



Identification of Glycolysis-Related Gene Signature and Prediction on Prognosis of Colon Cancer

Qingsen Ran^{1,2}, Manjing Li³, Jiayin Han³, Lifang Wang³, Han Wang^{3,4}, Shaobo Liu^{5*} and Yanping Wang^{1*}

¹Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Science, Beijing 100700, China

²Postdoctoral Research Station, China Academy of Chinese Medical Science, Beijing 100700, China

³Institute of Chinese Materia Medica, China Academy of Chinese Medical Science, Beijing 100700, China

⁴Affiliated Hospital of Changchun University of Chinese Medicine, Changchun 130021, China

⁵The Key Laboratory of Optimal Utilization of Natural Medicine Resources, School of Pharmaceutical Sciences, Guizhou Medical University, Guiyang 550025, China

Qingsen Ran and Manjing Li contributed equally to this work.

ABSTRACT

Identification of core biomarkers for cancer prevention and treatment is crucial. Tumor glycolysis is an important metabolic phenotype and many natural compounds from medicinal plants exhibit anti-glycation effects. Therefore, it is of potential significant to find biomarkers based on glycolysis-related genes in predicting the prognosis of colon cancer (CC), a common invasive gastrointestinal tumor. In this study, clinical and gene data were collected to identify glycolytic genes that significantly associated with overall survival (OS) rate of CC patients through gene set enrichment analysis (GSEA) and Cox regression models. The K-M method was used to determine the difference of OS rate with high and low risk scores. The accuracy of risk scores were determined based on receiver operating characteristic (ROC) curve. Hierarchical analysis and Cox regression were performed to assess the correlation between CC risk score and clinical symptoms. Moreover, qPCR was used to verify the expression level of prognostic genes in human colorectal cancer cell line (HCT116) and human colorectal epithelial cells (FHC cells). The results showed 202 glycolytic genes were found with statistical differences, and 4 glycolytic genes were identified, including Enolase 3 (*ENO3*), glycocalcin-1 (*GPC1*), Nucleolar Protein 3 (*NOL3*) and Stanniocalcin 2 (*STC2*). A prognostic risk score model was established, and the OS rate of high-risk patients was found significantly reduced ($p < 0.001$) compared to low-risk patients. The 5-year OS ROC area under curve (AUC) of the model was 0.75. qPCR results confirmed that glycolysis-related genes *ENO3*, *GPC1*, *NOL3* and *STC2* were significantly upregulated in HCT116 cells compared to FHC cells. In conclusion, our study identified four glycolytic genes that significant impact prognosis of CC, and established a prognostic risk prediction model, which can provide reference for prognosis assessment and efficient treatment in CC.

Article Information

Received 24 November 2021

Revised 25 December 2022

Accepted 18 March 2023

Available online 20 June 2023
(early access)

Authors' Contribution

Formal analysis, data curation, writing original draft preparation: QR and JH. Validation, visualization: ML, LW and HW. Methodology, data analysis, writing review and editing: SL and YW. All authors have read and agreed to the published version of the manuscript.

Key words

Colon cancer, Glycolysis, Biomarkers, Metabolic signature, Prognosis Judgment, Risk score model

INTRODUCTION

Identifying core biomarkers is crucial for cancer prevention and treatment, including colon cancer (CC), a common invasive gastrointestinal tumor with increasing incidence worldwide in recent years (Huang *et al.*, 2019; Lord and Hall, 2019). Although treatment regimens such as targeted therapy have greatly improved the clinical outcomes, most CC patients still have poor prognosis due to recurrence and metastasis (Gunter and Leitzmann,

* Corresponding author: liushaobo@gmc.edu.cn, ypwang@ircm.ac.cn
0030-9923/2023/0001-0001 \$ 9.00/0



Copyright 2023 by the authors. Licensee Zoological Society of Pakistan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

2006). Therefore, it is necessary to find biomarkers to predict the prognostic risk of distant metastasis in CC treatment for timely intervention and treatment.

In normal cells, cells can use a variety of metabolic pathways associated with biosynthesis in order to keep glucose in a relatively balanced state at all times. However, recent evidence has shown that the metabolism of cancer cells is very differently from that of the normal cells from which they originate. One of the characteristics of glucose metabolism in tumor cells is that they still choose to use glycolysis when the oxygen content is normal, which is called Warburg effect. Tumor cells mainly obtain energy through glycolysis, many natural compounds isolated from medicinal plants have shown promising anti-glycation effects. Glucose metabolism in cancer cells exhibits the feature of high aerobic glycolysis (Taieb *et al.*, 2019; Tao *et al.*, 2018). Enhanced glycolysis and glutamine metabolic reprogramming are the main signs of cancer, leading to a variety of metabolic reprogramming and hypoxia-induced resistance to chemotherapy drugs, which made the cancer cells to maintain a high proliferation rate and resistance to certain cell death signals. This kind of phenomenon has been found to be the new therapeutic targets and the focus of new cancer drugs.

According to the relevant researches, glycolysis is very important in proteus biosynthesis, hence conferring more advantages on cancer cells in the face of reduced nutrient supply (Abbaszadeh *et al.*, 2020; Feinberg *et al.*, 2017). The oncogenic regulation of glycolysis indicates the function of tumor glycolysis, and it is necessary to elucidate the underlying mechanism of glycolysis in CC (Jiang, 2017; Chen and Russo, 2012). Recent studies have shown that the clinicopathological characteristics such as age, metastasis and stage are not obvious and can not accurately predict the prognosis of cancer. Therefore, more and more mRNAs have been identified to determine the prognosis of CC (Bearn, 2017). However, these biomarkers are not sufficient to predict patient outcomes accurately and independently, especially the individual gene levels which may be influenced by various factors, thus restricting their application as effective independent prognostic indicators (Ji *et al.*, 2016).

In this study, we aimed to identify glycolysis-related genes in CC through gene set enrichment analysis (GSEA). Subsequently, we extracted meaningful mRNAs from CC patient samples and used whole genome expression data from the Cancer Genome Atlas (TCGA) to build a glycolysis-related prognosis model. Combined with multivariate data analysis, receiver operating characteristic (ROC) curve and independent prognosis analysis, the gene and protein expression levels in CC samples were explored. Additionally, the expression of related genes in

the human colorectal cancer cell line HCT116 was further verified by qPCR. The overall experimental design was shown in [Supplementary Figure S1](#). These analyses are expected to predict prognosis of CC, which provides a new target for the diagnosis of CC.

MATERIALS AND METHODS

Data collection

The CC patients' key features in TCGA (The Cancer Genome Atlas) data were downloaded through the data portal (<https://portal.gdc.cancer.gov/>). The clinical data including: survival, survival time, age, gender and cancer stages. Total 470 CC patients and 41 normal control samples in TCGA dataset were collected, and Perl code was used to preprocess the above raw data, including: transcriptome data processing, ID conversion, clinical information extraction, etc. (Mairers *et al.*, 2019).

Functional enrichment analysis

Gene Set Enrichment Analysis (GSEA) was used for enriching and analyzing the selected gene set, to find potential significant difference genes. Normalized enrichment score (NES) ≥ 1.0 and nominal $p \leq 0.05$ were set for the criteria of identifying significantly enriched gene sets in GSEA. Gene expression differences between CC patients and healthy samples were judged. The signal pathways related to glycolysis were further determined and glycolytic genes were identified. The relevant signal pathways were as follows: Glycolysis_by_fructose_2_6_bisphosphate_metabolism, Biocarta_glycolysis_pathway, Go_glycolytic_process, Hallmark_glycolysis, Kegg_glycolysis_gluconeogenesis, And Reactome_glycolysis.

The KEGG database was used for pathway enrichment analysis of these DEGs in an effort to gain high-level insights into their potential functional relevance. GO enrichment analyses were conducted to assess the enrichment of these DEGs in specific biological processes, molecular functions, and cell components (BPs, MFs, and CCs, respectively) (Liu *et al.*, 2022).

Glycolysis prognostic analysis and construction of risk scores model

The prognosis model of glycolysis was constructed by univariate Cox regression and multivariate Cox regression analyses. Cox algorithm was applied to calculate the relation of mRNA level and OS of tumor patients, and to determine the factors associate with glycolysis prognosis. For univariate Cox regression, hazard ratio (HR) >1 and Cox PV <0.05 were used as the screening criteria, multivariate Cox analysis was applied to assess results of univariate Cox analysis. The

final results were used as the related factors of glycolysis prognosis and these factors were also the weight of glycolysis prognosis model. The above analysis was conducted by Rstudio based on the R package. The screening of prognostic factors were divided into two types: risk factor ($HR > 1$) and protective factor ($0 < HR < 1$). As there was no protective prognostic factors screened, so it was ignored and only risk factor was focused. According to the Cox analysis, the formula of prognostic risk score was calculated as the sum of each risk coefficient (coef) value multiply by expression of corresponding prognostic factor (Chen *et al.*, 2020).

Cell culture

Human colorectal cancer cell line HCT116 and human normal colorectal mucosa epithelial cell (FHC) were obtained from China National Biomedical Cell Repository. The main reagents and equipment were: 1% penicillin-streptomycin (Solarbio, Beijing, China), 10% fetal bovine serum, 0.05% trypsin (Gibco, Waltham, MA, USA), RNAeasy animal RNA isolation kit (Qiagen, Hilden, Germany), and RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA).

HCT116 cells were cultured in RPMI-1640 medium supplemented with 1% penicillin-streptomycin and 10% fetal bovine serum, while FHC cells were grown in DMEM medium with the same conditions. Both cell lines were incubated at 37°C in a 5% CO₂ humidified incubator (Sanyo, Japan), and were passaged using 0.05% trypsin when they reached 80% confluency.

Validation of prognostic genes by qPCR assay

Total RNA was extracted using the RNAeasy animal RNA isolation kit with rotating columns according to the instructions. The extracted RNA was reverse-transcribed to cDNA using the RevertAid First Strand cDNA Synthesis Kit following standard protocols. The qPCR assay was performed on a Roche LightCycler 480 (Roche Diagnostics, Vilvoorde, Belgium). The protocol was as follows: initial incubation at 95°C for 10 min, followed by 50 cycles of 95°C for 15 s, annealing for 30 s, and elongation at 72°C for 30 s; and finally cooling at 40°C for 30 s. The annealing temperature for *ENO3*, *GPC1*, *NOL3*, *STC2* and β -actin was 61, 55, 56, 60 and 55°C, respectively. Three replicates of each sample were analyzed, and the relative expression of mRNA was quantified after normalization to β -actin. The primers are listed in Table I.

Statistical analysis

The relation of the survival data of every variable

were tested using K–M method, while that of clinical data and risk scores were measured using Cox analysis method (Zhang *et al.*, 2021). The significant expression of glycolysis prognostic factors in normal and tumor samples was detected using the student *t*-test, with a statistically different threshold set at 0.05. All the analysis was conducted using R-studio software (Version 3.6.2, Boston, MA, USA).

Table I. Primer sequences for qPCR used in this study.

Primer	Sequence (5' to 3')
<i>ENO3</i>	F GGCTTCGCACCCAACATCCTG R CCATGCCGATCACCACCTTGTC
<i>GPC1</i>	F GCGGTGATGGCTGTCTGGATG R AGGTCTTCTGTCCTTCCTGCTCTG
<i>NOL3</i>	F CGGAGGAGCCAGAGCCAGAG R GTTCTGCTTCAGCCTCGGGTTC
<i>STC2</i>	F GACCCTGGCTTTGGTGTGGC R GTGCTGGATCTCCGCTGTATTCTG
β -actin	F CACTGTCGAGTCGCGTCC R TCATCCATGGCGAACTGGTG

RESULTS AND DISCUSSION

Enrichment analysis of glycolysis and differential genes

The result revealed 3 significant functional groups of glycolysis genes, namely Glycolysis_by_fructose_26_bisphosphate_metabolism, Hallmark_glycolysis, And Reactome_glycolysis (Supplementary Fig. S2). Compared with normal samples, Glycolysis_by_fructose_2_6_bisphosphate_metabolism in tumor samples was significantly correlated ($p < 0.05$), while the other two showed a stronger significant correlation ($p < 0.01$). Perl code was applied to extract the expression levels of glycolysis gene that were significantly related in the functional group, followed by the application of the difference analysis filter.

Compared with the normal samples, a total of 202 significantly differentially expressed genes were found ($p < 0.05$), of which 54 were downregulated and 148 were upregulated in cancerous and paracancerous tissues. To confirm whether these differential genes were mainly involved in glycolysis, we conducted GO and KEGG analysis using Metascape. The result indicated that (Fig. 1), the most important biological process was associated with glycolysis and the most important metabolic process was related to glycometabolism.

Table II. Potential prognostic glycolysis genes.

Genes	HR	HR.95L	HR.95H	Cox P value
<i>NOL3</i>	1.08562788644099	1.04045984577194	1.13275674463334	0.000151092147214812
<i>ANKZF1</i>	1.09603376526771	1.03408003657031	1.16169926129817	0.00200961854957091
<i>IDUA</i>	1.07798480240408	1.00272060514574	1.15889832945567	0.0419981696582861
<i>PPFIA4</i>	2.91608494870452	1.45372674906806	5.8494840474748	0.00258370907825577
<i>STC2</i>	1.05029132248173	1.01116247625749	1.09093433348441	0.0113087367768567
<i>GPC1</i>	1.03151430929995	1.00073007654951	1.06324551967027	0.0447310436445428
<i>ENO3</i>	1.54819254587463	1.23185128752019	1.94577071387156	0.000178260609472247

Note: Four prognostic glycolysis genes were in bold.

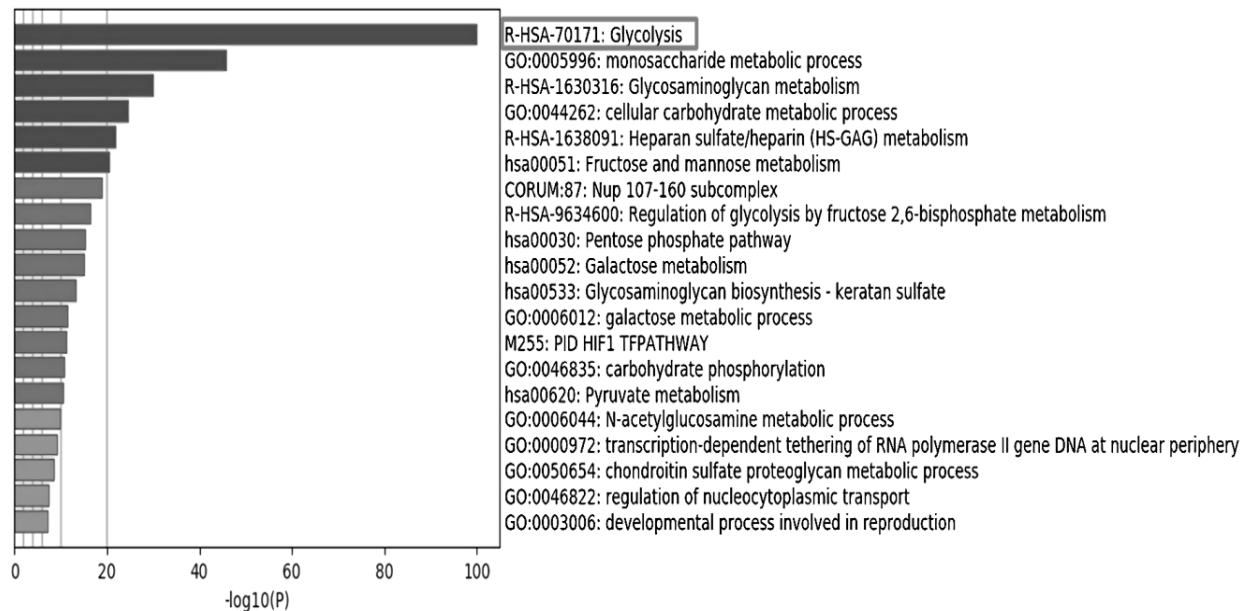


Fig. 1. GO and KEGG pathway enrichment analysis of significant glycolysis genes based on Metascape.

Note: the significance increased from bottom to top; marked in box indicated the most significant metabolic pathway.

Establishment of glycolysis prognosis model

The expression of the differential genes was analyzed in conjunction with the clinical survival data of tumor samples using R language. Single factor Cox regression analysis was performed on the differential gene expression of tumor samples, and the findings were shown in Table II. A total of 7 genes were screened as potential glycolysis prognostic genes. Subsequently, using the Cox regression analysis, four potential glycolysis prognostic genes were screened, including Enolase 3 (*ENO3*), glypican-1 (*GPC1*), Nucleolar Protein 3 (*NOL3*) and Stanniocalcin 2 (*STC2*). Using the above 4 screened genes, a glycolysis prognosis model was built. Notably, the criterion for accessing various risk was based on the mean value. Consequently, patients with values greater than median were regarded as

high-risk, whereas that with values below median were regarded as low-risk.

Analysis of glycolysis prognosis model

After the establishment of the glycolysis prognosis model, patients were divided into two risk groups, and their survival curves were analyzed. The survival rate of the two groups showed a downward trend over time, but the curves were significantly separated ($p < 0.01$), with eight-year survival rates of 30.1% and 55% (Fig. 2A). The ROC curve analysis revealed an area under curve (AUC) value of 0.75. The accuracy of prediction by the model was higher than 0.7, indicating that the diagnosis of glycolysis prognostic model had practical significance (Fig. 2B).

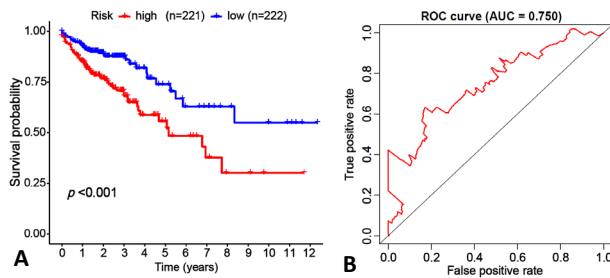


Fig. 2. Survival curve analysis (A) and receiver operating characteristic (ROC) curve analysis (B) of glycolysis prognosis model with colon cancer patients.

Notes: A: red indicates high-risk patients (n= 221) and blue indicates low-risk patients (n= 222), and with a significant difference in survival rate ($p < 0.001$). B: ROC curve analysis result was 0.750.

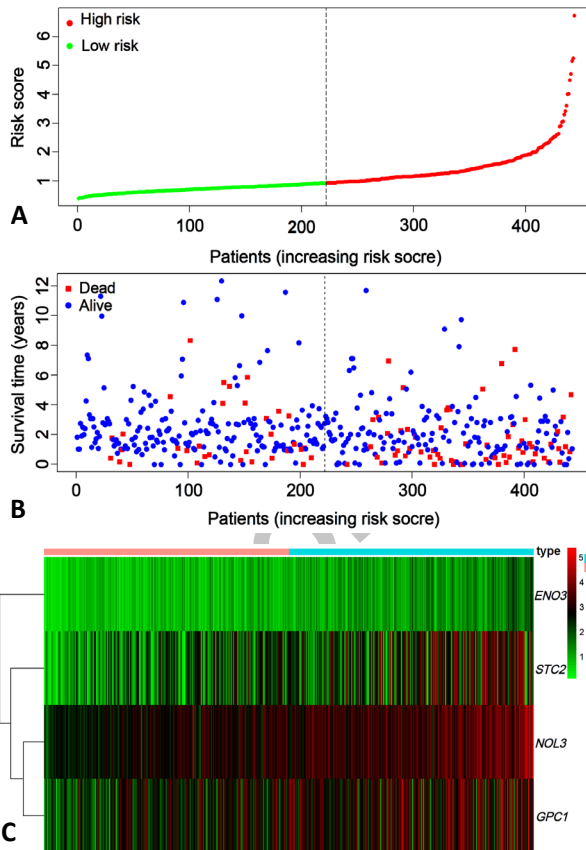


Fig. 3. Risk map for glycolysis prognosis model in colon cancer patients.

A: risk score. B: the survival number of low risk score patients was significantly more than that of high risk score patients. C: heat map analysis. *ENO3*, *GPC1*, *NOL3*, *STC2* genes expression in patients with high risk scores were significantly more than those with low risk scores.

The risk curve combinatorial graphs demonstrated that patients from left to right with low-risk scores were denoted with green and remained relatively stable over time, while those with high-risk scores were denoted with red and exhibited an increasing risk score over time, the overall risk score from left to right showing a gradually increasing (Fig. 3A). In Figure 3B, from left to right, the number of deaths increased and the maximum survival time decreased. Moreover, it can be seen that the genes level associated with prognosis of glycolysis from left to right was significantly deepened (Fig. 3C).

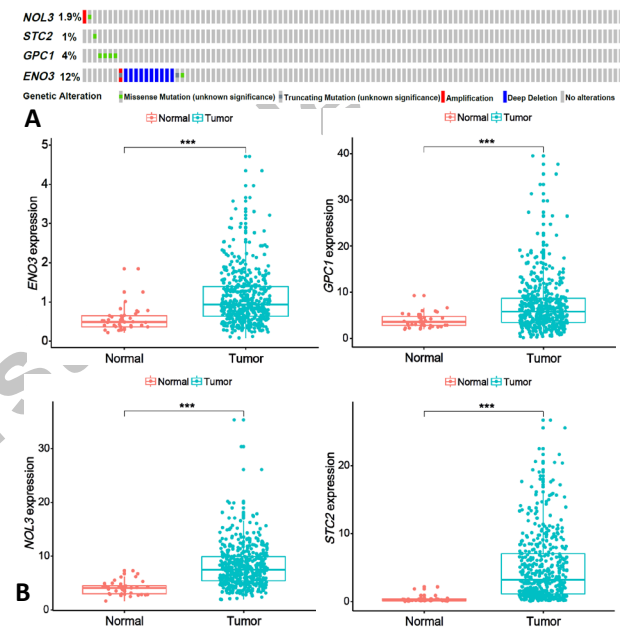


Fig. 4. Gene mutation analysis (A) and difference analysis (B) for model samples from TCGA.

Notes: A: 105 patients samples in cBioPortal online database; *NOL3*: gene amplification, 1 case; missense mutation, 1 case; *STC2*: missense mutation, 1 case; *GPC1*: missense mutation, 4 cases; *ENO3*: deep deletion, 10 cases; truncation mutation, 1 case; missense mutation, 1 case. B: Red color: normal samples; green color: tumor samples; ***: significant difference at $p < 0.001$.

To further analyze the mutations of the four selected genes for glycolysis prognosis of CC, clinical samples were analyzed through cBioPortal online database. Of the 105 patients studied, 1.9% had *NOL3* mutation, 1% had *STC2* mutation, 4% had *GPC1* mutation and 12% had *ENO3* mutation. Among them, 1 case of gene amplification (*NOL3*), 1 case of missense mutation; 4 cases of missense mutation (*GPC1*), 10 cases of deep deletion (*ENO3*), 1 case of truncation mutation, and 1 case of missense mutation (Fig. 4A). We further analyzed the expression of *ENO3*,

GPC1, *NOL3*, and *STC2* in tumor and normal samples for further verification. It was found that the expression level of the four genes were significantly up-regulated in tumor samples compared with normal samples ($p < 0.001$, Fig. 4B).

Independent prognostic analysis

To assess whether these mRNA markers on prognostic ability were independent of each other, various indicators including age, gender and tumor stage were established. The whole data set of prognostic was analyzed by single and multiple factors. The results showed that (Fig. 5A), in the single factor analysis, the p -value of age, tumor stage and risk score was less than 0.01, indicating that these three factors had a significant correlation with the survival status of patients. Moreover, the HR value was greater than 1, suggesting that these three factors were high-risk factors. That is, the order the patient, the higher the risk at tumor stage type, and the higher risk score, leading to the greater risk the patient is. The above results were further validated, confirming that age, tumor stage type and risk score can be independent of other clinical characteristics as independent prognostic factors (Fig. 5A).

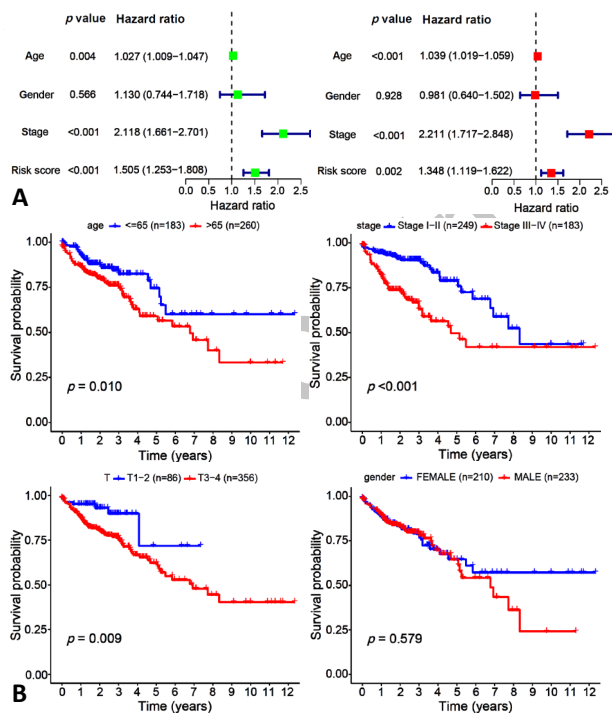


Fig. 5. Independent prognostic analysis (A) and clinical characteristics analysis (B) for the glycolysis prognosis model. A: uniForest and multForest analysis from left to right, the stage and risk scores were statistically significant (p -value < 0.05 , Hazard ratio > 1); B: 470 tumor samples with colon cancer based on TCGA.

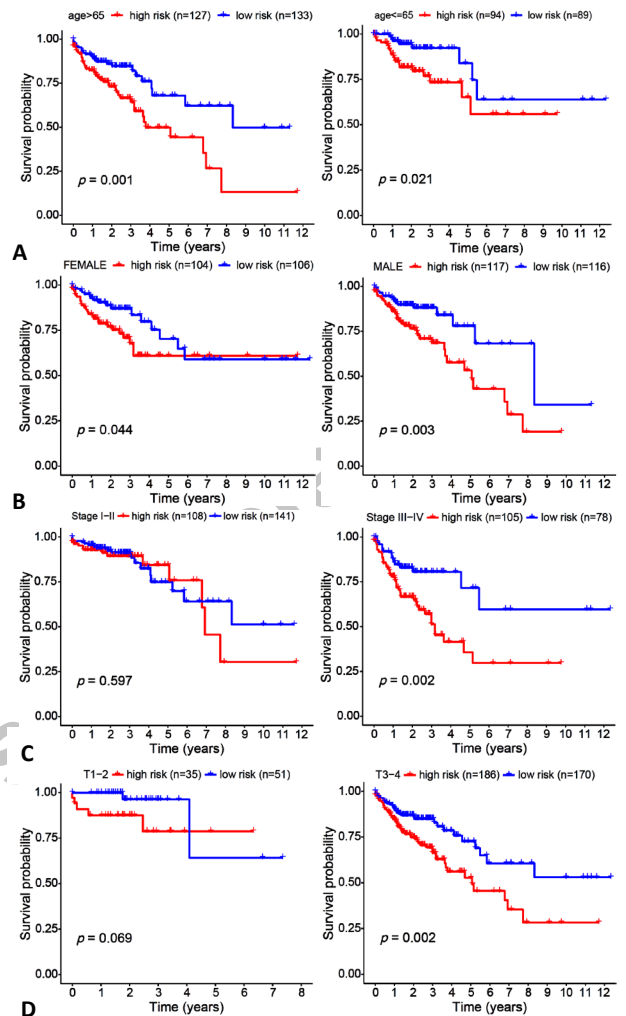


Fig. 6. Clinical grouping verification for colon cancer patients.

A: based on age; B: based on gender; C: based on stage; D: based on T.

Similarly, results were observed in the constructed clinical survival curve (Fig. 5B), the survival rate of patients over 65 years old was significantly different from that of patients aged under 65 years ($p = 0.01$). Moreover, compared to stage I, II in the tumor stage, CC stage III, IV were obviously different ($p = 0.01$). However, there was no significant difference observed in that of CC between the sexes.

Validation of clinical grouping model

To further verify whether the model was applicable to different populations, clinical grouping verification was carried out. As depicted in Figure 6A, there was a significant difference in the high-risk and low-risk groups

for both age groups > 65 years ($p = 0.001$) and ≤ 65 years ($p = 0.012$). With respect to gender groups (Fig. 6B), a significant difference between the high-risk and low-risk groups was observed among female patients ($p = 0.044$); the difference was more significant compared to male patients ($p = 0.003$). For the stage groups (Fig. 6C), there was no significant difference observed between the high-risk and low-risk groups in Stage I and II ($p = 0.597$), while a significant difference was observed in stage III and IV ($p = 0.002$). For T groups (Fig. 6D), there was no significant difference between the high-risk and low-risk groups in T1-2 group ($p = 0.069$); however, a significant difference was observed in T3-4 group ($p = 0.002$). These results suggested that the glycolysis prognosis factors in CC may not be well predicted in early diagnosis, and the glycolysis prognosis model may exhibit weaknesses in the early stage of CC.

Glycolysis-related genes *ENO3*, *GPC1*, *NOL3*, *STC2* in HCT116 and FHC cells

To identify whether these glycolysis-related genes *ENO3*, *GPC1*, *NOL3* and *STC2* were responsive in colon cancer, we studied the mRNA expression levels of these genes in FHC cells and HCT116 cells *in vitro* by qPCR. As shown in Figure 7, the mRNA expression levels of *ENO3*, *GPC1*, *NOL3* and *STC2* in HCT116 cells were significantly upregulated ($P < 0.01$) compared to those in FHC cells.

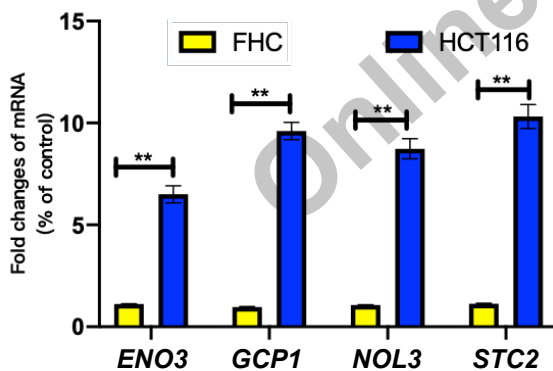


Fig. 7. Relative expression of *ENO3*, *GPC1*, *NOL3*, *STC2* in FHC and HCT116 cells.

FHC, human normal colorectal mucosa epithelial cells; HCT116, Human colorectal cancer cell line; Data were representative of three separate trials; ** $P < 0.01$.

DISCUSSION

CC is a common digestive system disease. Although the mortality of CC patients remains largely unchanged, the mortality of elderly patients (≥ 65 years old) is increasing

year by year (Abdel-Wahab *et al.*, 2019; Guillaumond *et al.*, 2014). The occurrence of diseases is usually accompanied with the abnormal expression of specific signal molecules (Song and Miao, 2022). Metabolism of tumor cells has a special way called the Warburg effect. Although under aerobic conditions, energy is still rapidly produced mainly through glycolysis, which is therefore a crucial energy source for tumor cells. Thus, glycolysis-related genes have significant associations with the prognosis of CC patients (Gendoo, 2020; Ibrahim *et al.*, 2018). Consequently, we developed a glycolytic-related genes model based on the characteristics of glycolytic metabolism of tumors to predict the prognosis of CC patients.

High-throughput technology has facilitated large-scale biological data research, with abundant genomic data have been extracted to identify new diagnostic biomarkers (Pan *et al.*, 2021; Chaneton and Gottlieb, 2012; Li *et al.*, 2014). Recent studies have constructed novel prognostic markers of gene expression levels or mutations based on RNA sequencing data, and identified them using Cox proportional hazards regression models (Ren *et al.*, 2020; Lai *et al.*, 2017). In this study, Cox regression analysis was carried out to evaluate the application value of four glycolytic gene combinations in predicting the prognosis of CC patients, and a correlation regression analysis model was established. The CC dataset in TCGA was used, and the tissue-related data were compared and analyzed. K-M analysis showed that there was a negative correlation between risk parameters and prognosis. This study used a statistical model composed of multiple prognostic markers, combined with the predicted effect of each component glycolytic gene to improve predictive accuracy, which is widely used because they are more accurate than single biomarkers in this field. However, a limitation of the study was the lack of data about cancer cell metastasis and recurrence in TCGA, therefore we only use OS rate to evaluate the patients with CC.

Using a bioinformatics approach, 4 genes that associated with cellular aerobic glycolysis (*ENO3*, *GPC1*, *NOL3*, *STC2*) were identified, and their prognostic value in CC were demonstrated. *ENO3* is mainly involved in glycolytic and glucose metabolism, and changes in *ENO3* gene expression can be observed in various tumor cells. Pyruvate kinase is a kind of crucial enzyme in glycolysis pathway, regulated by the enolase enzyme, which contributing in converting enol phosphate type ADP into ATP, pyruvic acid and pyruvate (Munir *et al.*, 2020; Amin *et al.*, 2019; Guillaumond *et al.*, 2014). *GPC1* is a member of the GPC family of glucosylphosphatidylinositol-anchored membrane-related heparan sulfate proteoglycans (HSPG), which includes six subtypes that mainly regulate cell growth and differentiation (Moutinho-Ribeiro *et al.*,

2020). However, limited research data has been conducted on the association between *GPC1* and other cancers, so the study of *GPC1* may need further exploration and analysis for CC. Nucleolus protein plays an important role in regulating tumor nucleolus proteins, improving the synthesis level of chromosomal structural proteins and promoting the early stage of tumor cells by accelerating the spindle division rate of cancer cells. However, although *NOL3* is highly expressed in many non-terminally differentiated tissue tumor cell lines, including liver cell lines (Huang *et al.*, 2018), the association between *NOL3* and tumor development remains unclear and needs to be further explored. *STC* is a glycoprotein hormone that is mostly involved in the physiological functions of the body through spontaneous secretion or paracrine. Changes in the concentration of *STC2* can significantly promote changes in cancer cells *in vivo*. *STC2* can affect the function of endoplasmic reticulum of cancer cells, and promote the damage of mitochondria, leading to abnormal regulation of tumor cell cycle with significantly increased changes in cell cycle (Yang *et al.*, 2020; Bironaite *et al.*, 2013; Ciriello *et al.*, 2012). Therefore, further studies on the genes and their prediction models are necessary to constantly identify new biomarkers and establish related prediction models. This holds great promise for the diagnosis, treatment and prognosis of cancer.

CONCLUSION

In this study, four glycolysis-related genes as biomarker in CC were identified, and a novel predictive risk score model was established based on their expression levels to predict the prognosis of CC patients. Moreover, their mRNA expressions were detected to further confirm their connection in CC. The identification of these glycolytic gene biomarkers for CC prognosis provides a new potential approach to improve the prediction and treatment of CC.

ACKNOWLEDGEMENT

Authors would like to express their gratitude to everyone who helped in research, and also extend their thanks to the reviewers for their valuable comments.

Funding

This research was funded by the Beijing Natural Science Foundation (7214291), the 71st batch of China Postdoctoral Science Foundation (2022M713495), National Natural Science Foundation of China (82104519), CACMS Innovation Fund (CI2021A04803),

China Academy of Chinese Medical Sciences Foundation (ZZ15-YQ-045), and Fundamental Research Funds for the Central Public Welfare Research Institutes (ZXKT 21018).

IRB approval and Ethical statement

The studies involving human participants were reviewed and approved by the Ethics Committee of China Academy of Chinese Medical Science and following all guidelines, regulations, legal, and ethical standards. Written informed consent for participation was not required for this study, the datasets analyzed in this study were publicly available.

Supplementary material

There is supplementary material associated with this article. Access the material online at: <https://dx.doi.org/10.17582/journal.pjz/20211124031155>

Statement of conflict of interests

The authors have declared no conflict of interests.

REFERENCES

- Abbaszadeh, Z., Çeşmeli, S. and Avcı, Ç.B., 2020. Crucial players in glycolysis: Cancer progress. *Gene*, **726**: 144158. <https://doi.org/10.1016/j.gene.2019.144158>
- Abdel-Wahab, A.F., Mahmoud, W. and Al-Harizy, R.M., 2019. Targeting glucose metabolism to suppress cancer progression: prospective of anti-glycolytic cancer therapy. *Pharmacol. Res.*, **150**: 104511. <https://doi.org/10.1016/j.phrs.2019.104511>
- Amin, S., Yang, P. and Li, Z., 2019. Pyruvate kinase M2: A multifarious enzyme in non-canonical localization to promote cancer progression. *Biochim. biophys. Acta Rev. Cancer*, **1871**: 331–341. <https://doi.org/10.1016/j.bbcan.2019.02.003>
- Bearne, S.L., 2017. The interdigitating loop of the enolase superfamily as a specificity binding determinant or flying buttress. *Biochim. biophys. Acta Rev. Cancer*, **1865**: 619–630. <https://doi.org/10.1016/j.bbapap.2017.02.006>
- Bironaite, D., Westberg, J.A., Andersson, L.C. and Venalis, A., 2013. A variety of mild stresses upregulate stanniocalcin-1 (STC-1) and induce mitohormesis in neural crest-derived cells. *J. neurol. Sci.*, **329**: 38–44. <https://doi.org/10.1016/j.jns.2013.03.011>
- Chaneton, B. and Gottlieb, E., 2012. Rocking cell metabolism: Revised functions of the key glycolytic regulator PKM2 in cancer. *Trends biochem. Sci.*, **37**: 309–316. <https://doi.org/10.1016/j.tics.2012.05.003>

- tibs.2012.04.003
- Chen, J.Q. and Russo, J., 2012. Dysregulation of glucose transport, glycolysis, TCA cycle and glutaminolysis by oncogenes and tumor suppressors in cancer cells. *Biochim. biophys. Acta Rev. Cancer*, **1826**: 370–384. <https://doi.org/10.1016/j.bbcan.2012.06.004>
- Chen, Q., Li, F., Gao, Y., Xu, G., Liang, L. and Xu, J., 2020. Identification of energy metabolism genes for the prediction of survival in hepatocellular carcinoma. *Front. Oncol.*, **10**: 1210. <https://doi.org/10.3389/fonc.2020.01210>
- Ciriello, J., Oiamo, T., Moreau, J., Turner, J. and Wagner, G., 2012. Effects of the calcium-regulating glycoprotein hormone stanniocalcin-1 within the nucleus of the solitary tract on arterial pressure and the baroreceptor reflex. *Neuroscience*, **207**: 88–102. <https://doi.org/10.1016/j.neuroscience.2012.01.044>
- Feinberg, A.E., Chesney, T.R., Acuna, S.A., Sammour, T. and Queresby, F.A., 2017. Oncologic outcomes following laparoscopic versus open resection of pT4 colon cancer. *Dis. Colon Rectum*, **60**: 116–125. <https://doi.org/10.1097/DCR.0000000000000641>
- Gendoo, D.M., 2020. Bioinformatics and computational approaches for analyzing patient-derived disease models in cancer research. *Comput. Struct. Biotechnol. J.*, **18**: 375–380. <https://doi.org/10.1016/j.csbj.2020.01.010>
- Guillaumond, F., Iovanna, J.L. and Vasseur, S., 2014. Pancreatic tumor cell metabolism: Focus on glycolysis and its connected metabolic pathways. *Arch. Biochem. Biophys.*, **545**: 69–73. <https://doi.org/10.1016/j.abb.2013.12.019>
- Gunter, M.J. and Leitzmann, M.F., 2006. Obesity and colorectal cancer: epidemiology, mechanisms and candidate genes. *J. Nutr. Biochem.*, **17**: 145–156. <https://doi.org/10.1016/j.jnutbio.2005.06.011>
- Huang, L., Jiang, X., Wang, Z., Zhong, X., Tai, S. and Cui, Y., 2018. Small nucleolar RNA host gene 1: A new biomarker and therapeutic target for cancers. *Pathol. Res. Pract.*, **214**: 1247–1252. <https://doi.org/10.1016/j.prp.2018.07.033>
- Huang, X.M., Yang, Z.J., Xie, Q., Zhang, Z.K., Zhang, H. and Ma, J.Y., 2019. Natural products for treating colorectal cancer: A mechanistic review. *Biomed. Pharmacother.*, **117**: 109142. <https://doi.org/10.1016/j.biopha.2019.109142>
- Ibrahim, B., McMahon, D.P., Hufsky, F., Beer, M., Deng, L., Le Mercier, P., Palmarini, M., Thiel, V. and Marz, M., 2018. A new era of virus bioinformatics. *Virus Res.*, **251**: 86–90. <https://doi.org/10.1016/j.virusres.2018.05.009>
- Ji, H., Wang, J., Guo, J., Li, Y., Lian, S., Guo, W., Yang, H., Kong, F., Zhen, L., Guo, L. and Liu, Y., 2016. Progress in the biological function of alpha-enolase. *Anim. Nutr.*, **2**: 12–17. <https://doi.org/10.1016/j.aninu.2016.02.005>
- Jiang, B., 2017. Aerobic glycolysis and high level of lactate in cancer metabolism and microenvironment. *Genes Dis.*, **4**: 25–27. <https://doi.org/10.1016/j.gendis.2017.02.003>
- Lai, X., Wang, M., Mcelyea, S.D., Sherman, S., House, M. and Korc, M., 2017. A microRNA signature in circulating exosomes is superior to exosomal glypican-1 levels for diagnosing pancreatic cancer. *Cancer Lett.*, **393**: 86–93. <https://doi.org/10.1016/j.canlet.2017.02.019>
- Li, Z., Yang, P. and Li, Z., 2014. The multifaceted regulation and functions of PKM2 in tumor progression. *Biochim. biophys. Acta Rev. Cancer*, **1846**: 285–296. <https://doi.org/10.1016/j.bbcan.2014.07.008>
- Liu, J., Jia, J., Wang, S., Zhang, J., Xian, S., Zheng, Z., Deng, L., Feng, Y., Zhang, Y. and Zhang, J., 2022. Prognostic ability of enhancer RNAs in metastasis of non-small cell lung cancer. *Molecules*, **27**: 4108. <https://doi.org/10.3390/molecules27134108>
- Lord, I. and Hall, N., 2019. Colorectal cancer: Features and investigation. *Medicine*, **47**: 400–404. <https://doi.org/10.1016/j.mpmed.2019.04.002>
- Maiers, M., Halagan, M., Gragert, L., Bashyal, P., Brelsford, J., Schneider, J., Lutsker, P. and Louzoun, Y., 2019. GRIMM: GRaph IMputation and matching for HLA genotypes. *Bioinformatics*, **35**: 3520–3523. <https://doi.org/10.1093/bioinformatics/btz050>
- Moutinho-Ribeiro, P., Costa-Moreira, P., Adem, B., Batista, I., Almeida, M., Barroca, H., Lopes, J., Carneiro, F., Melo, S.A. and Macedo, G., 2020. Exosomal glypican-1 for risk stratification of pancreatic cystic lesions: A case of pathological progression in the absence of any suspicious imaging finding. *Pancreatology*, **20**: 571–575. <https://doi.org/10.1016/j.pan.2020.01.015>
- Munir, J., Ngu, T.V., Ayudthaya, P.D.N. and Ryu, S., 2020. Downregulation of glypican-4 facilitates breast cancer progression by inducing cell migration and proliferation. *Biochem. biophys. Res. Commun.*, **526**: 91–97. <https://doi.org/10.1016/j.bbrc.2020.03.064>
- Pan, S., Niu, B., Tao, Y.N., Xia, G.L., Huang, L.F., Xu, J., Zhu, L. and Cao, H., 2021. Analysis of the relationship between preoperative serum tumor marker levels and tumor immune scores in patients with colorectal cancer based on nano-

- electrochemiluminescence. *Mater. Express*, **11**: 1476-1482. <https://doi.org/10.1166/mex.2021.2063>
- Ren, K.K., Ma, J.K., Zhou, B. Lin, X.Y., Hou, M.Y., Zhang, L., Yuan, L. and Ma, J., 2020. Identification of prognostic and immune-related lncRNAs in hepatocellular carcinoma based on the cancer genome atlas data mining. *Mater. Express*, **10**: 1189-1196. <https://doi.org/10.1166/mex.2020.1736>
- Song, H.X. and Miao, L., 2022. Clinical implications of miR-424 in diabetic retinopathy and its correlation with inflammatory cytokines. *Mater. Express*, **12**: 263-270. <https://doi.org/10.1166/mex.2022.2163>
- Taieb, J., André, T. and Auclin, E., 2019. Refining adjuvant therapy for non-metastatic colon cancer, new standards and perspectives. *Cancer Treat. Rev.*, **75**: 1–11. <https://doi.org/10.1016/j.ctrv.2019.02.002>
- Tao, J., Li, Y., Li, S. and Li, H.B., 2018. Plant foods for the prevention and management of colon cancer. *J. Funct. Fds.*, **42**: 95–110. <https://doi.org/10.1016/j.jff.2017.12.064>
- Yang, K., Yang, Y., Qi, C. and Ju, H., 2020. Effects of porcine STC-1 on cell metabolism and mitochondrial function. *Gen. comp. Endocrinol.*, **286**: 113298. <https://doi.org/10.1016/j.ygcen.2019.113298>
- Zhang, S., Li, J., Gao, H., Tong, Y., Li, P., Wang, Y., Du, L. and Wang, C., 2021. lncRNA profiles enable prognosis prediction and subtyping for esophageal squamous cell carcinoma. *Front. Cell Dev. Biol.*, **9**: 656554. <https://doi.org/10.3389/fcell.2021.656554>



Supplementary Material

Identification of Glycolysis-Related Gene Signature and Prediction on Prognosis of Colon Cancer

Qingsen Ran^{1,2}, Manjing Li³, Jiayin Han³, Lifang Wang³, Han Wang^{3,4}, Shaobo Liu^{5*} and Yanping Wang^{1*}

¹Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Science, Beijing 100700, China

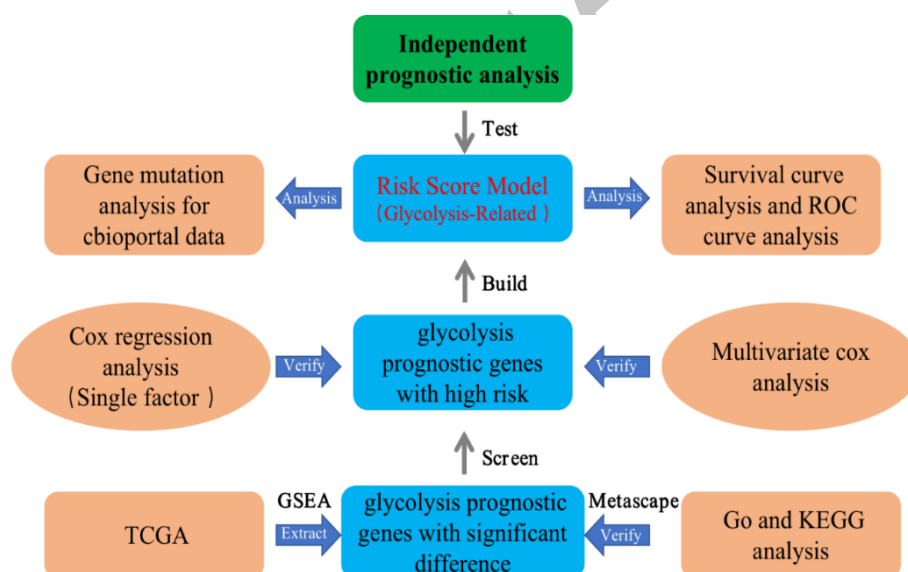
²Postdoctoral Research Station, China Academy of Chinese Medical Science, Beijing 100700, China

³Institute of Chinese Materia Medica, China Academy of Chinese Medical Science, Beijing 100700, China

⁴Affiliated Hospital of Changchun University of Chinese Medicine, Changchun 130021, China

⁵The Key Laboratory of Optimal Utilization of Natural Medicine Resources, School of Pharmaceutical Sciences, Guizhou Medical University, Guiyang 550025, China

Qingsen Ran and Manjing Li contributed equally to this work.



Supplimentary Fig. 1. The overall experimental design of this study.

* Corresponding author: liushaobo@gmc.edu.cn, ypwang@

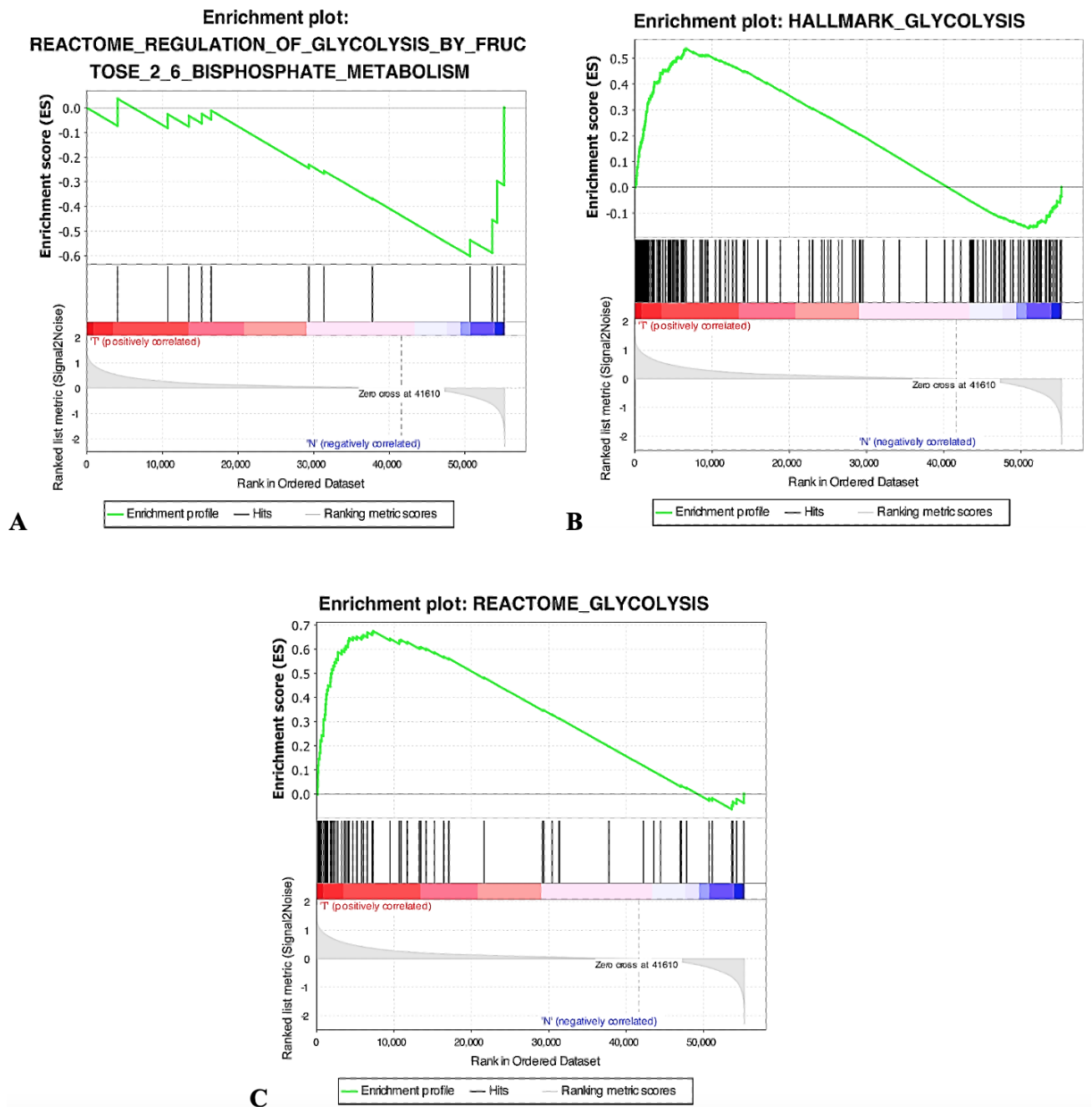
ircm.ac.cn

0030-9923/2023/0001-0001 \$ 9.00/0



Copyright 2023 by the authors. Licensee Zoological Society of Pakistan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



Supplementary Fig. 2. Significant enrichment analysis results of glycolysis metabolic.

A: GLYCOLYSIS_BY_FRUCTOSE_2_6_BISPHOSPHATE_METABOLISM; **B:** HALLMARK_GLYCOLYSIS; **C:** REACTOME_GLYCOLYSIS.

Notes: 41 normal and 470 tumor samples with colon cancer based on GSEA; NES: Normalized enrichment score according to requirements was greater than or equal to 1.0; Nominal p -value was less than or equal to 0.05; FDR: false discovery rate according to requirements was less than or equal to 0.25.