Comprehensive Analysis of the Association Between the Glycolysis Prognostic Risk Signature and Immune Infiltration in Endometrial Cancer

Xiao Yang

Peking University People's Hospital <https://orcid.org/0000-0002-5172-8267>

Xingchen Li

Peking University People's Hospital

Boqiang Shen

Peking University People's Hospital

Jingyi Zhou

Peking University People's Hospital

Yuan Cheng

Peking University People's Hospital

Yuanyuan Jia

Peking University People's Hospital

Jianliu Wang (✉ wjianliu1203@163.com)

Peking University People's Hospital

Primary research

Keywords: endometrial cancer, glycolysis, prognostic signature, WGCNA, bioinformatics, immune infiltration

Posted Date: June 30th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-37313/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Glycolysis is the primary mechanism of energy metabolism in tumor cells. The purpose of this study is to develop a glycolysis-related risk signature for endometrial cancer and analyze its relationship with immune function.

Methods: The mRNA expression profiling and clinical information of endometrial cancer were downloaded from TCGA database. GSEA was performed to screen out gene sets related to glycolysis, and the R software was used to screen DEGs. Univariate Cox and LASSO regression analyses were used to construct a glycolysis-related prognostic signature for endometrial cancer. WGCNA were performed to identify the potential mechanisms associated with the prognostic signature. Immune scores, stromal scores and ESTIMATE scores were calculated based on the R “ESTIMATE” algorithm. The “CIBERSORT” algorithm was used to evaluate the infiltration of immune cells. ROC curve, nomogram, gene alteration, coexpression analyses and PPIs were also performed.

Results: We identified a glycolysis-related gene signature (PFKM, PSMC4, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI) for predicting the prognosis of endometrial cancer. Based on this gene signature, the patients were divided into high- and low-risk subgroups. The overall survival of patients in the high-risk subgroup was significantly lower than that of the low-risk group. Multivariate Cox regression analysis results showed that the ten-gene signature was an independent prognostic factor for endometrial cancer. The ROC curve confirmed the accuracy of the prognostic signature (AUC=0.730). The immune scores and ESTIMATE scores of patients in the high-risk subgroup were lower than those of the low-risk subgroup, while the low immune score and ESTIMATE scores were closely related to the poor prognosis of patients. The high-risk group was associated with lower cellular immune infiltration of resting dendritic cells, resting memory T cells and Tregs, while the overall survival rate of resting dendritic cells and Tregs in the low-proportion group was significantly lower than that in the high-proportion group.

Conclusions: we constructed a glycolysis-related ten-gene signature to predict the survival and prognosis of endometrial cancer. Our findings may help to elucidate the mechanism of glycolysis and provide new ideas for targeted therapy and prognosis of endometrial cancer.

Background

Endometrial cancer is one of the most common gynecological malignancies. The incidence rate of this cancer has been increasing significantly and younger worldwide. The latest cancer statistics of the American Cancer Society showed that the number of new cases in the United States increased by 63230, with 11350 deaths being reported, and the incidence rate ranked fourth in female malignant tumors and sixth in deaths in 2018. It is estimated that the number of endometrial cancer patients in the United States will increase to 42.13 per 100000 cases in 2030 [1-2]. Although early diagnosis, hormone therapy, surgery, radiotherapy and chemotherapy can significantly improve the treatment effect of patients, the treatment for early patients with the need for fertility preservation, late relapse and relapse is still limited.

An analysis of clinical data from 276 patients with endometrial cancer showed that the 5-year disease-free survival rate and 5-year overall survival rate of patients with endometrial cancer were 82.3% and 81%, respectively, and the tumor recurrence rate and tumor-related mortality rate were 14.5% and 15.9%, respectively [3]. Therefore, the key to improving the therapeutic effect of tumors is to explore the biological characteristics of tumor cells, which are ubiquitous but different from normal cells, and to carry out specific interventions based on this characteristic.

Metabolic reprogramming is one of the most important characteristics of tumor cells. Compared with normal cells, tumor cells absorb more glucose with higher efficiency to produce energy and meet the needs of rapid growth. Even when the supply is sufficient, approximately 80% of glucose in tumor cells primarily produces ATP through aerobic glycolysis accompanied by a large amount of lactate, rather than the oxidative phosphorylation pathway, known as the “Warburg effect” [4]. An increasing number of studies have shown that the glycometabolism reprogramming of tumor cells is closely related to the occurrence, progression and chemotherapy resistance of tumors [5]. The glycolytic ability of endometrial cancer cells increased significantly, and the imbalance of multiple gene expression could promote the progression of endometrial cancer by promoting glycolysis [6]. Therefore, the screening of glycolytic genes related to the development of endometrial cancer is of considerable significance for the evaluation of prognosis and targeted therapy.

The tumor microenvironment consists of the extracellular matrix, soluble molecules and tumor stromal cells and is the cellular environment of tumor cells. Once the tumor microenvironment is formed, many immune cells, such as T cells, myelogenous inhibitory cells, and macrophages, are chemotactic to this point, forming the tumor microenvironment [7]. Immune cells and stromal cells are two primary types of nontumor components in the tumor microenvironment and have been proposed to be of considerable importance for tumor diagnosis and prognosis evaluation. Estimation of stromal and immune cells in malignant tumors using expression data (ESTIMATE) based on single sample GSEA is a tool for predicting tumor purity and the presence of infiltration stromal/immune cells in tumor tissues by analyzing gene expression data [8]. Based on the diversity of the immune microenvironment, research has created a set of methods to score the immune microenvironment to analyze the relationship between the immune microenvironment and prognosis of patients with colorectal cancer metastasis. The results showed that the metastasis with the smallest number of immune cells entering represented the worst immune microenvironment; therefore, tumor immune escape was most likely to occur [9]. It was reported that the infiltration of immune cells in tumors is closely related to the clinical outcome. The infiltrated immune cells are most likely to be used as drug targets to improve the survival rate of cancer patients. CIBERSORT is an analytical tool developed by [Newman et al](#) to provide an estimation of the abundances of member cell types in a mixed cell population by analyzing gene expression data [10]. Therefore, the establishment of an immune scoring model and analysis of tumor-infiltrating immune cells can effectively evaluate the prognosis of patients and provide guidance for tumor targeted therapy.

Many studies have reported the relationship between glycolysis and tumor immunity [11-13]. It has been reported that enhanced glycolysis of tumor cells could become the main obstacle of targeted treatment

of tumor immune cells by affecting the infiltration of immune cells in the tumor microenvironment, while interference with glycolysis of tumor cells could enhance the effective infiltration of antitumor immune cells [14]. It is suggested that further study of the relationship between glycolysis and tumor immunity is of strong significance for the effective targeted treatment of tumors. However, studies investigating glycolysis genes and their prognostic value and relationship with immune function in patients with endometrial cancer are limited. In this study, we used GSEA to screen out gene sets related to glycolysis, downloaded the mRNA expression profiling of endometrial cancer from The Cancer Genome Atlas (TCGA), screened glycolysis genes closely related to the prognosis of endometrial cancer based on bioinformatics, and analyzed the immune scores and immune cell infiltration related to the glycolysis-related gene signature. Our results may provide a reference for further research investigating endometrial cancer and individual treatment.

Materials And Methods

Data collection and preparation

We downloaded mRNA expression profiling (FPKM format) of endometrial cancer from TCGA database, including 552 endometrial cancer and 35 normal samples (<https://portal.gdc.cancer.gov/>). The corresponding clinicopathological information, including age, tumor stage, grade, metastasis, lymph node metastases, survival time and survival status, were downloaded from the TCGA data portal.

Screening glycolysis related genes and functional enrichment analysis

We downloaded the gene sets related to glycolysis from the Molecular Signatures Database (MSigDB) of the Gene set enrichment analysis (GSEA) website, including HALLMARK_GLYCOLYSIS, KEGG_GLYCOLYSIS_GLUONEOGENESIS, and REACTOME_GLYCOLYSIS (<http://software.broadinstitute.org/gsea/index.jsp>). Perl script was used to extract the expression matrix of glycolysis-related genes. The R "limma" package was used to screen the differentially expressed genes (DEGs), and the screening conditions were $|\log FC| > 0.5$ and false discovery rate (FDR) < 0.05 . Volcano plots and heatmap clustering were conducted using R software. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using the R "ClusterProfiler" package [15]. GSEA software was used to analyze the significantly enriched upregulated and downregulated signal pathways (using FDR < 0.05 as the cut-off criterion).

Construction of the prognostic glycolysis-related gene signature

A univariate Cox regression model was used to evaluate the relationship between glycolysis-related DEGs and the overall survival rate of endometrial cancer patients. $P < 0.01$ was considered to be statistically significant. According to the hazard ratio (HR) value, the genes related to overall survival were identified and further divided into protective genes ($0 < HR < 1$) and risk genes ($HR > 1$). LASSO regression was applied to select prognosis-associated genes and to develop the prognostic gene signature [16]. The risk score of the gene signature = (coef 1 × expression of gene 1) + (coef 2 × expression of gene 2) + (coef 3 × expression of

gene 3)+...+(coef n ×expression of gene n). Based on the risk score of this gene signature, the patients were divided into high- and low-risk subgroups for subsequent study. The overall survival rates of patients in the high- and low-risk subgroups were analyzed by using the R “survival” and “survminer” packages, and Kaplan-Meier (K-M) survival curves were drawn. The risk curve and survival state diagram were drawn by the R software package. Univariate and multivariate Cox regression analyses were used to analyze the prognostic factors of endometrial carcinoma, and receiver operating characteristic (ROC) curves were drawn by the R “survivalROC” package. The nomogram was constructed using R software to integrate multiple prediction indicators to show the relationship between the variables in the prediction model based on multivariate Cox regression analyses [17]. The expression of glycolysis-related genes in the gene signature was further validated at the protein level (The Human Protein Atlas database: <http://www.proteinatlas.org>).

A gene co-expression network was built by the WGCNA

WGCNA analysis uses the weighted value of correlation coefficient, which makes the connection between genes in the network obey the scale-free network distribution. This algorithm has more biological significance. We conducted the weighted correlation network analysis (WGCNA) to identify the potential mechanisms associated with the risk model. We first filtered out the best soft threshold by “WGCNA” R package to maintain sufficient connectivity and keep the gene network close to the scale-free topology. Secondly, we performed the analysis module associated with the risk model and other clinical factors. Furthermore, the key gene model was identified from the WGCNA analysis. Then we used the metascape online website (<https://metascape.org/gp/index.html>) to perform the GO and KEGG analysis about the key gene model related with the risk model (P Value cutoff: 0.01).

Estimation of immune and stromal scores related to gene signature

We used the “ESTIMATE” algorithm to calculate the immune scores (which capture the presence of stroma in tumor tissue), stroma scores (which capture the infiltration of immune cells in tumor tissue) and estimate scores (which infer tumor purity). According to the immune scores, stroma scores and ESTIMATE scores, the patients were divided into high- and low-score subgroups by using the median value as the threshold. The overall survival rate of patients in the low- and high-score subgroups was analyzed by the R “survival” package. The immune scores and stromal scores of patients in the high-risk and low-risk subgroups were also calculated by ESTIMATE, and the immune scores and stromal scores of the high-risk subgroup and low-risk subgroup were compared by the Wilcoxon test; P -values < 0.05 were considered to be significant.

Validation of the correlation of tumor-infiltrating immune cells

The mRNA expression profiling of endometrial cancer downloaded from TCGA was corrected to obtain suitable input files for the CIBERSORT software. Then, the gene matrix was transformed into a matrix of immune cells based on CIBERSORT software. The proportion of twenty-two kinds of immune cells was calculated by the deconvolution method of CIBERSORT software. The difference in tumor-infiltrating

immune cells between the high-risk subgroup and low-risk subgroup was screened by the R “limma” package, and the screening condition was $P < 0.05$. The cor.test in R was used to calculate and test the correlation coefficient between the glycolysis-related gene signature and immune cell infiltration, and the correlation graph was drawn by the R “ggcorrplot” package.

cBioPortal

The cBioPortal website (<https://www.cbioportal.org/>) developed by Memorial Sloan Kettering Cancer Center (MSK) is a comprehensive open network platform based on the TCGA database that integrates data mining, data integration and visualization [18]. The genetic alterations of ten glycolysis-related gene signatures were obtained from cBioPortal based on TCGA. Five hundred forty-eight endometrial cancer samples (TCGA, Firehose Legacy) were analyzed. Mutations and mRNA expression z-scores (RNA Seq V2 RSEM) with a z-score threshold ± 2 were selected.

Protein-protein interactions (PPIs) network construction

The STRING database (<http://string-db.org/>) is a commonly used software system for searching the interaction between proteins. The data in this database were obtained from experimental data, databases, text mining and bioinformatics prediction [19]. The PPI interaction network was constructed by using the STRING database to analyze coexpression, colocation and correlation between the ten genes in the gene signature and to screen the proteins that have the closest relationship with the ten genes with the minimum required interaction score: high confidence 0.700.

Results

Glycolysis-related gene sets differ significantly between endometrial cancer and normal samples

We used “glycolysis” as the keyword to search the gene sets related to glycolysis in the MSigDB database and found three gene sets related to glycolysis, namely, “METABOLISM_KEGG_GLYCOLYSIS_GLUONEOGENESIS”, “METABOLISM_HALLMARK_GLYCOLYSIS” and “METABOLISM_REACTOME_GLYCOLYSIS”. The mRNA expression profiling of endometrial cancer was downloaded from TCGA, including 35 normal and 552 cancer samples, and these data was analyzed as the flowchart (Figure 1). GSEA software was used to analyze the enrichment of three glycolysis gene sets in the endometrial cancer and normal groups. We found that there was a significant difference ($FDR < 0.01$) in the three glycolysis-related gene sets between the endometrial cancer group and the normal control group (Figure 2).

Identification of glycolysis-related mRNAs

Perl script was used to extract the expression matrix of the selected glycolysis genes from mRNA expression profiling data of endometrial cancer. We further screened the differentially expressed glycolysis genes in the endometrial cancer group and normal group using the R “limma” package (screening conditions: $FDR < 0.05$, $|\log FC| > 0.5$). The results showed that there were 156 DEGs, 128 of

which were upregulated, and 28 of which were downregulated (Figure 3A). The R "heatmap" package was used to draw the heatmaps (Figure 3B).

To further verify whether these differentially expressed genes are related to glycolysis, we used the R "ClusterProfiler" package to analyze GO and KEGG enrichment of these differentially expressed genes. The results showed that in the biological process (BP), the differentially expressed genes were mainly involved in such processes as the pyruvate metabolic process, the glycolytic process, and the oxidoreduction coenzyme metabolic process. In molecular function (MF), the differentially expressed genes are primarily involved in such processes as glucose binding, carbohydrate kinase activity, and sugar phosphatase activity. In the cell components (CC), the differentially expressed genes are mainly involved in the nuclear envelope and secreted granule lumen (Figure 2C). The results of KEGG enrichment analysis showed that the differentially expressed genes were mainly involved in such processes as glycolysis/gluconeogenesis, carbon metabolism, and the HIF-1 signaling pathway. These results suggested that these differentially expressed genes were indeed related to glycolysis.

Construction of the glycolysis-related ten-gene signature to predict patient outcomes

We performed univariate Cox regression analysis to screen glycolysis genes related to the prognosis of patients. A total of 11 genes closely related to survival and prognosis in patients were screened. $HR > 1$ represents a risk gene ($P < 0.05$) (Figure 4A). Furthermore, the prognostic gene signature was constructed by LASSO regression analysis, and 10 genes were screened to participate in the prognostic gene signature, namely, PFKM, PSMC4, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI (Figure 4B, C and table 1). The risk score of the gene signature = $(0.0420 \times \text{expression of PFKM}) + (0.0032 \times \text{expression of PSMC4}) + (0.0097 \times \text{expression of NUP85}) + (0.0138 \times \text{expression of PDHA1}) + (0.0028 \times \text{expression of CDK1}) + (0.0024 \times \text{expression of CLDN9}) + (0.0252 \times \text{expression of CENPA}) + (0.0004 \times \text{expression of GPI}) + (0.0253 \times \text{expression of NUP155}) + (0.0067 \times \text{expression of GPCI})$. Based on this gene signature, the risk score of each patient in the whole data set was calculated, and all patients in the whole data set were divided into high- and low-risk subgroups using the risk score median, as the threshold. Kaplan-Meier (K-M) analysis showed that the overall survival rate of the two subgroups was significantly different, and the overall survival rate of patients in the high-risk subgroup was significantly lower than that of the low-risk subgroup ($P < 0.05$) (Figure 4D). Then, patients were ranked according to the risk score, and the ten gene signatures were ranked according to the order of increasing risk score. The results showed that the number of deaths increased with increasing risk score, and the expression levels of PFKM, PSMC4, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI were positively correlated with the risk score, which further confirmed that PFKM, PSMC4, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI were risk genes (Figure 4-E, F, G). In addition, we analyzed the relationship between the ten genes and clinical stage and found that the expression of PFKM, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI increased with increasing clinical stage ($P < 0.05$) (Figure 5).

Table1.
The result of LASSO analysis in
endometrial cancer.

Gene	Coef
PFKM	0.0419514243508422
PSMC4	0.0032134000220372
NUP85	0.009721455192605
PDHA1	0.0138393006344833
CDK1	0.00279937912733931
CLDN9	0.00244161110689154
CENPA	0.0252333242412136
GPI	0.000396228842570881
NUP155	0.0252989772069441
GPC1	0.00672016797587528

To further evaluate whether the constructed ten-gene signature is an independent prognostic factor for endometrial cancer. We included clinical parameters, including age, stage, grade, metastasis, lymph node metastasis (LNM) and risk score, and analyzed these clinical factors with univariate and multivariate Cox regression analyses. The results of univariate Cox regression analyses showed that age, stage, grade, metastasis, lymph node metastasis (LNM) and risk score were all significantly correlated with patient survival ($P<0.05$) (Figure 6A). The results of multivariate Cox regression analysis showed that risk score was an independent prognostic factor of endometrial cancer, and age, grade and stage were also independent prognostic factors ($P<0.05$) (Figure 6B). These results indicated that the ten-gene signature could be used as an independent prognostic factor for endometrial cancer. To evaluate the clinical diagnostic ability of the ten-gene signature, we conducted ROC analysis. The results showed the risk score (area under the curve, AUC=0.730), stage (AUC=0.708), grade (AUC=0.667), age (AUC=0.630), LNM (AUC=0.597), and metastasis (AUC=0.567) (Figure 6C). ROC curve analysis was also performed to evaluate the diagnostic efficacy of the 3 factors (age, grade and stage) and 3 factors+riskScore (Figure 6D and E). The AUC of the survival assessment model was 0.807 of 3 factors and 0.822 of 3 factors+riskScore. It has been suggested that the ten-gene signature has great potential significance in predicting endometrial cancer prognosis. The nomogram plot was built based on prognostic factors in endometrial cancer (Figure 6F).

In order to better understand the network interaction between the risk score model and genes, we extracted mRNA expression profiling and clinical character information for WGCNA analysis. The DEGs between cancer and normal samples were chosen. For constructing a weighted gene network, the

threshold of the adjacency matrix should meet the criterion that the network is close to scale-free. The results showed that the threshold for network construction was selected as 3 (Figure 7A). These co-expression modules were then constructed and the similar modules were clustered, and finally eight gene modules were obtained (Figure 7B). The results of correlation analysis of gene modules with risk models and clinical traits showed that the turquoise module had the highest correlation with the gene-signature (Cor=0.65, $p=9e-64$ for risk; Cor=0.7, $p=4e-77$ for risk score) and the grade (Cor=0.49, $p=4e-33$) (Figure 7C). The turquoise module contained 604 genes, and we then analyzed these genes with Metascape. The GO and KEGG analysis were performed, and the top 20 clusters were chosen to construct a gene function clustering network (Figure 7D). These results indicated that the ten-gene signature was significantly associated with a functional gene module, which was also related with the clinical grade of patients. The gene module mainly enriched in cell division, regulation of cell cycle process and DNA replication, that were important biological processes in the tumor progression.

Gene set enrichment analysis (GSEA)

GSEA was used to analyze the enrichment signal pathways in the high-risk subgroup and low-risk subgroup. A total of 70 significantly enriched KEGG signaling pathways were screened (FDR<0.05). Many of these pathways are closely related to metabolism, including pyruvate metabolism, glycolysis gluconeogenesis and inositol phosphate metabolism. Some other pathways are closely related to metabolic diseases, including the insulin signaling pathway and type II diabetes mellitus. It has been reported that metabolic diseases, such as diabetes, hypertension and obesity, are important risk factors for endometrial cancer, and insulin resistance is closely related to endometrial cancer [20]. Additionally, some signaling pathways are closely related to the occurrence and development of tumors, such as the cell cycle, endometrial cancer, the ERBB signaling pathway, the MAPK signaling pathway, the mTOR signaling pathway and the Wnt signaling pathway (Figure 8).

Genetic alteration, coexpression, neighbor gene network analyses of gene signature

We analyzed the gene alteration of the ten-gene signature through the cBioPortal online website. The results showed that the expression alterations of PFKM, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPC1 in endometrial carcinoma samples were 5, 7, 4, 3, 5, 5, 7, 6 and 3%, respectively. Amplification and increased mRNA were the most common changes (Figure 9A). Coexpression analysis showed that NUP85, NUP155, CDK1 and CENPA had a strong correlation (Figure 9B). Furthermore, we used the STRING database to analyze the proteins coexpressed and interacting with the ten genes and constructed a PPI interaction network. Forty-eight proteins were screened in the PPI network (Figure 9C). In addition, we used immunohistochemistry results from the human protein atlas database to further verify the expression of the ten-gene signature. The results showed that the expression of PDHK1, NUP85, CDK1, CENPA, GPI, GPC1, PSMC4 and PFKM in endometrial cancer was higher than that in normal endometrium, and NUP155 was not detected in endometrial cancer and normal endometrium, although no data were found for CLDN9 (Figure 10).

Estimation of immune and stromal scores related to gene signature

Immune cells and stromal cells are two main types of nontumor components in the tumor microenvironment and have been proposed to be valuable for tumor diagnosis and prognosis evaluation. Therefore, to further reveal the relationship between the tumor microenvironment and the ten-gene signature, we estimated the infiltrating cells and tumor purity of 552 endometrial cancer samples in TCGA data set by the "ESTIMATE" algorithm. Immune scores and stromal scores are used to reflect the presence of immune cells and stromal cells, and single sample enrichment analysis was used to combine them to obtain ESTIMATE scores to represent the purity of the tumor. The immune scores, stromal scores and ESTIMATE scores of patients in the high-risk subgroup and low-risk subgroup were further analyzed. The results showed that the immune scores, stromal scores and ESTIMATE score of the high-risk subgroup were lower than those of the low-risk subgroup (Figure 11A-C). It is suggested that the purity of tumors in the high-risk group is low. We also found that the overall survival rate of patients with low immune scores was significantly lower than that of patients with high immune scores, and there was no significant difference in the overall survival rate of patients with low and high stroma scores, while the overall survival rate of patients with low ESTIMATE scores was also lower than that of patients with high ESTIMATE scores (Figure 11D-F). These results suggested that the poor prognosis of patients in the high-risk group may be closely related to the low immune score.

Correlation between Immune Cell Infiltration and Ten-gene Signature

To further verify the relationship between immune cell infiltration and the gene signature, we analyzed the proportion of twenty-two kinds of immune cells by using the "deconvolution method" of CIBERSORT software, and the samples were screened by $P < 0.05$. The infiltration of immune cells was compared between the high-risk subgroup and the low-risk subgroup. The results showed that compared with the low-risk subgroup, immune cells, such as activated dendritic cells, M1 macrophages, M2 macrophages, activated T memory cells and T follicular helper cells, increased significantly in the high-risk group, while dendritic cell resetting, T cell memory resetting and T regulatory (Treg) cells decreased significantly in the high-risk group (Figure 12A). Furthermore, K-M analysis was used to screen the immune cells closely related to the prognosis of patients. The results showed that the overall survival rate of patients in the high-proportion group of resting dendritic cells, activated NK cells and regulatory T cells (Tregs) was significantly higher than that of the low-proportion group (Figure 12B). We also analyzed the relationship between the ten-gene signature and immune cell infiltration. The results indicated that ten genes were negatively correlated with T cell regulation (Tregs), including CDK1 ($r = -0.41$), CENPA ($r = -0.34$), and NUP155 ($r = -0.22$). CENPA ($r = 0.26$) and PFKM ($r = 0.2$) were positively correlated with the infiltration of M1 macrophages. All ten genes were positively correlated with activated dendritic cells, such as CENPA ($r = 0.21$) (Figure 12C).

Discussion

Endometrial cancer is a common gynecological malignant tumor. Although surgery, radiotherapy, chemotherapy and hormone therapy can significantly improve the treatment effect, the clinical prognosis of patients with advanced and recurrent endometrial cancer is relatively poor, and the 5-year survival rate

of patients is low, which seriously threatens the life and health of women. In recent years, the role of metabolic reprogramming in tumors has been widely studied. Glycometabolism reprogramming is one of the characteristics that tumor cells are different from normal cells. Even under the condition of sufficient oxygen, tumor cells are more likely to use glycolysis for rapid energy supply. Therefore, studying the relationship between metabolic reprogramming and tumor development is becoming a new method for tumor diagnosis, prevention and treatment to elucidate the new mechanism of tumor development from the perspective of metabolic abnormalities and to intervene and correct the metabolic abnormalities of cells. At present, many studies have reported the relationship between glycolysis and endometrial cancer [21,22]. However, research on biomarkers related to glycolysis in endometrial cancer remains limited. It has been reported that clinical characteristics, such as age, stage, differentiation and lymph node metastasis, cannot accurately predict the prognosis of patients [23]. As a result, an increasing number of studies are exploring gene biomarkers, and many studies have found that developing multiple gene-related risk models can improve the prediction efficiency [24,25]. Therefore, the purpose of this study was to explore the risk biomarkers related to the prognosis of endometrial cancer and to analyze further their relationship with immune cell infiltration.

We first downloaded glycolysis-related gene sets from GSEA and screened differentially expressed genes in endometrial cancer and normal samples, including 128 upregulated genes and 28 downregulated genes. Furthermore, we used GO and KEGG enrichment analysis to verify the biological function and signaling pathways of differentially expressed genes. We found that these differentially expressed glycolysis genes are indeed mainly related to pyruvate metabolic processes, glycolytic processes and glycolysis/gluconeogenesis. Next, we used univariate Cox regression to initially screen genes related to the prognosis of endometrial cancer and further used LASSO regression analysis to screen and construct the prognostic gene signature. A total of ten mRNAs (PFKM, PSMC4, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI) significantly related to the overall survival of endometrial cancer were identified to construct the prognostic gene signature. Based on this gene signature, patients were divided into a high-risk subgroup and a low-risk subgroup according to the risk score. K-M analysis showed that the overall survival rate of patients in the high-risk group was significantly lower than that in the low-risk group. It is suggested that it is of great clinical significance to evaluate the prognosis of patients by calculating the risk score of endometrial cancer patients based on the gene signature.

To further verify the value of the ten-gene signature in clinical prognosis, we analyzed the relationship between ten genes and clinical stage and found that the expression of PFKM, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPCI increased with increasing clinical stage. To further evaluate whether the prognostic ability of the ten-gene signature is independent of other clinical parameters, including age, stage, grade, metastasis and lymph node metastasis, we conducted univariate and multivariate Cox regression analyses on the whole data set and found that the gene signature was an independent prognostic factor of endometrial cancer. Furthermore, ROC curve analysis was conducted to verify the prognostic and diagnostic value of the gene signature. The results showed that the diagnostic significance of the gene signature was the strongest (AUC = 0.730), which was better than that of age, stage, grade, metastasis and LNM. This result demonstrated that our gene signature has considerable

prognostic and clinical diagnostic value. In addition, we integrated multiple predictors (including risk scores) through a nomogram to effectively predict the 1-, 3- or 5-year survival rate of patients, which may help to plan short-term follow-up of individualized treatment.

Our GSEA enrichment analysis revealed that many signaling pathways were significantly enriched in the high-risk subgroup, including pathways related to metabolism and metabolic diseases, such as pyruvate metabolism, glycolysis gluconeogenesis, inositol phosphate metabolism, insulin signaling pathway and type II diabetes mellitus. It has been reported that tumor metabolic reprogramming, or the Warburg effect, has been considered one of the ten characteristics of cancer. Both malignant transformation and tumor development, including invasion and metastasis, require metabolic reprogramming [26]. Metabolic heterogeneity is an important reason for the failure of treatment to produce the same effect on cancer cells [27]. It has been reported that diabetes was closely related to the increased risk of endometrial cancer (RR = 1.72) [28]. High glucose and insulin resistance are important characteristics of diabetic patients. It has been confirmed that high glucose could increase glucose uptake and glycolysis activity by regulating the AMPK/mTOR/S6 and MAPK pathways and promote the proliferation and invasion of endometrial cancer cells [29]. High insulin levels are an independent influencing factor of endometrial cancer. Increased insulin and IGF-1 could activate downstream signaling pathways by binding with IR and IGF-1 receptor to promote the proliferation of endometrial cancer cells [30]. These results suggested that the poor prognosis of patients in the high-risk subgroup might be closely related to tumor metabolic reprogramming and the activation of metabolic disease-related pathways. In addition, other pathways closely related to tumorigenesis and development were also significantly enriched in the high-risk subgroup, such as the cell cycle, endometrial cancer, the ERBB signaling pathway, the MAPK signaling pathway, the mTOR signaling pathway, and the Wnt signaling pathway. Taken together, these results show that our gene signature is closely related to metabolic imbalance and provide a potential molecular mechanism for elucidating the relationship between the gene signature and endometrial cancer progression.

In our gene signature, most genes have been reported to be closely related to the occurrence and development of cancer. PFKM, the second rate-limiting enzyme in the glycolysis pathway, has been shown to be closely related to the increased risk of breast cancer [31]. PSMC4 is a member of the proteasome complex, which is responsible for recognizing ubiquitin-labeled substrates and ingesting them into the proteasome (19S regulatory complex). The overexpression of PSMC4 promoted the degradation of some key cell regulatory proteins, such as tumor suppressors, and further promoted the progression of tumors [32]. Therefore, inhibition of the proteasome is a promising cancer treatment strategy. The nucleoporins NUP155 and NUP85 were reported to be upregulated in hepatocellular carcinoma, accompanied by TP53 silencing and overexpression of cell cycle-related genes [33]. PDHA1 is the main regulatory site of PDH activity. PDHA1 regulates the deactivation or activation of PDH through phosphorylation and dephosphorylation and then affects the mitochondrial tricarboxylic acid cycle and glycolysis metabolic flow. It has been reported that the expression of PDHA1 is abnormal in a variety of tumors, and it is closely related to tumor invasion, drug resistance and prognosis by affecting tumor cell glucose metabolism. The upregulation of PDHA1 could promote the metastasis of cholangiocarcinoma

[34]. In contrast, another study reported that low expression of PDHA1 predicted poor prognosis in gastric cancer [35]. Several studies have shown that inhibition of CDK1 expression could significantly inhibit the proliferation and invasion of endometrial cancer cells [36-37]. CENPA and CDK1 were also identified as prognostic markers of lung cancer [38]. Overexpression of CLDN9 could promote tumor cell invasion through Tyk2/STAT3 signaling [39]. Upregulation of GPI-anchored proteins could promote tumor cell migration and progression by enhancing the ERBB signaling pathway [40]. GPCI has received growing interest in recent years due to its high capability of visualizing soft tissue, and GPCI has been reported to have potential value in the diagnosis of breast cancer [41]. We further verified the expression of the gene signature in endometrial carcinoma by using the immunohistochemistry results from the Human Protein Atlas database. The results showed that the expression of PDHK1, NUP85, CDK1, CENPA, GPI, GPC1, PSMC4 and PFKM in endometrial carcinoma was higher than that in normal endometrium, which was consistent with the prediction of the gene signature. There were frequent genetic alterations of ten gene signatures, and amplification and increased mRNA were the most common changes. The results of coexpression analysis showed that NUP85, NUP155, CDK1 and CENPA had a strong correlation. Next, we analyzed the PPI network to explore the correlation between the ten genes and explore the potential interacting proteins. A total of 48 proteins participated in the network construction. Among these proteins, we focused on the interaction between BUB1 and most model genes. In our previous study, we screened 21 genes closely related to the prognosis of endometrial cancer patients, including BUB1, and established a prognostic model for early warning of endometrial cancer based on a low-density chip of 492 tumor-related genes [42]. Moreover, The WGCNA results indicated that the ten-gene signature was significantly associated with a functional gene module that were mainly enriched in cell division, regulation of cell cycle process and DNA replication. It is noteworthy that there was significant correlation between the gene module and clinical grade of patients. Taken together, the above results suggest that the ten-gene signature may play an important role in the occurrence and development of endometrial cancer.

In the past, many studies have focused on the role of glycolysis in tumors. It has been reported that aerobic glycolysis in tumors constantly produces lactic acid, which provides energy for the tumor, and the increased lactic acid in the microenvironment could also affect the immunotherapy effect [43]. It has also been found that antitumor metabolism therapy combined with immunotherapy can effectively inhibit tumor growth [44]. To further explore the relationship between the glycolysis-related gene signature and immune cell infiltration and immune function, we first analyzed the immune scores, stromal scores and ESTIMATE scores of patients in the high-risk subgroup and low-risk subgroup based on our gene signature. We found that the immune scores, stromal scores and ESTIMATE scores of patients in the high-risk subgroup were significantly lower than those of the low-risk subgroup. At the same time, the overall survival rate of patients with low immune scores and estimated scores was significantly higher than that of patients with high scores. Some studies have shown that the more immune cells enter the tumor metastasis, the higher the immune score is, the higher the survival rate is and the lower the recurrence rate is. Metastasis with the smallest number of immune cells entering represented the worst immune microenvironment, and immune escape was most likely to occur under this condition [45]. These results suggested that the poor prognosis of patients in the high-risk subgroup might be closely related to

the low immune scores. However, whether the activation of glycolysis-related pathways affects the infiltration of immune cells warrants further investigation.

We further found that many immune cells, such as activated dendritic cells, M1 macrophages, M2 macrophages, memory activated T cells, and follicular helper T cells, were significantly higher in the high-risk subgroup than in the low-risk subgroup, while dendritic cell resting, memory resting T cells, and regulatory T cells (Tregs) were significantly lower in the high-risk subgroup. There was a positive correlation between the three immune cells and the overall survival rate of patients, including dendritic cell resetting, NK cell activation and T cell regulation (Treg). According to previous research, resting dendritic cells exist in most tissues and are activated to mature antigen-presenting cells under external stimulation. Antigen presentation by resting DCs could induce protective immunity [46]. Tregs play a key role in maintaining immune system homeostasis. Some studies have shown that the high density of Treg cells in tumors is related to the clinical prognosis of tumors, such as liver cancer and gastric cancer [47]. High proportions of Tregs among tumor-infiltrating CD4⁺ T cells were favorable [48]. In addition, we also found that the ten genes were negatively correlated with T cell regulation (Tregs), including CDK1 ($r=-0.41$), CDK1 ($r=-0.34$), and NUP155 (-0.22). Taken together, the above results suggest that the poor prognosis of patients in the high-risk subgroup may be closely related to the decrease in the proportions of dendritic cell resetting and T cell regulation (Tregs), while the glycolysis-related gene signature might be involved in the regulation of T cell regulation (Tregs), and the specific mechanism governing this phenomenon merits further investigation.

Conclusions

In this study, we identified a glycolysis-related ten-gene signature for the prognosis prediction of endometrial cancer patients based on TCGA data sets, where higher risk scores represent a worse prognosis. Moreover, our gene signature has strong value in the clinical diagnosis and prediction of endometrial cancer. In addition, a higher risk score is closely related to a lower immune score, which represents a poor prognosis. However, our research has several limitations. Specifically, it is necessary to perform a more independent cohort to verify the validity of our gene signature. Also, further molecular biological experiments are required to confirm the function of the identified genes in endometrial cancer *in vitro* or *in vivo*. Nevertheless, further research is needed to determine the relationship between the gene signature and metabolic therapy. In conclusion, our results may provide new prognostic molecular markers and clinical treatment targets for endometrial cancer patients and provide new insights for the accurate survival prediction of endometrial cancer patients.

Abbreviation

EC: Endometrial cancer; TCGA: The Cancer Genome Atlas; GSEA: Gene set enrichment analysis; DEGs: Differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; WGCNA: Weighted correlation network analysis; HR: Hazard ratio; ROC: Receiver operating characteristic; PPIs: Protein-protein interactions (PPIs).

Declarations

Availability of data and materials

The datasets analyzed for this study can be found in the TCGA (<https://portal.gdc.cancer.gov/>) database.

Authors' contributions

The study were designed by XY and JW. BS, JZ,YC and YJ downloaded and collected the data. XY and XL analyzed the data. XY wrote the paper. JW, XY and XL reviewed the draft. All authors read and approved the final manuscript.

Funding

This work is supported by The National Natural Science Foundation of China (grants :81672571 and 81874108) and Natural Science Foundation of Beijing (grants :7202213).

Acknowledgements

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018, 68(1):7-30.
2. Sheikh MA, Althouse AD, Freese KE, Soisson S, Edwards RP, Welburn S, Sukumvanich P, Comerici J, Kelley J, LaPorte RE *et al.* USA endometrial cancer projections to 2030: should we be concerned? *Future Oncol.* 2014, 10(16):2561-2568.
3. Tejerizo-Garcia A, Jimenez-Lopez JS, Munoz-Gonzalez JL, Bartolome-Sotillos S, Marqueta-Marques L, Lopez-Gonzalez G, Gomez JF. Overall survival and disease-free survival in endometrial cancer: prognostic factors in 276 patients. *Onco Targets Ther.* 2013, 9:1305-1313.

4. Hanahan D, Weinberg RA: Hallmarks of cancer. the next generation. *Cell*. 2011, 144(5):646-674.
5. Icard P, Shulman S, Farhat D, Steyaert JM, Alifano M, Lincet H. How the Warburg effect supports aggressiveness and drug resistance of cancer cells? *Drug Resist Updat*. 2018, 38:1-11.
6. Han X, Ren C, Yang T, Qiao P, Wang L, Jiang A, Meng Y, Liu Z, Du Y, Yu Z. Negative regulation of AMPK α 1 by PIM2 promotes aerobic glycolysis and tumorigenesis in endometrial cancer. *Oncogene*. 2019, 38(38):6537-6549.
7. Ren B, Cui M, Yang G, Wang H, Feng M, You L, Zhao Y. Tumor microenvironment participates in metastasis of pancreatic cancer. *Mol Cancer*. 2018, 17(1):108.
8. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA *et al*. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013, 4:2612.
9. Van den Eynde M, Mlecnik B, Bindea G, Fredriksen T, Church SE, Lafontaine L, Haicheur N, Marliot F, Angelova M, Vasaturo A *et al*. The Link between the Multiverse of Immune Microenvironments in Metastases and the Survival of Colorectal Cancer Patients. *Cancer Cell*. 2018, 34(6):1012-1026 e1013.
10. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol*. 2018, 1711:243-259.
11. Justus CR, Sanderlin EJ, Yang LV. Molecular Connections between Cancer Cell Metabolism and the Tumor Microenvironment. *Int J Mol Sci*. 2015, 16(5):11055-11086.
12. Cascone T, McKenzie JA, Mbofung RM, Punt S, Wang Z, Xu C, Williams LJ, Wang Z, Bristow CA, Carugo A *et al*. Increased Tumor Glycolysis Characterizes Immune Resistance to Adoptive T Cell Therapy. *Cell Metab*. 2018, 27(5):977-987 e974.
13. Deng Y, Yang J, Luo F, Qian J, Liu R, Zhang D, Yu H, Chu Y. mTOR-mediated glycolysis contributes to the enhanced suppressive function of murine tumor-infiltrating monocytic myeloid-derived suppressor cells. *Cancer Immunol Immunother*. 2018, 67(9):1355-1364.
14. Ganapathy-Kanniappan S. Taming Tumor Glycolysis and Potential Implications for Immunotherapy. *Front Oncol*. 2017, 7:36.
15. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012, 16(5):284-287.
16. Zhou Z, Huang R, Chai R, Zhou X, Hu Z, Wang W, Chen B, Deng L, Liu Y, Wu F. Identification of an energy metabolism-related signature associated with clinical prognosis in diffuse glioma. *Aging (Albany NY)*. 2018, 10(11):3185-3209.
17. Tang LQ, Li CF, Li J, Chen WH, Chen QY, Yuan LX, Lai XP, He Y, Xu YX, Hu DP *et al*. Establishment and Validation of Prognostic Nomograms for Endemic Nasopharyngeal Carcinoma. *J Natl Cancer Inst*. 2016, 108(1).
18. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E *et al*. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013, 6(269):pl1.

19. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019, 47(D1):D607-D613.
20. Yin XH, Jia HY, Xue XR, Yang SZ, Wang ZQ. Clinical analysis of endometrial cancer patients with obesity, diabetes, and hypertension. *Int J Clin Exp Med.* 2014, 7(3):736-743.
21. Han X, Ren C, Yang T, Qiao P, Wang L, Jiang A, Meng Y, Liu Z, Du Y, Yu Z. Negative regulation of AMPKalpha1 by PIM2 promotes aerobic glycolysis and tumorigenesis in endometrial cancer. *Oncogene.* 2019, 38(38):6537-6549.
22. Shim SH, Kim DY, Lee DY, Lee SW, Park JY, Lee JJ, Kim JH, Kim YM, Kim YT, Nam JH. Metabolic tumour volume and total lesion glycolysis, measured using preoperative 18F-FDG PET/CT, predict the recurrence of endometrial cancer. *BJOG.* 2014, 121(9):1097-1106; discussion 1106.
23. Zhao L, Jiang L, He L, Wei Q, Bi J, Wang Y, Yu L, He M, Zhao L, Wei M. Identification of a novel cell cycle-related gene signature predicting survival in patients with gastric cancer. *J Cell Physiol.* 2019, 234(5):6350-6360.
24. Tao C, Huang K, Shi J, Hu Q, Li K, Zhu X. Genomics and Prognosis Analysis of Epithelial-Mesenchymal Transition in Glioma. *Front Oncol.* 2020, 10:183.
25. Zeng Q, Sun S, Li Y, Li X, Li Z, Liang H. Identification of Therapeutic Targets and Prognostic Biomarkers Among CXC Chemokines in the Renal Cell Carcinoma Microenvironment. *Front Oncol.* 2019, 9:1555.
26. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab.* 2008, 7(1):11-20.
27. Yoshida GJ. Metabolic reprogramming: the emerging concept and associated therapeutic strategies. *J Exp Clin Cancer Res.* 2015, 34:111.
28. Saed L, Varse F, Baradaran HR, Moradi Y, Khateri S, Friberg E, Khazaei Z, Gharahjeh S, Tehrani S, Sioofy-Khojine AB *et al.* The effect of diabetes on the risk of endometrial Cancer: an updated a systematic review and meta-analysis. *BMC Cancer.* 2019, 19(1):527.
29. Han J, Zhang L, Guo H, Wysham WZ, Roque DR, Willson AK, Sheng X, Zhou C, Bae-Jump VL. Glucose promotes cell proliferation, glucose uptake and invasion in endometrial cancer cells via AMPK/mTOR/S6 and MAPK signaling. *Gynecol Oncol.* 2015, 138(3):668-675.
30. Mu N, Zhu Y, Wang Y, Zhang H, Xue F. Insulin resistance: a significant risk factor of endometrial cancer. *Gynecol Oncol.* 2012, 125(3):751-757.
31. Ahsan H, Halpern J, Kibriya MG, Pierce BL, Tong L, Gamazon E, McGuire V, Felberg A, Shi J, Jasmine F *et al.* A genome-wide association study of early-onset breast cancer identifies PFKM as a novel breast cancer gene and supports a common genetic spectrum for breast cancer at any age. *Cancer Epidemiol Biomarkers Prev.* 2014, 23(4):658-669.
32. Hellwinkel OJ, Asong LE, Rogmann JP, Sultmann H, Wagner C, Schlomm T, Eichelberg C. Transcription alterations of members of the ubiquitin-proteasome network in prostate carcinoma.

- Prostate Cancer Prostatic Dis. 2011, 14(1):38-45.
33. Beck M, Schirmacher P, Singer S. Alterations of the nuclear transport system in hepatocellular carcinoma - New basis for therapeutic strategies. *J Hepatol.* 2017, 67(5):1051-1061.
 34. Dan L, Wang C, Ma P, Yu Q, Gu M, Dong L, Jiang W, Pan S, Xie C, Han J *et al.* PGC1alpha promotes cholangiocarcinoma metastasis by upregulating PDHA1 and MPC1 expression to reverse the Warburg effect. *Cell Death Dis.* 2018, 9(5):466.
 35. Song L, Liu D, Zhang X, Zhu X, Lu X, Huang J, Yang L, Wu Y. Low expression of PDHA1 predicts poor prognosis in gastric cancer. *Pathol Res Pract.* 2019, 215(3):478-482.
 36. Yan GJ, Yu F, Wang B, Zhou HJ, Ge QY, Su J, Hu YL, Sun HX, Ding LJ. MicroRNA miR-302 inhibits the tumorigenicity of endometrial cancer cells by suppression of Cyclin D1 and CDK1. *Cancer Lett.* 2014, 345(1):39-47.
 37. Li L, Qu YW, Li YP. Over-expression of miR-1271 inhibits endometrial cancer cells proliferation and induces cell apoptosis by targeting CDK1. *Eur Rev Med Pharmacol Sci.* 2017, 21(12):2816-2822.
 38. Liu WT, Wang Y, Zhang J, Ye F, Huang XH, Li B, He QY. A novel strategy of integrated microarray analysis identifies CENPA, CDK1 and CDC20 as a cluster of diagnostic biomarkers in lung adenocarcinoma. *Cancer Lett.* 2018, 425:43-53.
 39. Liu H, Wang M, Liang N, Guan L. Claudin-9 enhances the metastatic potential of hepatocytes via Tyk2/Stat3 signaling. *Turk J Gastroenterol.* 2019, 30(8):722-731.
 40. Zhang XX, Ni B, Li Q, Hu LP, Jiang SH, Li RK, Tian GA, Zhu LL, Li J, Zhang XL *et al.* GPAA1 promotes gastric cancer progression via upregulation of GPI-anchored protein and enhancement of ERBB signalling pathway. *J Exp Clin Cancer Res.* 2019, 38(1):214.
 41. Li X, Chen Z, Zhang L, Zhu X, Wang S, Peng W. Quantitative characterization of ex vivo breast tissue via x-ray phase-contrast tomography. *J Xray Sci Technol.* 2019, 27(3):503-516.
 42. Zhou J, Yin F, Li X, Yao Y, Chen Y, Gao B, *et al.* The establishment of gene expression profiling-guided clinical precision treatment for patients with endometrial carcinoma[EB/OL]. Beijing:Sciencepaper Online [2017-05-16]. <http://www.paper.edu.cn/releasepaper/content/201705-1045>
 43. Bohn T, Rapp S, Luther N, Klein M, Bruehl TJ, Kojima N, Aranda Lopez P, Hahlbrock J, Muth S, Endo S *et al.* Tumor immunoevasion via acidosis-dependent induction of regulatory tumor-associated macrophages. *Nat Immunol.* 2018, 19(12):1319-1329.
 44. Gao F, Tang Y, Liu WL, Zou MZ, Huang C, Liu CJ, Zhang XZ. Intra/Extracellular Lactic Acid Exhaustion for Synergistic Metabolic Therapy and Immunotherapy of Tumors. *Adv Mater.* 2019, 31(51):e1904639.
 45. Van den Eynde M, Mlecnik B, Bindea G, Fredriksen T, Church SE, Lafontaine L, Haicheur N, Marliot F, Angelova M, Vasaturo A *et al.* The Link between the Multiverse of Immune Microenvironments in Metastases and the Survival of Colorectal Cancer Patients. *Cancer Cell.* 2018, 34(6):1012-1026 e1013.
 46. Probst HC, McCoy K, Okazaki T, Honjo T, van den Broek M. Resting dendritic cells induce peripheral CD8+ T cell tolerance through PD-1 and CTLA-4. *Nat Immunol.* 2005, 6(3):280-286.

47. Najafi M, Farhood B, Mortezaee K. Contribution of regulatory T cells to cancer: A review. J Cell Physiol. 2019, 234(6):7983-7993.
48. Choi HS, Ha SY, Kim HM, Ahn SM, Kang MS, Kim KM, Choi MG, Lee JH, Sohn TS, Bae JM *et al*. The prognostic effects of tumor infiltrating regulatory T cells and myeloid derived suppressor cells assessed by multicolor flow cytometry in gastric cancer patients. Oncotarget. 2016, 7(7):7940-7951.

Figures

Figure 1

The flow chart of the study design and analysis.

Figure 2

Enrichment plots of three glycolysis-related gene sets which were significantly differentiated between in normal and endometrial cancer tissues using GSEA. (FDR is the corrected p value of multiple hypothesis test, NES stands for normalized score enrichment score).

Figure 3

Identification of differentially expressed genes (DEGs) related to glycolysis of TCGA datasets between normal and endometrial cancer tissues and GO and KEGG pathway enrichment analysis of DEGs. A. Differential expression genes between two two groups. The red dot is the up-regulated gene, the green dot is the down regulated gene, and the black dot is the other genes without significant difference screened by the criteria of $|\text{Fold Change}| > 0.5$ and $\text{FDR} < 0.05$. B. Hierarchical clustering of differentially expressed genes in two groups. C. The GO functional enrichment analysis of differential gene includes three

domains: molecular function, biological process and cell composition. D. KEGG pathway analysis of differentially expressed genes.

Figure 4

Construction and validation of prognostic model prognosis-associated genes of endometrial cancer A. Results of univariate Cox analysis for the prognostic genes in endometrial cancer. B and C. LASSO regression model. D. Kaplan-Meier analysis for the ten glycolysis-related gene signatures related to risk

score predicts overall survival in patients with endometrial cancer. E. Gene signature-related risk score distribution in each patient. F. Survival days of patients in order of the value of risk scores. G. A heatmap of ten glycolysis-related gene signatures.

Figure 5

The correlation between ten glycolysis-related gene signature and clinical stages. The expression of PFKM, NUP85, PDHA1, CDK1, CLDN9, CENPA, GPI, NUP155 and GPC1 increased with the increase of clinical stage (P < 0.05) .

Figure 6

Forrest plot of the univariate and multivariate Cox regression analysis in endometrial cancer. A. The univariate Cox regression analysis in endometrial cancer. B. The multivariate Cox regression analysis in endometrial cancer. C. ROC curve analysis was performed to evaluate the diagnostic efficacy of the gene signature and other clinical characteristics. D and E. ROC curve analysis was performed to evaluate the diagnostic efficacy of the 3 factors (age, grade and stage) and 3 factors+riskScore. F. Nomogram is used

to show the relationship between the variables in the prediction model and predict the 1, 3 or 5-year overall survival rate of patients.

Figure 7

WGCNA were performed to identify the potential mechanisms associated with the prognostic signature. A. Screening of the soft threshold. B. Clustering dendrogram of genes in EC tissues. C. Correlation

analysis of gene modules with risk model and clinical traits. D. Enrichment clustering network analysis in the Metascape database.

Figure 8

The Fifteen representative significantly enriched KEGG pathways in high-risk by GSEA.

Figure 9

Gene alteration, coexpression analyses and protein-protein interactions were performed for the ten glycolysis-related gene signatures. A. Gene alteration of selected genes in patients with endometrial cancer from the cBioPortal website. B. Coexpression analyses of ten glycolysis-related gene signatures in endometrial cancer by the R “corrplot” package. C. Protein-protein interaction network of ten gene signatures and other closely related proteins.

Normal		PDHA1 Patient id: 3313 Endometrial stroma Staining: Not detected Intensity: Negative Quantity: None Glandular cells Staining: Not detected Intensity: Weak Quantity: <25%		NUP85 Patient id: 2941 Endometrial stroma Staining: Not detected Intensity: Weak Quantity: <25% Glandular cells Staining: Medium Intensity: Moderate Quantity: 75%~25%		CDK1 Patient id: 2361 Endometrial stroma Staining: Medium Intensity: Strong Quantity: <25% Glandular cells Staining: Medium Intensity: Strong Quantity: <25%
	PDHA1 Patient id: 2339 Tumor cells Staining: Medium Intensity: Moderate Quantity: >75%		NUP85 Patient id: 4205 Tumor cells Staining: Medium Intensity: Moderate Quantity: >75%		CDK1 Patient id: 1102 Tumor cells Staining: High Intensity: Strong Quantity: >75%
Normal		CENPA Patient id: 2941 Endometrial stroma Staining: Not detected Intensity: Weak Quantity: <25% Glandular cells Staining: Low Intensity: weak Quantity: 75%~25%		NUP155 Patient id: 3333 Endometrial stroma Staining: Not detected Intensity: Negative Quantity: None Glandular cells Staining: Not detected Intensity: Negative Quantity: None		GPI Patient id: 2361 Endometrial stroma Staining: Not detected Intensity: Weak Quantity: <25% Glandular cells Staining: Low Intensity: Weak Quantity: >75%
	CENPA Patient id: 1750 Tumor cells Staining: Medium Intensity: Moderate Quantity: >75%		NUP155 Patient id: 1750 Tumor cells Staining: Not detected Intensity: Negative Quantity: None		GPI Patient id: 3540 Tumor cells Staining: High Intensity: Strong Quantity: >75%
Normal		GPC1 Patient id: 2941 Endometrial stroma Staining: Medium Intensity: Moderate Quantity: 75%~25% Glandular cells Staining: Not detected Intensity: Negative Quantity: None		PSMC4 Patient id: 1978 Endometrial stroma Staining: Not detected Intensity: Weak Quantity: <25% Glandular cells Staining: Low Intensity: weak Quantity: 75%~25%		PFKM Patient id: 2004 Endometrial stroma Staining: Medium Intensity: Moderate Quantity: 75%~25% Glandular cells Staining: Medium Intensity: Moderate Quantity: >75%
	GPC1 Patient id: 3036 Tumor cells Staining: Low Intensity: weak Quantity: 75%~25%		PSMC4 Patient id: 2236 Tumor cells Staining: Medium Intensity: Moderate Quantity: >75%		PFKM Patient id: 1165 Tumor cells Staining: Medium Intensity: Moderate Quantity: >75%

Figure 10

The representative protein expression of ten glycolysis-related gene signature in endometrial cancer and normal tissue. Data were from the Human Protein Atlas (<http://www.proteinatlas.org>) database. There was no data of CLDN9 in the database.

Figure 11

Immune scores, stromal scores and ESTIMATE scores of ten glycolysis-related gene signatures were related to risk score. A. Immune scores for patients in the low- vs. high-risk group. B. The stromal scores for patients in the low- vs. high-risk group. C. The ESTIMATE scores for patients in the low- vs. high-risk group. D. Kaplan-Meier analysis of overall survival for patients with low vs. high immune scores. E. Kaplan-Meier analysis of overall survival for patients with low vs. high stromal scores. F. Kaplan-Meier analysis of overall survival for patients with low vs. high ESTIMATE scores.

Figure 12

A. Abundance ratios of the immune cells for patients in the low- vs. high-risk group. B. Kaplan-Meier analysis of the relationship between the infiltration of immunocytes and overall survival for patients (filter criteria: $P < 0.05$). C. Correlation between the ten glycolysis-related gene signatures and various immune cells. Red represents a positive correlation, and blue represents a negative correlation.