Identification and Validation of a Glycolysis Related Metabolic Reprogramming Gene Signature for Predicting Survival in Colon Cancer Patients

Gang Liu (✉ 20194155084@stu.suda.edu.cn)

Soochow University <https://orcid.org/0000-0003-3269-2470>

Xiaowang WU

Kunshan First People's Hospital

Jian Chen

Kunshan First People's Hospital

Primary research

Keywords: Colon cancer, Glycolysis, patients` prognosis, clinical feather, epithelial-to-mesenchymal transition

Posted Date: September 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-916034/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background

Colon cancer (CC) is one of the most common gastrointestinal malignant tumors with high mortality rate. Because of malignancy and easily metastasis feather, and limited treatments, the prognosis of CC remains poor. Glycolysis is a metabolic process of glucose in anoxic environments which is an important way to provide energy for tumor. The role of glycolysis in CC largely remains unknown and is necessary to be explored.

Method

In our study, we analyzed glycolysis related genes expression in CC, patients gene expression and corresponding clinical data were downloaded from GEO dataset, glycolysis related genes sets were collected from Msigdb. Through COX regression analysis, prognosis model based on glycolysis-related genes was established. The efficacy of gene model was tested by Survival analysis, ROC analysis and PCA analysis. Furthermore, the relationship between risk scores and clinical characteristic was researched.

Results

Our findings identified 13 glycolysis related genes (NUP107, SEC13, ALDH7A1, ALG1, CHPF, FAM162A, FBP2, GALK1, IDH1, TGFA, VLDLR, XYLT2 and OGDHL) consisted prognostic prediction model with relative high accuracy. The relationship between prediction model and clinical feathers were specifically studied, results showed age > 65years, TNM III-IV, T3-4, N1-3, M1 and high-risk score were independent prognostic risk factors with poorer prognosis. Finally, model genes were significantly expressed and EMT were activated in CC patients.

Conclusion

This study provided a new aspect to advance our understanding in the potential mechanism of glycolysis in CC.

Background

Colon cancer (CC) is third of gastrointestinal malignant tumor which derived from glandular epithelial cells with high mortality and increased incidence[1]. Despite the improvement of diagnosis and treatment for colon cancer in recent years, including surgery, radiotherapy, chemotherapy, neoadjuvant chemotherapy and immunotherapy, the 5-year relative survival rate remains poor, distant metastasis is

the main causes of colon cancer-associated death[2, 3]. Therefore, explore potential useful diagnosis and prognostic biomarkers of colon cancer patients are urgent needed.

Glycometabolism is a vital pathway in the living cells which is far different from the normal cells; The transformation of glucose utilization that from tumor cells oxidative phosphorylation to glycolysis is now regard as a major feature of cancer. This kind of energy metabolism change is regulated by gene changes and tumor microenvironment pressure, thus maintaining cell higher proliferation rate and resist to cells[4]. Activation of glycolysis can affect other phenotypic changes, for example glycolysis has a closely relationship with epithelial-to-mesenchymal transition (EMT)[5], Zhao et al reported that actived glycolysis promotes the pancreatic cancer cells stemness and epithelial to mesenchymal transition [6]. Hence, understanding the underlying interplay between glycolysis and caner progression is a critical goal in cancer research which can clear cancer pathogenesis and facilitate effective therapeutic approaches development.

In our study, GSEA of glycolysis-related pathway were conducted with 566 colon cancer and 19 normal control array data from GEO to identify different expressed glycolysis genes in CC patients. Then, Cox regression split patient into low- and high- risk groups according to glycolysis-related expression and risk score. A signature consisting of 13 genes were constructed and was closely related to poor patients' OS and high EMT environment. Risk score could serve as an independent pathological factor. Finally, the expression of the 13 signature genes were validated in our patient samples. Overall, our data uncovered a strong association between glycolysis metabolism and clinical prognosis in colon cancer.

Materials And Methods

Database download

The Affymetrix Human Genome U133 Plus 2.0 Array gene expression profile and corresponding clinical information data of 566 CC patients and 19 normal control tissue were downloaded from the GEO database (GSE143985). The clinical information includes survival time, survival status, sex, age, TNM staging, clinical T stage, clinical N stage and clinical M stage.

De Glycolysis Related Genes Identification

A total of 301 mRNAs of the 6 glycolysis related gene sets were selected from GSEA database (WP_COMPUTATIONAL_MODEL_OF_AEROBIC_GLYCOLYSIS, WP_GLYCOLYSIS_IN_SENESCENCE, REACTOME_REGULATION_OF_GLYCOLYSIS_BY_FRUCTOSE_2_6_BISPHOSPHATE_METABOLISM, REACTOME_GLYCOLYSIS, HALLMARK_GLYCOLYSIS, GO_GLYCOLYTIC_PROCESS). Then, the prognostic genes model was constructed based on the 217 differentially expressed glycolysis-related genes. We used the Wilcoxon to filtered differential expressed genes. The cut-offs were set as $\text{Log}_2(\text{fold change}) > 1$ and $\text{pValue} < 0.05$. Heatmap were conducted by Pheatmap package.

Construction Of The Glycolysis Related Genes Risk Model

Based on the 217 differently expressed glycolysis related genes of GSE143985, uniCOX and multiCOX analysis was conducted for identifying model genes. The risk score was obtained from the results of the multiplication between the Cox coefficient of multiCOX analysis and gene expression value. The calculating formula was listed: Risk score = (Cox coefficient of model gene A* A gene expression + Cox coefficient of model gene B* B gene expression+.....).

Then, all patients were spilited into low- and high-risk subgroups on the basis of median risk score. Cbioportal (<http://www.cbioportal.org/>) were used to screen model gene mutation statue and gene mutation site. Survival and survminer package were used for survival analysis. The ROC analysis was performed by survival ROC package. AUC value > 0.65 was regarded as an acceptable cutoff value. Risk score plots, PCA, survival status plots were also conducted to assume the prediction model between the low- and high-risk groups.

Independent Prognostic Value Of Risk Score

To assess the model prognostic value, we used univariate and multivariate Cox regression analyses for prognostic factors. Factors (Age, Garden, TNM stage, risk score) of which univariate and multivariate Cox regression analyses p value were all less than 0.05 were considered as an independent factor.

Clinical Application Of Gene Model

To exam our model efficiency in CC, we analyzed the correlation between the gene model and the clinical feather (age, gender, clinical TNM stage, clinical T stage, clinical N stage, and clinical M stage). All patients were clarified into two subgroups based on age (> 65and < = 65 years old), gender (female and male), TNM stage (TNM stage I-II and TNM stage III-IV), clinical T stage (T1–2 and T3–4), clinical M stage (M0 and M1), and clinical N stage (N0 and N1–3). Survival analyses were carried out between the two groups.

Function Enrichment Analyses Between Low- And High- Risk Group

Gene ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were conducted by org.Hs.eg.db R package and clusterProfiler R package. value. p.adjust < 0.05 were identified as significantly enriched.

Model Gene Validation By Western Blot Analysis

Model gene validation by Western blot analysis

CC patients' tissues and control were collected from Affiliated Kunshan Hospital of Jiangsu University. Patients' tissues were preserved in liquid nitrogen and lysed with RIPA Lysis Buffer (Strong) (MedChemExpress, China). BCA Protein Quantification Kit (Vazyme, China) was used to measure protein concentration. Proteins were transferred to nitrocellulose membrane (Millipore, USA) after separated by 10% SDS-PAGE. Then, incubated with primary antibodies: NUP107/SEC13/ALDH7A1/ALG1/CHPF/FAM162A/FBP2/GALK1/IDH1/TGFA/VLDLR/XYLT2/OGDHL (Abcam, USA, 1:1000), GAPDH (Abcam, USA, 1:3000). The membrane was washed and was incubated with HRP-labeled goat anti- rabbit IgG (1:5000, Sigma). Signals were visualized with BIO-RADXR.

Results

CC Patients Data Basic information

Overall, 566 colon cancer and 19 normal controls data on Affymetrix Human Genome U133 Plus 2.0 Array platform was downloaded from GEO datasets. The corresponding clinical characteristics data of the samples were shown in Table 1. The clinical features include gender, age, survival status, survival time, clinical tumor-node-metastasis (TNM) staging, clinical T stage, clinical N stage and clinical M stage.

Table 1
Clinical information of CC
patients from TCGA.

Vriables	Patients (566)
Fustat	
Alive	370
Death	191
unknow	5
Age (year)	
<=65	222
>65	344
Gender	
Male	256
Female	310
TNM Stage	
I	33
II	264
III	205
IV	60
0	4
T stage	
T1	11
T2	45
T3	367
T4	119
Tis	3
T0	1
unknow	20
M stage	
M0	482

Variables	Patients (566)
MI	61
unknow	23
N stage	
N0	302
N1	134
N2	98
N3	6
unknow	26

Filtered significant differences glycolysis gene between the CC samples and normal control

Six glycolysis-related genesets were collected from the MsigDB, including REACTOME_GLYCOLYSIS, WP_COMPUTATIONAL_MODEL_OF_AEROBIC_GLYCOLYSIS, WP_GLYCOLYSIS_IN_SENESCENCE, REACTOME_REGULATION_OF_GLYCOLYSIS_BY_FRUCTOSE_2_6_BISPHOSPHATE_METABOLISM, HALLMARK_GLYCOLYSIS and GO_GLYCOLYTIC_PROCESS. A total of 301 glycolysis-related gene in the 6 genesets were selected for the additional research, and 217 glycolysis-related genes were significant differentially expressed in CC samples compared with normal controls (Fig. 1).

Construction of the Glycolysis-Related Genes prognostic prediction model in CC Patients.

To construct the novel biomarkers model for predicting the CC patients' prognosis, uniCox and multiCox regression analysis on the differentially expressed glycolysis-related genes were performed. 13 risk genes (NUP107, SEC13, ALDH7A1, ALG1, CHPF, FAM162A, FBP2, GALK1, IDH1, TGFA, VLDLR, XYLT2, OGDHL) were identified ($P < 0.05$, Table 2). Then, prognostic gene model based on the 13 glycolysis-related genes divided patients into low- and high-risk group was established as follows: Risk score = $(-0.29 \times \text{NUP107 Expression} + -0.35 \times \text{SEC13 Expression} + 0.30 \times \text{ALDH7A1 Expression} + -0.29 \times \text{ALG1 Expression} + 0.45 \times \text{CHPF Expression} + -0.46 \times \text{FAM162A Expression} + 0.91 \times \text{FBP2 Expression} + -0.65 \times \text{GALK1 Expression} + -0.34 \times \text{IDH1 Expression} + 0.36 \times \text{TGFA Expression} + 0.18 \times \text{VLDLR Expression} + -1.28 \times \text{XYLT2 Expression} + -0.41 \times \text{OGDHL Expression})$. Then, the 13 gene alterations statue in CC tissue were analyzed, results showed that alterations of NUP107, SEC13, ALDH7A1, ALG1, CHPF, FAM162A, FBP2, GALK1, IDH1, TGFA, VLDLR, XYLT2, OGDHL were 3%, 0.8%, 1.9%, 1%, 1.9%, 0.2%, 2.3%, 1.5%, 1.7%, 0.4%, 4%, 4% and 4%, respectively (Fig. 2A). Many mutation occurred in gene domains, and the specific mutation sites were presented in Fig. 2B. Finally, these 13 genes expression in the CC patients and normal tissues was further analyzed, NUP107, SEC13, ALDH7A1, ALG1, CHPF, GALK1, XYLT2 and OGDHL were highly expressed in CC, however, FAM162A, FBP2, IDH1, TGFA and VLDLR showed downregulated ($P < 0.05$, Fig. 2C).

Table 2
13 genes were selected via uniCox regression analysis.

ID	HR	HR.95L	HR.95H	coxPvalue
NUP107	0.64	0.49	0.85	0.00
SEC13	0.64	0.43	0.95	0.03
ALDH7A1	0.70	0.50	0.97	0.03
ALG1	0.60	0.44	0.82	0.00
CHPF	1.38	1.03	1.85	0.03
FAM162A	0.59	0.40	0.88	0.01
FBP2	2.41	1.33	4.38	0.00
GALK1	0.38	0.24	0.60	0.00
IDH1	0.54	0.36	0.80	0.00
TGFA	1.75	1.26	2.43	0.00
VLDLR	1.17	1.02	1.35	0.03
XYLT2	0.28	0.13	0.61	0.00
OGDHL	0.56	0.41	0.75	0.00

Efficacy of Risk Score in GC Patients.

Based on the gene model, 280 CC patients and 281 CC were clarified into the high- and low-risk group by the median risk score (Fig. 3A). KM analysis were carried out between low- and high-risk group, high-risk group had obviously poorer prognosis ($P < 0.05$, Fig. 3B). ROC analysis was carried out, AUC were calculated which was 0.716 showed good prediction effect (Fig. 3C). The expression of the 13 genes were calculated, TGFA FBP2, CHPF, VLDLR were significant higher expressed in high-risk group, and NUP107, OGDHL, SEC13, IDH1, GALK1, FAM162A, ALG1, XYLT2 ALDH7A1 were downregulated ($P < 0.05$, Fig. 3D). Risk plot indicated that high-risk group was closely related with poor prognosis (Fig. 3E). PCA presented that high- and low-risk groups can be significantly distinguished via our model reduce the dimension of multiple genes expression (Fig. 3F). All the analysis results indicated that our gene model based on risk score had a good efficacy.

The Independent risk factor of Risk Score in CC patients and the its' relationship with patients Clinical Characteristics.

Univariate and multivariate independent prognostic analysis were carried out to identify independent prognostic factor including age, gender, TNM staging and risk score. Data showed that age, TNM staging and risk score were independent prognostic factors in CC patients, which all positively correlated to the poor survival prognosis. ($P < 0.05$, Fig. 4A-B). For more accuracy analysis the relationship between risk

factors and patient OS, firstly, a series Kaplan-Meier curve analysis were carried out. Results reveal that age > 65years, TNM III-IV, T3-4, N1-3 and M1 were positive with poor OS ($P < 0.05$, Fig. 4C). Then, the relationship between patients Clinical Characteristics and patient OS in low- and high-risk group were performed, the high-risk patients were positively associated with poor overall survival in the age $\leq$ 65years, age > 65years, Female, Male, TNM I-II, TNM III -IV, T1-2, T3-4, M0, M1, N0, N1-3 subgroups ($P < 0.05$, Fig. 4D). These results once again prove the reliability of our glycolysis-related genes model in CC.

Enrichment analysis of different genes between low- and high-risk group

9656 different expressed genes were screened out between low- and high-risk group. Then, GO and KEGG Enrichment analysis were conducted. Results of GO analysis indicated that genes were enriched on mitochondrion and ATP related pathway ($P < 0.05$, Fig. 5A). The results of KEGG were enriched on Carbon and Oxidative metabolism ($P < 0.05$, Fig. 5B). These results show that the gene model can correctly distinguish high-risk group from low-risk group, and verify the correctness of our model.

Role Of Emt In High Glycolysis Tumor Environment

Researches had demonstrated that glycolysis could promote EMT in cancers. For preliminary explore the role of EMT in CC, the expression of the biomarker of EMT (SNAI1, SNAI2, TWIST1, TWIST2 and ZEB2) were analysis in our study. SNAI1 and TWIST1 showed significantly higher expression in CC tissue ($P < 0.05$, Fig. 6A). Whereas, and SNAI1, SNAI2, TWIST1, TWIST2 and ZEB2 were all upregulated in high-risk group ($P < 0.05$, Fig. 6B). It suggested that EMT may play a role in high glycolysis tumor environment of CC patients. The relationship between glycolysis and EMT in CC need further deeper research.

Validation Of Model Genes Expression In Cc Patients' Samples

To lay the foundation for the next step of basic research, we validated the expression of all the model genes WB. Our results showed that NUP107, SEC13, ALDH7A1, ALG1, CHPF, GALK1, XYLT2 and OGDHL were highly expressed, whereas, FAM162A, FBP2, IDH1, TGFA and VLDLR were lowly expressed in CC patients ($P < 0.05$, Fig. 7). These gene expression level was consistent with the genes' expression in GSE143985.

Discussion

Colon cancer is a kind of common carcinomas in gastrointestinal tumor which jeopardizes human health and responsible for 1,148,515 new cases and 576,858 deaths in 2020 worldwide[7]. In the recent years, great strides have been occurred in the field of treatment in CC patients, including surgery, radiation, chemotherapy, adjuvant chemotherapy and immunotherapy, but the mortality of CC still remains high [7, 8]. Thus, it is necessary to discover novel biomarkers as potential therapeutic targets to improve the

prognosis of CC. Glycolysis is highly activated in tumor cells, more and more evidence has revealed that glycolysis metabolism reprogramming is closely related to the occurrence and development of tumor. [9]

In this study, we collected 6 glycolysis related gene sets from GSEA database for a total of 301 mRNAs expressions in 566 CC patients were selected for further analysis. Then, 217 differentially expressed glycolysis-related genes were screened out for prognostic genes model construction. After Cox regression analyses, a signature of 13 (NUP107, SEC13, ALDH7A1, ALG1, CHPF, FAM162A, FBP2, GALK1, IDH1, TGFA, VLDLR, XYLT2, OGDHL) glycolysis-related genes which all were independent risk factor with OS was established between CC patients and normal tissues. The mentioned signature could be considered as novel potential therapeutic targets. Reports reported that NUP37 promoted lung cancer cell proliferation and inhibited apoptosis[10], overexpressed TGFA, HPF induce cell proliferation and invasion in lung cancer [11, 12] Downregulated ALDH7A1 increased hepatocellular and renal clear carcinomas cell migration and invasiveness [13]. FBP2 over expression inhibits sarcoma cell growth [14]. VLDLR over-expression inhibited CC cells proliferation and migration[15]. Silencing of OGDHL contributed to hepatocellular carcinomas patients' survival by regulating glutamine metabolic pathways[16]. IDH1 mutation promotes glioma cells proliferation and migration of via EMT [17]. However, the specific role of SEC13, ALG1, FAM162A, GALK1 and XYLT2 in cancer have not been clear at present.

Prediction capability of the model was assessed though PCA, Kaplan-Meier curve analysis and ROC. Clinical variables and the risk score were also analyzed. PCA results showed our model specifically clarified samples into low- and high-risk group base on risk score. Kaplan-Meier curve analysis reveals high-risk score was related to poorer patients' prognosis. The expression of the signature genes was validated by PCR and WB which results were consentient with the genes' expression in GSE143985. These results suggested that our model based on risk score might be useful for predicting CC patients' prognosis. Our gene expression signature might serve as a workable tool for classification of CC patients and might provide the basis for personalized treatment.

For clinical variables, age > 65years, TNM III-IV, T3-4, N1-3, M1 and high-risk score were identified as independent prognostic risk factors. For evaluating the clinical efficacy of the model, the relationships between risk score and clinical feathers were analyzed. We found that high risk score in the age < = 65years, age > 65years, Female, Male, TNM I-II, TNM III -IV, T1-2, T3-4, M0, M1, N0, N1-3 subgroups was highly correlated with poor patiebt's OS. These results once again prove the gene model reliability in CC.

Study reported that EMT promotes CC cell invasion and migration, meanwhile involved in invasion and metastasis of colon cancer patients [18]. However, glycolysis also has a closely relationship with EMT in cancers[5]. In our study, we found that the risk score was positively associated with the expression of the EMT biomarker (SNAI1, SNAI2, TWIST1, TWIST2 and ZEB2). These results might suggest that high glycolysis metabolism is accompanied by changes in other regulatory pathways in tumor microenvironment.

Conclusions

To sum up, we identified and validated a 13 glycolysis related genes model of CC, including NUP107, SEC13, ALDH7A1, ALG1, CHPF, FAM162A, FBP2, GALK1, IDH1, TGFA, VLDLR, XYLT2 and OGDHL. This model could be positively related to CC poorest prognosis and EMT activation, which was a prognosis prediction instrument and provide novel aspects for CC therapy.

Declarations

Acknowledgements

The authors are grateful to Lei Shi for expert technical assistance and Zijiang Yang for statistical advices.

Authors' contributions

GZ raised the concept and designed this study. XW conducted the experiments and analyzed the data. JC wrote the manuscript and SZ revised the manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate

This study was approved by the Clinical Research Ethic Committee of Affiliated Kunshan Hospital of Jiangsu University. All patients were provided full information regarding the risks and benefits of the treatments, and all of the patients provided signed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Harada S, Morlote D. Molecular Pathology of Colorectal Cancer. *Adv Anat Pathol*. 2020; 27(1):20-6.
2. Mumprecht V, Detmar M. Lymphangiogenesis and cancer metastasis. *J Cell Mol Med*. 2009; 13(8A):1405-16.

3. Kaminski MF, Robertson DJ, Senore C, Rex DK. Optimizing the Quality of Colorectal Cancer Screening Worldwide. *Gastroenterology*. 2020; 158(2):404-17.
4. Abbaszadeh Z, Cesmeli S, Biray Avci C. Crucial players in glycolysis: Cancer progress. *Gene*. 2020; 726:144158.
5. Georgakopoulos-Soares I, Chartoumpekis DV, Kyriazopoulou V, Zaravinos A. EMT Factors and Metabolic Pathways in Cancer. *Front Oncol*. 2020; 10:499.
6. Zhao H, Duan Q, Zhang Z, Li H, Wu H, Shen Q, Wang C, Yin T. Up-regulation of glycolysis promotes the stemness and EMT phenotypes in gemcitabine-resistant pancreatic cancer cells. *J Cell Mol Med*. 2017; 21(9):2055-67.
7. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021; 71(3):209-49.
8. Labianca R, Beretta GD, Kildani B, Milesi L, Merlin F, Mosconi S, Pessi MA, Prochilo T, Quadri A, Gatta G et al: Colon cancer. *Crit Rev Oncol Hematol*. 2010; 74(2):106-33.
9. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab*. 2016; 23(1):27-47.
10. Huang L, Wang T, Wang F, Hu X, Zhan G, Jin X, Zhang L, Li Y. NUP37 silencing induces inhibition of cell proliferation, G1 phase cell cycle arrest and apoptosis in non-small cell lung cancer cells. *Pathol Res Pract*, 2020; 216(3):152836.
11. Cao C, Liu Y, Wang Q, Zhao J, Shi M, Zheng J: Expression of CHPF modulates cell proliferation and invasion in lung cancer. *Braz J Med Biol Res*. 2020; 53(5):e9021.
12. Wu H, Liu Y, Shu XO, Cai Q. MiR-374a suppresses lung adenocarcinoma cell proliferation and invasion by targeting TGFA gene expression. *Carcinogenesis*. 2016; 37(6):567-575.
13. Andrejeva D, Kugler JM, Nguyen HT, Malmendal A, Holm ML, Toft BG, Loya AC, Cohen SM. Metabolic control of PPAR activity by aldehyde dehydrogenase regulates invasive cell behavior and predicts survival in hepatocellular and renal clear cell carcinoma. *BMC Cancer*. 2018; 18(1):1180.
14. Huangyang P, Li F, Lee P, Nissim I, Weljie AM, Mancuso A, Li B, Keith B, Yoon SS, Simon MC. Fructose-1,6-Bisphosphatase 2 Inhibits Sarcoma Progression by Restraining Mitochondrial Biogenesis. *Cell Metab*. 2020; 31(5):1032.
15. Kim BK, Yoo HI, Lee AR, Choi K, Yoon SK. Decreased expression of VLDLR is inversely correlated with miR-200c in human colorectal cancer. *Mol Carcinog*. 2017; 56(6):1620-9.
16. Dai W, Xu L, Yu X, Zhang G, Guo H, Liu H, Song G, Weng S, Dong L, Zhu J et al: OGDHL silencing promotes hepatocellular carcinoma by reprogramming glutamine metabolism. *J Hepatol*. 2020; 72(5):909-23.
17. Lu J, Li D, Zeng Y, Wang H, Feng W, Qi S, Yu L: IDH1 mutation promotes proliferation and migration of glioma cells via EMT induction. *J BUON*. 2019; 24(6):2458-64.

18. Todosi AM, Gavrilescu MM, Anitei GM, Filip B, Scripcariu V: Colon cancer at the molecular level—usefulness of epithelial-mesenchymal transition analysis. Rev Med Chir Soc Med Nat Iasi. 2012; 116(4):1106-11.

Figures

Figure 1
Significantly different expressed glycolysis-related genes between CC and normal controls.

Figure 2

model genes mutation and expression level in CC. (A) Mutation frequency of model genes. (B) Model gene specific mutation domains. (C) Model gene expression between CC and normal control.

Figure 3

Text efficacy of gene model in CC Patients. (A) CC Patients distribution according to median risk score. (B) KM analysis between low- and high-risk group. (C) ROC analysis of risk score. (D) Model genes expression in low- and high- risk group. (E) The relationship between patients' survival statue and risk score. (F) PCA analysis between low- and high-risk group.

Figure 4

Analysis of risk factors. Univariate Cox (A) and Multivariate Cox regression analysis (B) of the relationship between risk core and other clinical features. (C) KM analysis of clinical features, included age, gender, TNM stage, T stage, N stage, M stage. (D) KM analysis of clinical features in low- and high-risk subgroup.

Figure 5

Enrichment analysis between low- and high-risk group. Go (A) and KEGG (B) Enrichment analysis.

Figure 6

The expression of EMT biomarkers. (A) SNAI1 and TWIST1 expression in CC patients and normal control. (B) SNAI1, SNAI2, TWIST1, TWIST2 and ZEB2 expression in low and high-risk group.

Figure 7

Model genes expression was confirmed by western blot.