Identification and validation of glycolysis-related signatures for prognosis in patients with breast cancer

Ming Zhang
Nanjing Medical University
Mengying Xing
Nanjing Medical University
Yunfei Ma
Nanjing Medical University
Bing Yao
Nanjing Medical University
Xiang Chen
The Affiliated Yixing Hospital of Jiangsu University
Changyan Ma (cyma@njmu.edu.cn)
Nanjing Medical University https://orcid.org/0000-0002-1609-4609

Research

Keywords: Glycolysis, Prognosis, Breast cancer, Risk-score, TCGA

Posted Date: June 22nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-35764/v1

Identification and validation of glycolysis-related signatures for prognosis in patients with breast cancer

Ming Zhang[1], Mengying Xing[1], Yunfei Ma[1], Bing Yao[1], Xiang Chen[2]*, Changyan Ma[1]*

[1] Department of Medical Genetics, Nanjing Medical University, Nanjing 211166, P.R. China

[2] Department of General Surgery, the Affiliated Yixing Hospital of Jiangsu University, Yixing 214200, P.R. China

*Correspondence: Xiang Chen, Department of General Surgery, the Affiliated Yixing Hospital of Jiangsu University, Yixing 214200, P.R. China, E-mail: staff984@yxph.com; Changyan Ma, Department of Medical Genetics, Nanjing Medical University, Longmian Road 101, Nanjing 211166, P.R. China, E-mail: cyma@njmu.edu.cn

Abstract

Background: Emerging evidence has demonstrated roles of glycolysis in the tumorigenesis and progression of human tumors. However, their underlying clinical implications have not been well elucidated in breast cancer. In present study, we aimed to generate a risk-score from glycolysis-related signatures to predict prognosis of patients with breast cancer.

Methods: We acquired mRNAs expression and clinical datasets in patients with breast cancer from The Cancer Genome Atlas (TCGA), then identifying glycolysis-related mRNAs by Gene Set Enrichment Analysis (GSEA), followed by construction of prognostic risk-score. The altered expression of glycolysis-related mRNAs was identified as candidates for further investigation. We constructed a risk-score from the prognostic glycolysis-related mRNAs by Cox regression. Receiver Operating Characteristic (ROC) and clinical subgroups analysis were performed to evaluate the values of risk-score to predict prognosis of breast cancer. Besides, we also compared the expression patterns of the signatures in breast cancer tissues and cell lines.

Results: Total of 1208 cases were obtained, including 112 normal tissues and 1096 tumor tissues. We found 4 glycolysis-related pathways significantly involved in breast cancer. And 298 mRNAs involved in the 4 pathways were defined as glycolysis-related mRNAs; of these, 241 dysregulated mRNAs were candidates for further exploration. Then we constructed a risk-score from the 5 candidates (IL13RA1, PGK1, SDC3, NUP43 and SDC1). The area under the curve (AUC) for the risk-score to predict prognosis was 0.729. Patients with high-risk score had poor prognosis among overall or clinical subgroups (P<0.05). And IL13RA1, PGK1, NUP43 and SDC1 were up-regulated in tumor tissues and cell lines (MDA-MB-231 and BT474) as compared to normal tissues and cell line (MCF-10A), while SDC3 was down-regulated.

Conclusions: We construct a risk-score based on 5 glycolysis-related signatures, which

can well predict prognosis in breast cancer. Additionally, our findings further unveil the molecular mechanisms of glycolysis in cancer, providing promising directions for the prognostic and therapeutic biomarkers for breast cancer.

Keywords: Glycolysis, Prognosis, Breast cancer, Risk-score, TCGA

Background

Breast cancer remains the leading lethal and most frequent diagnosed cancer among females worldwide. According to the global cancer statistics, approximately 2.1 million new cases are diagnosed with breast cancer, accounting for approximately 1 in 4 cancer-cases in females[1]. Breast cancer is defined as a heterogeneous disease, which is generally classified as Normal-like, Luminal A, Luminal B, HER2-enriched and Basal-like subtypes[2, 3]. Currently, multiple risk factors have been identified, including age at menarche, reproduction, intaking of exogenous hormone, nutrition according to cancer statistic 2018. Although great efforts dedicated to the early diagnosis and therapy, including surgery, targeted therapy, chemotherapy and radiotherapy, the 5-year survival rate of breast cancer remains unsatisfactory[4]. Statistically, the 5-year survival rate of patients with metastasis is less than 30%, as compared to that of 90% in overall breast cancer cases[5]. Therefore, it is of great significance to investigate the underlying molecular mechanisms of breast cancer and exploit promising prognostic and therapeutic biomarkers.

Altered energy metabolism is considered as one of the hallmarks of cancer, constituting the complexities of malignant tumor disease[6]. Of which, it is well accepted that tumor cells prefer to metabolize glucose by glycolysis, which is known as the Warburg effect (or aerobic glycolysis)[7, 8]. Briefly, tumor cells primarily acquire energy by processing the glucose into pyruvate even with sufficient oxygen, which produces large amounts of lactic acid and a small quantity of adenosine triphosphate (ATP). Emerging evidence has demonstrated that the Warburg effect exists in various types of tumors, favoring the proliferation, invasion and migration of tumor cells, or even inhibiting the efficiency of therapy[9, 10]. Although glycolysis produced ATP at low efficiency, the rate is faster than oxidative phosphorylation, which may contribute to the proliferation of tumor cells. Mechanically, the high glycolysis rate may produce large amounts of lipids, nucleotides and amino acids to aid the growth of tumor cells[9]. In addition, the metabolic molecules cause a consistent acidification of microenvironment to promote invasion[11].

Understanding the underlying mechanisms will allow us to generate novel diagnostic, prognostic and therapeutic biomarkers via modulating glycolytic pathways in human cancers. Currently, experimental evidence has demonstrated that glycolytic competence negatively impacts prognosis of patients treated with salvage paclitaxel-ramucirumab in gastric adenocarcinomas[12]. Korga et al. observed that inhibition of glycolysis may disrupt cellular antioxidant defense and sensitize human hepatocellular carcinoma cells (HepG2) to doxorubicin treatment[13]. However, a single gene signature cannot well meet the demands of diagnosis or therapeutic demands for human cancers. Herein, it is intriguing to exploit the comprehensive and efficient glycolysis-related gene signatures.

In the present study, we performed a comprehensive analysis of glycolysis-related mRNAs in breast cancer using the public available The Cancer Genome Atlas (TCGA) database[14]. We acquired mRNA expression dataset from TCGA database and identified the glycolysis-related mRNA involved in breast cancer using Gene Set Enrichment Analysis (GSEA). Of these, the dysregulated mRNAs were candidate for constructing the systematic risk signatures to predict the prognosis and diagnosis in patients with breast cancer. Interestingly, we found that the expression of these signatures was altered in breast tumor tissues and cell lines, and the glycolysis-related risk score can well predict patients with poor prognosis cases in breast cancer.

Materials and methods

Data collection from TCGA

We firstly consulted the public available TCGA database (https://portal.gdc.cancer.gov/database) to download the mRNA expression in breast cancer except for the expression data of males, and then sorted the expression files into matrix using Perl. The data downloaded were normalized with FPKM. Meanwhile, the level-3 clinical data were extracted (age, overall survival time, survival status T, N, M and clinical stage). We acquired the data from public available database, so informed consent was not required.

Cell lines and culture

Human normal breast epithelial cell line (MCF10A) and breast cancer cell lines (MDA-MB-231 and BT474) were obtained from the Chinese Academy of Sciences (Shanghai, China). MCF-10A cell lines were cultured in DMEM/F12 media (5% HS, 20 ng/mL EGF, 0.5 μg/mL Hydrocortisone, 10 μg/mL Insulin, 1% NEAA, 1% P/S). The cancer cell lines were cultured in DMEM media (10% FBS, 1% P/S) at 37 °C with 5% CO2.

RNA extraction and qRT-PCR

Total RNA was extracted from cells with TRIzol reagent (Takara, Dalian, China) followed by the manufacturer's protocols. The concentration of total RNA was measured by Nanodrop 2000 (Thermo Fisher Scientific, USA) and then reversed into cDNA using PrimeScript™ RT reagent Kit with gDNA Eraser (Takara, Dalian, China). qRT-PCR was performed with SYBR Premix Ex Taq™ (Yeasen, Shanghai, China). The relative expression of mRNAs was normalized to the internal control ACTB and calculated by $2^{-\Delta\Delta CT}$. The primers sequences involved in this study were showed in Table 1.

Gene Set Enrichment Analysis

To further identify the glycolysis-related mRNA involved in breast cancer, we performed GSEA (http://www.broadinstitute.org/gsea/index.jsp) based on the expression datasets in normal tissues and cancer tissues. Briefly, we divided cases into normal and cancer subgroups to perform GSEA. Total of 5 pathways, including GO_GLYCOLYTIC_PROCESS, KEGG_GLYCOLYSIS_GLUCONEOGENESIS, HALLMARK_GLYCOLYSIS, REACTOME_GLYCOLYSIS and BIOCARTA_GLYCOLYSIS_PATHWAY, were defined as glycolysis-related pathways. Afterwards, the mRNAs enriched in the significant glycolysis-related pathways were candidates for further analysis.

Statistical analysis

We applied the univariate Cox regression to identify the prognostic glycolysis-related mRNAs and then used multivariate Cox regression to select the candidates for further construction of a risk-score. After that, we defined the patients with risk-score more than median value as high-risk group, the rest as low-risk group. Next, we used Kaplan–Meier curves and log-rank test to validate the prognostic value of the risk-score and utilized the ROC curve to explore the risk-score to predict the prognosis of patients with breast cancer. The association between risk-score and the expression of candidate mRNA was displayed by heatmap. Besides, both univariate Cox regression and multivariable Cox regression analysis were performed to validate the risk-score to predict the prognosis of patients with breast cancer using predict function. we validated the risk-score model among different subtypes based on clinical characteristics. Additionally, we compared the expression levels of glycolysis-related genes in breast tissues and cell lines using Wilcox rank test. All statistical analysis was conducted using R software 3.6.1 and $P < 0.05$ was defined as statistically significant.

Results

Identification of altered glycolysis-related genes in breast cancer

To obtain the expression dataset and clinical characteristics of patients with breast cancer, we quired for TCGA database. In brief, total of 1208 cases were obtained, including 112 normal tissues and 1096 tumor tissues, and 1085 cases of clinical characteristics were acquired. To identify the glycolysis-related mRNA involved in breast cancer, we performed GSEA. The results indicated that GO_GLYCOLYTIC_PROCESS (NES=1.65, *P*=0.028), HALLMARK_GLYCOLYSIS (NES=2.07, *P*=0.002), BIOCARTA_GLYCOLYSIS_PATHWAY (NES=1.50, *P*=0.016) and REACTOME_GLYCOLYSIS (NES=2.05, *P*<0.001) were significantly enriched in tumor subgroup (Figure 1). And the 298 mRNAs involved in the four glycolysis-related pathways were defined as glycolysis-related mRNAs; of which, total of 241 mRNAs were dysregulated in patients with breast cancer and selected as candidates for further exploration.

Construction of prognostic glycolysis-related risk-score in breast cancer

To explore the values of glycolysis-related mRNAs in breast cancer, we dedicated to construct the prognosis model based on glycolysis-related signatures. Briefly, nine mRNAs were related to the prognosis in breast cancer by univariate Cox regression (IL13RA1, P4HA2, PGK1, SDC3, NUP43, SDC1, PGAM1 and RARS, $P<0.05$). These eight mRNAs were further selected to construct prognostic risk model. Afterwards, we applied the multivariable Cox regression to construct prognostic model, and then the highly correlated genes were removed using step function for optimization. Ultimately, five signatures (IL13RA1, PGK1, SDC3, NUP43 and SDC1) were recruited to construct risk-score. The coefficients for these five signatures in the model were: 0.007 for IL13RA1, 0.006 for PGK1, -0.024 for SDC3, 0.043 for NUP43 and 0.002 for SDC1, respectively. We finally generated a risk-score based this model for each patient. Meanwhile, patients with their risk-score more than median were classified into high-risk group, the rest into low-risk group (Figure 2A). As shown in Figure 2B, number of deaths grew as the risk-score increased. As compared to low-risk group, we also found that the expression of SDC3 were down-regulated in high-risk group, but NUP43, SDC1, IL13RA1 and PGK1 were up-expressed (Figure 2C). Besides, we found that expression of SDC3 were lower in tumor tissues and cell lines when compared to normal tissues and normal cell line, but that of NUP43, SDC1, IL13RA1 and PGK1were upregulated, which were consistent with expression patterns of patients with high and low risk (Figure 3). Furthermore, the area under the curve (AUC) was 0.729, indicating that our risk-score could well predicting the prognosis of patients with breast cancer (Figure 4A). Patients with high risk-score had poor prognosis as compared to that with low-risk score (Figure 4B).

Validation of risk-score from glycolysis-related genes based on clinical indicators

To further investigate the risk-score derived from five glycolysis-related signatures for predicting prognosis in breast cancer, we performed subgroup survival analysis. We

firstly found that age, T, N, M and clinical stage, were associated with prognosis and patients with high-risk score had poor prognosis (P<0.05, Figure 5). To well characterize the value of risk-score to predict prognosis in breast cancer, we performed subgroup analysis on basis of these clinical features. We found that cases of high-risk score in age subgroups, N subgroup, T subgroup, with-metastasis subgroup and clinical stage subgroup all had poor prognosis (P <0.05). However, the survival rate of patients with high-risk and low-risk-score were comparable in patients without-metastasis (P=0.660) (Figure 6). In addition, the univariate Cox regression suggested that higher age, clinical stage, T stage, N stage, M stage and risk-score were predicted with poor prognosis (P<0.05, Figure 7A). Also, the risk-score based glycolysis-related mRNAs was independent risk factors for prognosis in breast cancer according to multivariable Cox regression analysis indicated. (HR=1.395, 95%CI: 1.235-1.577, Figure 7B).

Comparison of the risk-score in different clinical subgroups

In this part of the analysis, we extracted the datasets of patients with complete clinical characteristics. Total of 897 patients were included, then we divided these cases of risk-score higher than median into high-risk group and that lower than median into low-risk group. As shown in Table 2, among patients older than 65 years, the number of high-risk cases was more than the low-risk. While, the risk-score distribution in T, N, M and clinical stage were consistent.

Discussion

Although the diagnosis and treatment of breast cancer have been widely investigated in recent years, its prognosis is still unsatisfactory. Statistically, approximately 30 to 40% of patients with early-stage breast cancer will undergo recurrence and metastasis after operation and progress to advanced cancer[15]. Hence, it is worthy to establish prognostic model for postoperative monitoring and treatment. To date, several prognostic factors have been identified, including age, clinical stage, and the type of breast cancer[16]. Advances in high-throughput sequencing techniques have accelerated

the understanding of the molecular roles in human cancers, suggesting that molecular indicators can be used as other indicators for diagnosis, prognosis and treatment. In this study, we constructed a prognostic risk-score based on glycolysis-related mRNAs in breast cancer by comprehensive bioinformatics analysis. Information of mRNA expression in breast cancer were acquired from TCGA database, then the glycolysis-related mRNAs were identified by GSEA. We constructed a risk-score derived from glycolysis-related mRNAs, which can well predict the prognosis of breast cancer. Additionally, these molecular indicators that were altered expressed in breast tumors and cell lines, are also expected to be targets for the treatment of breast cancer via inhibiting the glycolytic-related pathway.

It is well accepted that tumor cells rewire their metabolism to satisfy metabolic demands and adapt to environmental changes, especially for glucose metabolism[17, 18]. Accumulating evidence has demonstrated that dysregulated glucose metabolism is associated with carcinogenesis, progression and treatment of various cancers[19]. Firstly, tumor cells preferentially utilize the glycolysis for energy even with abundant oxygen[20]. Meanwhile, experimental evidence has suggested that the absorption of glucose in tumor cells is dramatically increased when compared to that of normal cells [21]. Currently, differences in glucose uptake rates between normal and tumor cells have been translated into clinical applications, of which Positron Emission Tomography involving 2-deoxy-2(18F)-fluoro-glucose glucose has been routinely performed to identify and classify tumors[22-25]. On basis of these observations, efforts to explore the glycolysis have facilitated the treatment of various tumors. Intriguing, some glycolytic inhibitors have been identified as direct anti-cancer activity at the bench or reached clinical trials. For instance, some antimetabolites, including fluorouracil, cytarabine (Ara-C) and methotrexate have been routinely used as chemotherapeutic agents for several years[26-29]. However, these direct anti-cancer agents lacks of specificity. Some indirect inhibitions of tumor metabolism have been identified, targeting at upstream regulators of metabolic pathways. For example, the hyper-activated

PTEN/PI3K/AKT/mTOR pathway, acting as central regulator of aerobic glycolysis, contributes to cancer metabolic switch and proliferation in bladder cancer[30]. Hu et al. found that AMPK inhibitor partly attenuates the malignant phenotype of pancreatic cancer cells by suppressing aerobic glycolysis[31]. In this study, we successfully identified four glycolysis-related pathways involved in breast cancer, of which expression of 241 mRNAs were altered including IL13RA1, P4HA2, PGK1, SDC3, NUP43, SDC1, PGAM1 and RARS, etc. Among which, PGK1 was also found to be upregulated in human colon cancers, whose O-GlcNAcylation coordinates glycolysis and TCA cycle to promote tumor growth[32]. Consistently, Sayyad et al. suggest that SDC1 promotes the migration of breast cancer across the blood-brain barrier through regulation of cytokines[33].

Many studies have revealed the clinical and molecular indicators for predicting the prognosis in human cancers. For example, the extent of the cancer at diagnosis is a key factor used to define treatment and to assess the chance of successful treatment outcome[34]. Of which, clinical stage codifies the extent of cancer, based on primary tumor (T), regional lymph nodes (N), and distant metastasis (M), providing the means to quantify prognosis for individual patients. Generally, it has been recognized that cancers with local, regional, and metastatic cancers have a worsening prognosis. In breast cancer, the molecular subtypes significantly affect the occurrence and prognosis of patients with brain metastasis[35]. However, these clinical factors cannot fully predict the prognosis of cancers. Increasing studies reveal that some molecules were independent risk factors for tumor prognosis. Sengal et al. reported that altered expression of FGFR2c appears as an independent prognostic indicator in endometroid endometrial cancers. They also found that FGFR2c can predict the prognosis of patients within grade 3 tumors, ESMO high-risk groups, as well as within the MMRd and p53wt subtypes, respectively[36]. Notably, the prediction of tumor prognosis by single mRNA is limited, so it is of great significance to construct comprehensive predict model. For example, the risk model based on systematic immune-related genes shows a superior

prognostic values in non-small cell lung cancer[37]. In our work, we constructed a risk-score model according to glycolysis-related mRNAs in breast cancer. Of which, the included signatures were screened by cox regression, including IL13RA1, PGK1, SDC3, NUP43 and SDC1. And multivariate Cox regression showed that the risk-score is an independent risk factor for prognosis in breast cancer. Meanwhile, our results also pointed out that age at the time of diagnosis is an independent risk factor for prognosis in breast cancer.

Conclusions

In all, we identified 241 glycolysis-related mRNAs with altered expression in breast cancer, and constructed risk-score based on 5 prognostic mRNAs to predict the outcomes of breast cancer patients. Additionally, our findings provided a novel insight to the target therapy of breast cancer.

Acknowledgements

Not applicable.

Abbreviations

IL13RA1	Interleukin-13 receptor subunit alpha-1
PGK1	Phosphoglycerate kinase 1
SDC3	Syndecan-3
NUP43	Nucleoporin 43
SDC1	Syndecan-1
PTEN	Phosphatase and tensin homolog
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
AKT	AKT serine/threonine kinase 1
mTOR	Mechanistic target of rapamycin kinase
AMPK	Catalytic subunit of AMP-activated protein kinase
P4HA2	Prolyl 4-hydroxylase subunit alpha 2

PGAM1	Phosphoglycerate mutase 1
RARS	Arginyl-tRNA synthetase
FGFR2c	Fibroblast growth factor receptor 2
GSEA	Gene Set Enrichment Analysis

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors consent to submit for publish.

Availability of data and material

The level-3 HTseq-FPKM of datasets of mRNA and corresponding clinical characteristics derived from 1096 breast cancer and 112 normal tissues were extracted from the TCGA database (https://portal.gdc.cancer.gov/). The data downloaded were normalized with FPKM.

Competing interests

The authors declare no competing interests.

Authors' contributions

All authors contributed to data analysis, drafting or revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding

This work was supported by grants from the National Natural Science Foundation of China (81872389), projects of Reinvigorating Medicine through Science and Education

from Jiangsu Health and Family Planning Commission, Women and Child Health Research project from Wuxi Health Commission (FYKY 201907).

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018, 68(6):394-424.

[2] Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. Nature, 2000, 406(6797):747-752.

[3] Comprehensive molecular portraits of human breast tumours. Nature, 2012, 490(7418):61-70.

[4] Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL, Siegel RL. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin, 2019, 69(5):363-385.

[5] Peart O. Metastatic Breast Cancer. Radiol Technol, 2017, 88(5):519m-539m.

[6] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell, 2011, 144(5):646-674.

[7] Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. Cell, 2008, 134(5):703-707.

[8] Warburg O. On respiratory impairment in cancer cells. Science, 1956, 124(3215):269-270.

[9] Lunt SY, Vander Heiden MG. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. Annu Rev Cell Dev Biol, 2011, 27:441-464.

[10] Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science, 2009, 324(5930):1029-1033.

[11] Smallbone K, Gavaghan DJ, Gatenby RA, Maini PK. The role of acidity in solid tumour growth and invasion. J Theor Biol, 2005, 235(4):476-484.

[12] Ruzzo A, Graziano F, Bagaloni I, Di Bartolomeo M, Prisciandaro M, Aprile G, Ongaro E, Vincenzi B, Perrone G, Santini D, Fornaro L, Vivaldi C, Tomasello G, Loupakis F, Lonardi S, Fassan M, Valmasoni M, Sarti D, Lorenzini P, Catalano V, Bisonni R, Del Prete M, Collina G, Magnani M. Glycolytic competence in gastric adenocarcinomas negatively impacts survival outcomes of patients treated with salvage paclitaxel-ramucirumab. Gastric Cancer, 2020.

[13] Korga A, Ostrowska M, Iwan M, Herbet M, Dudka J. Inhibition of glycolysis disrupts cellular antioxidant defense and sensitizes HepG2 cells to doxorubicin treatment. FEBS Open Bio, 2019, 9(5):959-972.

[14] Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn), 2015, 19(1a):A68-77.

[15] Gonzalez-Angulo AM, Morales-Vasquez F, Hortobagyi GN. Overview of resistance to systemic therapy in patients with breast cancer. Adv Exp Med Biol, 2007, 608:1-22.

[16] Waks AG, Winer EP. Breast Cancer Treatment. Jama, 2019, 321(3):316.

[17] Liberti MV, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? Trends Biochem Sci, 2016, 41(3):211-218.

[18] Tennant DA, Duran RV, Gottlieb E. Targeting metabolic transformation for cancer therapy. Nat Rev Cancer, 2010, 10(4):267-277.

[19] Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer,

2004, 4(11):891-899.
[20] Weinhouse S. The Warburg hypothesis fifty years later. Z Krebsforsch Klin Onkol Cancer Res Clin Oncol, 1976, 87(2):115-126.
[21] Warburg O, Wind F, Negelein E. THE METABOLISM OF TUMORS IN THE BODY. J Gen Physiol, 1927, 8(6):519-530.
[22] Engelman JA, Chen L, Tan X, Crosby K, Guimaraes AR, Upadhyay R, Maira M, McNamara K, Perera SA, Song Y, Chirieac LR, Kaur R, Lightbown A, Simendinger J, Li T, Padera RF, Garcia-Echeverria C, Weissleder R, Mahmood U, Cantley LC, Wong KK. Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. Nat Med, 2008, 14(12):1351-1356.
[23] Goel S, England CG, Chen F, Cai W. Positron emission tomography and nanotechnology: A dynamic duo for cancer theranostics. Adv Drug Deliv Rev, 2017, 113:157-176.
[24] Li P, Liu Q, Wang C, Wang T, Liu J, Huang G, Song S. Fluorine-18-fluorodeoxyglucose positron emission tomography to evaluate recurrent gastric cancer after surgical resection: a systematic review and meta-analysis. Ann Nucl Med, 2016, 30(3):179-187.
[25] Lee SM, Kim HS, Lee S, Lee JW. Emerging role of (18)F-fluorodeoxyglucose positron emission tomography for guiding management of hepatocellular carcinoma. World J Gastroenterol, 2019, 25(11):1289-1306.
[26] Cameron DA, Gabra H, Leonard RC. Continuous 5-fluorouracil in the treatment of breast cancer. Br J Cancer, 1994, 70(1):120-124.
[27] Metterle L, Nelson C, Patel N. Intralesional 5-fluorouracil (FU) as a treatment for nonmelanoma skin cancer (NMSC): A review. J Am Acad Dermatol, 2016, 74(3):552-557.
[28] Chhikara BS, Parang K. Development of cytarabine prodrugs and delivery systems for leukemia treatment. Expert Opin Drug Deliv, 2010, 7(12):1399-1414.
[29] Chabner BA, Roberts TG, Jr. Timeline: Chemotherapy and the war on cancer. Nat Rev Cancer, 2005, 5(1):65-72.
[30] Massari F, Ciccarese C, Santoni M, Iacovelli R, Mazzucchelli R, Piva F, Scarpelli M, Berardi R, Tortora G, Lopez-Beltran A, Cheng L, Montironi R. Metabolic phenotype of bladder cancer. Cancer Treat Rev, 2016, 45:46-57.
[31] Hu M, Chen X, Ma L, Ma Y, Li Y, Song H, Xu J, Zhou L, Li X, Jiang Y, Kong B, Huang P. AMPK Inhibition Suppresses the Malignant Phenotype of Pancreatic Cancer Cells in Part by Attenuating Aerobic Glycolysis. J Cancer, 2019, 10(8):1870-1878.
[32] Nie H, Ju H, Fan J, Shi X, Cheng Y, Cang X, Zheng Z, Duan X, Yi W. O-GlcNAcylation of PGK1 coordinates glycolysis and TCA cycle to promote tumor growth. Nat Commun, 2020, 11(1):36.
[33] Sayyad MR, Puchalapalli M, Vergara NG, Wangensteen SM, Moore M, Mu L, Edwards C, Anderson A, Kall S, Sullivan M, Dozmorov M, Singh J, Idowu MO, Koblinski JE. Syndecan-1 facilitates breast cancer metastasis to the brain. Breast Cancer Res Treat, 2019, 178(1):35-49.
[34] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol, 2010, 17(6):1471-1474.
[35] Xiao W, Li X, Yang A, Chen B, Zheng S, Zhang G, Deng W, Liao N. Analysis of Prognostic Factors Affecting the Brain Metastases Free Survival and Survival After Brain Metastases

in Breast Cancer. Front Oncol, 2020, 10:431.

[36] Sengal AT, Patch AM, Snell CE, Smith DS, Leung S, Talhouk A, Williams ED, McAlpine JN, Pollock PM. FGFR2c mesenchymal isoform expression is associated with poor prognosis and further refines risk stratification within endometrial cancer molecular subtypes. Clin Cancer Res, 2020.

[37] Sun L, Zhang Z, Yao Y, Li WY, Gu J. Analysis of expression differences of immune genes in non-small cell lung cancer based on TCGA and ImmPort data sets and the application of a prognostic model. Ann Transl Med, 2020, 8(8):550.

Figure legends

Figure 1 The enrichments of glycolysis-related pathways significantly differ between normal and neoplastic tissues in breast cancer. (P<0.05)

Figure 2 A risk-score based on five glycolysis-related signatures to predicts overall survival in patients with breast cancer. A) The risk-score distribution. B) Overall survival distribution of patients with different risk-score. C) Visualization expression pattern of the five glycolysis-related signatures in high-risk and low-risk patients.

Figure 3 Expression patterns of the five glycolysis-related signatures in breast tissues and cell lines. A) Expression patterns in breast tissues. B) Expression patterns in breast cell lines by qRT-PCR. (*P<0.05, ** P<0.01, *** P<0.001)

Figure 4 Values of a risk-score to predict the prognosis of patients with breast cancer. A) The Receiver Operating Characteristic Curve for the risk-score to predict prognosis of patients with breast cancer. B) Kaplan–Meier curves for prognostic value of the risk-score. (P<0.05)

Figure 5 Kaplan–Meier survival analysis for patients with high and low risk-score in breast cancer. (stage: clinical stage; T: primary tumor; N: lymph nodes and M: distant metastasis, P<0.05)

Figure 6 Kaplan–Meier survival analysis for patients with high and low risk-score among each clinical subgroup in breast cancer. (stage: clinical stage; T: primary tumor and N: lymph nodes, P<0.05)

Figure 7 Validation of prognostic glycolysis-related risk score by Cox regression. A) The univariate Cox regression analysis. B) The multivariable Cox regression analysis. (P<0.05)

487 **Table 1 The primer sequence for qRT-PCR involved in this study.**

mRNAs	Forward primers (5'→3')	Reverse primers (5'→3')
IL13RA1	TGAGTGTCTCTGTTGAAAACCTC	GGGGTACTTCTATTGAACGACGA
PGK1	GACCTAATGTCCAAAGCTGAGAAG	CAGCAGGTATGCCAGAAGCC
NUP43	TGGAGGGTTTGAAGGAGACCA	TGAAGCAGCGACAATTCTTTCC
SDC1	CCACCATGAGACCTCAACCC	GCCACTACAGCCGTATTCTCC
SDC3	TGGCGCAGTGAGAACTTCG	GAAGCGCATGGCTGTCTCA
β-actin	AGATGTGATCAGCAAGCAG	GCGCAAGTTAGGTTTTGTCA

488

Table 2 The distribution of risk-score among clinical characteristics.

Subgroup	Low-risk	High-risk	*P*-value
Age			
≤65	349	309	0.003
>65	100	139	
T stage			
T1-2	381	383	0.789
T3-4	68	65	
N stage			
N0	236	210	0.966
N1-3	238	213	
M stage			
M0	441	440	0.996
M1	8	8	
Clinical stage			
I-II	355	329	0.055
III-IV	95	119	

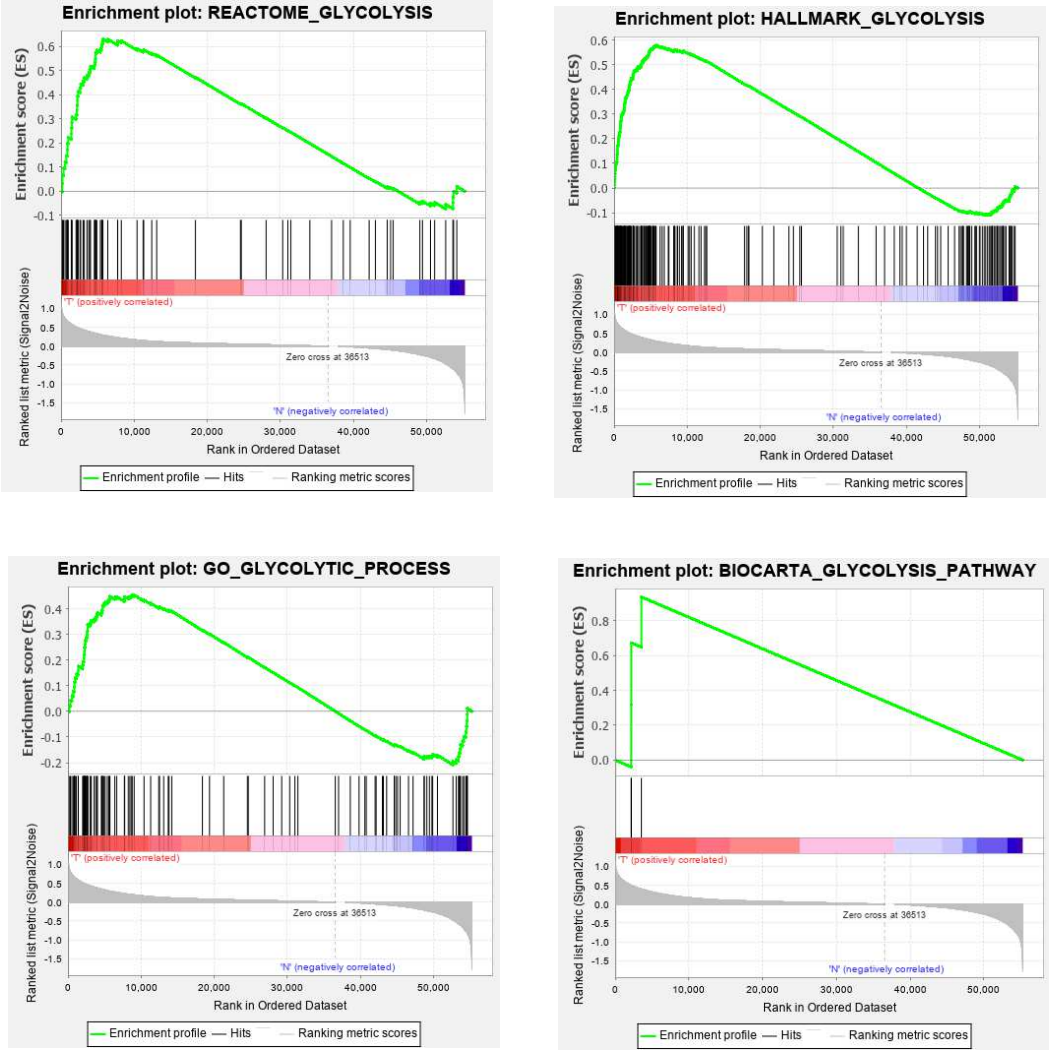

Enrichment plot: REACTOME_GLYCOLYSIS
Enrichment score (ES)
Ranked list metric (Signal2Noise)
'T' (positively correlated)
'N' (negatively correlated)
Zero cross at 36513
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
Enrichment plot: HALLMARK_GLYCOLYSIS
Enrichment score (ES)
Ranked list metric (Signal2Noise)
'T' (positively correlated)
'N' (negatively correlated)
Zero cross at 36513
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
Enrichment plot: GO_GLYCOLYTIC_PROCESS
Enrichment score (ES)
Ranked list metric (Signal2Noise)
'T' (positively correlated)
'N' (negatively correlated)
Zero cross at 36513
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
Enrichment plot: BIOCARTA_GLYCOLYSIS_PATHWAY
Enrichment score (ES)
Ranked list metric (Signal2Noise)
'T' (positively correlated)
'N' (negatively correlated)
Zero cross at 36513
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

Figure 1

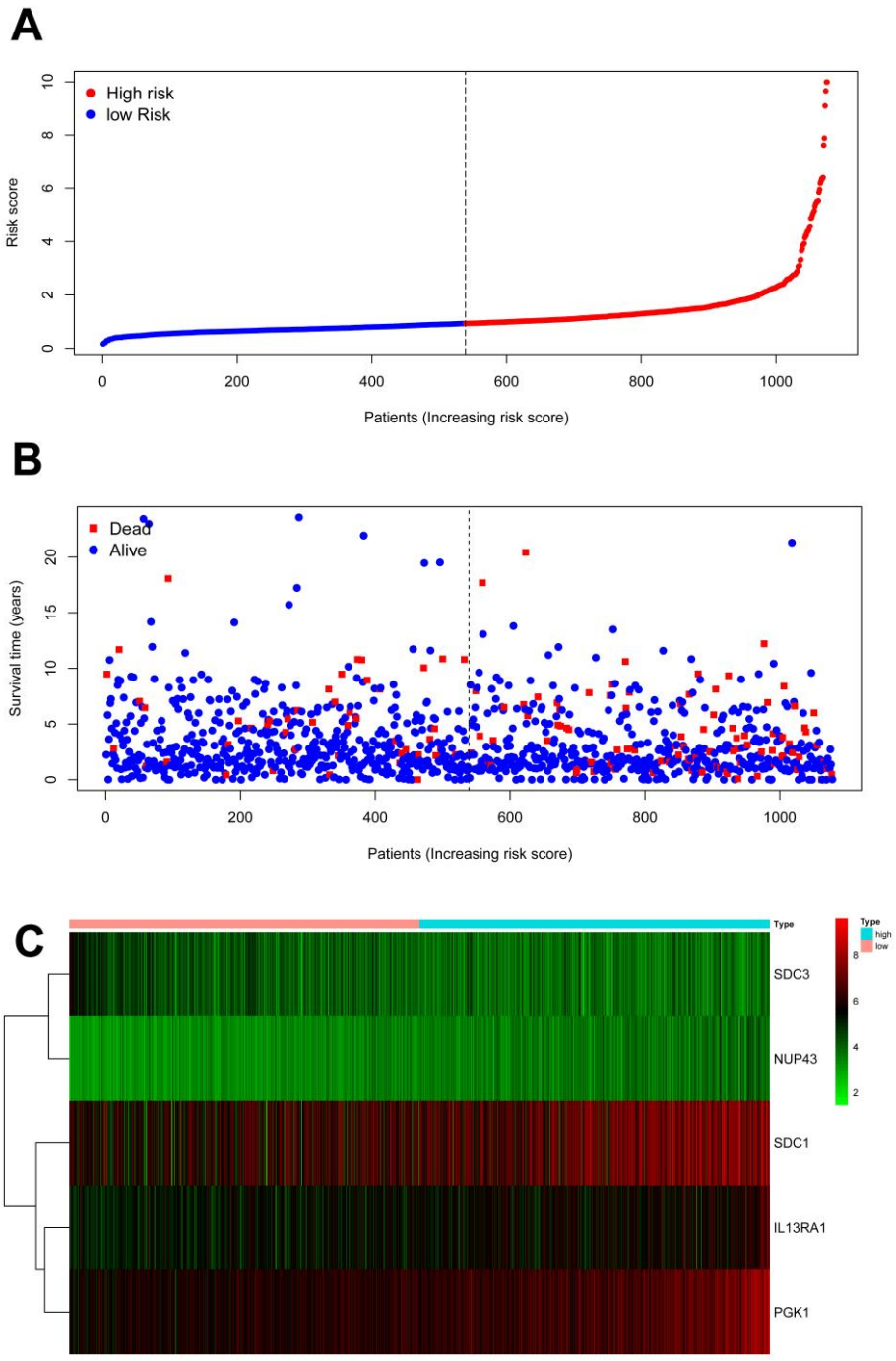

A
High risk
low Risk
Risk score
0
2
4
6
8
10
0
200
400
600
800
1000
Patients (Increasing risk score)
B
Dead
Alive
Survival time (years)
0
5
10
15
20
0
200
400
600
800
1000
Patients (Increasing risk score)
C
Type
SDC3
NUP43
SDC1
IL13RA1
PGK1
Type
high
low
8
6
4
2

Figure 2

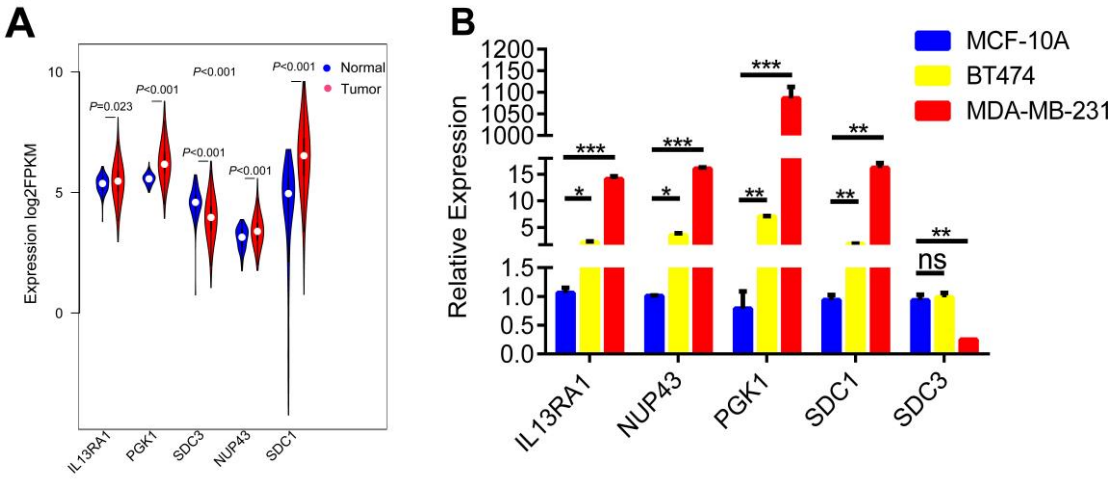
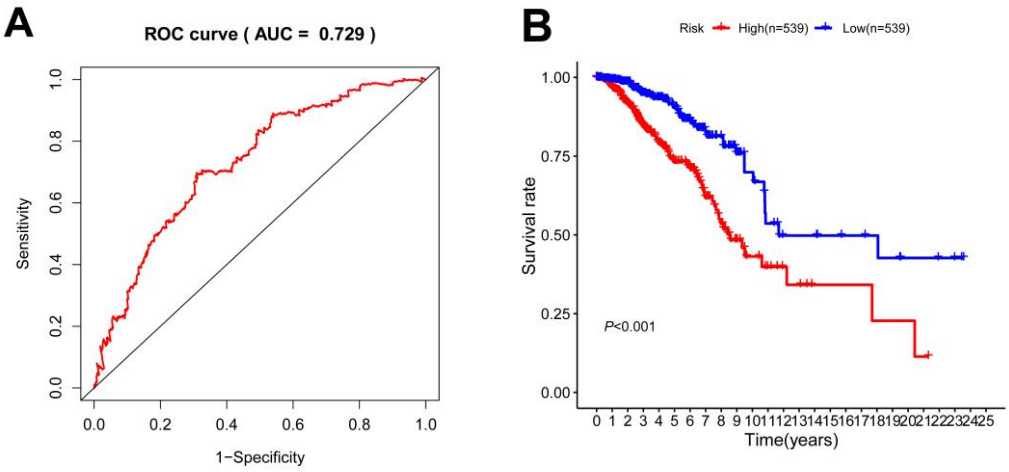

A
Expression log2FPKM
10
5
0
P=0.023
P<0.001
P<0.001
P<0.001
P<0.001
Normal
Tumor
IL13RA1
PGK1
SDC3
NUP43
SDC1
B
Relative Expression
1200
1150
1100
1050
1000
15
10
5
1.5
1.0
0.5
0.0
MCF-10A
BT474
MDA-MB-231
***
*
***
*
***
**
**
**
ns
IL13RA1
NUP43
PGK1
SDC1
SDC3

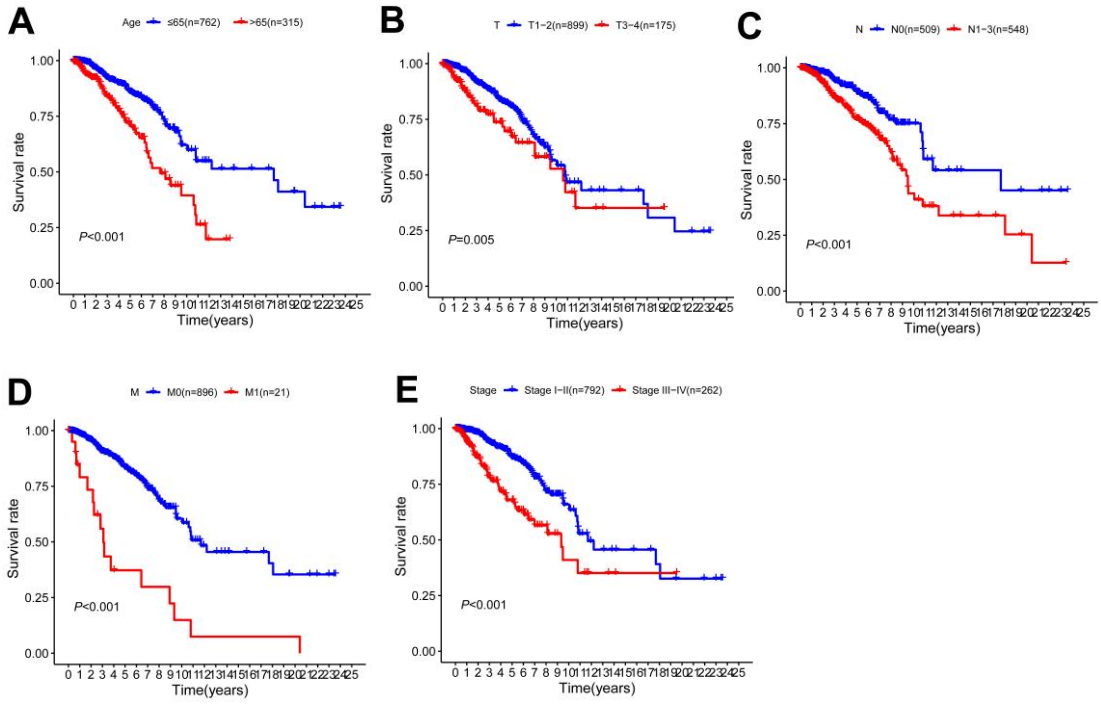
Figure 3



501

502 Figure 4

503



504

505 Figure 5

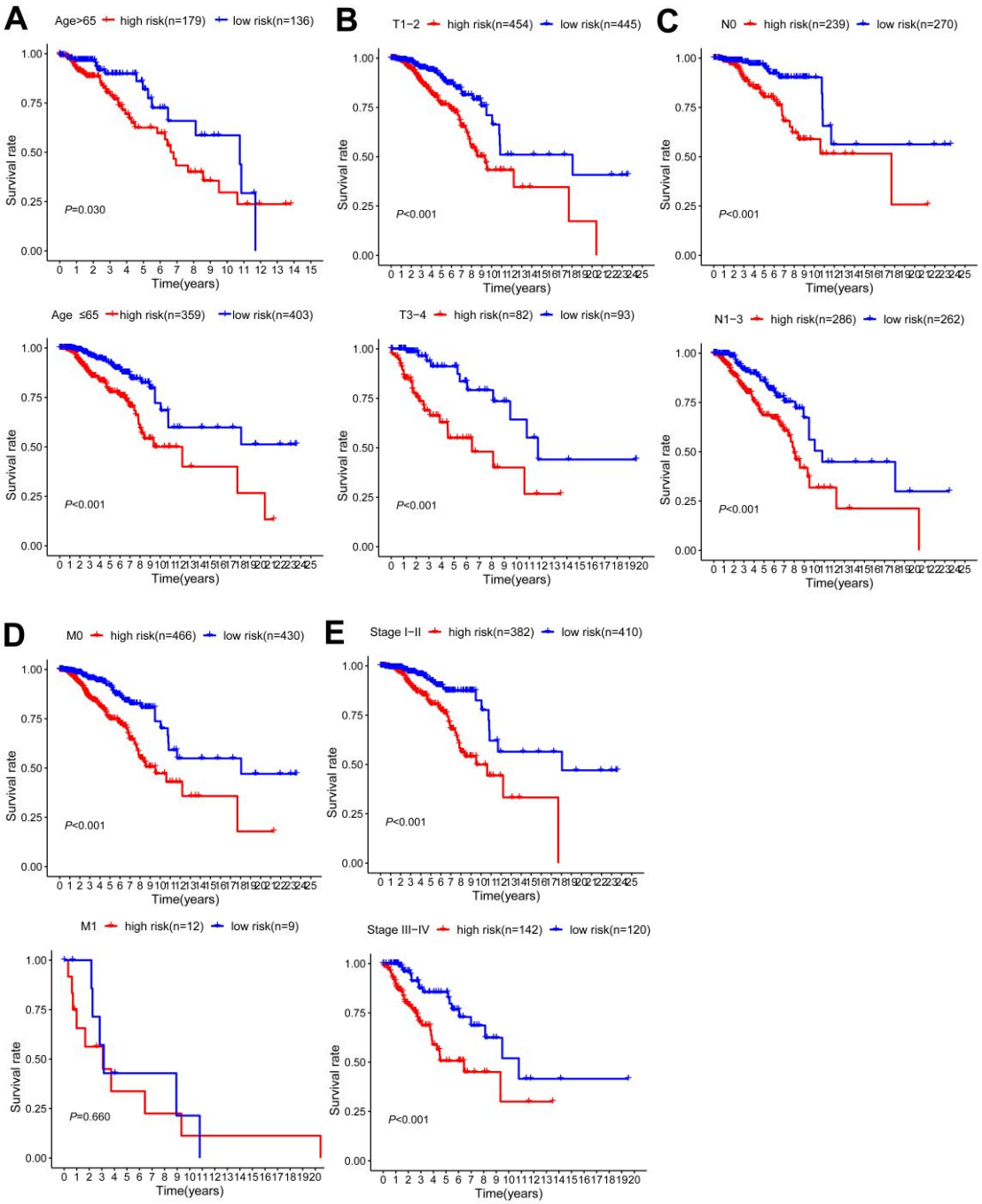

A
Age>65 high risk(n=179) low risk(n=136)
P=0.030
Age ≤65 high risk(n=359) low risk(n=403)
P<0.001
B
T1−2 high risk(n=454) low risk(n=445)
P<0.001
T3−4 high risk(n=82) low risk(n=93)
P<0.001
C
N0 high risk(n=239) low risk(n=270)
P<0.001
N1−3 high risk(n=286) low risk(n=262)
P<0.001
D
M0 high risk(n=466) low risk(n=430)
P<0.001
M1 high risk(n=12) low risk(n=9)
P=0.660
E
Stage I−II high risk(n=382) low risk(n=410)
P<0.001
Stage III−IV high risk(n=142) low risk(n=120)
P<0.001
Survival rate
Time(years)

Figure 6

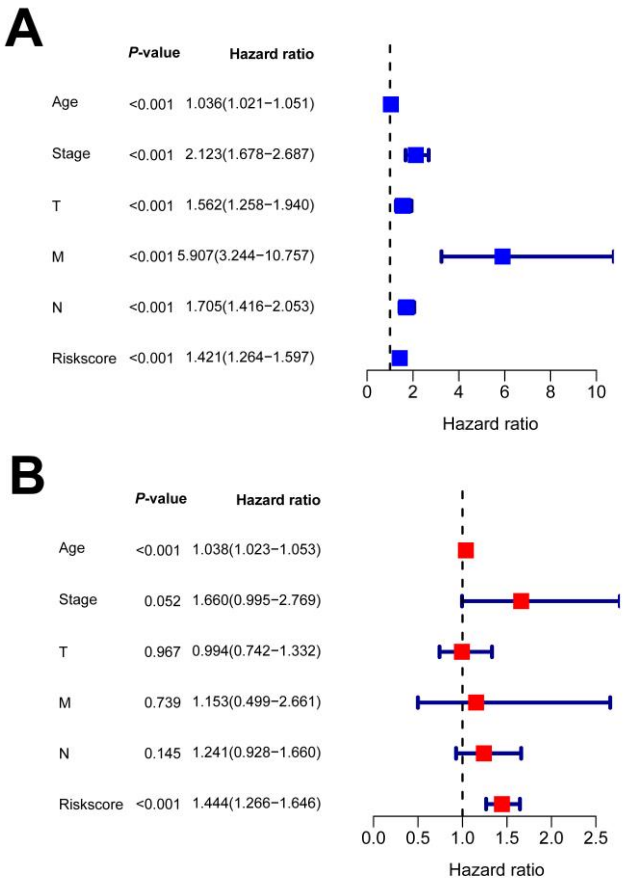

A
P-value
Hazard ratio
Age <0.001 1.036(1.021−1.051)
Stage <0.001 2.123(1.678−2.687)
T <0.001 1.562(1.258−1.940)
M <0.001 5.907(3.244−10.757)
N <0.001 1.705(1.416−2.053)
Riskscore <0.001 1.421(1.264−1.597)
0 2 4 6 8 10
Hazard ratio
B
P-value
Hazard ratio
Age <0.001 1.038(1.023−1.053)
Stage 0.052 1.660(0.995−2.769)
T 0.967 0.994(0.742−1.332)
M 0.739 1.153(0.499−2.661)
N 0.145 1.241(0.928−1.660)
Riskscore <0.001 1.444(1.266−1.646)
0.0 0.5 1.0 1.5 2.0 2.5
Hazard ratio

Figure 7

Figures

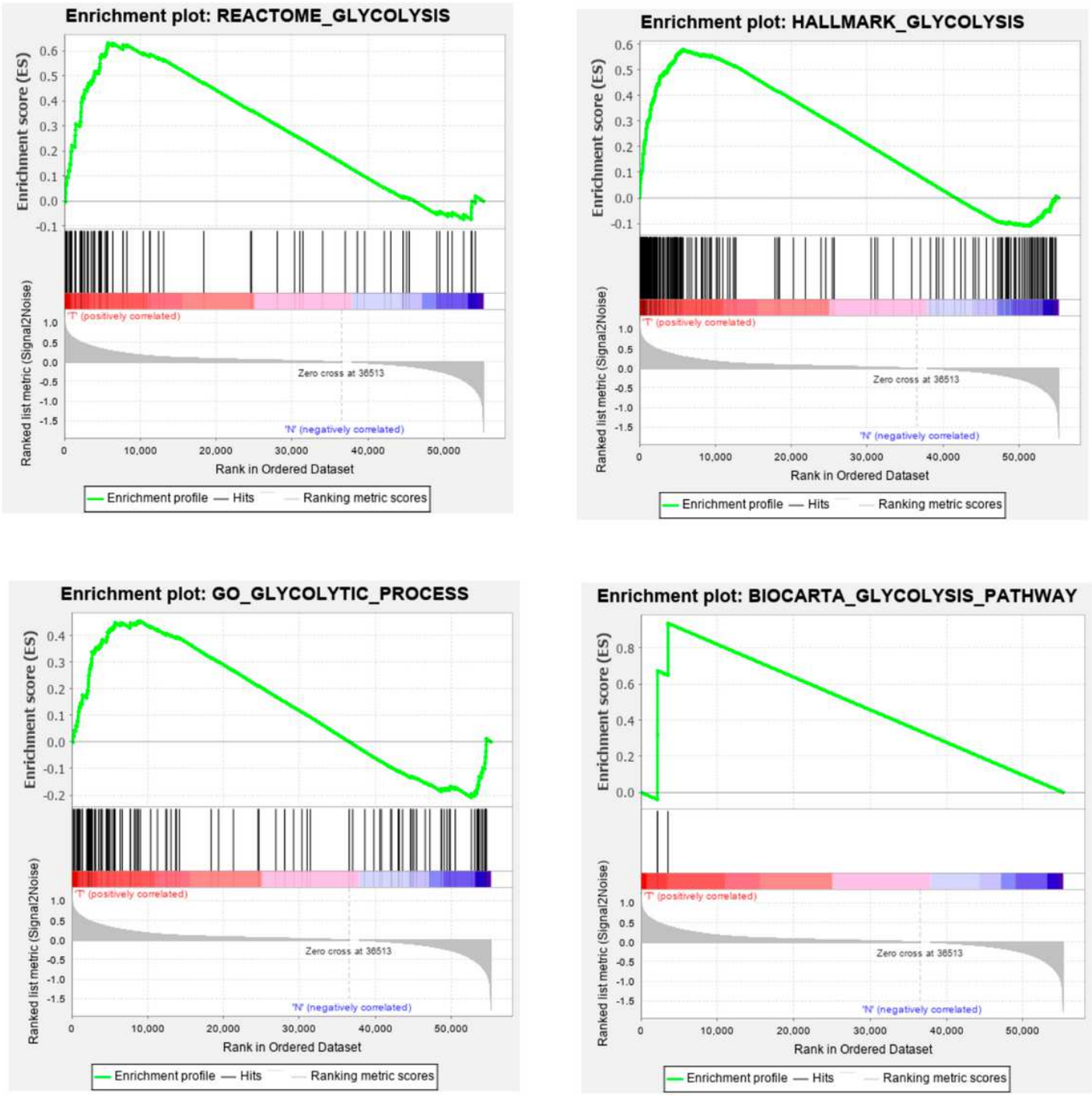


Figure 1

The enrichments of glycolysis-related pathways significantly differ between normal and neoplastic tissues in breast cancer. ($P<0.05$)

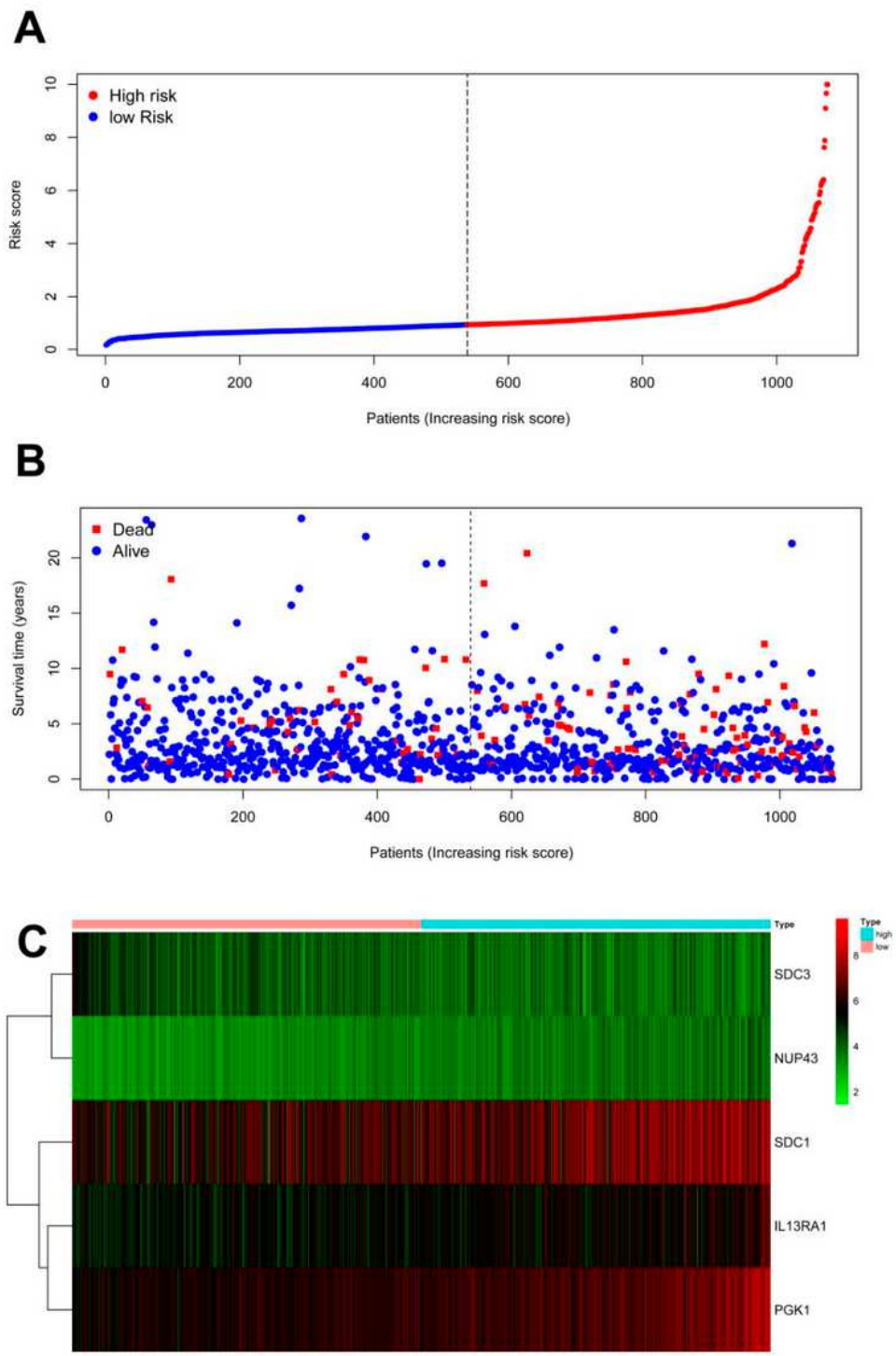


Figure 2

A risk-score based on five glycolysis-related signatures to predicts overall survival in patients with breast cancer. A) The risk-score distribution. B) Overall survival distribution of patients with different risk-score. C) Visualization expression pattern of the five glycolysis-related signatures in high-risk and low-risk patients.

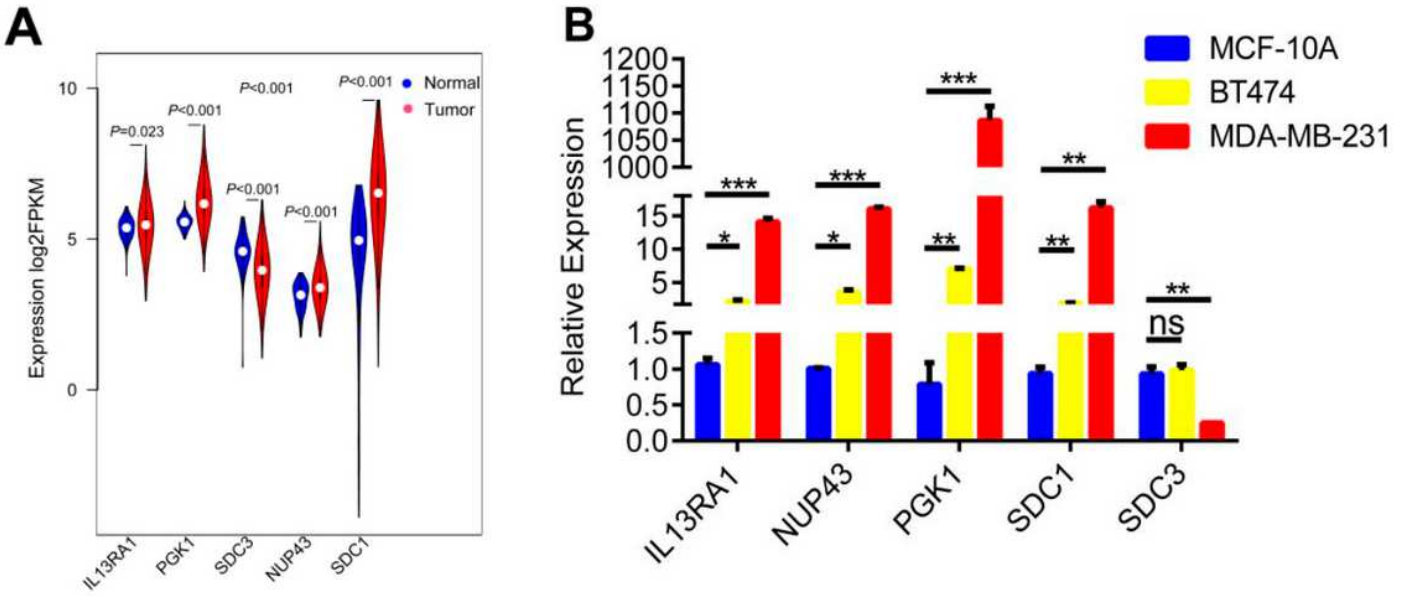


Figure 3

Expression patterns of the five glycolysis-related signatures in breast tissues and cell lines. A) Expression patterns in breast tissues. B) Expression patterns in breast cell lines by qRT-PCR. (*P<0.05, ** P<0.01, *** P<0.001)

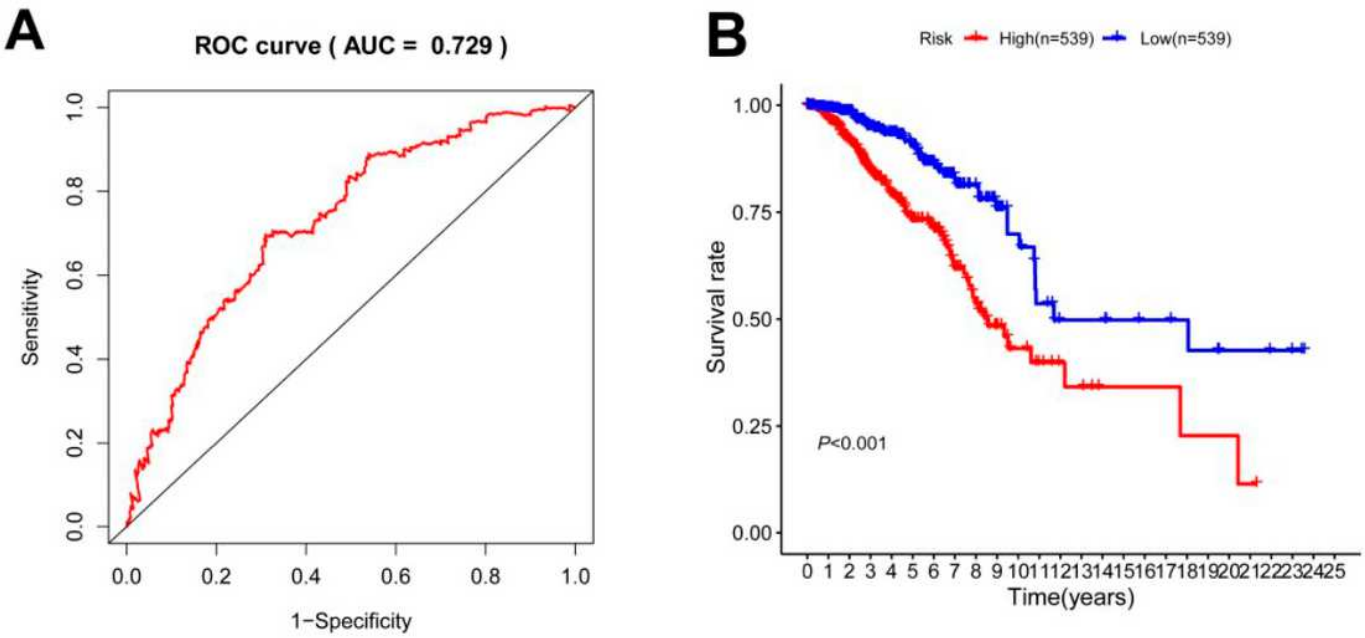


Figure 4

Values of a risk-score to predict the prognosis of patients with breast cancer. A) The Receiver Operating Characteristic Curve for the risk-score to predict prognosis of patients with breast cancer. B) Kaplan–Meier curves for prognostic value of the risk-score. (P<0.05)

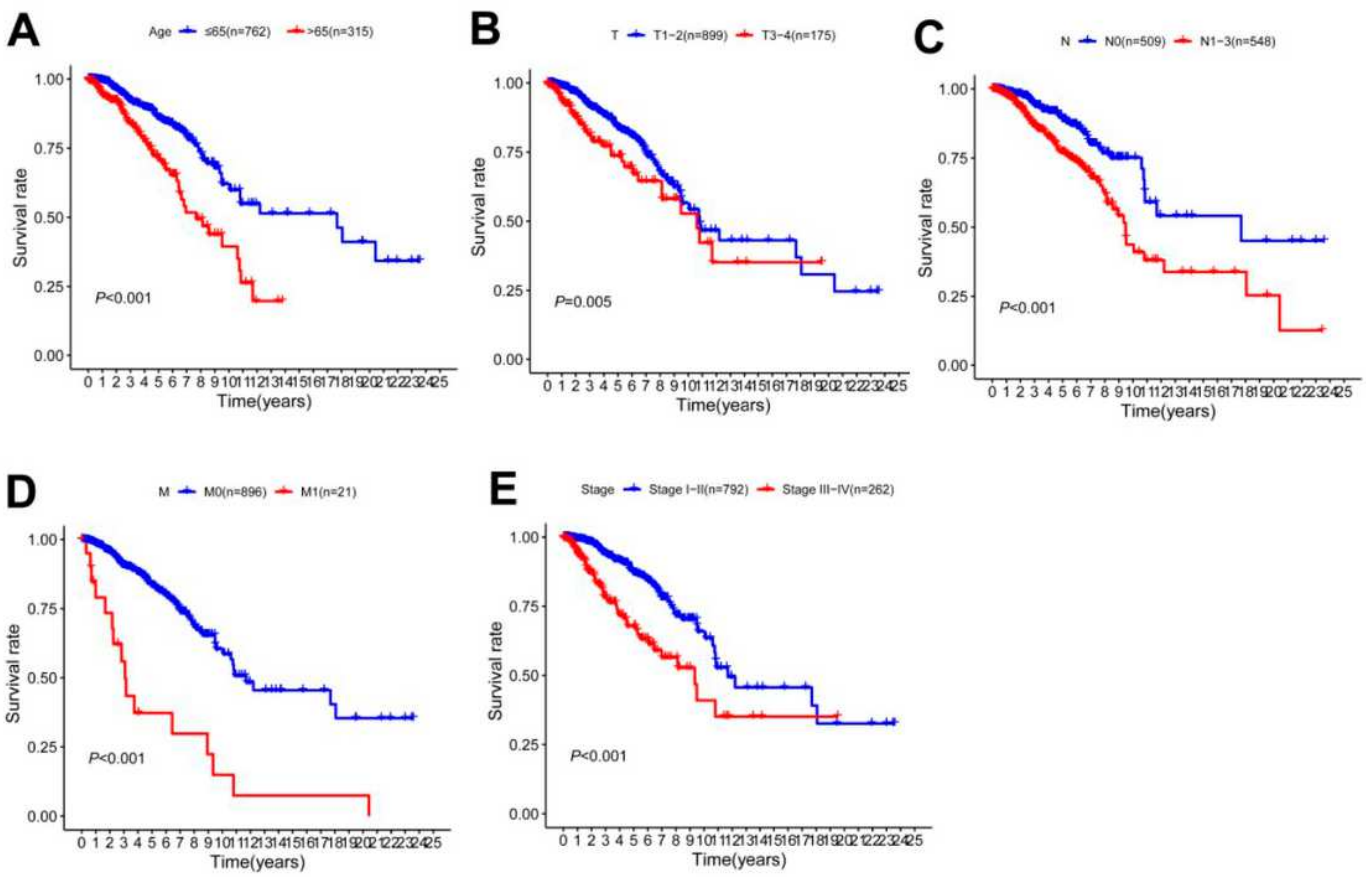


Figure 5

Kaplan–Meier survival analysis for patients with high and low risk-score in breast cancer. (stage: clinical stage; T: primary tumor; N: lymph nodes and M: distant metastasis, P<0.05)

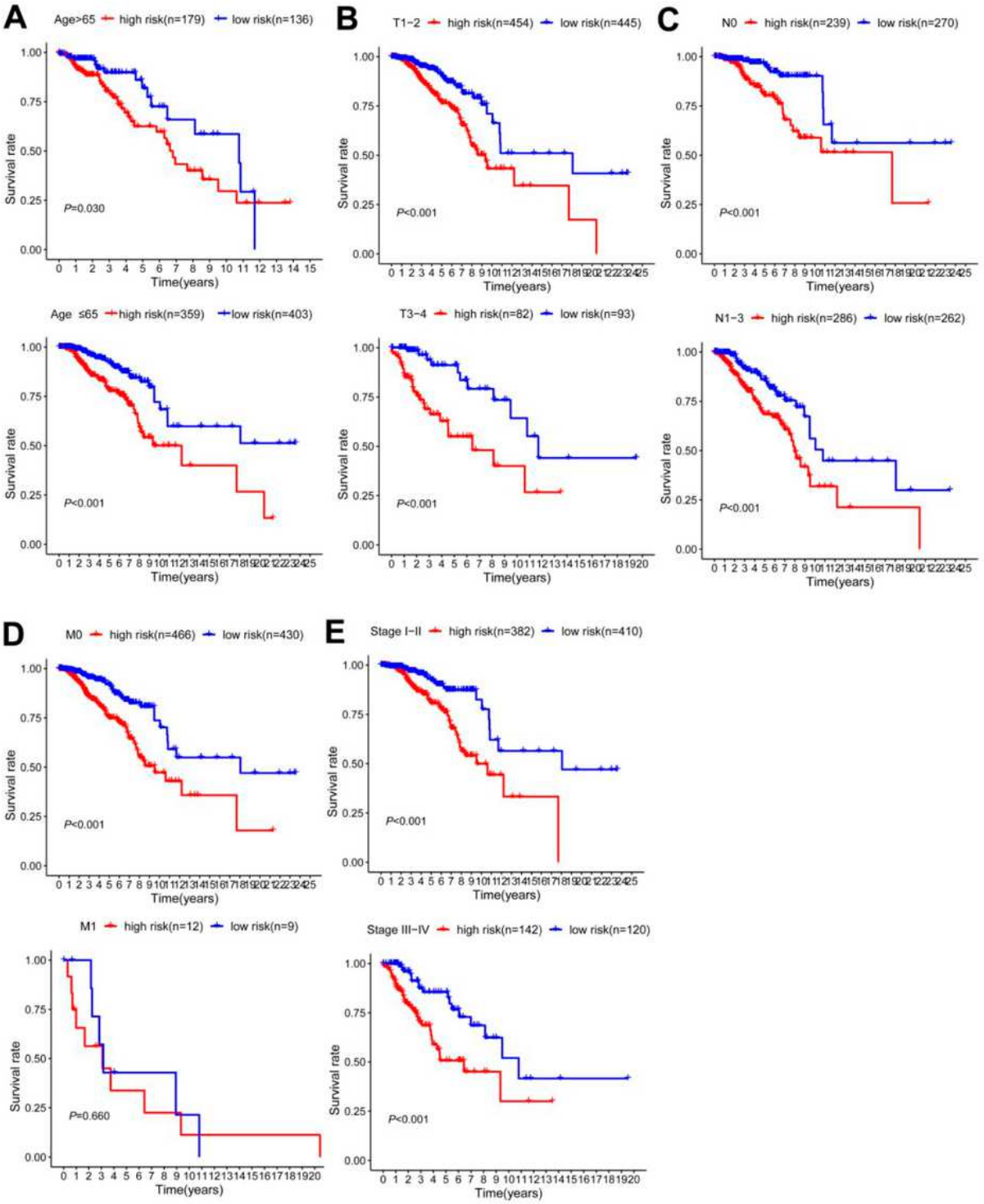


Figure 6

Kaplan–Meier survival analysis for patients with high and low risk-score among each clinical subgroup in breast cancer. (stage: clinical stage; T: primary tumor and N: lymph nodes, P<0.05)

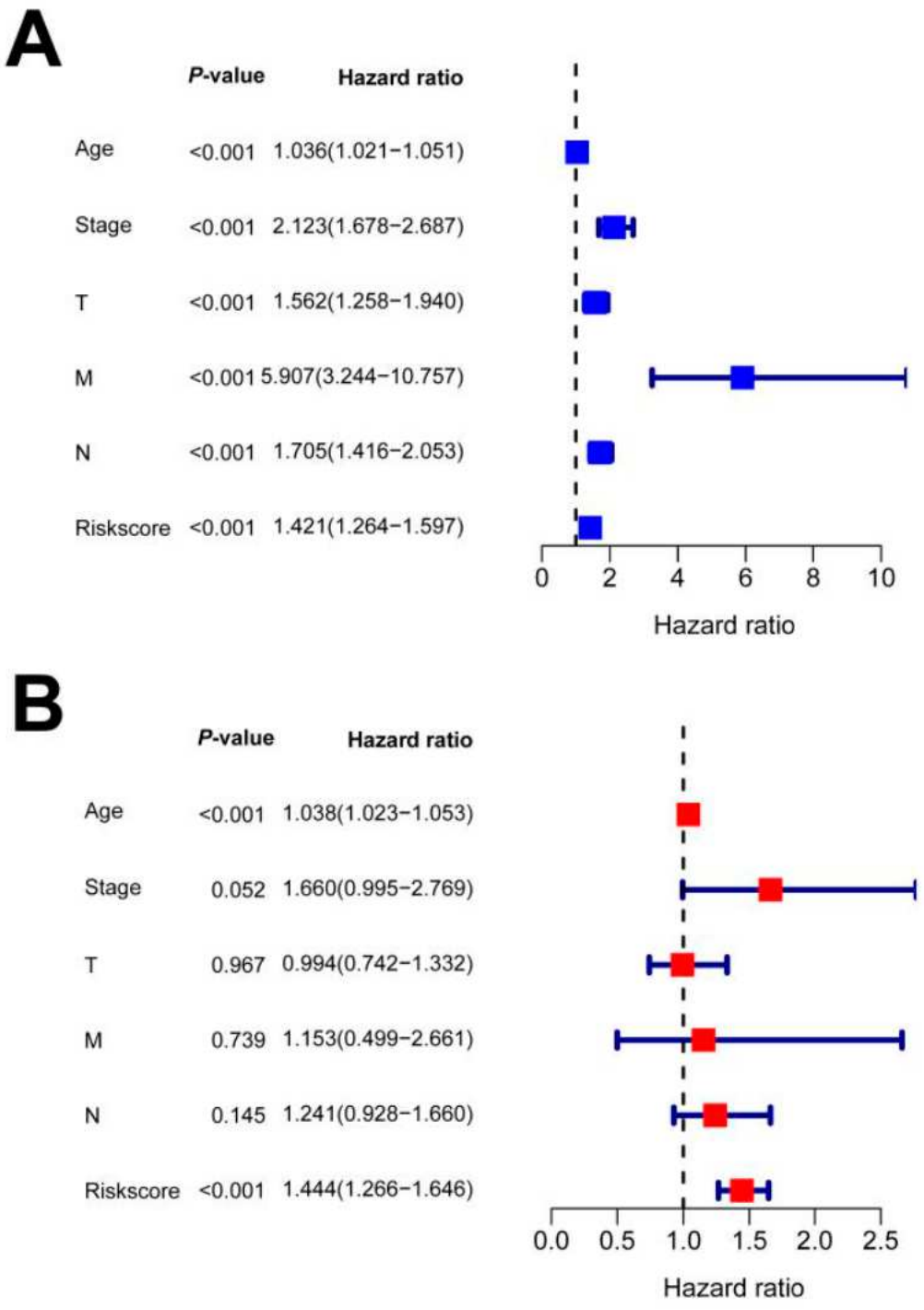


Figure 7

Validation of prognostic glycolysis-related risk score by Cox regression. A) The univariate Cox regression analysis. B) The multivariable Cox regression analysis. (P<0.05)

Supplementary Files

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

- Table0616.pdf