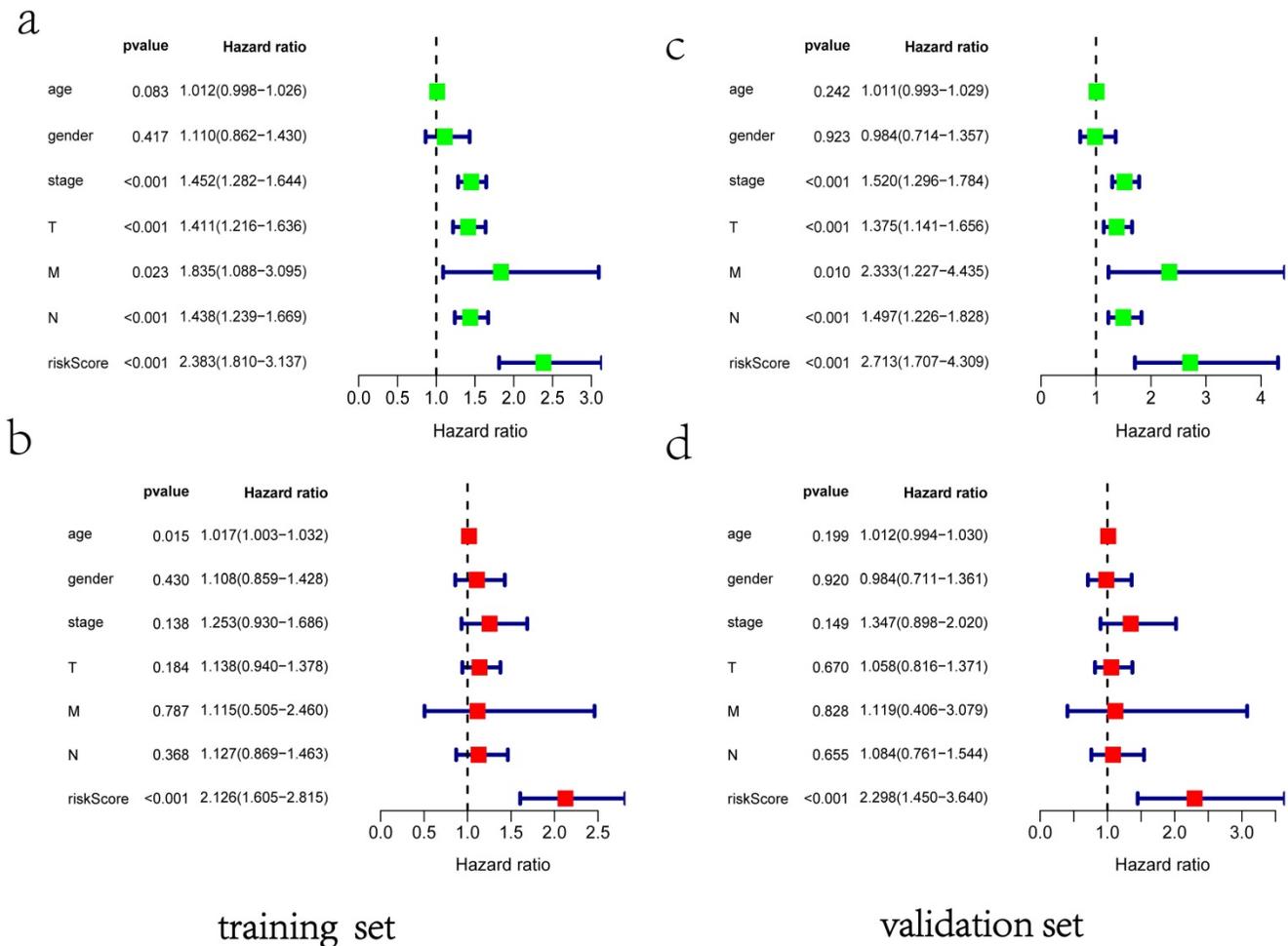
**Figure 2.** Differentially expressed glycolysis related genes. The red color indicates the higher gene expression value while the green color indicates the lower gene expression value. N indicates non-tumor tissues. T indicates tumor tissues. **a)** Heatmap; and, **b)** boxplot.







**Figure 3. The expression patterns of five genes in the signature.** a) Heatmap of five genes expression profiles. b-f) Different expression of five genes in the normal tissue and tumor tissue based on TCGA (\*\*\*) represents  $P < 0.001$ . g) Kaplan-Meier curve for NSCLC patients with high/low risk. h) Risk score distribution. i) Survival status distribution.



**Figure 4. The independent prognostic significance of the signature was demonstrated by univariate and multivariate regression analysis.**

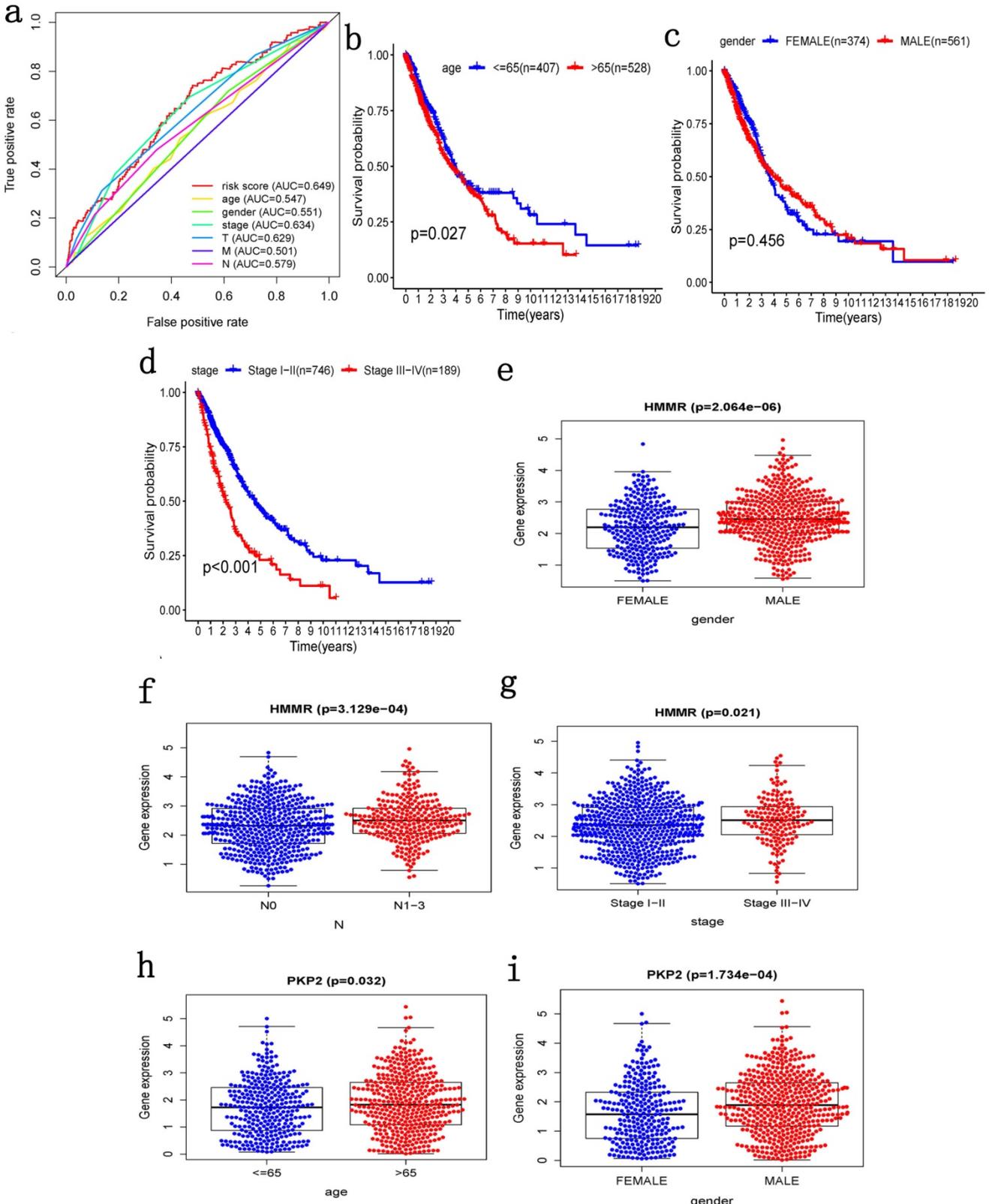
In recent years, the glycolysis mechanism in various diseases has caused wide attention [30-32]. Growing evidence show that glycolysis key genes play an important role in regulating cancer cell metabolism and may be a potential therapeutic option [33,34]. In lung cancer, glucose metabolism feature has also been further elucidated [35-37]. For instance, significantly upregulated OTUB2 in NSCLC stimulated the Warburg effect and was closely related

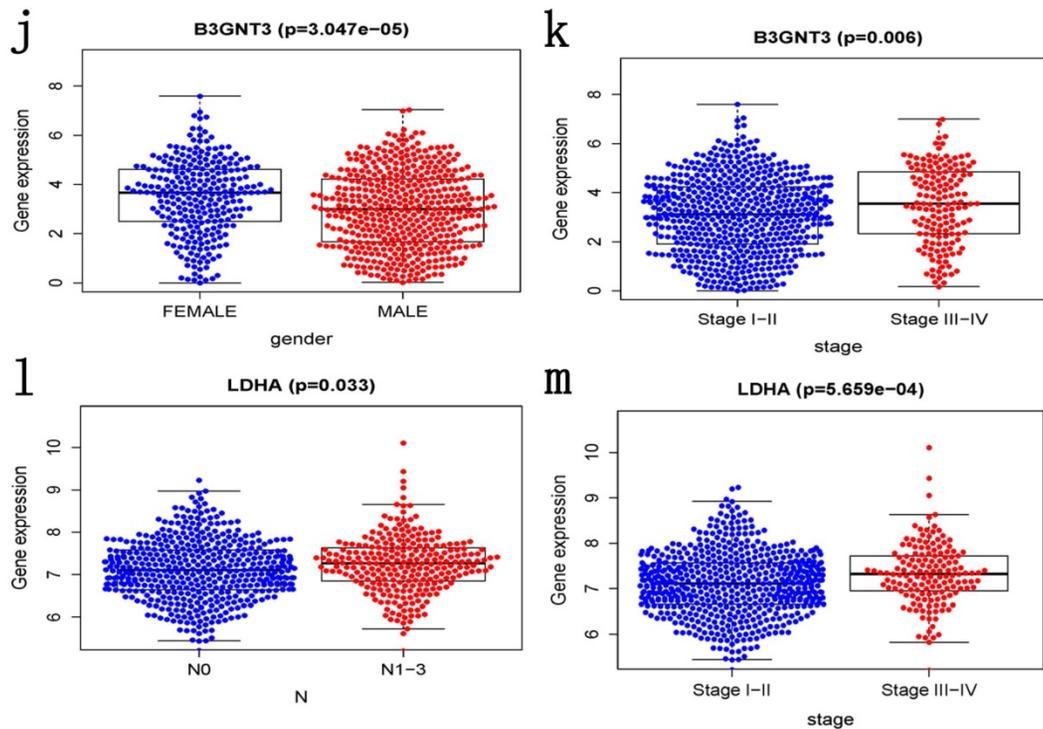
to metastasis, advanced tumor stages, poor survival, and recurrence [38]. Transcription factor BACH1 stimulates glycolysis-dependent lung cancer metastasis by increasing glucose uptake, glycolysis rates, and lactate secretion [39].

In this study, we established a new predictive signature including five glycolysis related genes. Univariate and multivariate cox regression analysis determined the independent prognostic effect for

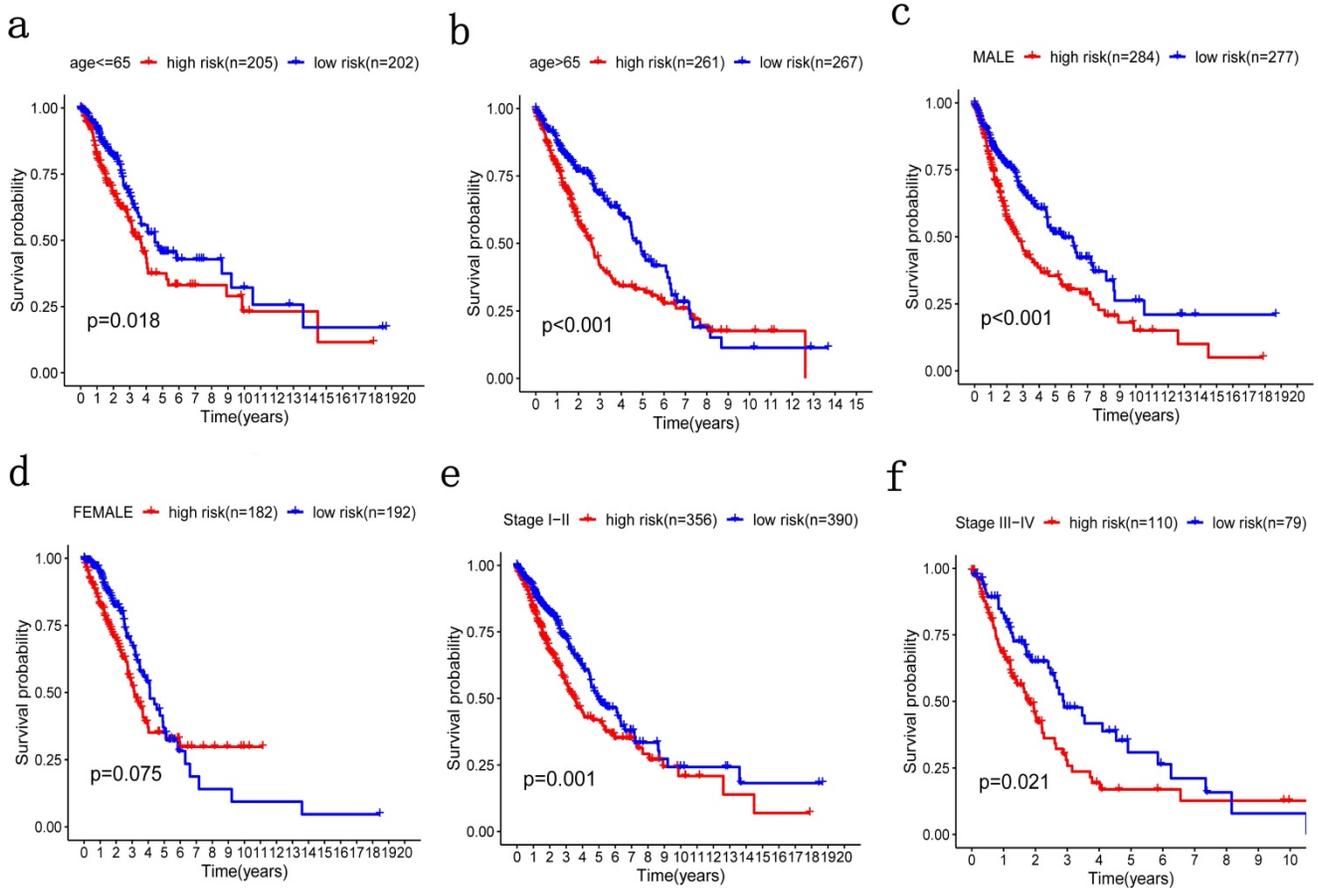
predicting the risk of NSCLC. ROC curve and stratification analysis also proved the prognostic significance. More importantly, we established a validation set to verify the reliability of this signature. Similar studies have been published before. Zhang et

al. and Liu et al. respectively identified glycolysis-related gene signature for predicting survival of lung adenocarcinoma patients [40,41]. However, their signatures were not validated internally, which is the advantage of this study.

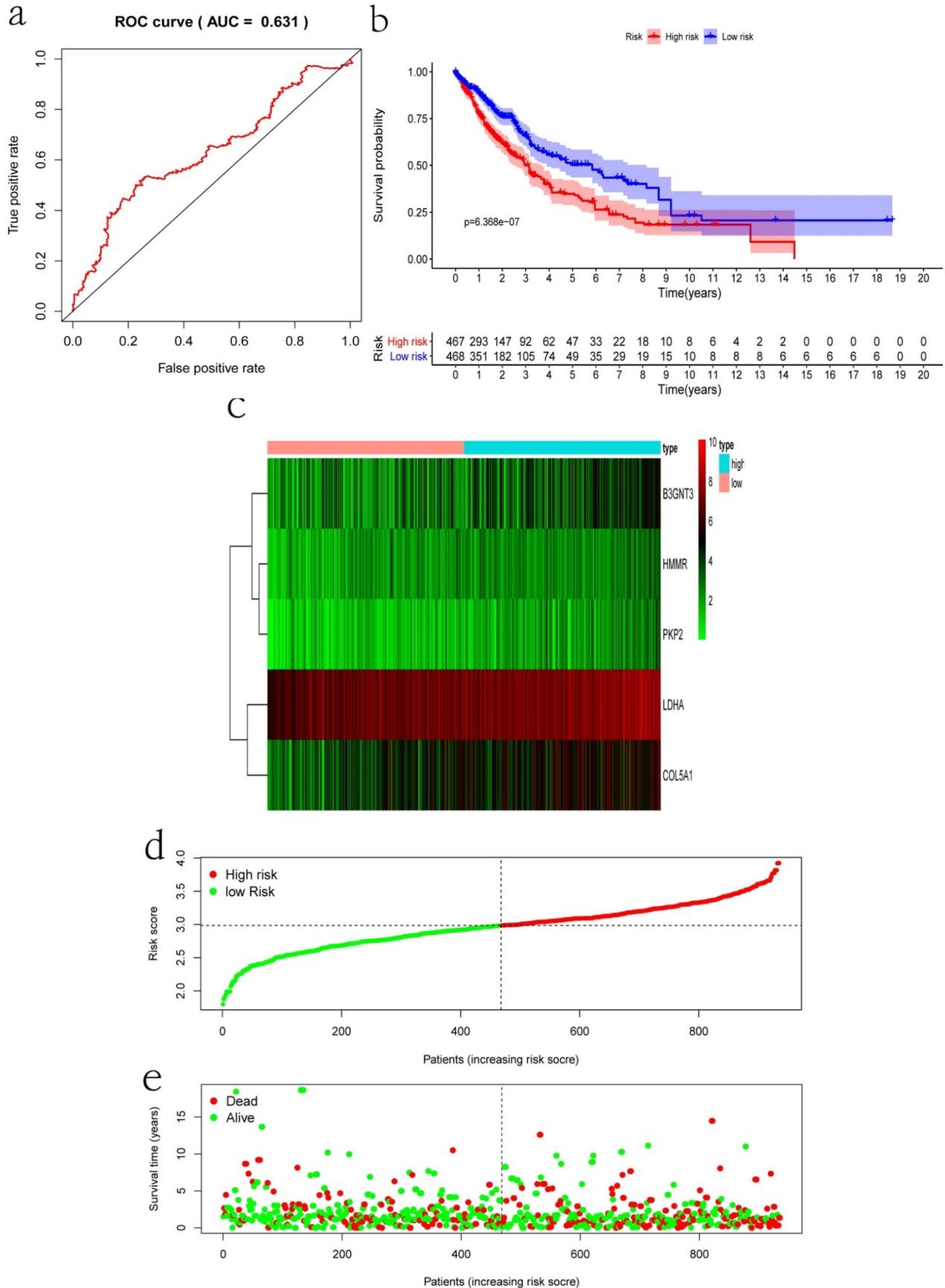




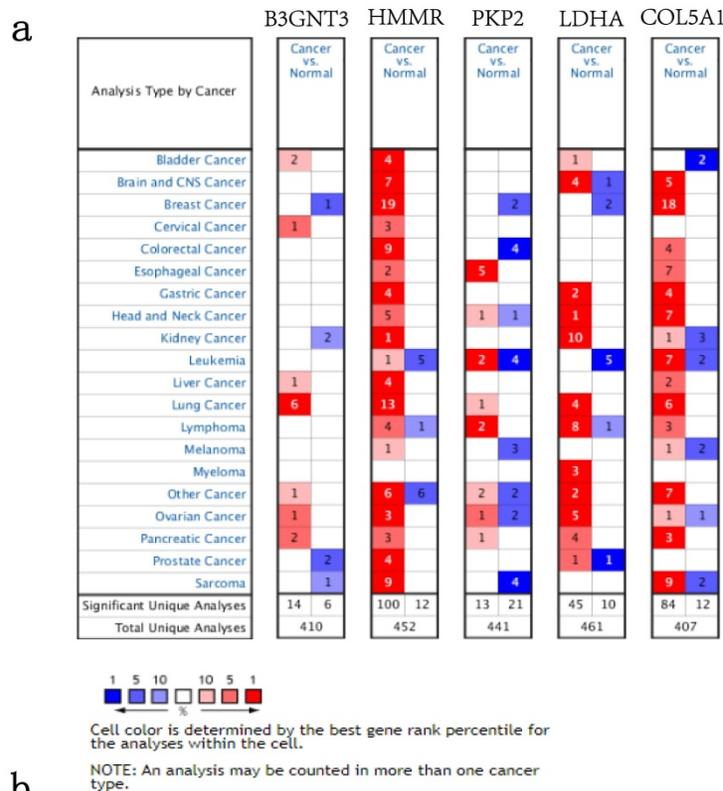
**Figure 5. Clinical correlation analysis.** a) The ROC analysis of OS for the signature and various clinical parameters. b-d) The statistical connection between signature and different clinical pathological factors. e-m) Five genes in signature with different expression among some clinical parameters.



**Figure 6. Patients were grouped by different pathological parameters for stratification analysis.** a-b) Age; c-d) gender; and, e-f) stage.



**Figure 7. Verification of the prognostic signature with the validation set. a) ROC curve; b) Kaplan-Meier curve; c) Heatmap of five genes expression profiles in validation set; d) Risk score distribution; e) Survival status distribution.**



**Figure 8. The expression and alterations of the five prognostic genes. a.** The expression profiles of the five genes in the Oncomine database. **b.** The alteration proportion for the five genes in 1144 clinical samples of non-small cell lung cancer in the cBioPortal database.

For the five signature genes, there have been some studies on its role in cancer. LDHA, a key enzyme which catalyzes the formation of lactic acid from pyruvic acid, is a key gene in aerobic glycolysis and is widely seen as a therapeutic target for cancer [42]. Phosphorylated LDHA could promote head and neck cancer and breast cancer cells invasion and metastasis [43]. Besides, LDHA inhibitors can be used together with other chemotherapy drugs to play a synergistic role in anti-tumor [44]. HMMR plays an important role in neural development by correcting spindle position. HMMR knockout mice suffer defective neural development [45]. Furthermore, the overexpression of HMMR in numerous tumor types is also associated with tumor relapse and propagation. HMMR also supports the cancer stem cell properties in gastric cancer and glioblastoma [46,47]. HMMR might be a potential prognostic marker and therapeutic target [48,49]. Some studies have suggested that the B3GNT3 were significantly overexpressed in various tumors [50,51]. The

expression level of B3GNT3 in patients with NSCLC or early-stage cervical cancer was associated with unfavourable clinicopathological parameters. These patients with high B3GNT3 expression had a shorter OS and disease-free survival (DFS) compared with those with low expression [52,53]. In contrast, B3GNT3 predicts a favorable cancer behavior of neuroblastoma [54]. In addition, down-regulation of B3GNT3 can enhance cytotoxic T-cell-mediated anti-tumor immunity in triple-negative breast cancer [55]. COL5A1 is collagen family member that encodes an alpha chain for one of the low abundance fibrillar collagens. Upregulated COL5A1 indicated poor prognosis in breast cancer, clear cell renal cell carcinoma, lung adenocarcinoma and tongue squamous cell carcinoma [56-59]. PKP2, a member of the arm-repeat protein family, is important for the assembly of junctional proteins. But as far as we know, the mechanism of PKP2 in cancer is still unclear.

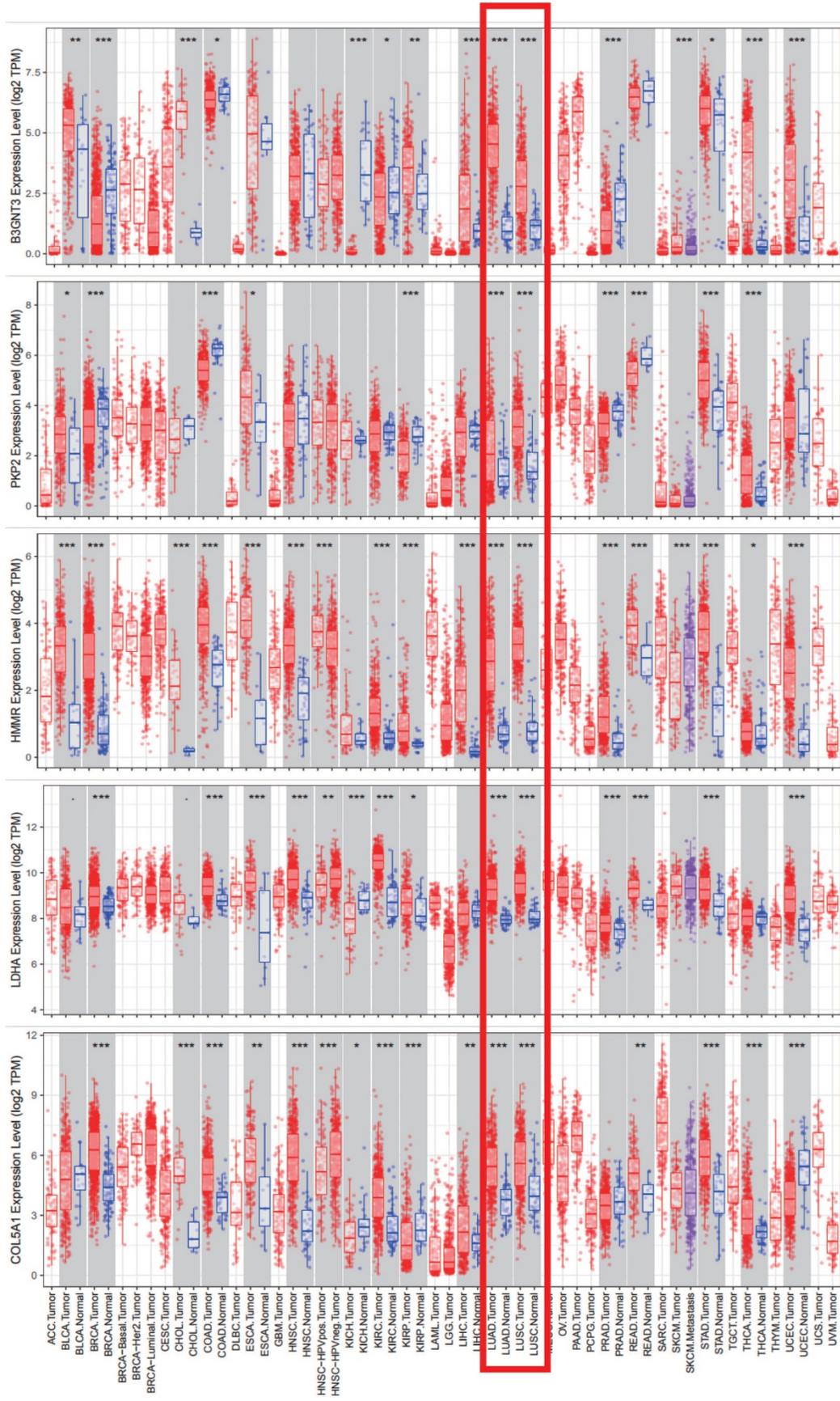


Figure 9. Five signature genes expression data from TIMER database.

The risk score was based on mRNA expression which could be easily acquired. The signature can effectively evaluate the prognosis of NSCLC patients and may be a complement to CT and pathological methods. Despite practical independent prognosis of our results, our study leaves much to be desired. First, our study only involved 200 glycolysis related genes, not the entire mRNA expression profiles. Second, validating the signature in a larger independent external set is necessary. Third, cell and animal experiments are needed to clarify specific mechanism of the five mRNAs in the regulation of tumor glycolysis. Nevertheless, the present study provides a novel suggestion for predicting prognosis of NSCLC patients from the perspective of glycolysis.

## Conclusion

In summary, we recognized five glycolysis related genes associated with OS of NSCLC patients, and constructed a risk score signature to make a survival prediction. Our risk score model could distinguish NSCLC patients with different survival outcomes, which may contribute to the clinical decision of individualized treatment program.

## Abbreviations

NSCLC: Non-small cell lung cancer; TCGA: The Cancer Genome Atlas; GSEA: Gene Set Enrichment Analysis; OS: Overall survival; ROC: Receiver operating characteristic; NES: Normalized enrich score; FDR: False discovery rate; TIMER: Tumor Immune Estimation Resource; DFS: Disease-free survival.

## Acknowledgements

This work was supported by grants from the Major Scientific and Technological Innovation Project of Shandong Province (2018CXGC1212), the Science and Technology Foundation of Shandong Province (2014GSF118084), the CSCO-Qilu Cancer Research Fund (Y-Q201802-014), the Medical and Health Technology Innovation Plan of Jinan City (201805002) and the National Natural Science Foundation of China (81372333).

## Competing Interests

The authors have declared that no competing interest exists.

## References

- Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394-424.
- Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015; 65:87-108.
- Gridelli C, Rossi A, Carbone DP, et al. Non-small-cell lung cancer. *Nat Rev Dis Primers.* 2015;1:15009.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018;68:7-30.
- Pataer A, Kalhor N, Correa AM, et al. Role of FDG PET/CT in the eighth edition of TNM staging of non-small cell lung cancer. *J Thorac Oncol.* 2012; 7: 825-32.
- Ashworth AB, Senan S, Palma DA, et al. An individual patient data of oligometastatic non-small-cell metaanalysis of outcomes and prognostic factors after treatment lung cancer. *Clin Lung Cancer.* 2014; 15:346-55.
- Belczacka L, Latosinska A, Metzger J, et al. Proteomics biomarkers for solid tumors: current status and future prospects. *Mass Spectrom Rev.* 2019;38:49-78.
- Aftimos PG, Barthelemy P, Awada A. Molecular biology in medical oncology: diagnosis, prognosis, and precision medicine. *Discov Med.* 2014; 17:81-91.
- Newsheer S, Aziz K, Panayiotidis MI, Georgakilas AG. Molecular markers for cancer prognosis and treatment: have we struck gold? *Cancer Lett.* 2012;327:142-52.
- Vargas AJ, Harris CC. Biomarker development in the precision medicine era: lung cancer as a case study. *Nat Rev Cancer.* 2016;16:525-37.
- Ganapathy-Kanniappan S. Molecular intricacies of aerobic glycolysis in cancer: current insights into the classic metabolic phenotype. *Crit Rev Biochem Mol Biol.* 2018;53:667-682.
- Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and Beyond. *Cell.* 2008;134:703-7.
- Koppenol WH, Bounds PL, Dang CV. Otto Warburg's contributions to current concepts of cancer metabolism. *Nat Rev Cancer.* 2011;11:325-37.
- Li Z, Zhang H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell Mol Life Sci.* 2016;73:377-92.
- Akins NS, Nielson TC, Le HV. Inhibition of glycolysis and glutaminolysis: an emerging drug discovery approach to combat cancer. *Curr Top Med Chem.* 2018;18:494-504.
- Matsuura K, Canfield K, Feng W, Kurokawa M. Metabolic regulation of apoptosis in cancer. *Int Rev Cell Mol Biol.* 2016;327:43-87.
- Ganapathy-Kanniappan S, Geschwind JFH. Tumor glycolysis as a target for cancer therapy: progress and prospects. *Mol Cancer.* 2013;12:152.
- Ganapathy-Kanniappan S. Linking tumor glycolysis and immune evasion in cancer: emerging concepts and therapeutic opportunities. *Biochim Biophys Acta Rev Cancer.* 2017;1868 :212-220.
- Gomes AS, Ramos H, Soares J, Saraiva L. p53 and glucose metabolism: an orchestra to be directed in cancer therapy. *Pharmacol Res.* 2018 May;131:75-86.
- Krzyslak A, Wojcik-Krowiranda K, Forma E, et al. Expression of GLUT1 and GLUT3 glucose transporters in endometrial and breast cancers. *Pathol Oncol Res.* 2012;18:721-8.
- Ma Y, Yu C, Mohamed EM, et al. A causal link from ALK to hexokinase II overexpression and hyperactive glycolysis in EML4-ALK-positive lung cancer. *Oncogene.* 2016;35:6132-6142.
- Li XB, Gu JD, Zhou QH. Review of aerobic glycolysis and its key enzymes – new targets for lung cancer therapy. *Thorac Cancer.* 2015;6:17-24.
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102:15545-50.
- Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43:e47.
- Steyerberg EW, Jr FEH, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54:774-81.
- Rhodes DR, Yu J, Shanker K et al. ONCOMINE: A cancer microarray database and integrated data-mining platform. *Neoplasia.* 2004;6:1-6.
- Li T, Fan J, Wang B, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 2017;77:e108-e110.
- Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2:401-4.
- Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013;6:pl1.
- Tanner LB, Goglia AG, Wei MH. Four key steps control glycolytic flux in mammalian cells. *Cell Syst.* 2018;7:49-62.e8.
- San-Millán I, Brooks GA. Reexamining cancer metabolism: lactate production for carcinogenesis could be the purpose and explanation of the warburg effect. *Carcinogenesis.* 2017;38:119-133.
- Jia D, Lu M, Jung KH, et al. Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. *Proc Natl Acad Sci U S A.* 2019;116:3909-3918.
- Israelsen WJ, Heiden MG. Pyruvate kinase: function, regulation and role in cancer. *Semin Cell Dev Biol.* 2015;43:43-51.
- Graziano F, Ruzzo A, Giacomini E. Glycolysis gene expression analysis and selective metabolic advantage in the clinical progression of colorectal cancer. *Pharmacogenomics J.* 2017;17:258-264.
- Luengo A, Abbott KL, Davidson SM, et al. Reactive metabolite production is a targetable liability of glycolytic metabolism in lung cancer. *Nat Commun.* 2019; 10:5604.
- Faubert B, Li KY, Cai L, et al. Lactate metabolism in human lung tumors. *Cell.* 2017;171(2):358-371.e9.
- Hensley CT, Faubert B, Yuan Q, et al. Metabolic heterogeneity in human lung tumors. *Cell.* 2016;164:681-94.

38. Li J, Cheng D, Zhu M. OTUB2 stabilizes U2AF2 to promote the Warburg effect and tumorigenesis via the AKT/mTOR signaling pathway in Non-Small Cell Lung Cancer. *Theranostics*. 2019 ;9:179-195.
39. Wiel C, Gal KL, Ibrahim MX, et al. BACH1 stabilization by antioxidants stimulates lung cancer metastasis. *Cell*. 2019;178:330-345.e22.
40. Zhang L, Zhang Z, Yu Z. Identification of a novel glycolysis-related gene signature for predicting metastasis and survival in patients with lung adenocarcinoma. *J Transl Med*.2019;17:423.
41. Liu C, Li Y, Wei M, et al. Identification of a novel glycolysis-related gene signature that can predict the survival of patients with lung adenocarcinoma. *Cell Cycle*. 2019;18:568-579.
42. Pathria G, Scott DA, Feng Y, Lee JS, et al. Targeting the Warburg effect via LDHA inhibition engages ATF4 signaling for cancer cell survival. *EMBO*. 2018;37(20):e99735.
43. Jin L, Chun J, Pan C, et al. Phosphorylation-mediated activation of LDHA promotes cancer cell invasion and tumour metastasis. *Oncogene*. 2017;36:3797-3806.
44. Ooi AT, Gomperts BN. Molecular pathways: targeting cellular energy metabolism in cancer via inhibition of SLC2A1 and LDHA. *Clin Cancer Res*. 2015;21:2440-4.
45. Connell M, Chen H, Jiang J, et al. HMMR acts in the PLK1-dependent spindle positioning pathway and supports neural development. *Elife*. 2017;6:e28672.
46. Tilghman J, Wu H, Sang Y, et al. HMMR maintains the stemness and tumorigenicity of glioblastoma stem-like cells. *Cancer Res*. 2014; 74 :3168-79.
47. Zhang H, Ren L, Ding Y, et al. Interdenominational motility receptor confers resistance to chemotherapy via TGFβ/Smad2-induced epithelial/mesenchymal transition in gastric cancer. *FASEB J*.2019;33:6365-6377.
48. Yang D, Ma Y, Zhao P, et al. Systematic screening of protein-coding gene expression identified HMMR as a potential independent indicator of unfavorable survival in patients with papillary muscle-invasive bladder cancer. *Biomed Pharmacother*. 2019;120:109433.
49. Bidadi B, et al. Pathway-based analysis of genome-wide association data identified SNPs in HMMR as biomarker for chemotherapy-induced neutropenia in breast cancer patients. *Front Pharmacol*. 2018;9:158.
50. Barkeer S, Chugh S, Karmakar S, et al. Novel role of O-glycosyltransferases GALNT3 and B3GNT3 in the self-renewal of pancreatic cancer stem cells. *BMC Cancer*. 2018;18:1157.
51. Sun Y, Liu T, Xian L, et al. B3GNT3, a direct target of miR-149-5p, promotes Lung cancer development and indicates poor prognosis of lung cancer. *Cancer Manag Res*. 2020;12:2381-2391.
52. Zhang W, Hou T, Niu C, et al. B3GNT3 expression is a novel marker correlated with pelvic lymph node metastasis and poor clinical outcome in early-stage cervical cancer. *PLoS One*. 2015;10:e0144360.
53. Gao LW, Zhang H, Zhang B, et al. B3GNT3 overexpression is associated with unfavourable survival in non-small cell lung cancer. *J Clin Pathol*. 2018;71:642-647.
54. Ho WL, Che ML, Chou CH, et al. B3GNT3 expression suppresses cell migration and invasion and predicts favorable outcomes in neuroblastoma. *Cancer Sci*. 2013;104:1600-8.
55. Li CW, Lim SO, Chung EM, et al. Eradication of triple-negative breast cancer cells by targeting glycosylated PD-L1. *Cancer Cell*. 2018;33:187-201.
56. Wu M, Sun Q, Mo CH, et al. Prospective molecular mechanism of COL5A1 in breast cancer based on a microarray, RNA sequencing and immunohistochemistry. *Oncol Rep*. 2019;42:151-175.
57. Feng G, Ma HM, Huang HB, et al. Overexpression of COL5A1 promotes tumor progression and metastasis and correlates with poor survival of patients with clear cell renal cell carcinoma. *Cancer Manag Res*. 2019;11:1263-1274.
58. Liu W, Wei H, Gao Z, et al. COL5A1 may contribute the metastasis of lung adenocarcinoma. *Gene*. 2018;665:57-66.
59. Chen HC, Tseng YK, Shu CW, et al. Differential clinical significance of COL5A1 and COL5A2 in tongue squamous cell carcinoma. *J Oral Pathol Med*. 2019;48:468-476.