RESEARCH

Identification of a novel glycolysis-related signature to predict the prognosis of patients with breast cancer

Menglin He[†], Cheng Hu[†], Jian Deng, Hui Ji and Weiqian Tian[*]

Abstract

Background: Breast cancer (BC) has a high incidence and mortality rate in females. Its conventional clinical characteristics are far from accurate for the prediction of individual outcomes. Therefore, we aimed to develop a novel signature to predict the survival of patients with BC.

Methods: We analyzed the data of a training cohort from the Cancer Genome Atlas (TCGA) database and a validation cohort from the Gene Expression Omnibus (GEO) database. After the applications of Gene Set Enrichment Analysis (GSEA) and Cox regression analyses, a glycolysis-related signature for predicting the survival of patients with BC was developed; the signature contained *AK3*, *CACNA1H*, *IL13RA1*, *NUP43*, *PGK1*, and *SDC1*. Furthermore, on the basis of expression levels of the six-gene signature, we constructed a risk score formula to classify the patients into high- and low-risk groups. The receiver operating characteristic (ROC) curve and the Kaplan-Meier curve were used to assess the predicted capacity of the model. Later, a nomogram was developed to predict the outcomes of patients with risk score and clinical features over a period of 1, 3, and 5 years. We further used Human Protein Atlas (HPA) database to validate the expressions of the six biomarkers in tumor and sample tissues, which were taken as control.

Results: We constructed a six-gene signature to predict the outcomes of patients with BC. The patients in the high-risk group showed poor prognosis than those in the low-risk group. The area under the curve (AUC) values were 0.719 and 0.702, showing that the prediction performance of the signature is acceptable. Additionally, Cox regression analysis revealed that these biomarkers could independently predict the prognosis of BC patients with BC without being affected by clinical factors. The expression levels of all six biomarkers in BC tissues were higher than that in normal tissues; however, *AK3* was an exception.

Conclusion: We developed a six-gene signature to predict the prognosis of patients with BC. Our signature has been proved to have the ability to make an accurate prediction and might be useful in expanding the hypothesis in clinical research.

Keywords: Bioinformatics, Glycolysis, Gene signature, Breast cancer, Prognosis

Introduction

In 2020, breast cancer (BC) was estimated to account for 30% of cancers in females, in which 15% of the cases lead to death [1, 2]. So far, the effects of novel biomarkers on diagnosis and prognosis are far from satisfactory [3]. Moreover, despite considering the option of lymph node dissection for the treatment of BC, there are still some limitations in clinical practice [4, 5]. Therefore, there

*Correspondence: yfy0066@njucm.edu.cn
[†]Menglin He and Cheng Hu contributed equally to this work.
Department of Anesthesiology, Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Chinese Medicine, No. 155 Hanzhong Road, Qinhuai District, Nanjing 210029, Jiangsu, China

He *et al. World J Surg Onc*   (2021) 19:294

Page 2 of 14

is a need to construct an effective prognostic model for predicting the outcomes of patients with BC in clinical practice.

With the supply of adequate oxygen, normal cells metabolize glucose to pyruvate via glycolysis that is further used to produce adenosine triphosphate (ATP) through mitochondrial oxidative phosphorylation; whereas in the case of anaerobic or hypoxic conditions, normal cells ferment glucose to lactate [6]. However, even after a sufficient supply of oxygen, tumor cells mainly use the mechanism of glycolysis to produce energy and thus have a high glycolysis rate. This phenomenon is known as the Warburg effect, which is found in all types of cancer [7]. Many studies have proved that glycolysis can accelerate the proliferation, invasion, and migration of certain tumor cells and enhance drug resistance [8, 9]. Therefore, these glycolysis-related genes and proteins can be used as targets for prognosis or treatment in patients with BC. For instance, pieces of evidence have proved that key enzymes in the aerobic glycolytic pathway including hexokinase (HK), phosphofructokinase (PFK), and pyruvate kinase (PK) can serve as potential therapeutic targets for anti-tumor [10]. A previous study reported that the HKII, one isoenzyme of HK, is overexpressed in tumor cells [11]. PFK-1, a rate-limiting enzyme in glycolysis, promotes the conversion of fructose 6-phosphate to fructose 1,6-bisphosphate and adenosine diphosphate (ADP) [12]. This conversion is facilitated by 6-phosphofructo 2-kinase/fructose 2, 6-bisphosphatase (PFKFB). As a subtype of PFKFB, PFKFB3 has been found in patients with BC in high expression levels and was further implicated to a poor prognosis in clinical practice [13]. PKM2 is an isoenzyme of PK and is mainly located in muscles [14]. It has been reported that the overexpression of PKM2 is related to worse overall survival (OS) and progression-free survival (PFS) in patients with BC [14]. Among the 14 different types of glucose transporter protein (GLUT), the notable GLUT1 subtype was found to have a close correlation with the progression of tumor [15]. Li et al. found that the overexpression of GLUT1 is associated with more severe outcomes in patients with head and neck squamous cell carcinoma (HNSCC) that was further ascribed to the activation of the nuclear factor kappa B (NF-κB) signaling pathway [16]. Moreover, the expression level of GLUT1 increased significantly in the kidney cancer [17], ovarian cancer [18], and liver cancer [19].

With the development of bioinformatics, biomarkers are used in comparatively more comprehensive studies to predict the prognosis of patients [20, 21]. Nine glycolysis-related genes (GRGs) were identified as being closely relevant to worse outcomes in endometrial cancer [22] and glioblastoma [23], respectively. Another set of 11 GRGs prognostic model was constructed and served as

an effective tool to predict prognosis and guide clinical practice [24]. Circ-FOXM1 has been demonstrated that it can contribute to cell proliferation and glycolysis in melanoma [25]. These studies made an effort to probe the effect of glycolysis, but the role of glycolysis-related genes involved in BC is far from understood. Therefore, it is of great importance to profoundly understand the effect of glycolysis on tumorigenesis and tumor progressions.

In this study, we conducted a six-GRGs risk signature to predict the prognosis of patients with BC by integrating high-throughput data. This prognostic model might become a part of clinical prognostic features and provides a novel insight into BC study.

Materials and methods
Data processing
The overall design of this study is shown in Fig. 1. The gene expression and relevant clinicopathological data of patients with BC were downloaded from Genomic Data Commons (GDC) database (https://portal.gdc.cancer.gov/). A total of 1066 BC samples and 112 normal samples were included in our study, which was set as a training cohort. Whereas 198 BC samples in the *GSE25065* dataset were obtained from the Gene Expression Omnibus (GEO) of the National Center of Biotechnology Information (NCBI) database [26] and were set as a validation cohort, clinical information, including age, American Joint Committee on Cancer (AJCC) stage, grade, and TNM staging were presented (Table 1). Kaplan-Meier plotter (http://kmplot.com/analysis/) was performed to evaluate the prediction capacity of biomarkers.

Gene set enrichment analysis
Gene Set Enrichment Analysis (GSEA) (http://www.broadinstitute.org/gsea/index.jsp) was applied to determine whether there is any significant difference in the expression levels between BC tissues and normal tissues. Eight gene-sets related to glycolysis were downloaded from the Molecular Signature Database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp). Each gene set had 1000 permutations to get a normalized enrichment score (NES), false discovery rate (FDR), and normalized P value (P), wherein NES $\geq$ 1.5, FDR < 0.1, and P < 0.05 were set to screen the eligible gene set for the following analysis.

Differential expression analysis and model construction of glycolysis-related genes
Firstly, we performed univariate Cox proportional hazard regression analysis to determine the relationship between glycolysis-related genes and OS. Genes with P<0.05 were deemed as candidate prognostic genes for BC. Then, multivariate Cox proportional hazard regression analysis
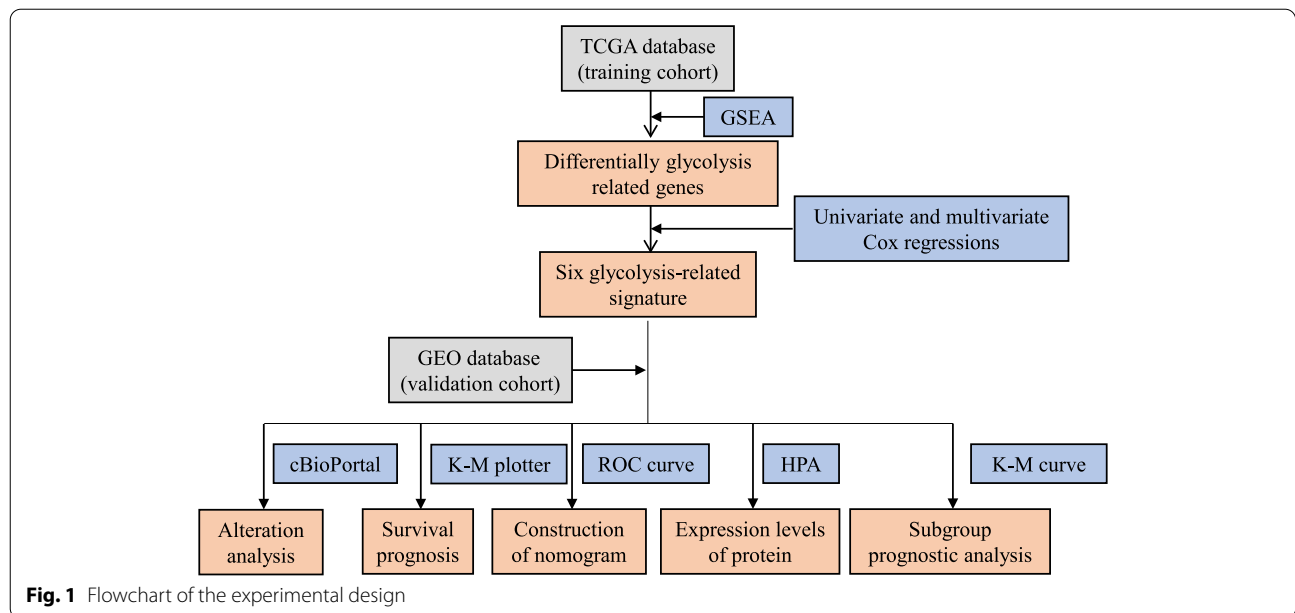
He *et al. World J Surg Onc*   (2021) 19:294

Page 3 of 14

Fig. 1 Flowchart of the experimental design

Table 1 Clinical characters in patients with BC in TCGA dataset

Characters	Subgroup	Ratio
Age	≤ 60	581 (54.5%)
	> 60	465 (43.62%)
	Unknown	20 (1.87%)
AJCC stage	I	172 (16.35%)
	II	597 (56.75%)
	III	239 (22.72%)
	IV	20 (1.9%)
	Unknown	24 (2.28%)
T	I	270 (25.66%)
	II	614 (58.37%)
	III	129 (12.26%)
	IV	36 (3.42%)
	Unknown	3 (0.29%)
M	0	875 (83.17%)
	1	22 (2.09%)
	Unknown	155 (14.73%)
N	0	483 (45.91%)
	1	354 (33.65%)
	2	119 (11.31%)
	3	76 (7.22%)
	Unknown	20 (1.9%)
Status	Dead	138 (13.12%)
	Alive	914 (86.88%)

was used to establish the prognostic model based on the abovementioned candidate prognostic genes. The risk score was calculated using the following formula: Risk score$= \sum_{i=1}^{n} coef * id$. The median value of the risk score

was set as cut-off, and samples were divided into high- or low-risk groups. We conducted the Kaplan-Meier survival analysis to assess the difference in survival between the two groups using the R package survival. The time-dependent ROC curve was used to assess the accuracy of the model through the R package survival ROC. Similarly, the same signature also generates risk scores in each patient in the *GSE25065* dataset that could be used to validate the performance of the model.

The glycolysis-related signature is an independent prognostic factor for breast cancer

Univariate, and multivariate Cox regression analysis, along with data stratification analysis was performed to assess whether the risk score was independent of the clinical characters, including age and grade. *P* < 0.05 was considered as statistically significant.

cBioPortal analysis

The cBioportal for Cancer Genomic database (http://www.cbiop ortal.org) was used to analyze the forms and ratios of alterations in biomarkers in BC. All steps were performed complying with the instructions of the cBioPortal.

Construction and validation of nomogram

The nomograms were developed to explore the prognosis of patients with BC in 1-, 3-, and 5-year overall survival. We further used calibration curves and concordance (C_index) to assess the predictive capacity of the nomogram, both in training and validation cohorts.

He *et al. World J Surg Onc* (2021) 19:294

Page 4 of 14

Human Protein Atlas

The immunohistochemical (IHC) images in the Human Protein Atlas (HPA) database (https://www.proteinatlas.org/) were applied to validate the expression of biomarkers in tumor and normal tissues.

Statistical analysis

R software and Perl languages (http://www.perl.org/) were used to perform all the statistical analyses. Before the analysis, the gene expression data were normalized via log2 transformation. The threshold of $P < 0.05$ was considered significantly different.

Results

Screening genes with gene set enrichment analysis

We downloaded a dataset including the gene expression and clinical information on 1066 BC patients and 112 normal controls from THE Cancer Genome Atlas (TCGA). Eight glycolysis-related gene sets were downloaded, and a total of 410 genes were obtained. We used the data mentioned above and GSEA to analyze the possibility of significant differences in gene-sets between BC samples and normal samples and obtained four significantly enriched gene-sets in our study (Fig. 2). Next, we sorted out 284 glycolysis-related genes through the integration of these four gene-sets.

Construction of a six-gene signature as a prognostic indicator

Univariate analysis was applied to analyze the 284 glycolysis-related genes, and eight genes (*AK3*, *CACNA1H*, *IL13RA1*, *NUP43*, *PGAM1*, *PGK1*, *P4HA2*, and *SDC1*) were obtained. Furthermore, multivariate analysis was performed to screen genes associated with the survival of patients with BC. Finally, six genes were identified as prognostic genes, with each gene having a coefficient value (Table 2). Risk score = coefficient gene 1 × expr (gene 1) + coefficient gene 2 × expr (gene 2) + ... + coefficient gene n × expr (gene n), where expr is the expression of the corresponding gene. The median risk
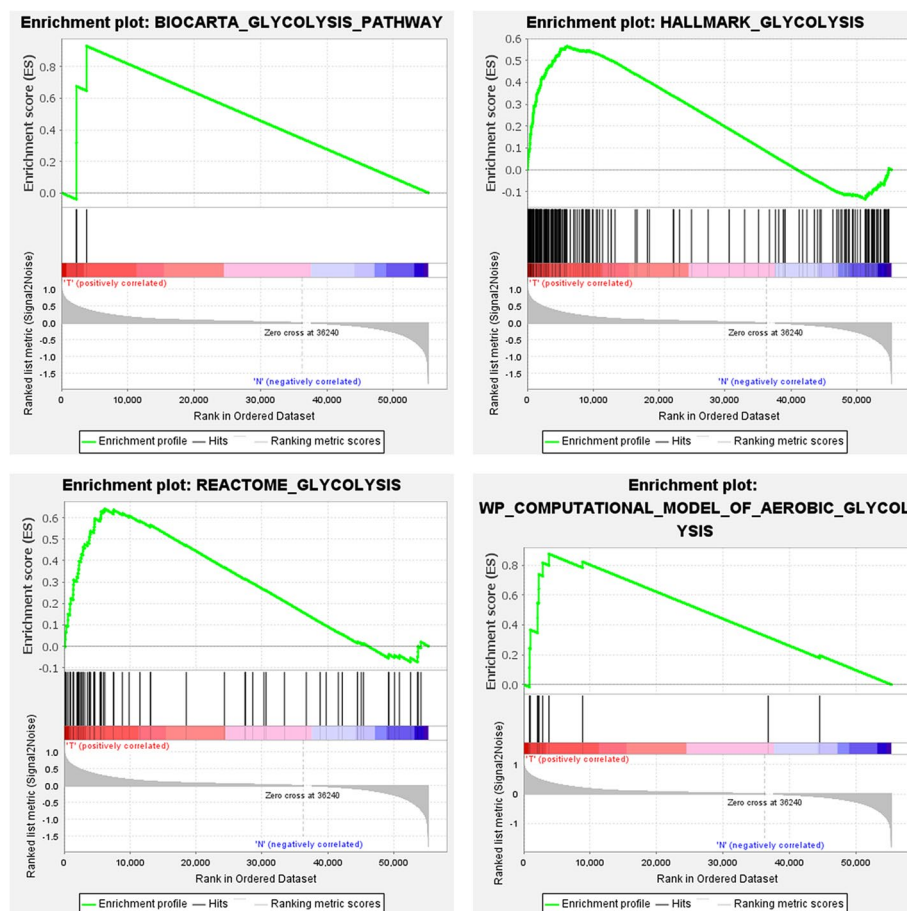

Fig. 2 Enrichment of four glycolysis-related gene sets in GSEA. **A** BIOCARTA_GLYCOLYSIS_PATHWAY. **B** HALLMARK_GLYCOLYSIS. **C** REACTOME_GLYCOLYSIS. **D** WP_COMPUTATIONAL_MODEL_OF_AEROBIC_GLYCOLYSIS

He *et al. World J Surg Onc*   (2021) 19:294

Page 5 of 14

Table 2 Univariate and multivariate Cox regression analysis for glycolysis-related genes and respective coefficient

Gene	Univariate Cox regression		Gene	Multivariate Cox regression		
	HR (95% CI)	P value		HR (95% CI)	P value	Coefficient
AK3	0.964 (0.931-0.997)	0.034	AK3	0.975 (0.942-1.009)	0.145	−0.0255
CACNA1H	1.018 (1.005-1.032)	0.008	CACNA1H	1.017 (1.004-1.031)	0.013	0.01724
IL13RA1	1.009 (1.002-1.016)	0.010	IL13RA1	1.005 (0.998-1.013)	0.139	0.00545
NUP43	1.054 (1.022-1.088)	***	NUP43	1.049 (1.016-1.084)	0.003	0.04819
PGAM1	1.038 (1.006-1.070)	0.019	PGK1	1.006 (1.003-1.009)	***	0.00594
PGK1	1.008 (1.005-1.010)	***	SDC1	1.002 (1.000-1.004)	0.012	0.00215
P4HA2	1.069 (1.024-1.116)	0.002				
SDC1	1.002 (1.001-1.004)	0.002				

HR hazard ratio, *CI* confidence interval

***$p < 0.001$

score was set as the threshold, and patients were divided into high- and low-risk groups. The Kaplan-Meier survival curve demonstrated that the survival rate of the high-risk group was significantly lower than that of the low-risk group (Fig. 3A). The area under the curve (AUC) value of the dependent ROC for the six-mRNA signature was 0.719, indicating that the signature could accurately predict the survival of patients with BC. The consistent results were obtained in the validation cohort (Fig. 3B). Moreover, the risk scores, survival status, and expression levels of biomarkers were drawn in both training and validation cohorts (Fig. 3C, D).

The prediction of biomarkers on prognosis was validated and used as an independent prognostic indicator in patients with breast cancer

Kaplan-Meier plotter database was used to assess the prognostic ability of biomarkers. Patients were divided into two groups based on the expression levels of biomarkers, and the OS of patients were analyzed. We found that the high expression of the other five biomarkers is related to shorter OS, while that of AK3 indicates a better OS (Fig. 4).

To assess the prognostic ability of biomarkers and to analyze whether they can be independent of clinical features, including age and grade, univariate and multivariate Cox regression analyses were performed. The results demonstrated that biomarkers could independently predict the survival of patients both in training and validation cohorts (Fig. 5A, B).

Analysis of the biomarkers

Firstly, we analyzed the genetic alterations of biomarkers in BC by using the cBioPortal database. The results showed that 280 (12.88%) of the 2173 patients had genetic alterations. Among the six biomarkers, CACNA1H

became the most commonly altered gene with a rate of 13% (Fig. 6A), and 258 (92.14%) of 280 cases had amplification that was the most frequent gene alternation in BC (Fig. 6B).
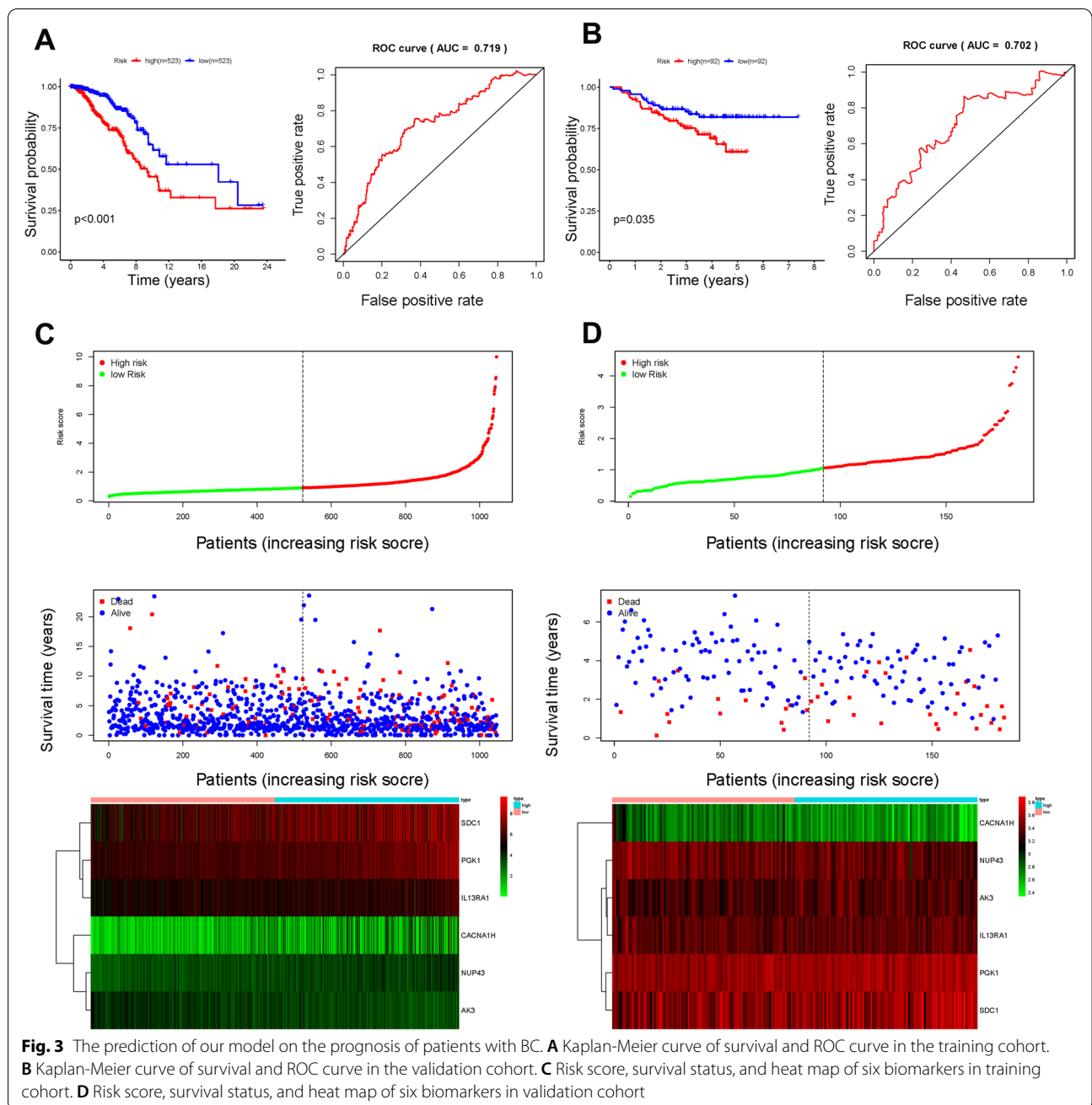
We further assessed possible significant differences in expression levels of biomarkers between tumor samples and normal samples from our TCGA dataset. The results showed that *CACNA1H*, *IL13RA1*, *NUP43*, *PGK1*, and *SDC1* were highly expressed in tumor samples, while expression of *AK3* was significantly decreased (Fig. 6C).

Development of a nomogram to predict the prognosis of patients in breast cancer with a risk score and clinical characters

To further evaluate the prognostic ability of our model and its clinical parameters, we presented the model with a nomogram both in TCGA and GEO datasets (Fig. 7A, B). Every indicator got a point, and the total number of points was calculated by the sum of all points that could predict the prognosis of patients in 1-, 3-, and 5-year survival. In training and validation cohorts, the C_index values were 0.764 and 0.677, respectively. The results of calibration curves demonstrated that the nomogram had a good accuracy of prediction on the prognosis of patients with BC (Fig. 7C, D).

Correlation between the risk score and clinical characters

Firstly, we stratified the patients into the subgroups of age (≥ 60 and < 60), AJCC stage (I + II and III + IV), T stage (T1-2 and T3-4), N stage (N0 and N1-3), and M stage (M0 and M1). The Kaplan-Meier analysis was applied by log-rank test to assess the prediction capacity of multiple clinical characters on patients with BC. The results showed that the performances of these characters were well satisfied (Fig. 8A-E). Furthermore, we further found that the prognosis of high-risk patients in

He *et al. World J Surg Onc*   (2021) 19:294

Page 6 of 14

Fig. 3 The prediction of our model on the prognosis of patients with BC. **A** Kaplan-Meier curve of survival and ROC curve in the training cohort. **B** Kaplan-Meier curve of survival and ROC curve in the validation cohort. **C** Risk score, survival status, and heat map of six biomarkers in training cohort. **D** Risk score, survival status, and heat map of six biomarkers in validation cohort

these subgroups was poorer than that of low-risk patients (Fig. 8F-N).
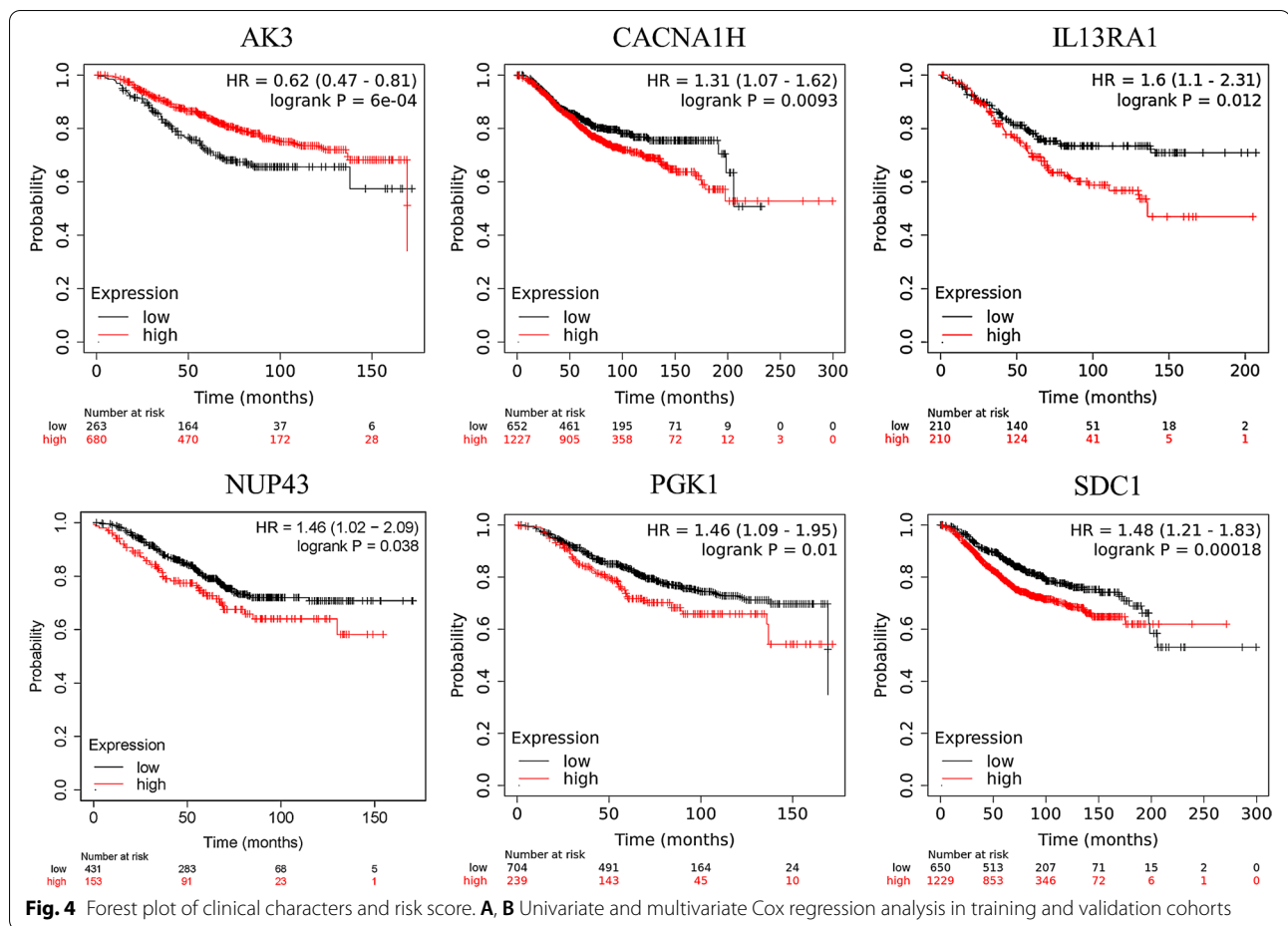
The protein expression levels of biomarkers in breast cancer tissues

To validate the significant differences in protein expression levels of biomarkers between BC tissues and normal tissues, we performed an HPA database to assess the IHC images. Consistently, the results demonstrated that the protein expression levels of *CACNA1H, NUP43,*

PGK1, and *SDC1* were higher in BC tissues compare with normal tissues, while the expression level of *AK3* was comparatively lower (Fig. 9). Unfortunately, the data of IL13RA1 was not available in the HPA database.

Discussion

The prediction on prognosis in BC is far from satisfactory, due to the complication of its phenotypes and molecular mechanisms. Previous studies revealed that clinicopathological characters, including age, gender,

He *et al. World J Surg Onc*   (2021) 19:294

Page 7 of 14

Fig. 4 Forest plot of clinical characters and risk score. **A**, **B** Univariate and multivariate Cox regression analysis in training and validation cohorts
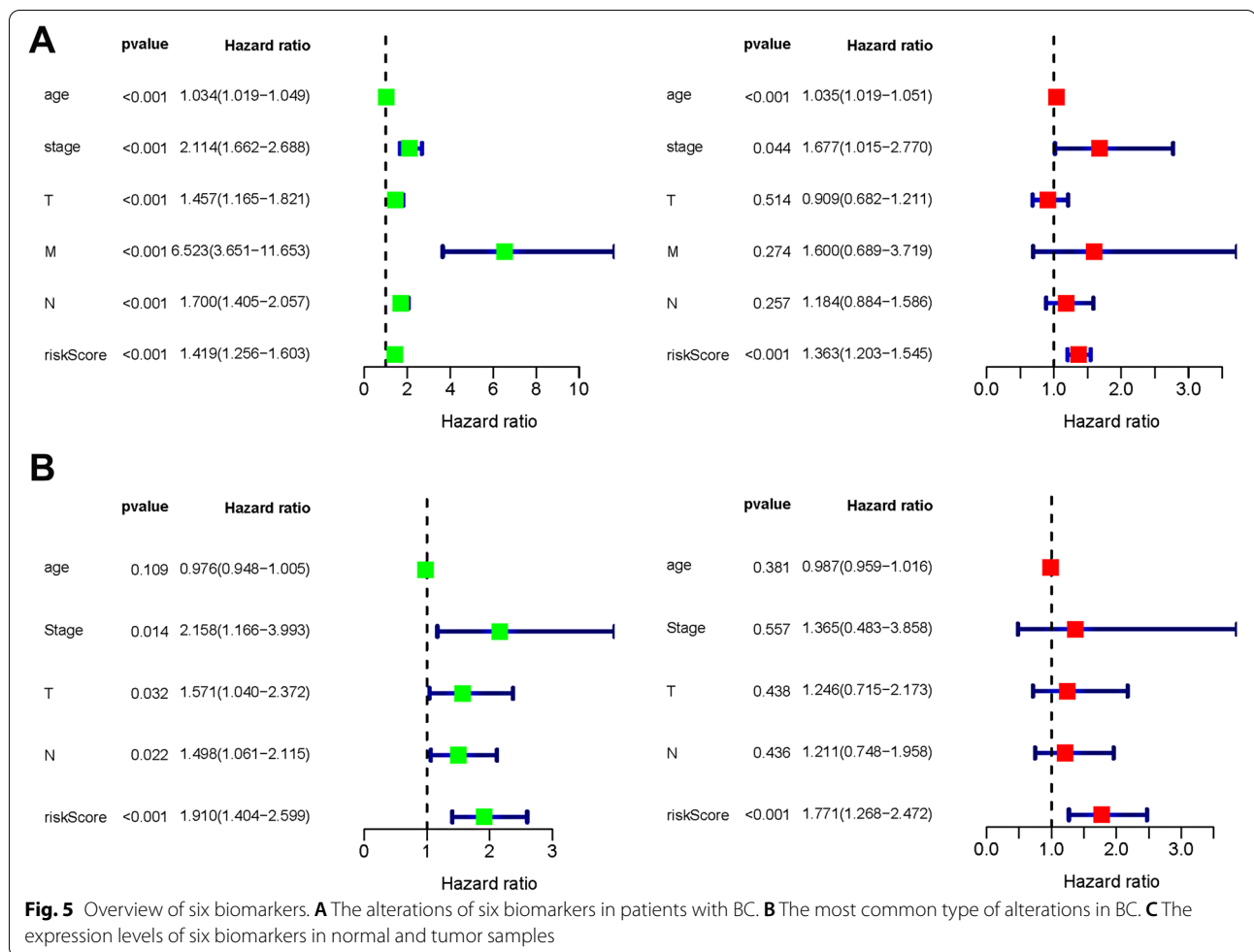
and metastatic diagnosis are insufficient to accurately assess the outcomes of patients with cancer [27]. Therefore, with the development of bioinformatics, an increasing number of novel biomarkers have been constructed to get involved in the diagnosis and prediction of cancer [28, 29]. These signatures are associated with metabolism [30], immune [31], and DNA methylation [32]. The standards to judge the model are convenient, credible, and precise.

Tumor cells have the ability of unlimited proliferation and distal metastasis, and this process needs cellular metabolism to meet the requirement of adequate energy [33]. ATP and nutrients are provided by aerobic glycolysis in tumor cells that form an acidic microenvironment to promote the mutation and invasion of tumor cells [34]. Recent studies have made efforts to clarify the role of glycolysis-related genes in tumor. Ang Li et al. [27] developed a glycolysis-related signature for predicting survival of pancreatic adenocarcinoma (PAAD) that provided a novel therapy target for PAAD. Lei Zhang et al. [35] also found another similar signature to predict survival in patients with

lung adenocarcinoma. Yuchao Liu et al. [36] explored the role of glycolysis in HNSCC and found that high expression levels of identified glycolysis-related genes are correlated to poor prognosis in patients with HNSCC. The abovementioned studies emphasized the importance of multiple genes rather than a single gene, which might be interfered with many factors and provide worse predictive effects. Therefore, we established a signature to predict the prognosis of survival in patients with BC, consisting of six genes.

In our study, firstly, we downloaded microarray data and clinical information from TCGA and GEO datasets, where the former was used as a training cohort, while the latter was defined as a validation cohort. Then, GSEA was applied to identify mRNAs associated with glycolysis that further found differentially expressed glycolysis-related genes. A signature was constructed through univariate and multivariate Cox regression that contained six glycolysis-related genes (*AK3*, *CACNA1H*, *IL13RA1*, *NUP43*, *PGK1*, and *SDC1*) and that would be used in the following analysis. We divided the training cohort into two groups based on

He *et al. World J Surg Onc*   (2021) 19:294

Page 8 of 14

Fig. 5 Overview of six biomarkers. **A** The alterations of six biomarkers in patients with BC. **B** The most common type of alterations in BC. **C** The expression levels of six biomarkers in normal and tumor samples

the expression levels of these six biomarkers. Survival analysis showed a significant difference between high- and low-risk groups. The AUC value of the model was greater than 0.7 in both training and validation cohorts. Additionally, this risk score, age, and the stage could act as an independent prognostic index. We further developed a nomogram to predict the prognosis of patients with BC, and the results validated that our model can be used as a prognostic indicator with a good performance. The consistent results were obtained in the validation cohort, demonstrating that the GRGs model had high accuracy and could predict the prognosis of patients with BC. We further explored whether this risk score could stratify clinically defined groups of patients into subgroups of age > 60, age ≤ 60, T1-2, T3-4, N0, T1-3, M0, M1, stage I-II, and stage III-IV. The results showed that all subgroups could be efficiently predicted by the risk score except the M1 subgroup, due to limited samples. However, when compared with traditional clinical features, the prediction model in this

study had similar, even better clinical application capacity. Furthermore, the HPA database was used to explore any existence of statistical significance in the expression of six biomarkers between BC and normal tissues. We found that four of six biomarkers were highly expressed, whereas *AK3* was significantly decreased in BC tissues. The results stated above suggested that our prediction model could have better reliability and accuracy, providing a possibility to be a prospective prognostic indicator for patients with BC.

To make sense of this signature, we analyzed each biomarker. *AK3* regulates the homeostasis of adenine nucleotide composition and has been proved to have the ability of anti-tumor [37]. *AK3* could promote the BC cell migration, and its decreased expression level was found to be related to a worse prognosis in patients with BC [38]. Consistently, the expression of *AK3* was also significantly downregulated in our study. A novel glycolysis-related gene signature, including
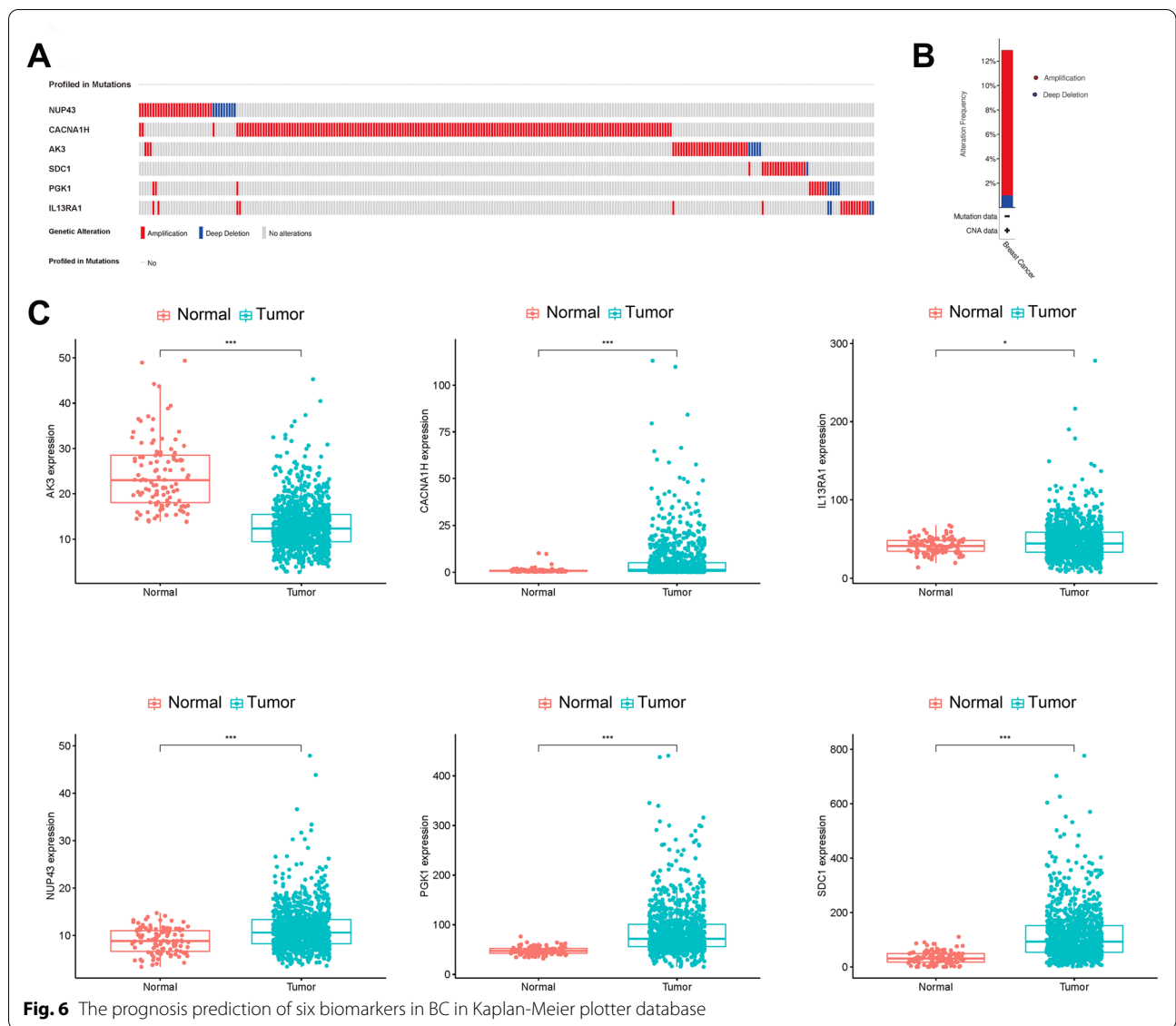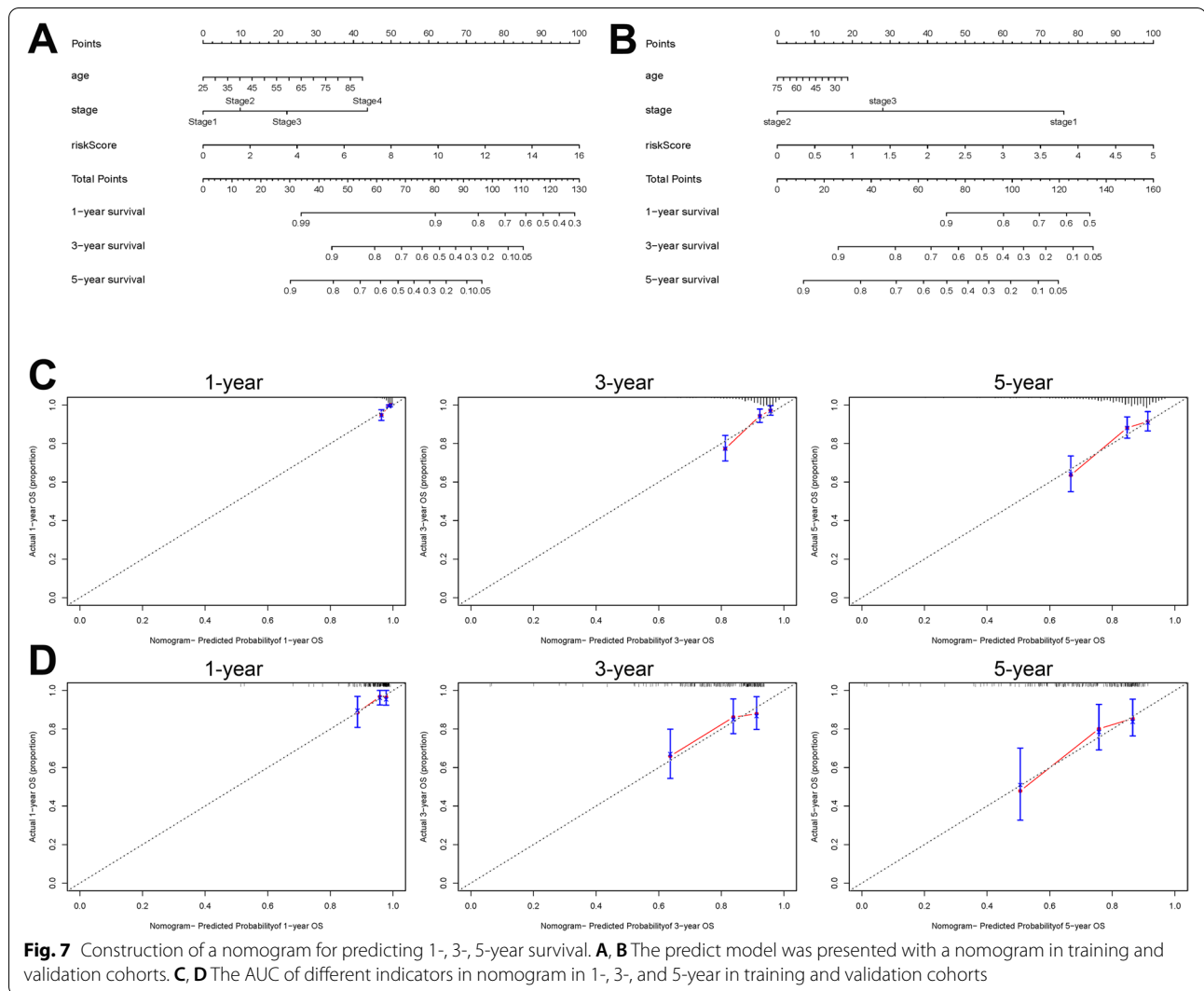
He *et al. World J Surg Onc*   (2021) 19:294

Page 9 of 14

Fig. 6 The prognosis prediction of six biomarkers in BC in Kaplan-Meier plotter database

CACNA1H, might provide a new indicator of prediction on patients with invasive BC [24]. Another study showed that *CACNA1H* was highly expressed in luminal A and B subtypes, with the basal subtype expressing significantly lower levels of *CACNA1H* [39]. IL-13 modulates the proliferation and status of lymphocytes and has been explored in tumor research [40]. As a subtype of IL-13, high *IL13RA1* expression has been found to have a poor prognosis in patients with invasive BC and was associated with HER2- and a high Ki-67 index, showing its role of potential prognostic marker in BC [41]. In luminal A and HER2+ BC, the expression level of *NUP43* was increased and

it further predicted poor prognosis [42]. A previous study demonstrated that *NUP43* plays the role of promoter in the progression of tumors [43]. Yan Zhang et al. developed glycolysis-related genes to predict the outcomes of patients in clear cell renal cell carcinoma (ccRCC); *PGK1* is one of these genes, whose high expression predicted poor prognosis in patients with ccRCC [44]. *PGK1* expression was also significantly increased in BC tissues and associated with HER2+ and positive status of ER [45]. Syndecan-1 (SDC), a heparin sulfate proteoglycan, has been found to correlate with tumor progressions [38]. Additionally, another study revealed that high *SDC1* expression is

He *et al. World J Surg Onc* (2021) 19:294

Page 10 of 14

Fig. 7 Construction of a nomogram for predicting 1-, 3-, 5-year survival. **A**, **B** The predict model was presented with a nomogram in training and validation cohorts. **C**, **D** The AUC of different indicators in nomogram in 1-, 3-, and 5-year in training and validation cohorts

associated with increased risked of age and HER2 in patients with BC [46]. However, more investigations should be done to clarify the biological functions of these biomarkers and their interactions in BC.

Moreover, we compared the performance of our signature with others. The AUC value is an indicator to evaluate the prognostic accuracy, which means the larger the AUC value, the better the predictive ability of the biomarkers. The AUC values in this study were 0.719 and 0.702 that are higher than that in a six-gene signature associated with tumor mutation burden [47], a prognostic signature based on eight DNA repair-related genes [48], and a seven RNA signature [49]. This demonstrated that our glycolysis-related gene signature has a better performance on prediction.

Conclusions

We constructed a glycolysis-related gene signature with *AK3, CACNA1H, IL13RA1, NUP43, PGK1*, and *SDC1* and identified them as potential prognostic biomarkers. The expression levels of these biomarkers have significant differences between BC and normal samples. Our signature can be used to predict the prognosis of patients with BC conveniently and precisely. However, there are some limitations in this study: (1) More datasets should be taken into consideration to guarantee the accuracy of results; (2) the mechanism of the six predicted genes should be explored profoundly; (3) experiments are needed to validate the results obtained from the bioinformatics analysis.
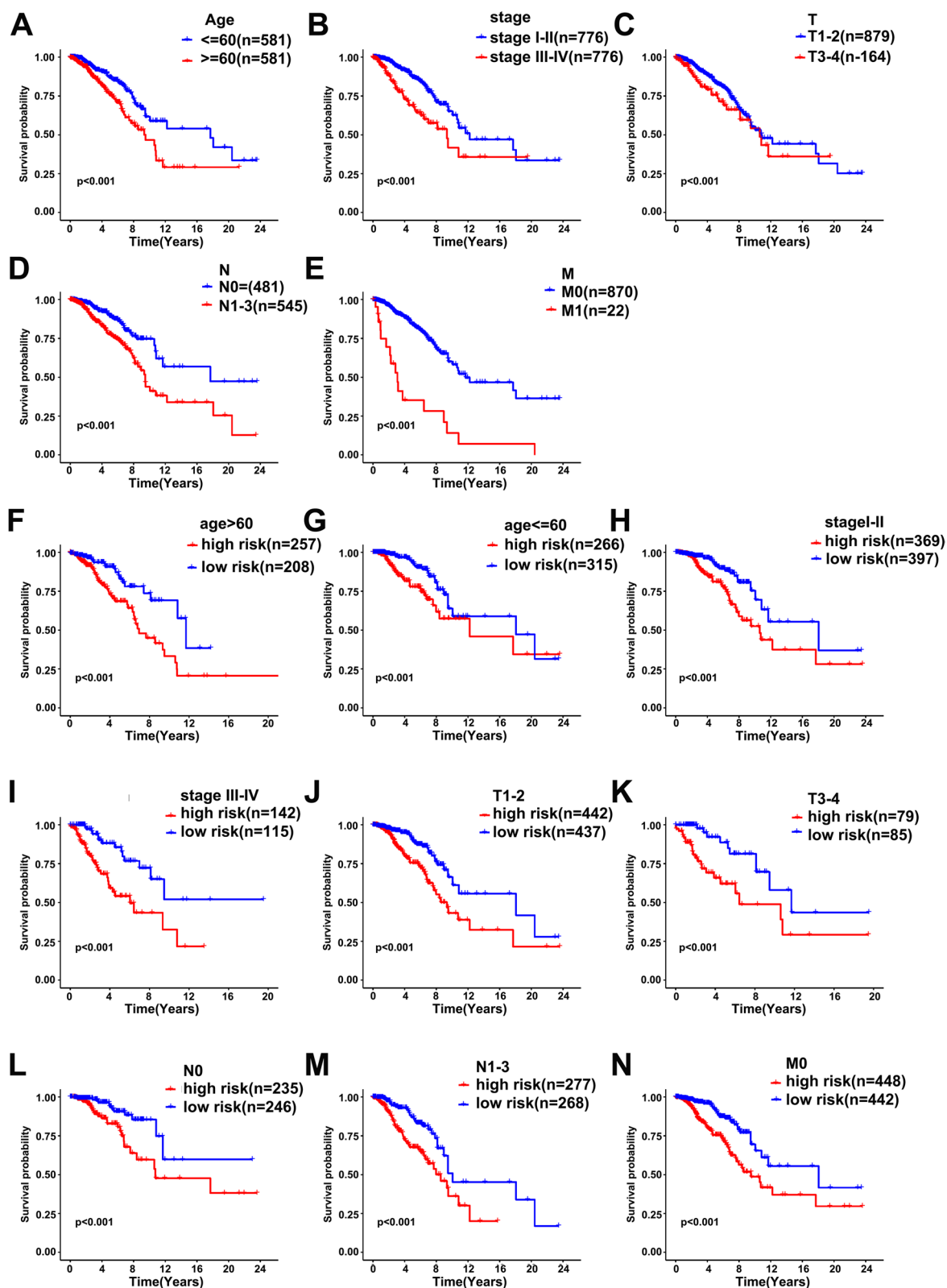
He *et al. World J Surg Onc*   (2021) 19:294

Page 11 of 14

Fig. 8 Kaplan-Meier survival analysis for BC patients with diverse clinical characters of (**A**) age, (**B**) AJCC Stage, (**C**) T stage, (**D**) N stage, (**E**) M stage. The outcomes of patients with different risk scores in subgroups of (**F**) age> 60, (**G**) age ≤ 60, (**H**) AJCC stage I-II, (**I**) AJCC stage III-IV, (**J**) T stage 1-2, (**K**) T stage 3-4, (**L**) N stage 0, (**M**) N stage 1-3, (**N**) M stage 0

He *et al. World J Surg Onc*   (2021) 19:294

Page 12 of 14
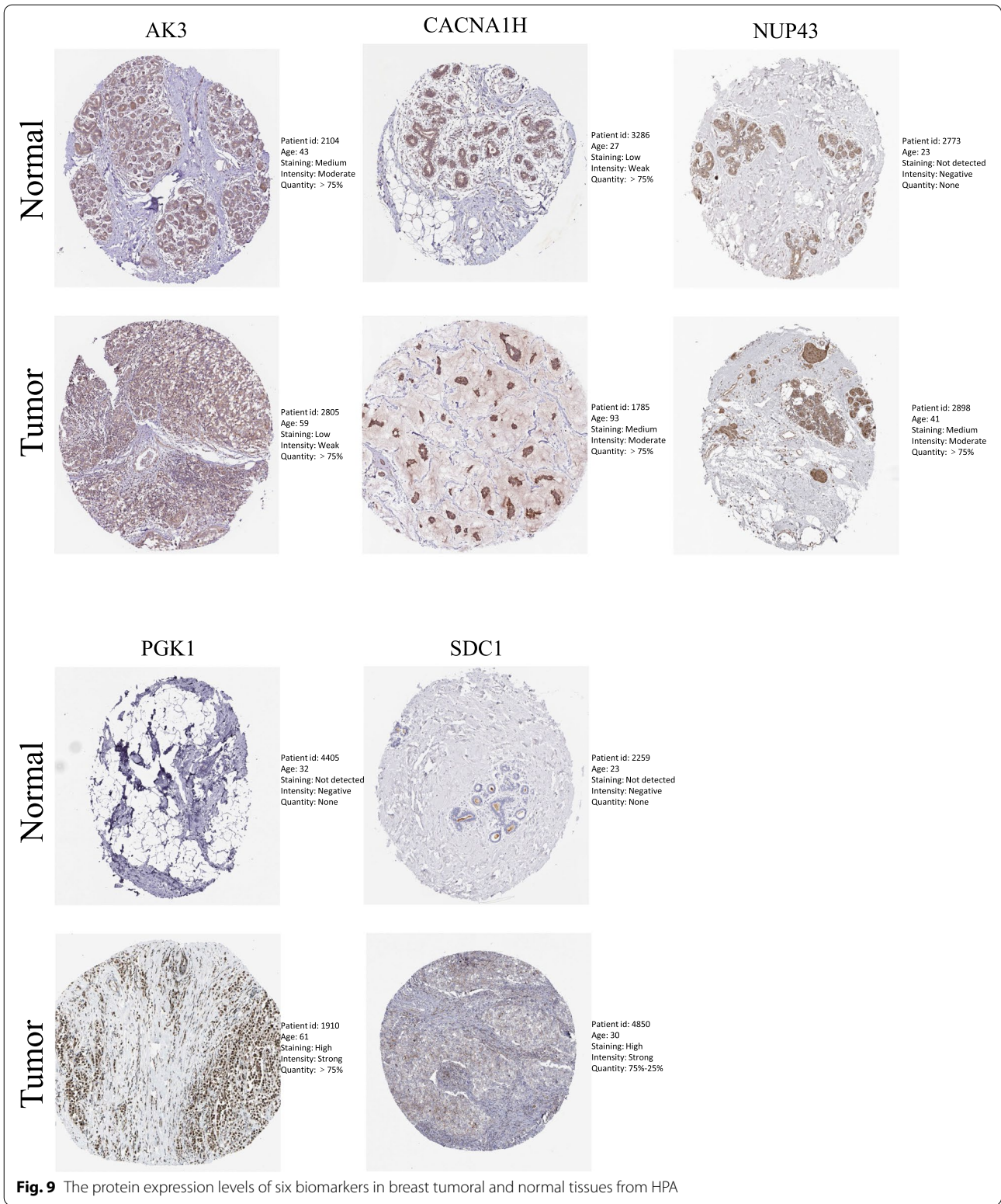
Fig. 9 The protein expression levels of six biomarkers in breast tumoral and normal tissues from HPA

He *et al. World J Surg Onc* (2021) 19:294

Page 13 of 14

Abbreviations

BC: Breast cancer; ROC: Receiver operating characteristic; HPA: Human Protein Atlas; AUC: Area under the curve; ATP: Adenosine triphosphate; HK: Hexokinase; PFK: Phosphofructokinase; PK: Pyruvate kinase; ADP: Adenosine diphosphate; PFKFB: 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase; OS: Overall survival; PFS: Progression free survival; GLUT: Glucose transporter protein; HNSCC: Head and neck squamous cell carcinoma; GRGs: Glycolysis-related genes; GDC: Genomic Data Commons; GEO: Gene Expression Omnibus; NCBI: National Center of Biotechnology Information; AJCC: American Joint Committee on Cancer; GSEA: Gene set enrichment analysis; NES: Normalized enrichment score; IHC: Immunohistochemical; HR: Hazard ratio; PAAD: Pancreatic adenocarcinoma.

Authors' contributions

WT designed the study. MH and CH searched databases and performed data analysis. JD and HJ wrote the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

All the data referred to this study are available from TCGA and GEO databases. TCGA: https://portal.gdc.cancer.gov/; GSE25065: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE25065.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69:7–34.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70:7–30.
3. Kim J. In silico analysis of differentially expressed genesets in metastatic breast cancer identifies potential prognostic biomarkers. World J Surg Oncol. 2021;19:188.
4. Hashem T, Abdelmoez A, Rozeka AM, Abdelazeem H. Intra-mammary lymph nodes, an overlooked breast cancer prognostic tool? World J Surg Oncol. 2021;19:114.
5. Peng Y, Liu M, Li X, Tong F, Cao Y, Liu P, et al. Application of the ACOSOG Z0011 criteria to Chinese patients with breast cancer: a prospective study. World J Surg Oncol. 2021;19:128.
6. Wong TL, Ng KY, Tan KV, Chan LH, Zhou L, Che N, et al. CRAF methylation by PRMT6 regulates aerobic glycolysis-driven hepatocarcinogenesis via ERK-dependent PKM2 nuclear relocalization and activation. Hepatology. 2020;71:1279–96.
7. Qiu Z, Wang L, Liu H. Hsa_circ_0001982 promotes the progression of breast cancer through miR-1287-5p/MUC19 axis under hypoxia. World J Surg Oncol. 2021;19:161.
8. Mao L, Wu X, Gong Z, Yu M, Huang Z. PDIA6 contributes to aerobic glycolysis and cancer progression in oral squamous cell carcinoma. World J Surg Oncol. 2021;19:88.
9. Liu Y, Li H, Ye X, Ji A, Fu X, Wu H, et al. Hsa_circ_0000231 knockdown inhibits the glycolysis and progression of colorectal cancer cells by regulating miR-502-5p/MYO6 axis. World J Surg Oncol. 2020;18:255.
10. 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.
11. Yang T, Ren C, Qiao P, Han X, Wang L, Lv S, et al. PIM2-mediated phosphorylation of hexokinase 2 is critical for tumor growth and paclitaxel resistance in breast cancer. Oncogene. 2018;37:5997–6009.
12. Okar D, Manzano A, Navarro-Sabatè A, Riera L, Bartrons R, Lange AJT. PFK-2/FBPase-2: maker and breaker of the essential biofactor fructose-2,6-bisphosphate. Trends in biochemical sciences. 2001;26:30–5.
13. Peng F, Li Q, Sun JY, Luo Y, Chen M, Bao Y. PFKFB3 is involved in breast cancer proliferation, migration, invasion and angiogenesis. Int J Oncol. 2018;52:945–54.
14. Dayton TL, Jacks T, Vander Heiden MG. PKM2, cancer metabolism, and the road ahead. Embo Rep. 2016;17:1721–30.
15. Evans A, Bates V, Troy H, Hewitt S, Holbeck S, Chung YL, et al. Glut-1 as a therapeutic target: increased chemoresistance and HIF-1-independent link with cell turnover is revealed through COMPARE analysis and metabolomic studies. Cancer Chemother Pharmacol. 2008;61:377–93.
16. Li S, Yang X, Wang P, Ran X. The effects of GLUT1 on the survival of head and neck squamous cell carcinoma. Cell Physiol Biochem. 2013;32:624–34.
17. Chan DA, Sutphin PD, Nguyen P, Turcotte S, Lai EW, Banh A, et al. Targeting GLUT1 and the Warburg effect in renal cell carcinoma by chemical synthetic lethality. Sci Transl Med. 2011;3:94ra70.
18. Semaan A, Munkarah AR, Arabi H, Bandyopadhyay S, Seward S, Kumar S, et al. Expression of GLUT-1 in epithelial ovarian carcinoma: correlation with tumor cell proliferation, angiogenesis, survival and ability to predict optimal cytoreduction. Gynecol Oncol. 2011;121:181–6.
19. Amann T, Maegdefrau U, Hartmann A, Agaimy A, Marienhagen J, Weiss TS, et al. GLUT1 expression is increased in hepatocellular carcinoma and promotes tumorigenesis. Am J Pathol. 2009;174:1544–52.
20. Li Y, Shan F, Chen J. Lipid raft-mediated miR-3908 inhibition of migration of breast cancer cell line MCF-7 by regulating the interactions between AdipoR1 and Flotillin-1. World J Surg Oncol. 2017;15:69.
21. Du Y, Miao Z, Wang K, Lv Y, Qiu L, Guo L. Expression levels and clinical values of miR-92b-3p in breast cancer. World J Surg Oncol. 2021;19:239.
22. Wang ZH, Zhang YZ, Wang YS, Ma XX. Identification of novel cell glycolysis related gene signature predicting survival in patients with endometrial cancer. Cancer Cell Int. 2019;19:296.
23. Chen C, Shi Y, Li Y, He ZC, Zhou K, Zhang XN, et al. A glycolysis-based ten-gene signature correlates with the clinical outcome, molecular subtype and IDH1 mutation in glioblastoma. J Genet Genomics. 2017;44:519–30.
24. Zhang D, Zheng Y, Yang S, Li Y, Wang M, Yao J, et al. Identification of a novel glycolysis-related gene signature for predicting breast cancer survival. Front Oncol. 2020;10:596087.
25. Tian S, Han G, Lu L, Meng X. Circ-FOXM1 contributes to cell proliferation, invasion, and glycolysis and represses apoptosis in melanoma by regulating miR-143-3p/FLOT2 axis. World J Surg Oncol. 2020;18:56.
26. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA. 2011;305:1873–81.
27. Zhao L, Jiang L, He L, Wei Q, Bi J, Wang Y, et al. Identification of a novel cell cycle-related gene signature predicting survival in patients with gastric cancer. J Cell Physiol. 2019;234:6350–60.
28. Xu Q, Zhang Q, Dong M, Yu Y. MicroRNA-638 inhibits the progression of breast cancer through targeting HOXA9 and suppressing Wnt/beta-cadherin pathway. World J Surg Oncol. 2021;19:247.
29. Li Y, Lu S, Zhang Y, Wang S, Liu H. Loco-regional recurrence trend and prognosis in young women with breast cancer according to molecular subtypes: analysis of 1099 cases. World J Surg Oncol. 2021;19:113.
30. Yu YZ, Mu Q, Ren Q, Xie LJ, Wang QT, Wang CP. miR-381-3p suppresses breast cancer progression by inhibition of epithelial-mesenchymal transition. World J Surg Oncol. 2021;19:230.
31. Xu H, Xiong C, Chen Y, Zhang C, Bai D. Identification of Rad51 as a prognostic biomarker correlated with immune infiltration in hepatocellular carcinoma. Bioengineered. 2021;12:2664–75.
32. Li C, Long Q, Zhang D, Li J, Zhang X. Identification of a four-gene panel predicting overall survival for lung adenocarcinoma. BMC Cancer. 2020;20:1198.

He *et al. World J Surg Onc*   (2021) 19:294

Page 14 of 14

33. Li A, Hou S, Chen J, Jiang Y. Development and validation of a novel glycolysis-related risk signature for predicting survival in pancreatic adenocarcinoma. Clin Chim Acta. 2021;518:156–61.
34. Peppicelli S, Bianchini F, Calorini L. Extracellular acidity, a "reappreciated" trait of tumor environment driving malignancy: perspectives in diagnosis and therapy. Cancer Metastasis Rev. 2014;33:823–32.
35. 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.
36. Liu Y, Yin S. A novel prognostic index based on the analysis of glycolysis-related genes in head and neck squamous cell carcinomas. J Oncol. 2020;2020:7353874.
37. Wu Z, Wen Z, Li Z, Yu M, Ye G. Identification and prognostic value of a glycolysis-related gene signature in patients with bladder cancer. Medicine (Baltimore). 2021;100:e23836.
38. Qin WY, Feng SC, Sun YQ, Jiang GQ. MiR-96-5p promotes breast cancer migration by activating MEK/ERK signaling. J Gene Med. 2020;22:e3188.
39. Pera E, Kaemmerer E, Milevskiy MJG, Yapa K, O'Donnell JS, Brown MA, et al. The voltage gated Ca(2+)-channel Cav3.2 and therapeutic responses in breast cancer. Cancer Cell Int. 2016;16:24.
40. Suzuki A, Leland P, Joshi BH, Puri RK. Targeting of IL-4 and IL-13 receptors for cancer therapy. Cytokine. 2015;75:79–88.
41. Park MH, Kwon HJ, Kim JR, Lee B, Lee SJ, Bae YK. Elevated interleukin-13 receptor alpha 1 expression in tumor cells is associated with poor prognosis in patients with invasive breast cancer. Ann Surg Oncol. 2017;24:3780–7.
42. Tian C, Zhou S, Yi C. High NUP43 expression might independently predict poor overall survival in luminal A and in HER2+ breast cancer. Future Oncol. 2018;14:1431–42.
43. Liu YY, Jiao WY, Li T, Bao YY. MiRNA-409-5p dysregulation promotes imatinib resistance and disease progression in children with chronic myeloid leukemia. Eur Rev Med Pharmacol Sci. 2019;23:8468–75.
44. Zhang Y, Chen M, Liu M, Xu Y, Wu G. Glycolysis-related genes serve as potential prognostic biomarkers in clear cell renal cell carcinoma. Oxid Med Cell Longev. 2021;2021:6699808.
45. Sun S, Liang X, Zhang X, Liu T, Shi Q, Song Y, et al. Phosphoglycerate kinase-1 is a predictor of poor survival and a novel prognostic biomarker of chemoresistance to paclitaxel treatment in breast cancer. Br J Cancer. 2015;112:1332–9.
46. Olah C, Tschirdewahn S, Hoffmann MJ, Krafft U, Hadaschik B, Nyirady P, et al. Soluble syndecan-1 levels are associated with survival in platinum-treated bladder cancer patients. Diagnostics (Basel). 2020;10:864.
47. Wang F, Tang C, Gao X, Xu J. Identification of a six-gene signature associated with tumor mutation burden for predicting prognosis in patients with invasive breast carcinoma. Ann Transl Med. 2020;8:453.
48. Zhang D, Yang S, Li Y, Yao J, Ruan J, Zheng Y, et al. Prediction of overall survival among female patients with breast cancer using a prognostic signature based on 8 DNA repair-related genes. JAMA Netw Open. 2020;3:e2014622.
49. Liu L, Chen Z, Shi W, Liu H, Pang W. Breast cancer survival prediction using seven prognostic biomarker genes. Oncol Lett. 2019;18:2907–16.