RESEARCH

Predictive significance of glycolysis-associated lncRNA profiles in colorectal cancer progression

Rui Mao[1†], Chenxin Xu[2,3†], Quanzheng Zhang[4†], Zheng Wang[5], Yanjun Liu[2,3*], Yurui Peng[2,3*] and Ming Li[6*]

Abstract

Background The Warburg effect is a hallmark characteristic of colorectal cancer (CRC). Despite extensive research, the role of long non-coding RNAs (lncRNAs) in influencing the Warburg effect remains incompletely understood. Our study aims to identify lncRNAs that may modulate the Warburg effect by functioning as competing endogenous RNAs (ceRNAs).

Methods Utilizing bioinformatics approaches, we extracted glycolysis-associated gene data from the Kyoto Encyclopedia of Genes and Genomes (KEGG) and identified 101 glycolysis-related lncRNAs in CRC. We employed Univariable Cox regression, Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis, and Multivariable Cox regression to develop a prognostic model comprising four glycolysis-linked lncRNAs. We then constructed a prognostic nomogram integrating this lncRNA model with other relevant clinical parameters.

Results The prognostic efficacy of our four-lncRNA signature and its associated nomogram was validated in both training and validation cohorts. Functional assays demonstrated significant glycolysis and hexokinase II (HK2) inhibition following the silencing of RUNDC3A − AS1, a key lncRNA in our prognostic signature, highlighting its regulatory importance in the Warburg effect.

Conclusions Our research illuminates the critical role of glycolysis-centric lncRNAs in CRC. The developed prognostic model and nomogram underscore the pivotal prognostic and regulatory significance of the lncRNA RUNDC3A − AS1 in the Warburg effect in colorectal cancer.

Keywords Warburg effect, Colorectal cancer, lncRNA, Prognostic signature, Nomogram

†Rui Mao, Chenxin Xu and Quanzheng Zhang contributed equally to this work.

*Correspondence:
Yanjun Liu
liuyanjun_001@163.com
Yurui Peng
2232554370@qq.com
Ming Li
docliming0818@163.com

Full list of author information is available at the end of the article

Introduction

Colorectal cancer (CRC) occupies a significant position among gastrointestinal malignancies, ranking as the fourth most common cause of cancer-related deaths worldwide. Annually, it accounts for over 550,000 fatalities and registers more than 1 million new cases [1]. Despite considerable research endeavors, the complex etiology of CRC continues to be only partially understood, thereby limiting the potential for early diagnosis and effective intervention [2]. As a result, the prognosis for many CRC patients remains relatively poor [3]. This underscores the urgent need for in-depth research into CRC's multifaceted mechanisms, aimed at discovering innovative diagnostic and prognostic biomarkers.

A hallmark of cancerous cells is their relentless proliferation, not merely due to unchecked growth but also because of metabolic alterations. These adaptations are crucial for meeting the energy and macromolecular demands of cells, particularly in the hypoxic and nutrient-deprived microenvironment of tumors. Such metabolic flexibility, known as metabolic reprogramming, has recently been recognized as a critical aspect of cancer biology. In this context, aerobic glycolysis, famously known as the Warburg effect, plays a pivotal role in this metabolic transformation [4]. The essence of the Warburg effect is characterized by the abnormal activity of key glycolytic enzymes [5], notably hexokinase II (HK2). HK2, functioning downstream of oncogenes, significantly contributes to the maintenance of this altered metabolic state [6].Long non-coding RNAs (lncRNAs), which are transcripts exceeding 200 nucleotides, uniquely intervene in cellular processes without encoding proteins [7]. Their pivotal role in the evolution and progression of colorectal cancer (CRC) has been established, particularly as competing endogenous RNAs (ceRNAs). These ceRNAs modulate target gene expression by sponging microRNAs (miRNAs) [8]. Research by Meng et al. has shown that the lncRNA LINC00525 influences the Warburg effect in CRC by activating Hypoxia-Inducible Factor 1-alpha (HIF-1α) through the miR-338-3p/UBE2Q1/β-catenin signaling pathway [9]. While Zhu J and colleagues have linked glycolysis-related genes with CRC prognosis [10], a comprehensive catalog of lncRNAs driving the Warburg effect is still lacking. Furthermore, prognostic assessments using glycolysis-related lncRNAs in CRC have not been adequately explored.

In our study, we extracted a list of glycolysis-related genes from the KEGG database and integrated gene expression profiles with relevant clinical data of CRC patients from The Cancer Genome Atlas (TCGA). Utilizing Gene Set Variation Analysis (GSVA), we determined the enrichment scores of the glycolysis pathway, thereby revealing the relationship between glycolytic activity and CRC prognosis. Building on this foundation, we developed a ceRNA network to identify glycolysis-related lncRNAs, leading to the creation and validation of a glycolysis-focused prognostic signature. We highlighted a specific lncRNA from this signature to emphasize its significant role in CRC glycolysis, suggesting new potential targets for therapeutic intervention in CRC.

Materials and methods

Data acquisition and preprocessing

RNA expression datasets, accompanied by clinical data for CRC patients, were sourced from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). Sample inclusion criteria were as follows: (1) histologically confirmed CRC; (2) mRNA expression profile data and clinical information were available simultaneously; and (3) only samples with an overall survival (OS) time longer than 30 days were involved when survival analysis was performed. LncRNAs and mRNAs quantification followed the reference catalog provided by the GENCODE (GRCh38) from the Genome Research Project of ENCyclopedia of DNA Elements (https://www.gencodegenes.org/). Long non-coding RNA gene annotation The reference table is provided in the supplementary material. Glycolysis-centric genes were collated from the KEGG's hsa00010 gene dataset (https://www.genome.jp/kegg/pathway/hsa/hsa00010.html). The "caret" package in R was employed to randomly bifurcate eligible patients into two cohorts: Training and Validation.

Differentially expressed analysis

Employing the R "limma" package, we discerned differentially expressed lncRNAs (DElncRNAs), miRNAs (DEmiRNAs), and glycolysis-associated mRNAs (DEmRNAs) between CRC and adjacent non-tumor tissues. The delineation criteria were an absolute logFC value of ≥1 and an adjusted P-value of <0.05. Resultant differential expression patterns were visualized via heatmaps, generated using the "pheatmap" R package.

Gene set variation analysis (GSVA)

GSVA represents a non-parametric and unsupervised approach to gene set enrichment, facilitating the estimation of gene set activity specific to individual samples based on gene expression data. Contrasting with conventional enrichment analyses like Gene Set Enrichment Analysis (GSEA), GSVA uniquely operates on an individual sample basis, thus offering a novel methodology to assess variations in gene set expression within a singular sample context. GSVA was conducted utilizing the "GSVA" R package. Post-analysis, patient data was bifurcated, based on the median enrichment score of the KEGG glycolysis pathway, into two distinct clusters.

The construction of the ceRNA network

We consulted the miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/), miRDB (http://www.mirdb.org/), and TargetScan (http://www.targetscan.org/) databases to pinpoint miRNAs potentially modulating glycolysis-related DEmRNAs. Only miRNAs concurrently acknowledged across these three databases were deemed as "predicted miRNAs." The StarBase platform (http://starbase.sysu.edu.cn/) facilitated the identification of interactions between lncRNAs and our predicted miRNAs, yielding a subset of anticipated lncRNAs. Ensuing this, we juxtaposed these predicted entities with DEmiRNAs and DElncRNAs from TCGA CRC datasets, retaining only coinciding miRNAs, lncRNAs, and their interaction duos for subsequent analyses. The curated glycolysis-focused ceRNA interaction network was visualized via the Cytoscape 3.6.1 software.

The construction of the glycolysis-related lncRNA-based prognostic signature

Univariate and multivariate Cox regression, LASSO analyses were used to select the independent risk glycolysis-related lncRNAs. Multivariate COX regression was used to identify corresponding coefficients of CRC prognostic signature which was calculated according to the following formula:

$$\text{Riskscore} = \sum regression\, coefficient\,(lncRNA_i) \times expression\, value\,(lncRNA_i)$$

,using "glment", "survminer" and "survival" packages of R. According to the equation above, risk score of each patient was calculated separately in the training and validation cohorts. Patients were subsequently divided into high- and low-risk groups with the median risk score as the cut-off point. Kaplan-Meier curves and the log-rank test were used to compare the survival outcomes of the two groups. Receiver operating characteristic (ROC) curve analysis and Harrell's concordance index were employed to assess the accuracy and precision of the survival predictions according to the risk scores.

Development and validation of the prognostic nomogram with signature and clinicopathological parameters

Risk score and the clinicopathological parameters are selected as candidate predictors in this model. Univariate and Multivariate COX regressions were analysed. All the identified independent prognostic parameters were utilized to develop a prognostic nomogram for predicting the 1-, 3-, and 5-year survival outcomes of CRC patients using the "rms" package of R. Calibration plots at 1-, 3-, and 5-year were constructed to graphically evaluate the discriminative ability of the nomogram. Decision curve analysis (DCA) was performed using the code downloaded from MSKCC (https://www.mskcc.org/). The discrimination performance of the nomogram was quantitatively assessed by the concordance index, ROC curve, and DCA.

RNA extraction and quantitative reverse transcription PCR (qRT-PCR) analysis

TRIzol reagent (Invitrogen, Carlsbad, CA, USA) was used to isolate total RNA. Then the total RNA was reverse transcribed into cDNA with random primers using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Penzberg, Germany). The relative quantification of lncRNA RUNDC3A−AS and mRNA HK2 was examined using the SYBR green PCR Master Mix (Qiagen, Hilden, Germany), with internal GAPDH as control. The data were processed using a $2-\Delta\Delta Ct$ method. All reactions were run in triplicate independently. The primers used are presented in supplementary Table S1.

Cell culture and reagents

CRC cell lines, including SW480, SW620 were obtained from Cell Bank of Type Culture Collection, Chinese Academy of Sciences (Shanghai, China). SW480 and SW620 cells were maintained in RPMI 1640 medium (Invitrogen) supplemented with 10% FBS at 37 ℃ in a humidified atmosphere of 5% CO2. Knockdown experiments were conducted using siRNAs targeting lncRNA RUNDC3A−AS1 and negative control (siNC) with Lipofectamine RNAiMAX reagent (Invitrogen), according to manufacturer's instructions. The sequence of siRNA is 5'-GTCTCTACATTGATTCAGATTTG-3'.

Glucose consumption and lactate production assays

Glucose was quantified by glucose assay kits (Applygen Technologies, Beijing, China) and glucose consumption was calculated by subtracting the amount of glucose present in the supernatants of cell culture medium. Lactic acid produced in the culture medium was measured using an assay kit (Nanjing Jiancheng Bioengineering Institute, Nanjing, China) according to the manufacturer's protocol.

Protein extraction and western-blot analysis

SW480 and SW460 cells in 6-well plates were lysed with radio immunoprecipitation assay (RIPA) buffer (Thermo Scientific, Waltham, MA, USA) with a protease inhibitor cocktail (Roche, IN, USA). Equal amounts of protein were separated by 10% SDS-PAGE gel electrophoresis and transferred onto a polyvinylidene difluoride membrane. The membranes were probed with antibodies anti-HK2 (ab104836; Abcam) or GAPDH (ab9484; Abcam). Then, the probed membranes were incubated with horseradish peroxidase-conjugated secondary anti-body. To

visualize signals, enhanced chemiluminescence (ECL) substrates (Bio-Rad) was conducted, with GAPDH as an endogenous protein for normalization.

Colony formation assays
Cell colony formation rate was detected using colony formation assays. 0.5×103 colorectal cancer cells were plated in a 60-mm plate and cultured for 7 days. Then, 10% formaldehyde was used for fixing plates for 5 min and 1% crystal violet was used for stained for 30 s. Finally, counting the number of colonies.

Statistical analysis
Statistical analyses were performed with R Project. We checked the distributions of metrological data fitted the normal distribution using the Kolmogorov–Smirnov test. Comparisons between the two groups were performed using the non-parametric Mann-Whitney U-test and t-test, where applicable. Chi-square test was used to detect the differences in the clinicopathologic characteristics of patients in two different groups. Survival was analyzed using Kaplan-Meier survival analysis. A two-side test $P < 0.05$ was indicated statistical significance.

Results

Dysregulated glycolysis plays an important role in the development of CRC
First, we calculated the glycolysis pathway enrichment score of every CRC patient in the TCGA RNA-seq dataset by using GSVA method to explore the association of glycolysis activity and clinical outcome. Patients were

divided into two groups with the median score of glycolysis pathway as the cutoff point. Patients in group1 had a higher enrichment score of glycolysis pathway than patients in group2. The heatmap of the two groups showed significant differences in glycolysis-related genes expression (Fig. 1A). Subsequently, we performed Chi-square tests to see if there exist any differences in clinical characteristics between these 2 groups of patients. The results showed that patients in group1 had the characteristics of lower stage grade, lower N, M stage, and higher possibility of being alive. Moreover, Kaplan-Meier curve analysis revealed that the OS time of patients in group1 was significantly longer than that of patients in group2 (Fig. 1B). Wilcox test was performed to further detect the difference in GSVA enrichment score between patients with different clinical characteristics (Fig. 1C-E). The results suggested lower glycolysis pathway enrichment score was associated with a higher tumor stage, N and M stage ($p < 0.05$). These results indicated that glycolysis GSVA enrichment score was closely correlated with clinical outcomes of patients with CRC, suggesting an essential biological role of dysregulated glycolysis in CRC development.

Construction of the CeRNA network
A total of 568 CRC samples and 44 adjacent normal samples were downloaded to screen DElncRNAs and DEmRNAs; 446 CRC samples and 45 adjacent normal samples were utilized to screen DEmiRNAs. 62 glycolysis-related genes were extracted from KEGG glycolysis gene set, and 11 of which were differentially expressed in CRC samples.

Fig. 1 Glycolysis pathway plays an important role in CRC. **(A)** The heatmap of glycolysis-related genes expression in TCGA datasets ranked based on GSVA score of glycolysis pathway and the relationship between the GSVA score of glycolysis pathway and CRC-related clinical characteristics. *$p < 0.05$; **$p < 0.01$; **$p < 0.001$. **(B)** Kaplan–Meier curves of OS in TCGA-CRC patients who were were divided into two groups based on the glycolysis pathway GSVA score. **(C-E)** Patients with different clinical characteristics showed significantly different glycolysis pathway GSVA scores

2550 DElncRNAs and 212 DEmiRNAs were screened from TCGA CRC datasets.

We then tried to find predicted targeting miRNAs of the glycolysis-related DEmRNAs from TargetScan, miRDB, and miRTarBase databases. After intersecting with the DEmiRNAs, only 11 miRNAs remained (Fig. 2A). Subsequently, 319 target lncRNAs were predicted for these 11 miRNAs by StarBase databases, which were then overlapped with 2449 DElncRNAs collected from the TCGA, resulting in 101 lncRNAs shared (Fig. 2B). The top 10 up-regulated and down-regualted Warburg-effect related lncRNAs are showed in Tables S1 and Table S2 respectively. Based on the combination of the above data, we established the ceRNA network using cytscope (Fig. 2C).

Establishment of glycolysis-related lncRNA signature

506 selected CRC patients were divided randomly into two equal parts, one for the training cohort, the other for internal validation cohort. 101 glycolysis-related lncRNAs were initially subjected to univariate regression analysis, with the OS of the training cohorts as the dependent outcomes. The result suggested that 10 glycolysis-related lncRNAs were significantly correlated with the OS of patients in training cohort (Fig. 3A). The

LASSO regression analysis was performed to select the most valuable predictive genes with nonzero regression coefficients (Fig. 3B-C). Finally, a 4-lncRNA signature (*CRYZL2P−SEC16B, RUNDC3A−AS1, HOTAIR, and ALMS1−IT1*) was identified using multivariate COX regression (Fig. 3D).

Then, the risk score of each patient was calculated using the formula based on the corresponding coefficients: Risk score = (expr $CRYZL2P−SEC16B \times 1.1191$) + (expr $RUNDC3A−AS1 \times 8.1389$) - (expr $HOTAIR \times 0.2031$) -(expr $ALMS1−IT1 \times 0.5462$).

All patients were divided into high- or low-risk groups based on the median risk score (Fig. 3E-F). The Kaplan-Meier log-rank test confirmed that patients in the high-risk group tended to have shorter OS time in training cohort ($p=9.608e-05$; Fig. 3G). The AUCs for 1-year, 2-year and 3-year OS were 0.7, 0.741, 0.732 respectively (Fig. 3H). The clinicopathological features were then compared between the high risk-score group and low-risk score group. We found the two groups in tumor stage (stage) ($P<0.05$), metastasis status (M) ($P<0.05$) and N state ($P<0.05$) were all significantly different (Fig. 3I). We separate the patients with stage II colorectal cancer into high- or low-risk groups based on the median risk score, in training cohort (Figure S1A-B).
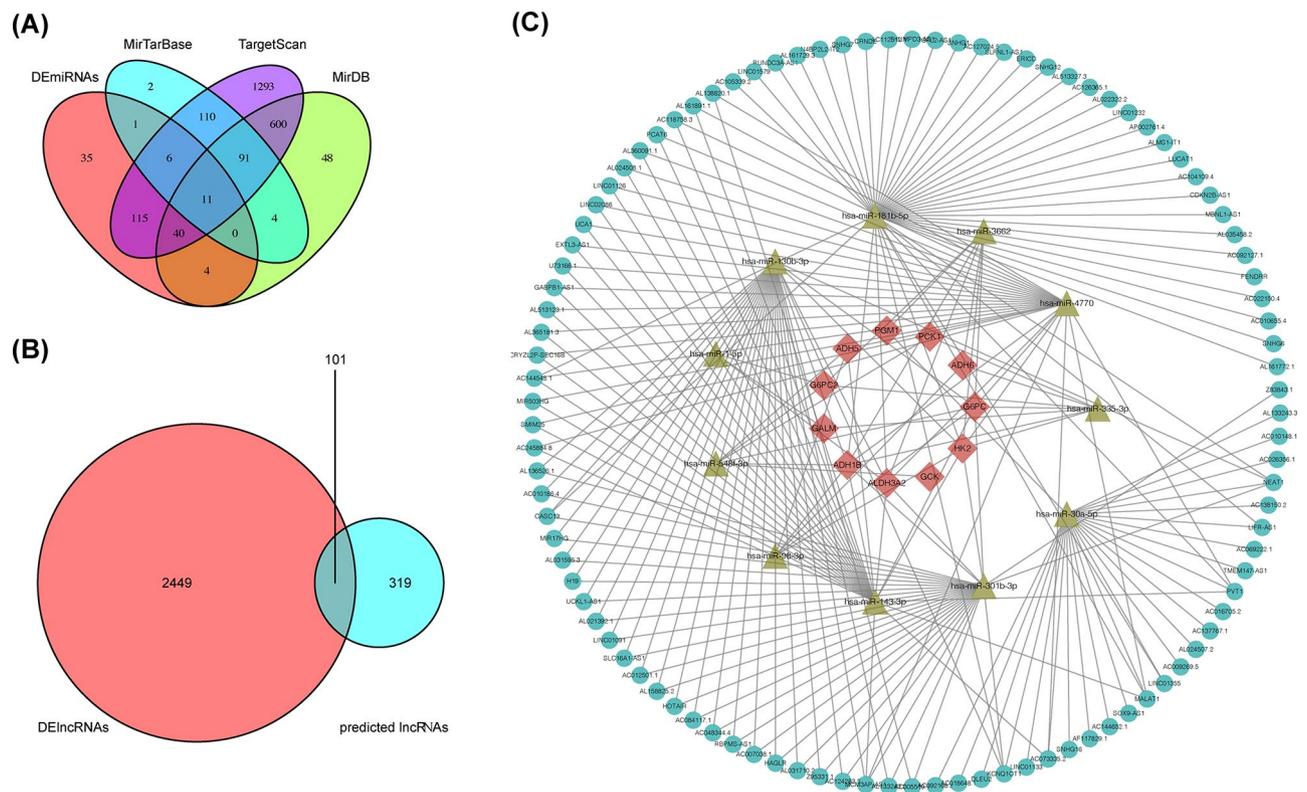
Fig. 2 Screen out glycolysis-related lncRNAs by constructed ceRNA network. **(A)** 11 target miRNAs were obtained by taking the intersection of the predicted-miRNAs and DEmiRNAs; **(B)** 101 glycolysis-related lncRNAs were obtained by taking the intersection of the predicted-lncRNAs and DElncRNAs; **(C)** Construction of a ceRNA network
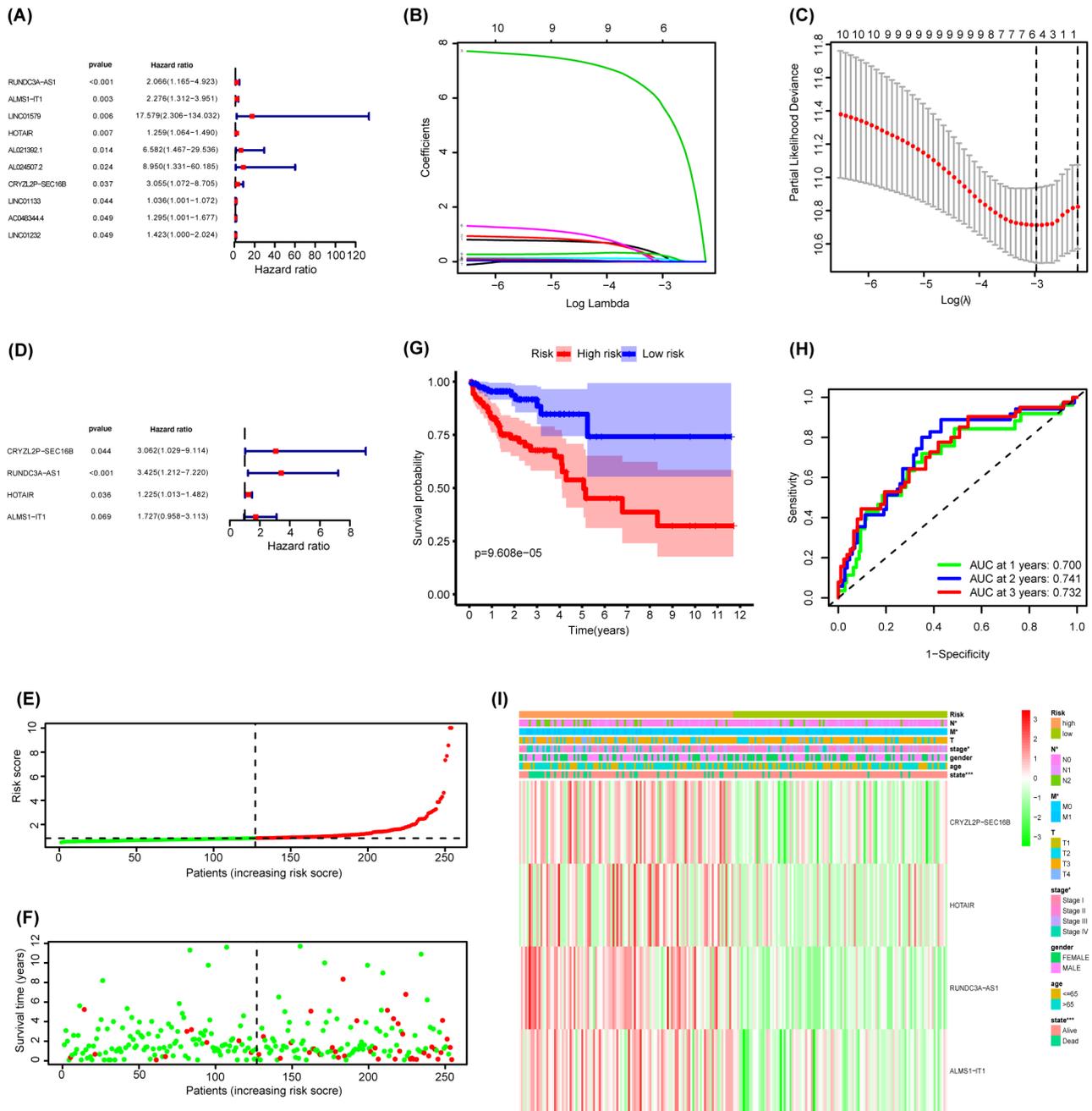
Fig. 3 Establishment of Glycolysis-Related lncRNA Signature. **(A)** 10 glycolysis-related lncRNAs were extracted correlated with prognosis of CRC patients by univariable Cox regression analysis; **(B)** LASSO coefficient profiling; **(C)** Cross-validation for turning parameter screening in the LASSO regression model; **(D)** Multivariable Cox regression analysis was conducted to select lncRNAs independently affected OS; **(E-F**, **I)** The distribution of risk score, survival time and lncRNA expression of CRC patients stratified by risk score; **(G)** Kaplan-Meier survival curves for high and low-risk score groups.; **(H)** Kaplan-Meier survival analysis of CRC patients stratified by risk score; **(I)** ROC curve of predicted 1-, 2-, 3-year OS according to risk score

The Kaplan-Meier log-rank test confirmed that patients with stage II colorectal cancer in different risk group have different prognosis (*P*<0.05, Figure S1C). The AUCs for 1-year, 2-year and 3- year OS were 0.703, 0.689 and 0.637 respectively (Figure S1D).

Construction of a nomogram

As tumor stage, age and gender may have an impact on the prognosis of patients with CRC. We took the factors above and the risk score as candidates consisting of nomogram. The results of univariate and multivariate Cox-regression analyses showed that, for patients with CRC, the prognostic signature (*P*<0.001), age (*P*=0.004)

and stage ($P<0.001$) are independent prognostic parameters (Fig. 4A-B), which we selected as the composition of the nomogram (Fig. 4C). Then, the total risk score was calculated according to each independent prognostic parameter in the nomogram. Based on the median cutoff of the total risk score we classified the patients into high- and low-risk groups. Patients in high-risk group had a poorer OS time than in the low-risk group ($P=9.778e-07$, Fig. 4D). We use ROC and calibration curve to evaluate the predictive accuracy of the nomogram. ROC analysis showed that the AUC values for 1-, 3-, and 5-year survival of CRC patients predicted by the total score were 0.835, 0.830, and 0.816, respectively (Fig. 5A-C), which also indicated that the nomogram had a better predictive accuracy than any of a single clinical parameter. The gray lines of the calibration curve presented the actual observation, and the red lines are the nomogram prediction of OS time (Fig. 5D-F), from which we can see a strong consistency. Similarly, the decision curves showed that the nomogram maximized clinical net benefits for patients in the prediction of OS time at 1-, 3-, 5-year (Fig. 5G-I).

Validation of the signature and nomogram

The performance of the signature and nomogram model was validated using the validation cohort. First, the formula mentioned above was used to calculate the risk score of each patient in the internal validation cohorts. The risk score being an independent factor of the prognosis of CRC was proved again by the validation cohort

(Fig. 6A-B). We separated patients in the validation cohort into high- and low- risk groups according to the same method mentioned above. The OS time of different groups showed significant difference (Fig. 6C). ROC analysis showed that the AUC values for 1-, 3-, and 5-year survival of CRC patients predicted by the total score were 0.713, 0.749, and 0.700, respectively (Fig. 6D-F) calibration curve was performed, indicating the good predictive power of the signature and nomogram (Fig. 6G-I).

Patients with stage II colorectal cancer in validation cohort were also separated into high and low-risk groups (FigureS1E-F). The OS time of different groups showed significant difference (Figure S1G). In the validation cohort, the AUCs for 1-year, 2-year and 3- year OS were 0.753, 0.702 and 0.629 respectively (Figure S1H).

Effect of *lncRNA RUNDC3A − AS1* on CRC cells

To validate the lncRNAs in the signature we constructed were glycolysis-related, we performed the glucose consumption and lactic acid production assays. In the risk score formula, lncRNA *RUNDC3A−AS1* has the maximum coefficients, indicating that the expression of lncRNA *RUNDC3A−AS1* has the highest impact among our signature on the OS of CRC patients. So, we chose to exam the effect of lncRNA *RUNDC3A−AS1* on the Warburg effect of CRC. We achieved lncRNA *RUNDC3A−AS1* knockdown to investigate its cellular function by transfecting SW480 and SW620 cells with si-*RUNDC3A−AS1*(lncRNA-KD), as confirmed by qPCR (Fig. 7A). LncRNA *RUNDC3A−AS1* silence reduced the
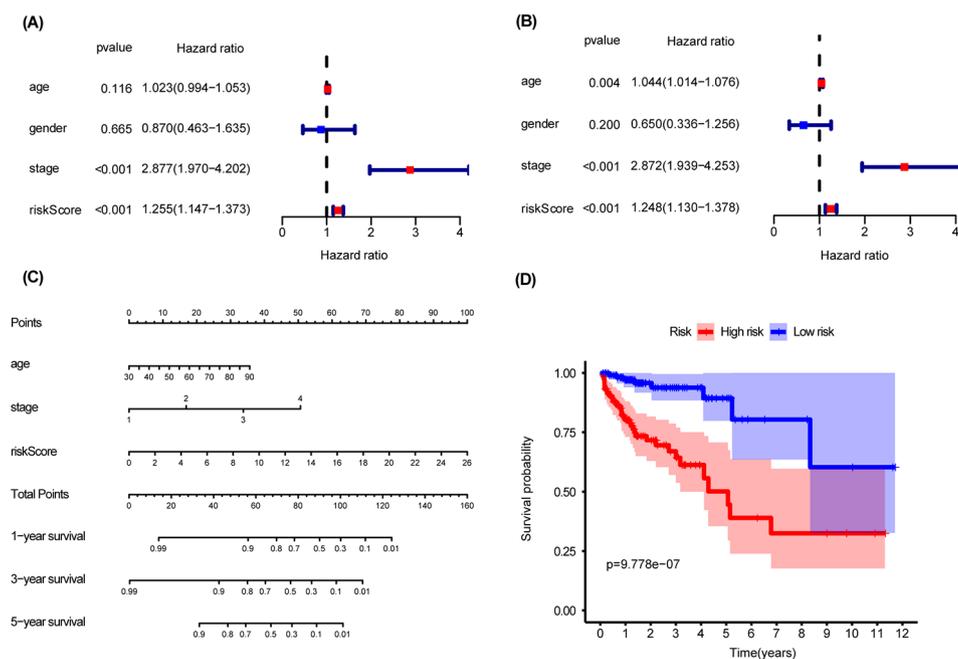
Fig. 4 Construction of a Nomogram. (**A**-**B**) Univariable and Multivariable Cox regression analyses were performed to select independent factors influencing CRC prognosis; (**C**)The predictive nomogram; (**D**) Kaplan-Meier curves of CRC patients in train cohort stratified by the nomogram
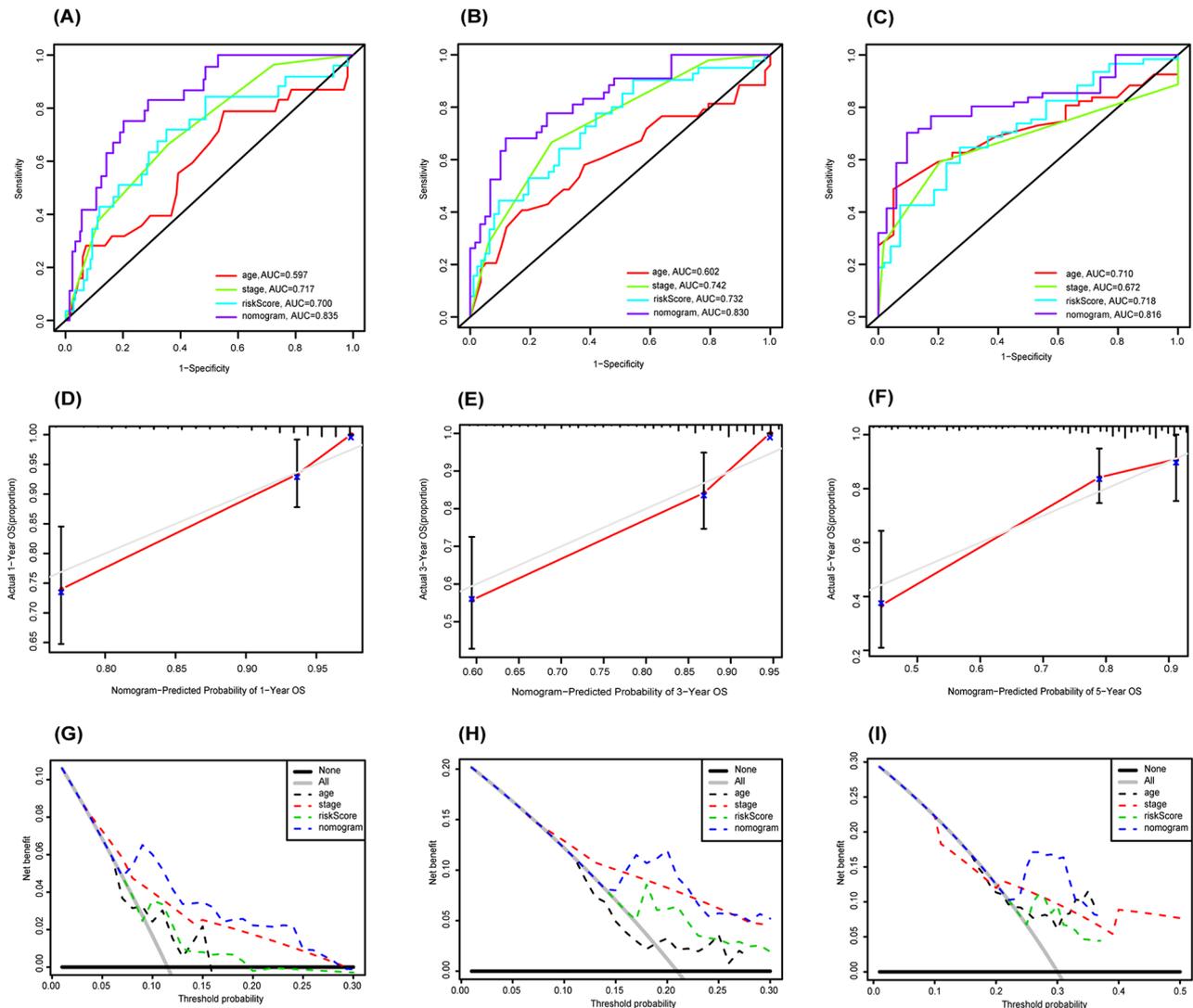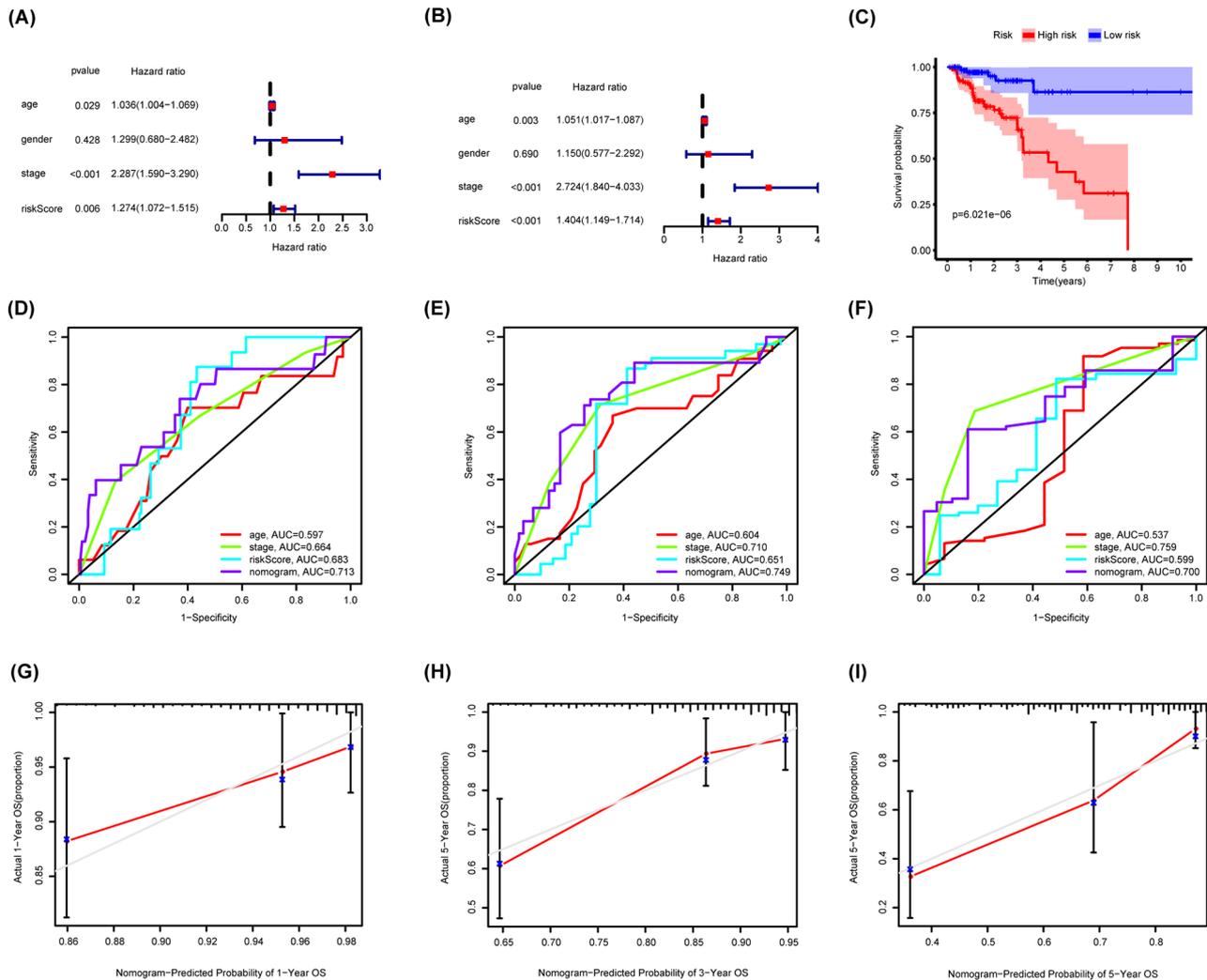
Fig. 5 Evaluating the predictive accuracy of the nomogram. (**A**-**C**) ROC curve of predicted 1-, 3-, 5-year OS according to nomogram and other clinical characteristics. (**D**-**F**) Nomogram calibration plot of 1-, 3-, and 5-year OS in the train cohort. (**G**-**I**) Depict the 1-, 3-, and 5-year Decision Curve Analysis (DCA) curves, respectively

lactate production while increased the glucose level in the CRC cells culture medium (Fig. 7B-C). Consistently, in lncRNA-KD SW480 and SW620 cells, the mRNA and protein levels of *HK2* were significantly down-regulated (Fig. 7D-E), indicating a positive correlation between *HK2* and lncRNA *RUNDC3A−AS1* expression in CRC cells. In addition, we conducted colony formation assays to detect the colony formation ability of lncRNA *RUNDC3A−AS1* in colorectal cancer. Knockdown of lncRNA *RUNDC3A−AS1* significantly decreased the mean colony number in the colony formation assay (*P* < 0.001) (Fig. 7F-G). These findings indicate that *lncRNA RUNDC3A−AS1* knockdown could suppress the glycolysis and colony formation ability in CRC cells.

Discussion

The proliferation of tumor cells fundamentally relies on an adequate supply of biological macromolecules and energy. Consequently, enhanced glycolysis emerges as a distinct hallmark of cancer cells [11]. Our research has shown that colorectal cancer (CRC) patients exhibiting significantly elevated glycolysis activity are associated with advanced tumor stages and poorer prognoses, as determined through the calculation of Gene Set Variation Analysis (GSVA) enrichment scores for each patient. Glucose serves as the primary energy substrate in the human body, metabolized through pathways including glycolysis and aerobic oxidation. During the oncogenesis and progression of CRC, notable expression changes occur in key enzymes involved in the glycolysis pathway [12]. For instance, glucose transport protein 1 (GLUT1), a key

Fig. 6 Validation of the nomogram. (**A**-**B**) Univariable and Multivariable Cox regression analyses showed that the risk signature was an independent factor for prognosis in the validation cohort; (**C**) Kaplan-Meier curves of CRC patients in the validation cohort stratified by the nomogram; (**D**-**F**) ROC curve of predicted 1-, 3-, 5-year OS according to nomogram and other clinical characteristics in the validation cohort; (**G**-**I**) Nomogram calibration plot of 1-, 3-, and 5-year OS in the validation cohort

facilitator of glucose transport, is upregulated in CRC, thereby enhancing glucose absorption [13]. Conversely, the knockdown of pyruvate kinase M2 (PKM2) in cancer cells leads to a reduction in glucose uptake [14, 15]. Hexokinase II (HK2) plays a pivotal role in tumorigenesis by facilitating the Warburg effect [16]. These variations form a complex network that regulates cancer cell glucose metabolism, with cancer cells exhibiting increased glycolytic activity even in oxygen-rich environments [4]. Collectively, this evidence underscores the critical role of dysregulated glycolysis in colorectal cancer.Recent investigations have illuminated the role of long noncoding RNAs (lncRNAs) in modulating the expression of protein-coding genes and influencing cancer-related biological processes, including the Warburg effect [17–19]. Nonetheless, the mechanisms through which lncRNAs regulate the Warburg effect in colorectal cancer (CRC)

remain inadequately explored. Given that one of the primary mechanisms by which lncRNAs exert their regulatory effects is through acting as endogenous microRNA (miRNA) sponges [8, 20], starting with glycolysis-related differentially expressed mRNAs to construct a competing endogenous RNA (ceRNA) network presents an efficient strategy for identifying glycolysis-associated lncRNAs. Employing this approach, we identified 101 lncRNAs potentially involved in modulating the Warburg effect in CRC, marking a pioneering effort in comprehensively screening for glycolysis-related lncRNAs.

To pinpoint lncRNAs with a strong correlation to prognosis, we conducted univariable, multivariable, and LASSO regression analyses. This process culminated in the establishment of a prognostic signature consisting of four lncRNAs (CRYZL2P−SEC16B, HOTAIR, ALMS1−IT1, and RUNDC3A−AS1). A risk score
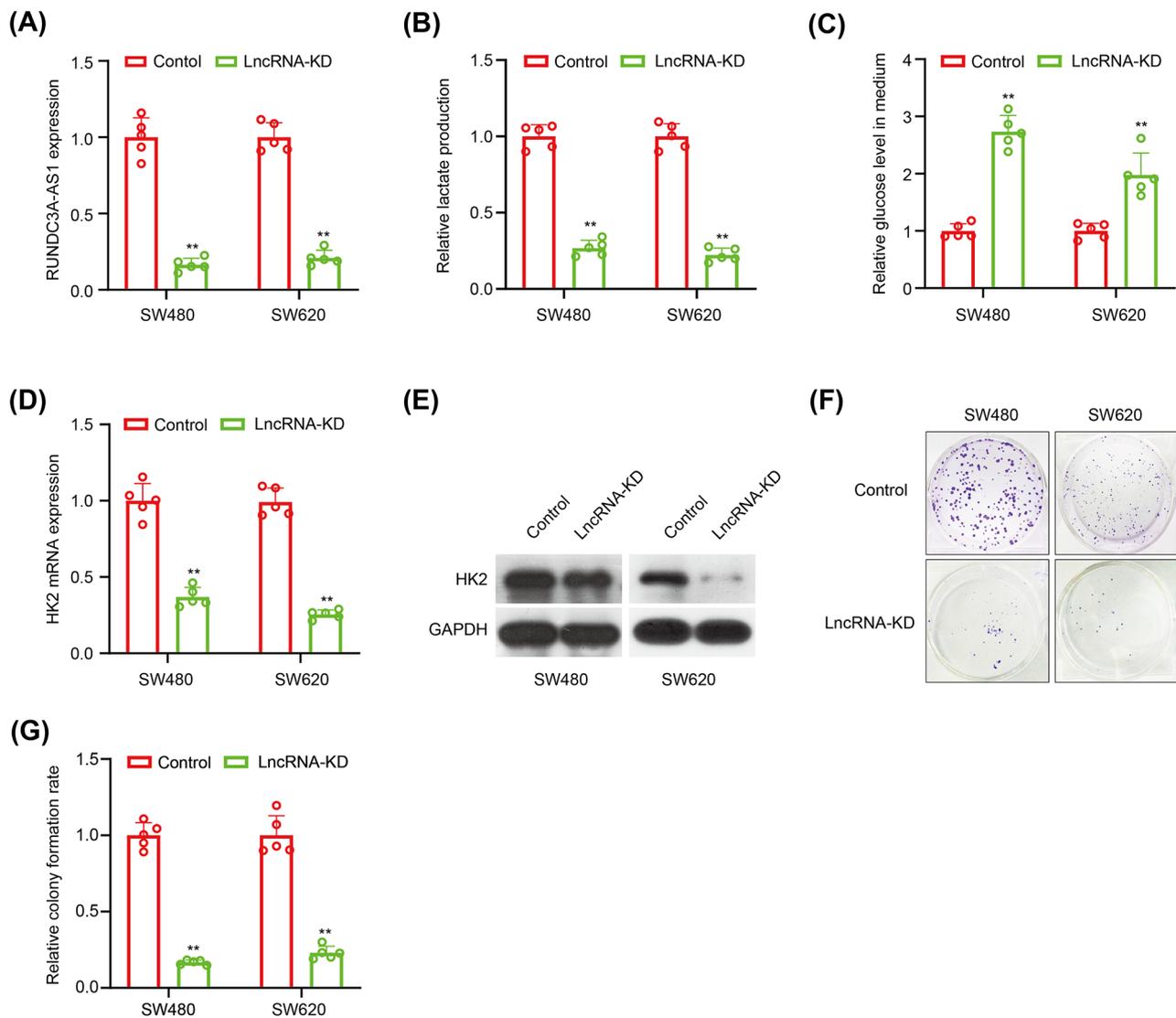
Fig. 7 Effect of lncRNA RUNDC3A−AS1 on CRC cells. **(A)** The silence of lncRNA RUNDC3A−AS1 was achieved in SW480 and SW460 cells by transfection of si-RNA for lncRNA RUNDC3A−AS1, as confirmed by real-time PCR; **(B-C)** The effects of si- RUNDC3A−AS1 inhibitor on glucose consumption and cellular lactate production levels in CRC cells; **(D-E)** HK2 mRNA expression and protein levels in lncRNA RUNDC3A−AS1-silenced colorectal cancer cells examined by real-time PCR and western-blotting analysis. **(F-G)** Downregulation of lncRNA *RUNDC3A−AS1* reduced the mean colony number in the colony formation assay. $N_{case} = 5$, $N_{control} = 5$

derived from this signature was calculated, aligning with our hypothesis, emerged as an independent prognostic factor (Fig. 4). Acknowledging the significance of clinicopathological features in CRC patient outcomes, we developed a nomogram that integrates these features with our signature to enhance prognostic accuracy. The efficacy of this signature was validated through receiver operating characteristic (ROC) curves, calibration curves, and decision curves, showcasing its innovative approach in employing glycolysis-related lncRNAs. Compared to existing models [21–25], our prognostic model demonstrated superior predictive reliability in both training and validation cohorts. A notable limitation of current staging systems is their inability to account for the clinical

outcome heterogeneity among patients within the same stage, especially evident in stage II CRC. Our findings reveal that the prognosis significantly diverges between high-risk and low-risk groups among stage II CRC patients based on the risk score, underscoring the utility of our signature in identifying high-risk individuals within this subgroup.Every lncRNA identified in our prognostic signature serves as an independent prognostic factor for colorectal cancer (CRC) patient outcomes. The regulatory role of lncRNA CRYZL2P−SEC16B is yet to be reported, setting a precedent for future investigations. Within our model, several lncRNAs have been previously associated with various cancers. For instance, Hu et al. highlighted that the overexpression of lncRNA

Mao *et al. BMC Medical Genomics*   (2024) 17:112

Page 11 of 12

HOTAIR enhances glycolysis in hepatocellular carcinoma (HCC) cells [25, 26], aligning with our predictions. ALMS1-IT1 was found by Luan et al. to potentially accelerate the malignant progression of lung adenocarcinoma by activating the cyclin-dependent kinase pathway [26, 27]. RUNDC3A−AS1 exhibited significant differential expression in papillary thyroid carcinoma, playing a crucial role in cancer histotype determination [27, 28]. Yet, its involvement in modulating the Warburg effect in cancer has not been elucidated.

In our signature, lncRNA RUNDC3A−AS1 had the highest coefficient, indicating its paramount influence on CRC prognosis. This led us to investigate RUNDC3A−AS1's function in CRC glycolysis further. Silencing RUNDC3A−AS1 markedly reduced lactate production and glucose uptake in CRC cell lines in vitro, highlighting its potential regulatory role. Given that hexokinase, especially HK2, is pivotal in glycolysis and is overexpressed in various malignancies [15, 16, 29, 30], we examined HK2 expression post-silencing RUNDC3A−AS1. A decrease in HK2 expression in CRC cells upon downregulating RUNDC3A−AS1 underscores its probable regulatory mechanism involving HK2 and glycolysis in CRC, warranting further mechanistic studies.

In summary, our research represents the inaugural comprehensive screening of glycolysis-related lncRNAs in CRC. Utilizing these lncRNAs, we developed a four-lncRNA signature and a predictive nomogram for CRC outcomes, validated by extensive analyses for its high predictive accuracy. By downregulating lncRNA RUNDC3A−AS1, we demonstrated its significant role in modulating the Warburg effect in CRC cells, contributing novel insights into CRC glycolysis regulation.

Abbreviations

CRC	colorectal adenocarcinoma
TCGA	The Cancer Genome Atlas
GEO	Gene Expression Omnibus
ROC	receiver operating characteristic
OS	overall survival
AUC	the area under the curve
KEGG	Kyoto Encyclopedia of Genes and Genomes
qRT-PCR	Real-time quantitative reverse transcription polymerase chain reaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-024-01862-2 .

Supplementary Material 1

Author contributions

RM, CX, and QZ designed the whole study. CX, PY, and QZ conducted the bioinformatics analyses. ML, KZ and RM conducted the in vitro experiments. CX and RM wrote the manuscript. MW, RM, and ZW helped to improve the manuscript. CX, RM, and QZ contributed equally to this work. ML, YL and PY contributed equally to this work. All authors contributed to the article and approved the submitted version.

Data availability

All data and materials are available from the corresponding author on reasonable request. Interested parties should contact Rui Mao, 218102100@csu.edu.cn for data access.

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Ethics Review Board of the Third People's Hospital of Chengdu (record #:2018S75; Chengdu, Sichuan, China), and was conducted in accordance with the Chinese ethical guidelines for human genome/gene research. informed consent/waiver of informed consent (Institutional Ethics Review Board of the Third People's Hospital of Chengdu) was obtained from all subjects and/or their legal guardian(s).

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.

Author details

[1]Hunan Key Laboratory of Aging Biology, Xiangya Hospital, Central South University, Changsha, China
[2]Center of Gastrointestinal and Minimally Invasive Surgery, Department of General Surgery, The Second Affiliated Hospital of Chengdu, The Third People's Hospital of Chengdu, Affiliated Hospital of Southwest Jiaotong University, Chongqing Medical University, NO.82 Qinglong Road, Chengdu, Sichuan 610031, China
[3]Center of Obesity and Metabolism disease, Department of General surgery, The Second Affiliated Hospital of Chengdu, The Third People's Hospital of Chengdu, Affiliated Hospital of Southwest Jiaotong University, Chongqing Medical University, Chengdu 610031, China
[4]Department of Critical Care Medicine, Chengdu Third People's Hospital, Chengdu 610031, China
[5]Department of Colorectal Surgery, National Clinical Research Center for Cancer, Cancer Hospital, National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[6]Department of hepatobiliary surgery, The Second Affiliated Hospital of Chengdu, The Third People's Hospital of Chengdu, Affiliated Hospital of Southwest Jiaotong University, Chongqing Medical University, NO.82 Qinglong Road, Chengdu, Sichuan 610031, China

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. 21492. PubMed PMID: 30207593. IF: 254.7 Q1 B1.
2. Sveen A, Kopetz S, Lothe RA. Biomarker-guided therapy for colorectal cancer: strength in complexity. Nat Rev Clin Oncol. 2020;17(1):11–32. https://doi.org/10.1038/s41571-019-0241-1. Epub 2019/07/11.

3.   Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. Nat Rev Gastroenterol Hepatol. 2019;16(12):713–32. https://doi.org/10.1038/s41575-019-0189-8. Epub 2019/08/29.

4.   Warburg O. On the origin of cancer cells. Science. 1956;123(3191):309–14. https://doi.org/10.1126/science.123.3191.309. Epub 1956/02/24.

5.   Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74. https://doi.org/10.1016/j.cell.2011.02.013. Epub 2011/03/08.

6.   Fang Y, Shen ZY, Zhan YZ, Feng XC, Chen KL, Li YS, et al. CD36 inhibits β-catenin/c-myc-mediated glycolysis through ubiquitination of GPC4 to repress colorectal tumorigenesis. Nat Commun. 2019;10(1):3981. https://doi.org/10.1038/s41467-019-11662-3. Epub 2019/09/06.

7.   Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. Cell. 2013;152(6):1298–307. https://doi.org/10.1016/j.cell.2013.02.012. Epub 2013/03/19.

8.   Wang L, Cho KB, Li Y, Tao G, Xie Z, Guo B. Long Noncoding RNA (lncRNA)-Mediated Competing Endogenous RNA Networks Provide Novel Potential Biomarkers and Therapeutic Targets for Colorectal Cancer. *Int J Mol Sci* (2019) 20(22). Epub 2019/11/21. https://doi.org/10.3390/ijms20225758. PubMed PMID: 31744051. IF: 5.6 Q1 B2; PubMed Central PMCID: PMCPMC6888455. IF: 5.6 Q1 B2.

9.   Meng F, Luo X, Li C, Wang G. LncRNA LINC00525 activates HIF-1α through miR-338-3p / UBE2Q1 / β-catenin axis to regulate the Warburg effect in colorectal cancer. Bioengineered. 2022;13(2):2554–67. PMID: 35156520. IF: 4.9 Q1 B4; PMCID: PMC8973709. IF: 4.9 Q1 B4.

10. Zhu J, Wang S, Bai H, Wang K, Hao J, Zhang J et al. Identification of Five Glycolysis-Related Gene Signature and Risk Score Model for Colorectal Cancer. *Front Oncol* (2021) 11:588811. Epub 2021/03/23. https://doi.org/10.3389/fonc.2021.588811. PubMed PMID: 33747908. IF: 4.7 Q2 B3; PubMed Central PMCID: PMCPMC7969881. IF: 4.7 Q2 B3.

11. Liberti MV, Locasale JW. The Warburg Effect: how does it Benefit Cancer cells? Trends Biochem Sci. 2016;41(3):211–8. https://doi.org/10.1016/j.tibs.2015.12.001. Epub 2016/01/19.

12. Schell JC, Olson KA, Jiang L, Hawkins AJ, Van Vranken JG, Xie J et al. A role for the mitochondrial pyruvate carrier as a repressor of the Warburg effect and colon cancer cell growth. *Mol Cell* (2014) 56(3):400– 13. https://doi.org/10.1016/j.molcel.2014.09.026. Epub 2014/12/03. IF: 16.0 Q1 B1. PubMed PMID: 25458841. IF: 16.0 Q1 B1; PubMed Central PMCID: PMCPMC4268416. IF: 16.0 Q1 B1.

13. Zhang ZJ, Zhang YH, Qin XJ, Wang YX, Fu J, Circular. RNA circDENND4C facilitates proliferation, migration and glycolysis of colorectal cancer cells through miR-760/GLUT1 axis. Eur Rev Med Pharmacol Sci. 2020;24(5):2387–400. Epub 2020/03/21. https://doi.org/10.26355/eurrev_202003_20506. IF: 3.3 Q3 B4. PubMed PMID: 32196590. IF: 3.3 Q3 B4.

14. Wong N, Ojo D, Yan J, Tang D. PKM2 contributes to cancer metabolism. Cancer Lett. 2015;356(2 Pt A):184–91. https://doi.org/10.1016/j.canlet.2014.01.031. Epub 2014/02/11. PubMed PMID: 24508027. IF: 9.7 Q1 B1.

15. Zhang W, Zhang X, Huang S, Chen J, Ding P, Wang Q, et al. FOXM1D potentiates PKM2-mediated tumor glycolysis and angiogenesis. Mol Oncol. 2021;15(5):1466–85. https://doi.org/10.1002/1878-0261.12879. Epub 2020/12/15.

16. Shangguan X, He J, Ma Z, Zhang W, Ji Y, Shen K, et al. SUMOylation controls the binding of hexokinase 2 to mitochondria and protects against prostate cancer tumorigenesis. Nat Commun. 2021;12(1):1812. https://doi.org/10.1038/s41467-021-22163-7. Epub 2021/03/24.

17. Liu H, Luo J, Luan S, He C, Li Z. Long non-coding RNAs involved in cancer metabolic reprogramming. Cell Mol Life Sci. 2019;76(3):495–504. https://doi.org/10.1007/s00018-018-2946-1. Epub 2018/10/21.

18. Liao M, Liao W, Xu N, Li B, Liu F, Zhang S et al. LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. *EBioMedicine* (2019) 41:200– 13. Epub 2019/02/24. https://doi.org/10.1016/j.ebiom.2019.01.035. PubMed PMID: 30796006. IF: 11.1 Q1 B1; PubMed Central PMCID: PMCPMC6444057. IF: 11.1 Q1 B1.

19. Li N, Zhan X, Zhan X. The lncRNA SNHG3 regulates energy metabolism of ovarian cancer by an analysis of mitochondrial proteomes. Gynecol Oncol. 2018;150(2):343–54. https://doi.org/10.1016/j.ygyno.2018.06.013. Epub 2018/06/21.

20. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. Nature. 2014;7483505. https://doi.org/10.1038/nature12986. Epub 2014/01/17.

21. Song W, Ren J, Wang C, Ge Y, Fu T. Analysis of Circular RNA-Related Competing Endogenous RNA Identifies the Immune-Related Risk Signature for Colorectal Cancer. *Front Genet* (2020) 11:505. Epub 2020/06/26. https://doi.org/10.3389/fgene.2020.00505. PubMed PMID: 32582276. IF: 3.7 Q2 B3; PubMed Central PMCID: PMCPMC7283524. IF: 3.7 Q2 B3.

22. Song W, Fu T. Circular RNA-Associated, competing endogenous RNA network and prognostic nomogram for patients with Colorectal Cancer. Front Oncol. 2019;9(1181). https://doi.org/10.3389/fonc.2019.01181. Epub 2019/11/30. PubMed PMID: 31781492. IF: 4.7 Q2 B3; PubMed Central PMCID: PMCPMC6857072. IF: 4.7 Q2 B3.

23. Long J, Xiong J, Bai Y, Mao J, Lin J, Xu W, et al. Construction and investigation of a lncRNA-Associated ceRNA Regulatory Network in Cholangiocarcinoma. Front Oncol. 2019. https://doi.org/10.3389/fonc.2019.00649. 9:649. Epub 2019/08/27.

24. Yu S, Hu C, Cai L, Du X, Lin F, Yu Q, et al. Seven-gene signature based on glycolysis is closely related to the prognosis and Tumor Immune infiltration of patients with gastric Cancer. Front Oncol. 2020;10:1778. https://doi.org/10.3389/fonc.2020.01778. PubMed PMID: 33072557. IF: 4.7 Q2 B3; PubMed Central PMCID: PMCPMC7531434. IF: 4.7 Q2 B3. Epub 2020/10/20.

25. Hou C, Cai H, Zhu Y, Huang S, Song F, Hou J. Development and Validation of Autophagy-Related Gene Signature and Nomogram for Predicting Survival in Oral Squamous Cell Carcinoma. *Front Oncol* (2020) 10:558596. Epub 2020/11/13. https://doi.org/10.3389/fonc.2020.558596. PubMed PMID: 33178587. IF: 4.7 Q2 B3; PubMed Central PMCID: PMCPMC7596585. IF: 4.7 Q2 B3.

26. Hu M, Fu Q, Jing C, Zhang X, Qin T, Pan Y. LncRNA HOTAIR knockdown inhibits glycolysis by regulating miR-130a-3p/HIF1A in hepatocellular carcinoma under hypoxia. Biomed Pharmacother. 2020;125:109703. https://doi.org/10.1016/j.biopha.2019.109703. Epub 2020 Feb 13. PMID: 32062551. IF: 7.5 Q1 B2.

27. Luan T, Zhang TY, Lv ZH, Guan BX, Xu JY, Li J, Li MX, Hu SL. The lncRNA ALMS1-IT1 may promote malignant progression of lung adenocarcinoma via AVL9-mediated activation of the cyclin-dependent kinase pathway. FEBS Open Bio. 2021;11(5):1504–15. Epub 2021 Apr 3. PMID: 33683834. IF: 2.6 Q4 B4; PMCID: PMC8091588. IF: 2.6 Q4 B4.

28. Guo K, Chen L, Wang Y, Qian K, Zheng X, Sun W, Sun T, Wu Y, Wang Z. Long noncoding RNA RP11-547D24.1 regulates proliferation and migration in papillary thyroid carcinoma: identification and validation of a novel long noncoding RNA through integrated analysis of TCGA database. Cancer Med. 2019;8(6):3105–19. Epub 2019 May 1. PMID: 31044550. IF: 4.0 Q2 B2; PMCID: PMC6558462. IF: 4.0 Q2 B2.

29. Xu S, Herschman HR. A tumor agnostic therapeutic strategy for hexokinase 1-Null/Hexokinase 2-Positive cancers. Cancer Res. 2019;79(23):5907–14. https://doi.org/10.1158/0008-5472.Can-19-1789. Epub 2019/08/23.

30. Yang L, Yan X, Chen J, Zhan Q, Hua Y, Xu S et al. Hexokinase 2 discerns a novel circulating tumor cell population associated with poor prognosis in lung cancer patients. *Proc Natl Acad Sci U S A* (2021) 118(11). Epub 2021/04/11. https://doi.org/10.1073/pnas.2012228118. PubMed PMID: 33836566. IF: 11.1 Q1 B1; PubMed Central PMCID: PMCPMC7980452. IF: 11.1 Q1 B1.

Publisher's Note