MEDS Basic Medicine (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/medbm.2024.020210
ISSN 2616-2091 Vol. 2 Num. 2

Construction of Colorectal Cancer Prognostic Model Utilizing Mitochondrial Energy Metabolism-Related Genes

Kai Wang[1,2,a,#], Peng Zhu[1,b,#], Zhenghui Yang[3,c], Guoping Sun[1,d,*]

[1]Department of Clinical Laboratory, Pingshan Hospital, Southern Medical University (Pingshan District People's Hospital of Shenzhen), Shenzhen, China

[2]Department of Shu Lan International Medical, Zhejiang Shu Ren University, Hangzhou, China

[3]Hangzhou QuanMei Dental Clinic, Hangzhou, China

[a]WangKai@zjsru.edu.cn, [b]b00427@zjsru.edu.cn, [c]630447078@qq.com, [d]1131156578@qq.com

[*]Corresponding author

[#]Co-first author

Keywords: Mitochondrial energy metabolism, Colorect cancer, Prognosis model, Tumor microenvironment

Abstract: The objective of this study was to construct a prognostic model and medicine therapeutic response by utilizing mitochondrial energy metabolism-related genes (MMRGs), thus establishing a risk score for colorectal cancer (CRC). Based on the TCGA-CRC and GEO data set, MMRGs expression levels were identified by clustering analysis. 10 differential expression genes were used to construct RiskScore by Cox regression. GSE 39582 data set was used for validation. The clinical characteristics, survival characteristics, SNV, CNV, methylation, immune features, and potential benefits of chemotherapy drugs were analyzed for two risk groups. RiskScore was constructed based on the genes ACOX1, ATP6V1G2, COX7A1, CPT2, DLAT, ECGS1, ECI2, NDUFA1, PPA2, and SUCLG2. Patients in the low risk group exhibited a superior overall survival. In addition, Univariate Cox regression analysis and Multivariate Cox regression analysis demonstrated that the risk score, stage and lymphatic invasion can serve as the independent prognostic factors. Mitochondrial Energy -Related Genes was a promising biomarker that can be used to distinguish CRC prognosis, immune features, and sensitivity to chemotherapy drugs.

1. Introduction

Colorectal cancer (CRC) is the second significant contributor to cancer-related mortality globally. The high incidence and mortality due to its related with bad lifestyles , unhealthy dietary factors ,smoking , alcohol consumption genetic and other factors[1-2] While with the varity development of clinic treatments ,the incidence of CRC has been declining in recent years, there has been a significant increase in the death rate of CRC individuals in China[3].In light of this, researchers persist in undertaking novel studies to enhance the efficacy of CRC treatment and prognostic outcomes for individuals .In many studies, advanced clinical manifestation and molecular characteristics is related with colorectal cancer[1-2],including biomarkers ,mutations , methylation, Cellular senescence or signals biomarker signatures[4-6].

Cellular senescence was demonstrated that was connected with CRC or changed of metabolic pathways to inhibit cancer progression[6].Senescent cells are characterized by need sufficient energy like mitochondrial to synthesize various metabolish pathway[7].Recently studies have shown that Mitochondrial disorders have been discoved in varity tumor cells and have demonstrated various mitochondrial proteins as pathological mechanism part to play a important role[8].In support of this. To change the number of mitochondrial content and activity, can affect the growth of cell differentiation and toward oxidative phosphorylation in metabolism[8].The mitochondrial energy metabolism pathway can regulate stem cell functions through many mechanisms, including citrate cycle tca cycle, fatty acid metabolism , ketone body metabolism , glycolysis gluconeogenesis, oxidative phosphorylation[9].Mitochondrial dysfunction was demonstrated that was involved in cellular senescence, and may changed the metabolic pathways to favor carcinogenesis[10].Such as a high expression gene GAPDHs in the mitochondrial glycolytic pathway was associated with worse clinical outcome in stage III melanoma [10]. Accumulating evidence has proven molecular testing can help to diagnose and intervent CRC, improve patient prognosis, reduce mortality, and lower the economic burden of disease on individuals[5,11]. Hence, investigations into the metabolic aspects of mitochondrial are anticipated to offer valuable insights for prognostic predictions and optimization of treatment strategies for individuals with CRC.

Nevertheless, at present, there is a lack of research exploring the correlation between mitochondrial energy metabolism pathway-related genes (MMRGs) and the prognosis of CRC patients. Therefore, differentiating clustering features and establishing features related to MMRGs may be effective in predicting the prognosis and immunotherapy response of CRC patients.We plan to utilize CRC samples from public databases to study the linkage between MMRGs and clinical features and prognosis of CRC. We established a prognostic model through regression analysis and conducted an independent analysis of the prognosis. We also conducted tumor mutation analysis, immunotherapy response, and chemotherapy drug sensitivity analysis, aiming to contribute

additional data support toward improving the prognosis and treatment options available for individuals with CRC.

2. Method

2.1 Data source and acquisition

TCGA-COAD was downloaded in the TCGA database from University of California Santa Cruz (UCSC) was used to get the latest clinical follow-up information and corresponding STAD, COAD, and READ gene expression data. We downloaded GSE39582 data from National Center for Biotechnology Information (NCBI), containing 562 samples, respectively. The following steps were performed for pretreatment: 1) Remove samples with no clinical data or uncompleted data. 2) Remove samples of normal tissue. 3) Remove the gene with a transcript per million (TPM) <0.01 in half the samples.

Mitochondrial energy metabolism pathway-related genes (MMRGs) were derived from KEGG and MsigDB databases(Version 7.5.1). We downloaded MMRG corresponding to pathways containing keywords such as KEGG _ GLYCOLYSIS _ GLUCONEOGENESIS, KEGG_ CITRATE _ CYCLE _ TCA _ CYCLE, KEGG _ OXIDATIVE_PHOSPHORYLATION, REACTOME _ KETONE _ BODY _ METABOLISM, KEGG _FATTY_ ACID _ METABOLISM, respectively. It was verified that the differentially gene expression level of mitochondrial metabolism in rectal canner data obeyed a normal distribution, subsequently, we matched the clinical feather and gene expression (include genomic alteration , methylation analysis , survival expression ,immunotherapy response) , and then verificate the independent risk factor and classifed those differentlly express genes with expression above the average level as high expression group, and those differently express genes with expression below the average level as low expression group to construct nomogram predict the dignosis value.

2.2 Mitochondrial energy metabolism pathway-related genes (MMRGs) express analysis of rectal cancer

Differential analysis of the GEO dataset and TCGA-CRC dataset (log2 (TPM + 1)) was conducted to screen the DEGs using the R ``limma" package with FDR < 0.05 as the threshold. According to the type of dataset. We divide the sample into adjacent and cancerous tissues, and calculate the differentially expressed genes using the Bayes method of limma (V3.50.3) package. Data between the two groups were compared using the Wilcoxon rank sum test. The R clusterProfiler(V4.2.2)"pheatmap" and ``ggplot2″ software packages were used to plot heat maps and GSEA anylysis to explore and understand the mitochondrial metabollism in rectal cancer.All analyses in this study were performed under R cluster Profiler version 4.2.2

2.3 Genomic alteration and methylation analysis

According the MMRG, we used maftools(v2.10.05)to analyze the genomic alterations of mitochondrial metabolism in rectal canner , evaluate the expression of MMRG correlated with copy-number variation (CNV) and DNA methylation, and used ggplot2(v3.3.5) Fisher's statistical test method to visualize the genomic alteration of DEGs. The DEGs was selected for genomic profle change analysis, including mutations, putative CNAs from genomic identifcation of signifcant targets of mitochondrial metabolism in rectal canner, and structural variants. Additionally, we analyzed the Spearman's correlations between DNA methylation level and its expression in MMRG by using R ChAMP package (version 2.26.0).

2.4 Correlation analysis between MMRG and clinical characteristics

According to the MMRG ,we used T test method to assess the differentially gene expression level between paracancer and cancerous tissues;We used Kruskal-Wallis rank test to value the connection between DEGs and clinical characteristics(including Age, Gender, Grade,BMI, histological type,TNM stage, Tumor stage,and survival) .We categorized DEGS into the high-risk group and the low-risk group, based on the median values of gene expression values and to calculate the difference between two group by using software R survival package(v3.3-1). Then,the K-M curve was drawn to evaluate the prognostic value of the risk model by using software R survminer package(v0.4.9).

2.5 Immune microenvironment and immune relative analysis

To investigate whether the immune microenvironment relative with MMRG and has the ability to assist in affecting mitochondrial metabolism pathway in rectal canner.We used softwareCIBERSORT, GSVA(v1.42.0),xCell(v1.1.0)and estimate(v1.0.13) to calculate the correlation between the expression of the MMRG core genes and the immune infiltration score. Further research was conducted to determine whether the signature affects the clinical outcomes of patients with comparable immune infiltration gene expression levels.

2.6 Constructure and Validate MMRGs differentilly gene express signature

In order to understand the gene and protein interaction networks of MMRG were constructed using online STRING database (http://string-db.org) and using software STRINGdb(v2.6.5) to complish the gene-set enrichment analysis to understand the biological processes of MMRG. In order to achieve the differentially gene express signature ,we used software glmnet(v4.1-4)to constructure LASSO-Cox model.Further,the differentially gene express signature of MMRG was

identified the independence of the clinical factors(P value<0.05) (age,stage,gender,tumor location,perineural invasion,venous invasion,lymphatic invasion)and the prognostic risk model through the univariate Cox regression analysis and the multivariate Cox regression analysis by software R packagesurvival(v3.3-1).

2.7 Statistical analysis

All data analysis was conducted in the R environment (version 4.1.0). The Student's T test or one-way ANOVA was used to analyze group differences in variables with normal distribution; the chi-square test was used to compare differences between categorical variables; and the Wilcoxon rank test was used to compare group differences in variables with unknown distribution. Unless otherwise stated, P value<0.05 was regarded as statistically significant.* represent p <0.05, * *represent p <0.01, * * *represent p <0.001, * * * * represent p <0.0001 and ns, $p \geq 0.05$.

3. Results

3.1 Identification of MMRGs between adjacent and cancerous tissues

Based on TCGA-CRC expression profile data, we used the R limma package gene set enrichment analysis (GSEA) to compare the difference between the adjacent and cancerous tissues of mitochondrial energy metabolism gene expression, and the genes were ranked according to the statistical value (logFC), and the unregulated genes ranked before and downregulated ranked the last. Then, enrichment analysised the preranked genes. The results showed that a total of 114 Mitochondrial energy metabolism pathway-related genes (MMRGs) was obtained and all pathways(including Citrate cycle tca cycle, Fatty acid metabolism, Ketone body metabolism, Glycolysis gluconeogenesis, Oxidative phosphorylation) were significantly enriched (p <0.05),and most pathways were concentrated in upregulated regions, these genes were significantly more expressed in adjacent tissues than cancerous tissues . These analysis results suggested that mitochondrial energy metabolism activity was reduced in cancer cells.

3.2 Analysis of the association between SNV, CNV, methylation and different clinical features based on MMRGs

Our study of somatic alteration analysis was TCGA-CRC MMRGs by using software maftools(v2.10.05).The gene base mutations replace model almost was C>T, T>C, C>A, and the proportion of base transition is significantly higher than base tranversion(Fig 1A). Only OGDHL and PC gene sets in the MMRGs have associated gene mutations, which are generally rare with frequencies ranging 10%, where OGDHL has the highest mutation frequency(Fig 1B) .We obtained

differentially methylated sites (DMP) significantly enriched by using the R package ChAMP (V2.26.0), along with their corresponding MMRGs related target genes(Fig 1C). Differentially methylation sites were screened against a threshold deltaBeta> 0.1 and a corrected p-value <0.05.

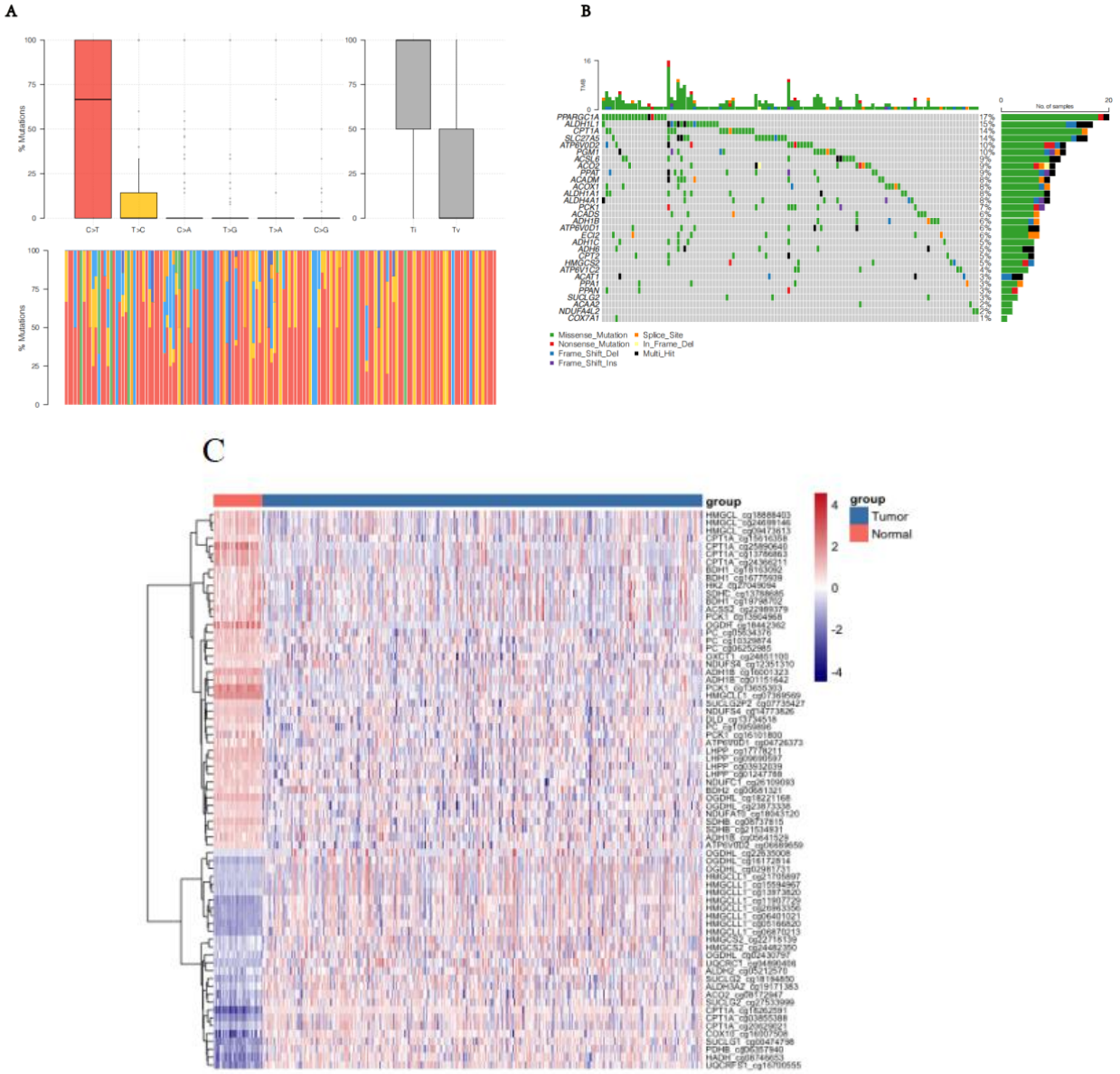


Figure 1: (A)CNV plots depict the distribution of base mutations in genomic variants ;(B) The SNV plots depict the distribution of mutations in GGM-related genes, along with classifications of the various SNV types;(C) DNA methylation differences in adjacent and cancerous tissues of MMRGs

Kruskal-Wallis Rank sum test was used to analyze the correlation between MMRGs expression level and clinical characteristics (age, tumor stage, tumor location, lymphatic invasion, perineural invasion, venous invasion). Based on the results, the clinic opathological characteristics were depicted in Fig.3A, highlighting signifcant diferences in tumor location, T stage,pathologic stage,perineural invasion ,venous invasion ,lymphatic invasion between two risk stratifcations of the

TCGA-CRC (all p<0.05). In addtion , The number of the MMRGs specific biomarkers were 28 in tumor location,and the highlight expression of MMRGs biomarkers were ACOX1, NDUFV2, OGDHL, PCK1 (Fig. 2A); the number of MMRGs specific biomarkers associated with pathologic stage were 64,the high expression of MMRGs specific biomarkers were ACADM, ACADS, ACADSB (Fig.2B); the number of MMRGs specific biomarkers associated with perineural invasion were 4 (Fig.2C), the number of MMRGs specific biomarkers associated with venous invasion were 12,and the highlight expression of MMRGs biomarkers were ACADSB, ATP6V0C, COX6B2, MT-ND4 (Fig. 2D) the number of MMRGs specific biomarkers associated with lymphatic invasion were 36,and the highlight expression biomarkers were UQCRC1, SUCLG1, SDHB, ATP6V1G2, COX5A, ACADSB (Fig.2E) (all p<0.05).

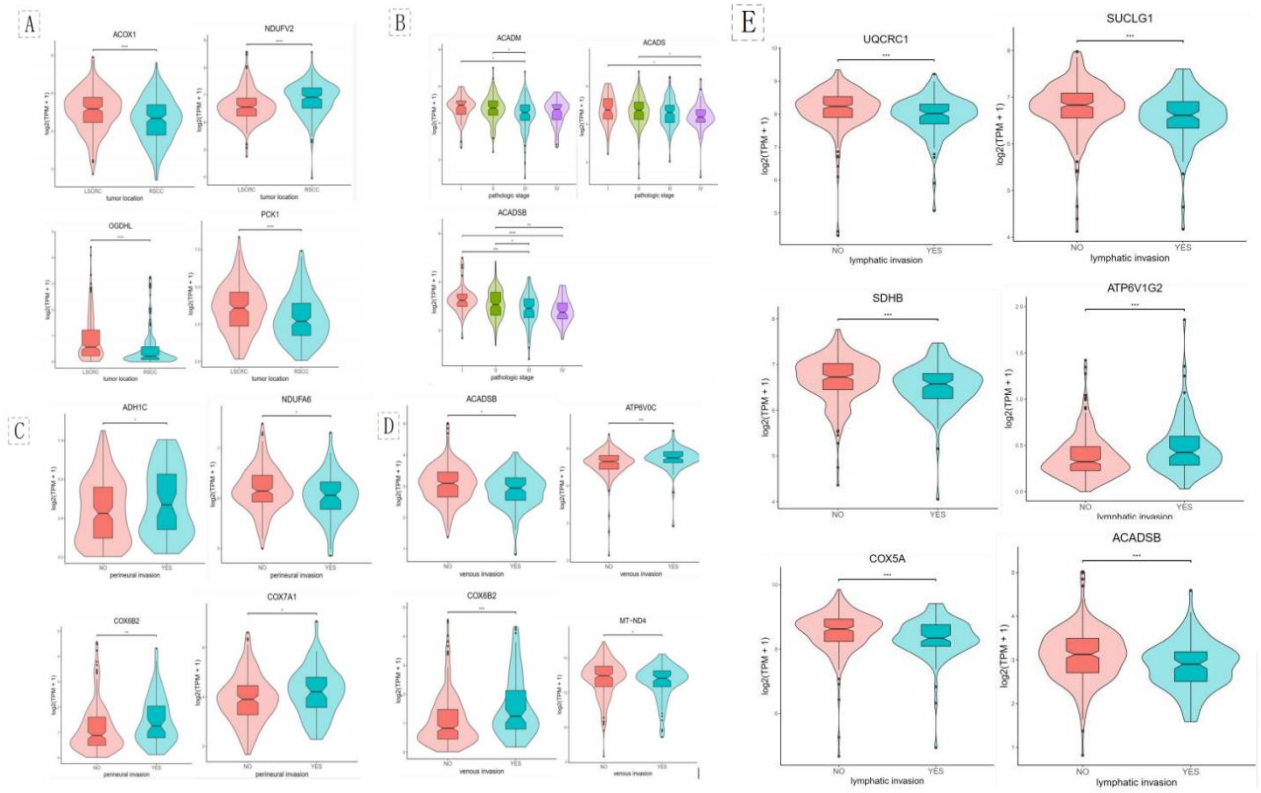


Figure 2: Heatmap illustrating the variation in expression levels of the genes in the CRC signature and the specifc CRC-related biomarkers between the two risk stratifcations.(A)the highlight expression MMRGs associated with tumor location; (B)the highlight expression MMRGs associated with pathologic stage; (C)the highlight expression MMRGs associated with perineural invasion ; (D)the highlight expression MMRGs associated with venous invasion (E)the highlight expression MMRGs associated withlymphatic invasion.

3.3 Survival analysis of the expression levels of the MMRGs

Based on the median value of the gene expression, samples were divided into high and low

expression groups to delve deeper into the potential prognostic value of MMRGs, and we utilized R package survival to analyze their association with patient survival across CRC. The results, presented in Fig. 3 and supplementary revealed a complex landscape of relationships between MMRGs expression and survival outcomes. Overall survival (OS) showed different prognostic effects depending on the different gene. In this results the high expression of PGAM2, COX7A1a and ATP6V1G2 predicted poorer overall survival (OS) in MMRGs. And the opposite result is the low expression of other MMRGs were associated with poorer overall survival(OS).

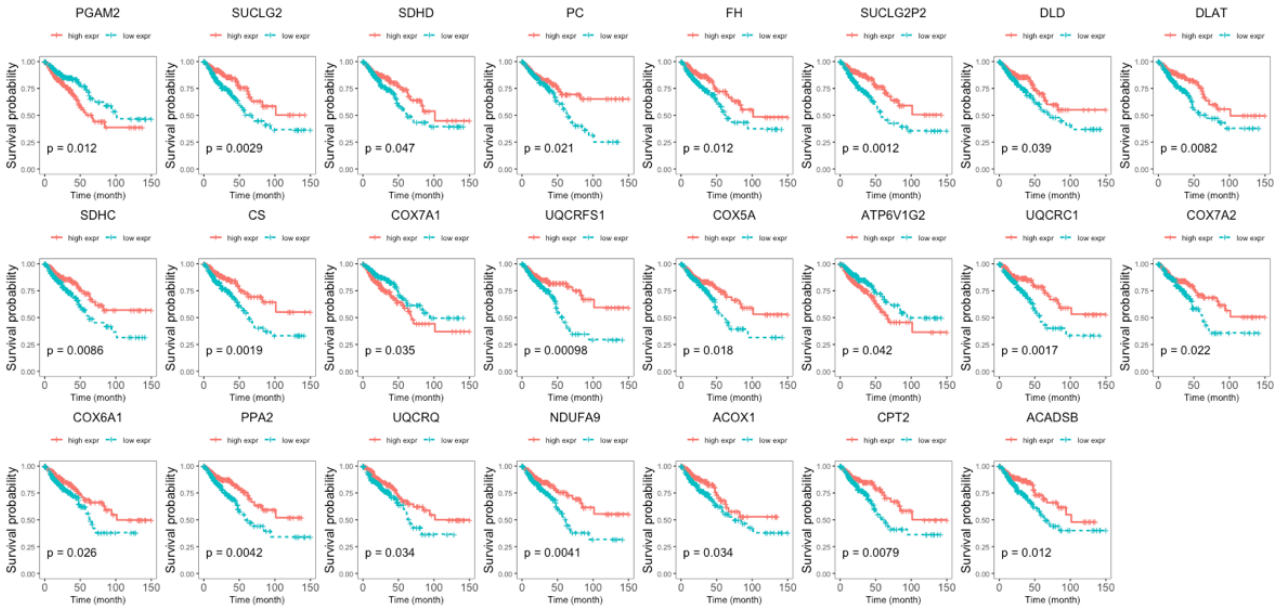
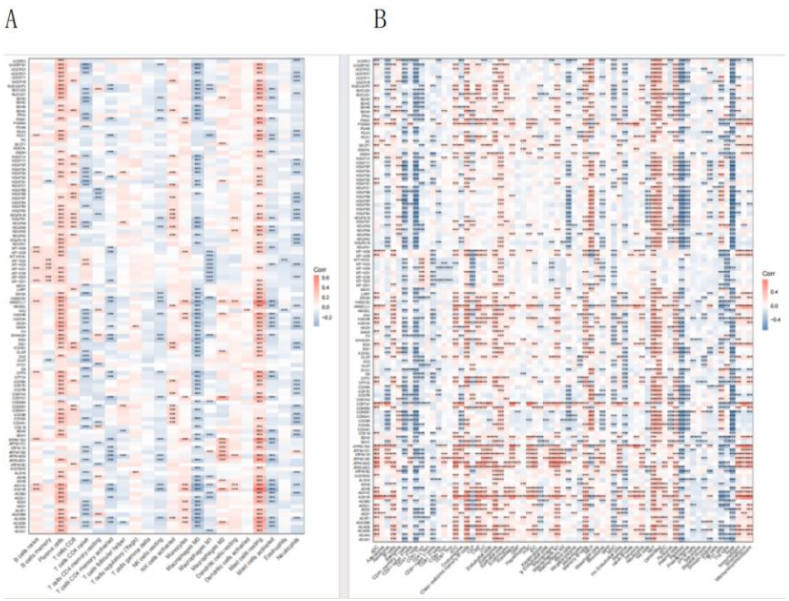


Figure 3: Kaplan-Meier plots depicting correlation of MMRGs expression with OS in CRC.

3.4 Evaluation of MMGRs and Immune enviroment Levels in TCGA-CRC gene set



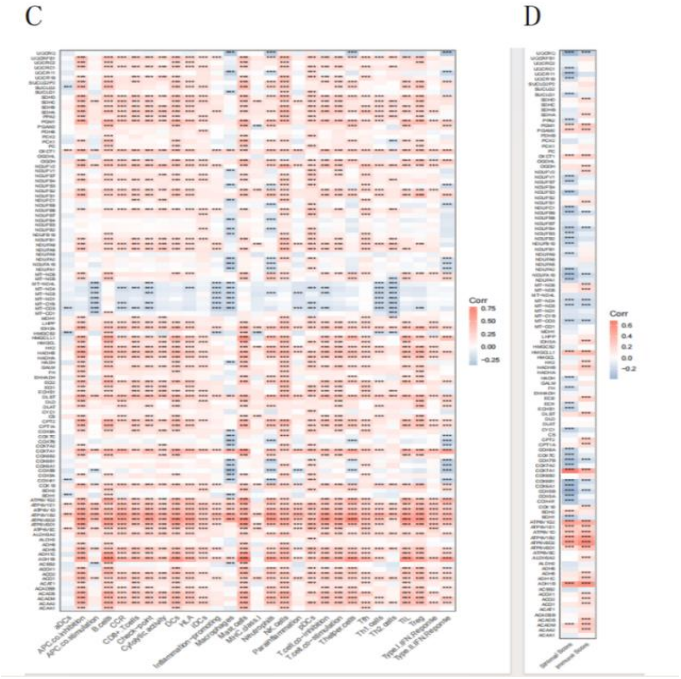


Figure 4: The relationship between immune cell infiltration risk scores and MMRGs gene expression were visualized on the different tool software.(A)The results from software CIBERSORT;(B)The results from software xCell(v1.1.0);(C)The results from software estimate(v1.0.13) ;(D)The results from software GSVA (v1.42.0) .

We assessed any possible association between the risk model and immune-associated biomarkers and observed a significant relationship by using online software CIBERSORT,xCell(v1.1.0), estimate(v1.0.13) and the ssgsea method of GSVA (v1.42.0) .MMRGs in TCGA-CRC gene set were assessed to observe whether there was an association between MMRGs expression and immune infiltration microenvironment. The results of online software CIBERSORT showed that Spearman's correlation analysis revealed that the most MMRGs expression was positively associated with immune infiltration, including Plasma cell, resting mast cell infiltration ; In contrast, this group was negatively associated with activatedMast cells, macrophages M0, neutrophils, initial CD4 + T cells, and memory-activated CD4 + T cellsion risk cores in the tumor(Fig 4.A) .Similar results were obtained by software xCell(v1.1.0).As shown in Fig.4 C, MMRGs had a significantly higher level of correlation between immune matrix score and immune score was found by following estimate (v1.0.13) software (Fig 4.B). Furthermore, based on the TCGA-CRC gene expression data and the manually collected immune infiltration gene set (PMID: 30594216), the immune infiltration status of TCGA-CRC patients was calculated using the ssgsea method of GSVA (v1.42.0),the results showed that most of the MMRG core gene expression showed a high positive correlation with the immune infiltration scores of 29 immune cells, but a small number of mitochondrial genes were negatively correlated with the immune infiltration scores of immune cells(Fig 4.D).These results suggested that we observed a significant relationship between MMGRs and various immune

cells levels and the immune enviroment may affect the the expression of MMRGS.

3.5 Molecular networks and functional enrichment of MMRGs

To unlock the MMRGs family's functional roles, we explored their molecular networks and enriched pathways using bioinformatic tools. Using STRING (https://string-db.org/), we predicted protein-protein interaction networks and analyzied this combined network, revealing likely partners for MMRGs, respectively. By analyzing this combined network, we unraveled enrichment patterns in biology process analysis. The results showed showed positive correlation between most MMRGs, but the correlation between mitochondrial genes and nuclear genes was low and negative correlation, and there were strong positive correlation between mitochondrial genes ($P<0.05$) (Fig.5A).The protein interaction network diagram of MMRGs showed that the type of genes are clearly divided into three parts. The first part is several proteins related to ATP enzyme hydrogen ion transport, which have interaction with other parts; the second part is ubiquinone redox-related protein, whose protein interaction is complex, and the third part is other proteins, with no obvious rules($P<0.05$)(Fig. 5B). Furthermore, the results of MMRGs enrichment analysis showed that genes are mainly enriched in metabolic related of generation of precursor metabolites, energyand oxidation-reduction process, drug metabolic process, cellular respiration, small molecule metabolic process (Table 1).

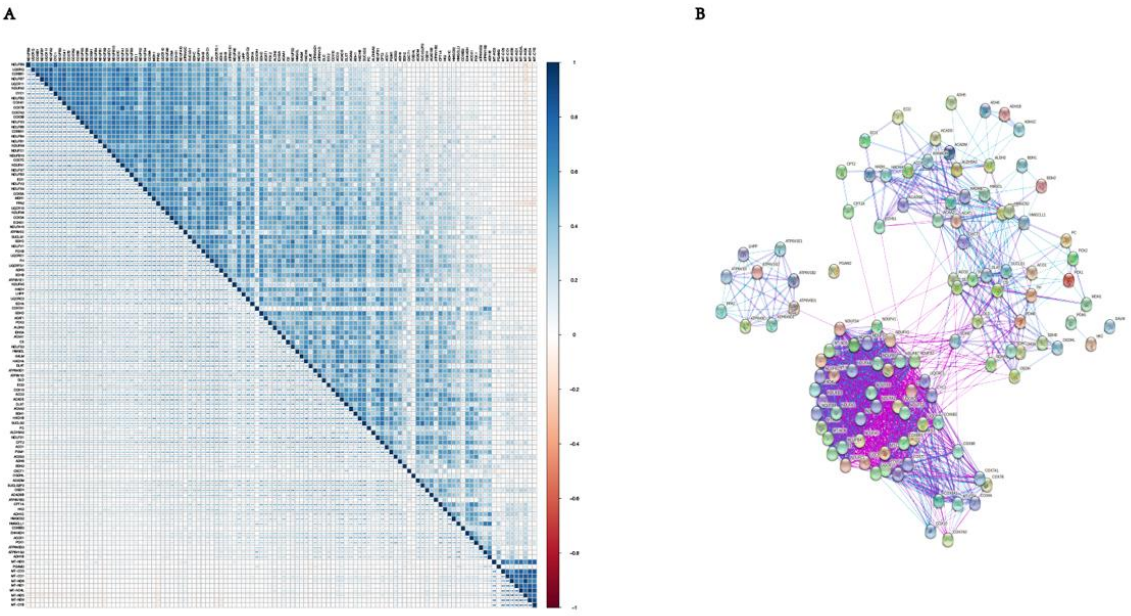


Figure 5: (A) The plot of correlation analysis of MMRGs; (B) The molecular interaction networks of MMRGs;

Table 1: The enrichment analysis results of the MMRGs in biology progress

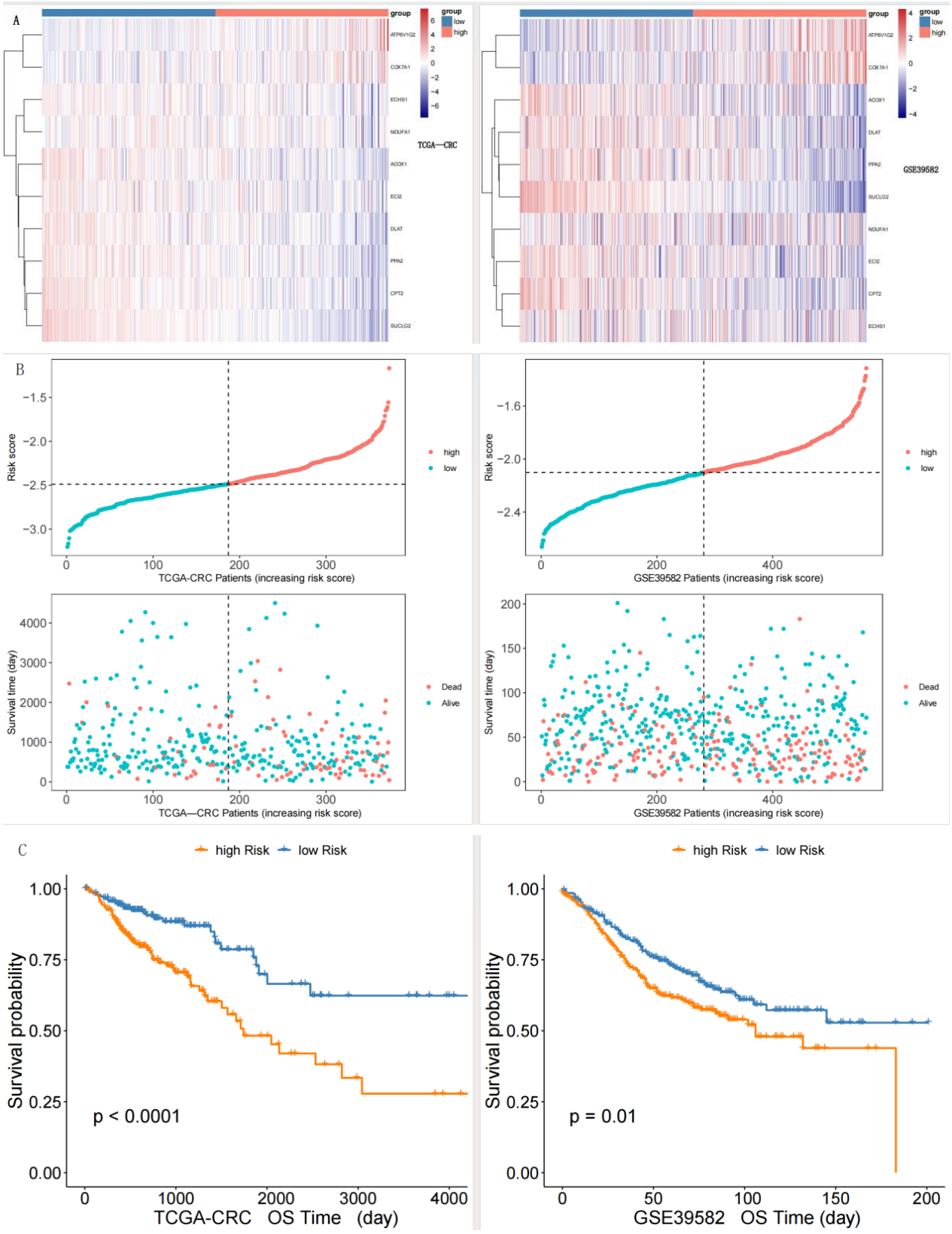
term	description	count	fdr
GO.0006091	generation of precursor metabolites and energy	83 of 400	1.96E-107
GO.0055114	oxidation-reduction process	95 of 932	1.62E-101
GO.0017144	drug metabolic process	80 of 622	7.66E-88
GO.0045333	cellular respiration	60 of 153	1.44E-87
GO.0044281	small molecule metabolic process	98 of 1779	8.52E-82

3.6 Validation of the Prognostic Model

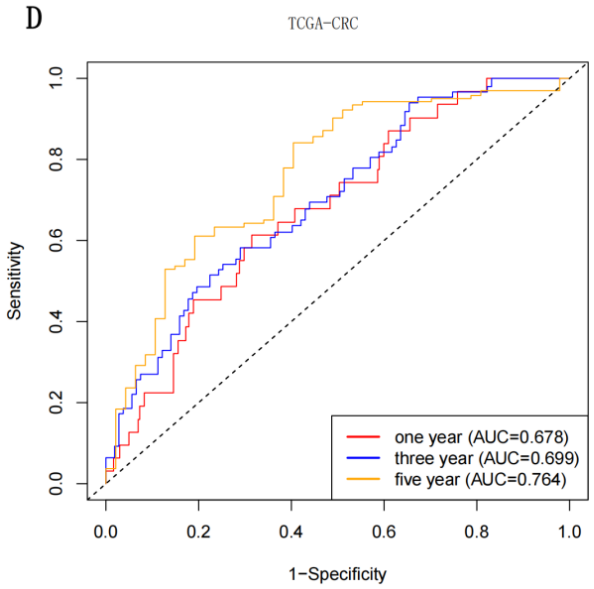
To examine the MMRGs express for the prognostic prediction of rectal cancer, 23 MMRGs of TCGA-CRC, which exhibited associations with overall survival in rectal cancer patients, underwent LASSO-Cox regression analysis. We tested the reliability of the prognostic risk score model in the GSE39582 data set. This analysis led to the establishment of a signature comprising ten genes, as follows (1):

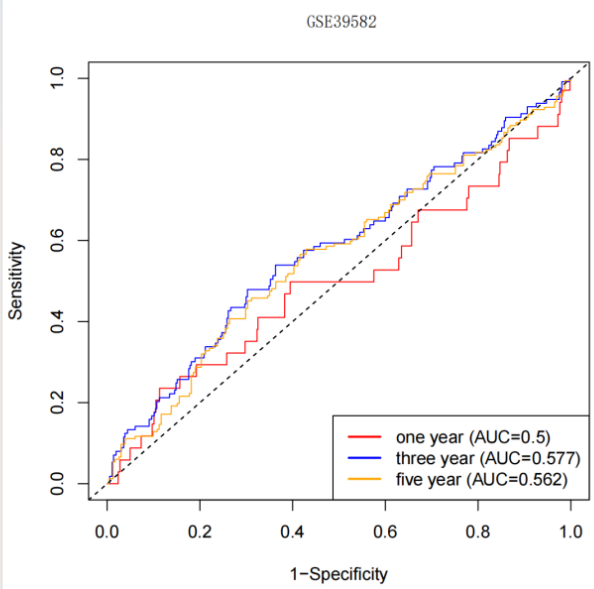
MMRGS=ACOX1 expression * -0.12684 + ATP6V1G2 expression * 0.36042 + COX7A1 expression * 0.02761 + CPT2 expression *-0.07471 + DLAT expression * -0.03181 + ECGS1 expression * 0.06075 + ECI2 expression * -0.0480 + NDUFA1 expression * 0.03525 + PPA2 expression * -0.0706 + SUCLG2 expression * -0.028727. (1)

The TCGA-CRC and GSE39582 MMRGs, was used to cut-of median score value to classify all cases into high- and low-risk groups. Expression levels of these ten candidate genes of TCGA-CRC and GSE 39582 in high- and low-risk groups were visualized in a heatmap (Fig. 6A) .Based on LASSO-Cox regression model, the higher MMRG scores the lower survival status and survival time in TCGA-CRC and GSE 39582 (Fig. 6B).Kaplan–Meier survival analysis demonstrated the same result of TCGA-CRC set and GSE39582 of rectal cancer patients with higher risk scores experienced worse prognosis (Fig. 6C). AUC values of TCGA-CRC set under the ROC curve were 0.678 at 1 year, 0.669 at 3 years, and 0.764 at 5 years (Fig. 6D), and AUC values of GSE 39582 set under the ROC curve were 0.5 at 1 year, 0.577 at 3 years, and 0.562 at 5 years, both results indicating that the prognostic model of MMRGs had a good predictive ability.

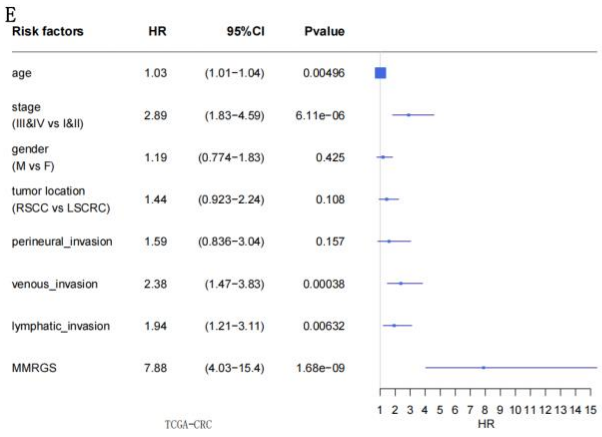

A
group
low
high
ATP6V1G2
COX7A1
ECHS1
NDUFA1
ACOX1
ECI2
DLAT
PPA2
CPT2
SUCLG2
TCGA—CRC
GSE39582
B
Risk score
TCGA-CRC Patients (increasing risk score)
GSE39582 Patients (increasing risk score)
Survival time (day)
Dead
Alive
C
high Risk
low Risk
Survival probability
p < 0.0001
p = 0.01
TCGA-CRC OS Time (day)
GSE39582 OS Time (day)

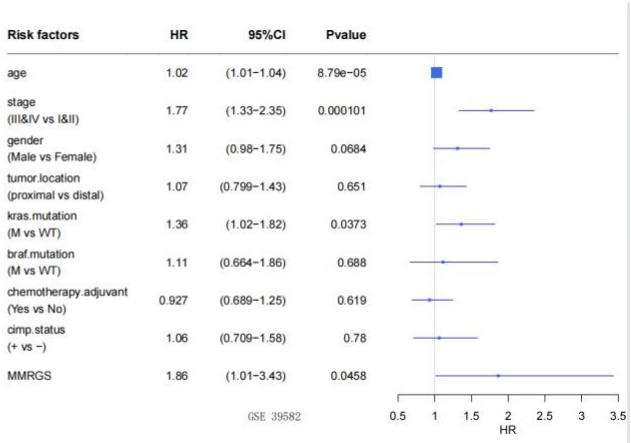
D


TCGA-CRC
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
one year (AUC=0.678)
three year (AUC=0.699)
five year (AUC=0.764)

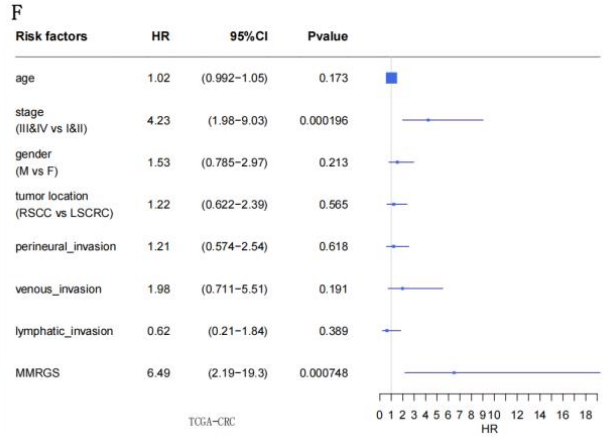

GSE39582
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
one year (AUC=0.5)
three year (AUC=0.577)
five year (AUC=0.562)

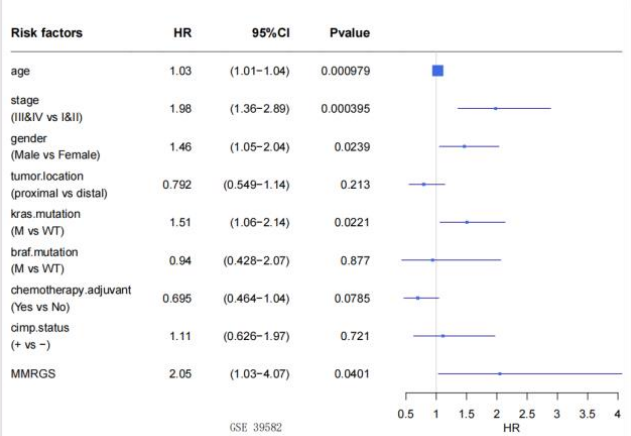
E


Risk factors HR 95%CI Pvalue
age 1.03 (1.01−1.04) 0.00496
stage (III&IV vs I&II) 2.89 (1.83−4.59) 6.11e−06
gender (M vs F) 1.19 (0.774−1.83) 0.425
tumor location (RSCC vs LSCRC) 1.44 (0.923−2.24) 0.108
perineural_invasion 1.59 (0.836−3.04) 0.157
venous_invasion 2.38 (1.47−3.83) 0.00038
lymphatic_invasion 1.94 (1.21−3.11) 0.00632
MMRGS 7.88 (4.03−15.4) 1.68e−09
TCGA-CRC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
HR


Risk factors HR 95%CI Pvalue
age 1.02 (1.01−1.04) 8.79e−05
stage (III&IV vs I&II) 1.77 (1.33−2.35) 0.000101
gender (Male vs Female) 1.31 (0.98−1.75) 0.0684
tumor.location (proximal vs distal) 1.07 (0.799−1.43) 0.651
kras.mutation (M vs WT) 1.36 (1.02−1.82) 0.0373
braf.mutation (M vs WT) 1.11 (0.664−1.86) 0.688
chemotherapy.adjuvant (Yes vs No) 0.927 (0.689−1.25) 0.619
cimp.status (+ vs −) 1.06 (0.709−1.58) 0.78
MMRGS 1.86 (1.01−3.43) 0.0458
GSE 39582
0.5 1 1.5 2 2.5 3 3.5
HR

F


Risk factors HR 95%CI Pvalue
age 1.02 (0.992−1.05) 0.173
stage (III&IV vs I&II) 4.23 (1.98−9.03) 0.000196
gender (M vs F) 1.53 (0.785−2.97) 0.213
tumor location (RSCC vs LSCRC) 1.22 (0.622−2.39) 0.565
perineural_invasion 1.21 (0.574−2.54) 0.618
venous_invasion 1.98 (0.711−5.51) 0.191
lymphatic_invasion 0.62 (0.21−1.84) 0.389
MMRGS 6.49 (2.19−19.3) 0.000748
TCGA-CRC
0 1 2 3 4 5 6 7 8 9 10 12 14 16 18
HR


Risk factors HR 95%CI Pvalue
age 1.03 (1.01−1.04) 0.000979
stage (III&IV vs I&II) 1.98 (1.36−2.89) 0.000395
gender (Male vs Female) 1.46 (1.05−2.04) 0.0239
tumor.location (proximal vs distal) 0.792 (0.549−1.14) 0.213
kras.mutation (M vs WT) 1.51 (1.06−2.14) 0.0221
braf.mutation (M vs WT) 0.94 (0.428−2.07) 0.877
chemotherapy.adjuvant (Yes vs No) 0.695 (0.464−1.04) 0.0785
cimp.status (+ vs −) 1.11 (0.626−1.97) 0.721
MMRGS 2.05 (1.03−4.07) 0.0401
GSE 39582
0.5 1 1.5 2 2.5 3 3.5 4
HR

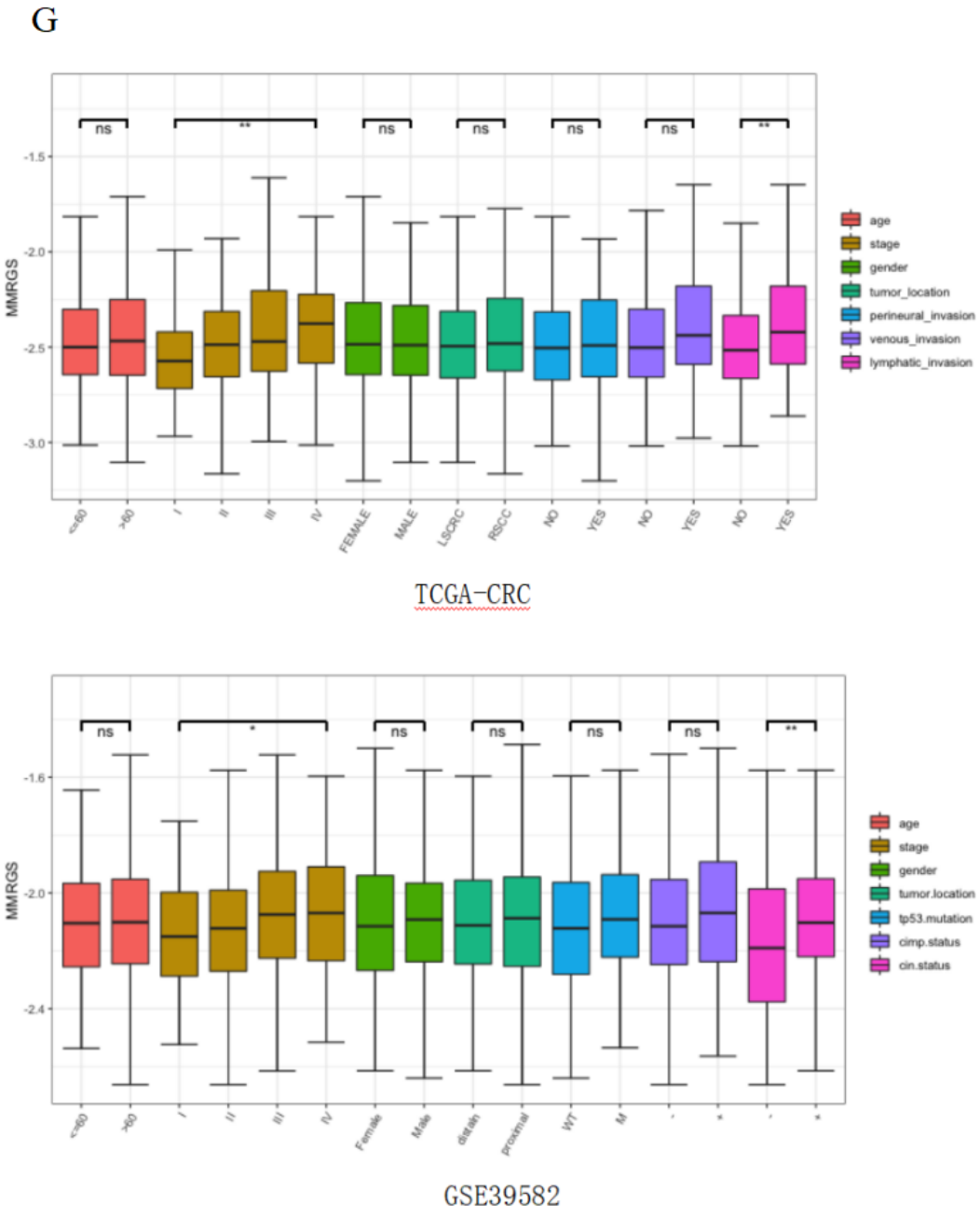


Figure 6: (A) A heatmap displays the differences in the expression levels of the ten genes between the high-risk and low-risk groups of TCGA-CRC set and GSE39582 set. (B)Model scores associated with patient survival status (OS) and survival time (OS time) of TCGA-CRC set and GSE39582 set.(C)Survival analysis indicates that in both the TCGA-CRC set and GSE39582 set, patients with high-risk scores have signifcantly worse prognoses compared to those with low-risk scores. (D) AUC curves of 1/3/5 year survival of patients in the TCGA-CRC set and GSE 39582 set.(E) Univariate Cox regression analysis results of risk score and clinical factors.(F)Multivariate Coxregression analysis results of risk score and clinical factors.(G)The plot between MMRGS and other clinical features in the TCGA-CRC set and GSE39582 set.

Therefore, we used R package survival (v3.3-1) to evaluate the independence of the risk model, we performed a univariate Cox regression analysis and a multivariate Cox regression analysis on clinical factors and risk score. As shown in (Fig.6 E,F) we established a nomogram based on these independent factors and demonstrated that the risk score, stage and lymphatic invasion can serve as the independent prognostic factors($p<0.05$). This reaffirms the influence of varied MMRGs expression patterns on prognosis and supports the use of signature-derived risk scores as clinical prognostic factors.In the GSE39582 data set, MMRGS was significantly associated with tumor stage and chromosomal instability status (CIN)($P<0.05$)(Fig. 6 G). Patients with advanced tumor stage and positive chromosomal instability had higher MMRGs scores. These findings aligned with the aforementioned survival analysis, indicating that an elevated risk score is associated with tumor stage, implying an unfavorable prognosis.

4. Discussion

Previous studies have demonstrated that DEGs between tumor and peritumor tissues in patients with CRC can be used to assess patient prognosis [12]. Consistent with this, the present study found that MMRGs was obtained between tumor and peritumor tissues and all pathways(including Citrate cycle tca cycle, Fatty acid metabolism, Ketone body metabolism, Glycolysis gluconeogenesis, Oxidative phosphorylation) were significantly enriched ($p <0.05$).

Previous research has found that mutations and methylaton in DNA are commonly found in cancer cells and regulated tumor metabolism and progression to promote the tumor of bioenergetics and biosynthesis[13-14].In our study,the same results is that we obtained mutations sites and differentially methylated sites (DMP) significantly enriched by using the R package ChAMP (V2.26.0), along with their corresponding MMRGs related target genes(Fig 3C). Increasing evidence has shown that the different regions of CRC originate from different factors, the rate of different anatomical regions of colorectal cancer associated with embryonic origins ,prognostic factors ,clinical presentations[15-16].In our study, Kruskal-Wallis Rank sum test ,immune infiltration analysis and protein-protein interaction (PPI) network identified key immune signature gene sets associated with CRC , and MMRGs expression level was associated with the clinical characteristics (age, tumor stage, tumor location, lymphatic invasion, perineural invasion, venous invasion, survival rate).This finding aligns with previous reports this study reveals unbalanced distribution of genes in the immune environment, the previous reported that the migration of immune cells was coordinated with infiltration ,we hypothesize that the immune enviroment may affect the the expression of MMRGS, and this process might regulate in tumor micro environment[17].

To better predict the risk of mortality and prognosis of patients with CRC, the present study selected 114 MMRGs from TCGA datasets of patients with CRC, and GEO datasets for the validation results. From the result, we identified 10 prognostic genes (namely ACOX1, ATP6V1G2,

COX7A1, CPT2, DLAT, ECGS1, ECI2, NDUFA1, PPA2, SUCLG2) based on single-factor and multi-factor Cox analyses. Therefore a prognostic model was built using the comprehensive risk scores of the aforementioned 10 genes, and the results showed that the predictive effectiveness of the model was good when using GEO datasets. Consistent with this, the clinical data showed that the risk scores of these 10 genes were independent predictive factors when using both TCGA and GEO datasets.Previous studies have found that ACOX1 is significantly regulated in CRC tissues and dephosphorylation of ACOX1 promotes β-catenin palmitoylation to drive colorectal cancer progression[18]. ATP6V1G2 was associated With Polymorphic Variations in Breast Cancer Patients[19]. COX7A1 may connected with methylation and suppressed the viability of human non-small cell lung cancer cells via regulating autophagy[20-21].The regulation of CPT2 was found downregulation in cancer tissue[22-23].The downregulation of CPT2 promoted proliferation and inhibited apoptosis through p53 pathway in colorectal cancer[23].And CPT2 downregulation triggered stemness and oxaliplatin resistance in colorectal cancer via activating the ROS/Wnt/β-catenin-induced glycolytic metabolism[24].DLAT as an immunological and prognostic biomarker correlated with Prognosis, Chemoresistance, and Immune Infiltration in variety of cancer[25-26].A previous study revealed ECI2 was connected with epigenetic and genetic changes ECI2 in high-grade serous ovarian carcinoma[27]. Downregulation of NDUFA1 and other oxidative phosphorylation-related genes was a consistent feature of basal cell carcinoma[28]. PPA2 was connected sudden cardiac death and the Mitochondrial Inorganic Pyrophosphatase[29]. SUCLG2 participated in epigenetic expression and regulated cancer express[30].According this results, ATP6V1G2, ECI2, COX7A1, SUCLG2 was connected with epigenetic regulation. And the expression of ACOX1, PPA, NDUFA1 may regulate the mitochondria-related pathways in CRC. Therefore, according to drug sensitivity of TCGA and GEO MMRGs ,We found drug target prediction of the model's 10 genes, evaluated the potential therapeutic value of these genes, and screened for possible drug .

These results suggested that MMRGs was a critical gene affecting the proliferation and migration of tumor cells in CRC. In conclusion, the present study developed and constructed a prognostic model with potential clinical application. The functional role of the key gene, ACOX1, ATP6V1G2, COX7A1, CPT2, DLAT, ECGS1, ECI2, NDUFA1, PPA2, SUCLG2 in CRC was identified, providing evidence for its potential application as a prognostic biomarker and therapeutic target in CRC. However, the present study had some limitations. Our study may have inherent biases and inaccuracies, lack external experimental to validate ATP6V1G2, ECI2, COX7A1, SUCLG2 can lead to CRC gene mutation or other change.

5. Conclusion

In conclusion, our study provided a systematic analysis of mitochondrial energy

metabolism-gene in colorectal cancer, covering their expression, clinical feature, prognostic significance, biological functions, immune microenviroment, survival prognosis. Notably, we identified mitochondrial energy metabolism as an essential factor in colorectal cancer, playing a crucial role as a key determinant of colorectal cancer prognosis. Our study provides new potential targets for prognosis and therapeutic prediction in colorectal cancer. The nomogram constructed based on the risk score and clinical variables is more benefcial for clinicians to assess the survival prognosis of CRC. These findings may have some limitation to consider, as our study being bioinformatics research based on data collected from public databases.

Data availability

All relevant data are within the manuscript and its Additional files

Acknowledgments

Financial Disclosure: This study was supported by grants from by the Natural Science Foundation of Shen Zhen techology plan (JCYJ20210324135205016) and ShenZhen PingShan healthy system technology program (No:2023202)

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021; 71(3):209–249. doi: 10.3322/caac.21660.

[2] Lian JW, Liu YC, Yu HP. Progress in the global prevalence, risk factors and attributable disease burden of colorectal cancer. The Chinese Journal of Cancer Prevention and Treatment, 2024, 16 (01): 1-9.

[3] Li, JB, Qiu, ZY, Deng, YX, et al. Factors associated with positive predictive value of preliminary screening in a two-step screening strategy for colorectal neoplasms in China. Discov Oncol. 2022; 13 (1): 4. doi: 10.1007/s 12672-022-00463-8

[4] Erben, V, Poschet, G, Schrotz-King, P, et al. Comparing Metabolomics Profiles in Various Types of Liquid Biopsies among Screening Participants with and without Advanced Colorectal Neoplasms. Diagnostics (Basel). 2021; 11 (3): doi: 10.3390/diagnostics11030561

[5] Li, N, Zhou, YY, Lu, M, et al. [Participation rate and detection of colorectal neoplasms based on multi-round fecal immunochemical testing for colorectal cancer screening in the Chinese population]. Chinese Journal of Oncology2023; 45 (12): 1041-1050. doi: 10.3760/cma.j.cn112152-20230221-00073

[6] Choi, YW, Kim, YH, Oh, SY, et al. Senescent Tumor Cells Build a Cytokine Shield in Colorectal Cancer. Adv Sci (Weinh). 2021; 8 (4): 2002497. doi: 10.1002/advs.202002497

[7] Shin, JS, Kim, TG, Kim, YH, et al. Senescent tumor cells in colorectal cancer are characterized by elevated enzymatic activity of complexes 1 and 2 in oxidative phosphorylation. J PATHOL TRANSL MED. 2023; 57 (6): 305-314. doi: 10.4132/jptm.2023.10.09

[8] Shi, Y, Kang, Q, Zhou, H, et al. Aberrant LETM1 elevation dysregulates mitochondrial functions and energy

metabolism and promotes lung metastasis in osteosarcoma. GENES DIS. 2023; 11 (3): 100988. doi: 10.1016/j.gendis. 2023.05.005

[9] Yang, S, Liu, L, Liu, X, et al. The mitochondrial energy metabolism pathway-related signature predicts prognosis and indicates immune microenvironment infiltration in osteosarcoma. MEDICINE. 2023; 102 (46): e36046. doi: 10.1097/MD.0000000000036046

[10] Falkenius, J, Lundeberg, J, Johansson, H, et al. High expression of glycolytic and pigment proteins is associated with worse clinical outcome in stage III melanoma. MELANOMA RES. 2013; 23 (6): 452-460. doi: 10.1097/CMR.0000000000000027

[11] Zeng, T, Huang, Z, Yu, X, et al. Combining methylated SDC2 test in stool DNA, fecal immunochemical test, and tumor markers improves early detection of colorectal neoplasms. Front Oncol. 2023; 13 1166796. doi: 10.3389/fonc.2023.1166796

[12] Chou, CK, Yang, PC, Tsai, PY, et al. Methylglyoxal Levels in Human Colorectal Precancer and Cancer: Analysis of Tumor and Peritumor Tissue. Life (Basel). 2021; 11 (12): doi: 10.3390/life11121319

[13] Liu, Y, Sun, Y, Guo, Y, et al. An Overview: The Diversified Role of Mitochondria in Cancer Metabolism. Int J Biol Sci. 2023; 19 (3): 897-915. doi: 10.7150/ijbs.81609

[14] Cui, H, Rong, W, Ma, J, et al. DNA N6-Adenine methylation in HBV-related hepatocellular carcinoma. GENE. 2022; 822 146353. doi: 10.1016/j.gene.2022.146353

[15] Zhao, K, Li, H, Zhang, B, et al. Factors influencing advanced colorectal neoplasm anatomic site distribution in China: An epidemiological study based on colorectal cancer screening data. Cancer Med. 2023; 12 (24): 22252-22262. doi: 10.1002/cam4.6722

[16] Saadati, HM, Okhovat, B, Khodamoradi, F. Incidence and Risk Factors of Colorectal Cancer in the Iranian Population: a Systematic Review. J GASTROINTEST CANC. 2021; 52 (2): 414-421. doi: 10.1007/s12029-020-00574-x

[17] Liu, WQ, Li, WL, Ma, SM, et al. Discovery of core gene families associated with liver metastasis in colorectal cancer and regulatory roles in tumor cell immune infiltration. Transl Oncol. 2021; 14 (3): 101011. doi: 10.1016/j.tranon.2021.101011

[18] Zhang, Q, Yang, X, Wu, J, et al. Reprogramming of palmitic acid induced by dephosphorylation of ACOX1 promotes β-catenin palmitoylation to drive colorectal cancer progression. Cell Discov. 2023; 9 (1): 26. doi: 10.1038/s41421-022-00515-x

[19] Todorova, VK, Makhoul, I, Dhakal, I, et al. Polymorphic Variations Associated With Doxorubicin-Induced Cardiotoxicity in Breast Cancer Patients. ONCOL RES. 2017; 25 (8): 1223-1229. doi: 10.3727/096504017X14876245096439

[20] Rönn, T, Poulsen, P, Hansson, O, et al. Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. DIABETOLOGIA. 2008; 51 (7): 1159-1168. doi: 10.1007/s00125-008-1018-8

[21] Zhao, L, Chen, X, Feng, Y, et al. COX7A1 suppresses the viability of human non-small cell lung cancer cells via regulating autophagy. Cancer Med. 2019; 8 (18): 7762-7773. doi: 10.1002/cam4.2659

[22] Zhang, X, Zhang, Z, Liu, S, et al. CPT2 Down-Regulation Promotes Tumor Growth and Metastasis Through Inducing ROS/NFκB Pathway in Ovarian Cancer SSRN. 2020; doi: 10.2139/ssrn.3736010

[23] Liu, F, Li, X, Yan, H, et al. Downregulation of CPT2 promotes proliferation and inhibits apoptosis through p53 pathway in colorectal cancer. CELL SIGNAL. 2022; 92 110267. doi: 10.1016/j.cellsig.2022.110267

[24] Li, H, Chen, J, Liu, J, et al. CPT2 downregulation triggers stemness and oxaliplatin resistance in colorectal cancer via activating the ROS/Wnt/β-catenin-induced glycolytic metabolism. EXP CELL RES. 2021; 409 (1): 112892. doi: 10.1016/j.yexcr.2021.112892

[25] Zhou, C, Jin, L, Yu, J, et al. Integrated analysis identifies cuproptosis-related gene DLAT and its competing

endogenous RNAs network to predict the prognosis of pancreatic adenocarcinoma patients. MEDICINE. 2024; 103 (9): e37322. doi: 10.1097/MD.0000000000037322

[26] Fang, Z, Wang, W, Liu, Y, et al. Cuproptosis-Related Gene DLAT as a Novel Biomarker Correlated with Prognosis, Chemoresistance, and Immune Infiltration in Pancreatic Adenocarcinoma: A Preliminary Study Based on Bioinformatics Analysis. CURR ONCOL. 2023; 30 (3): 2997-3019. doi: 10.3390/curroncol30030228

[27] Němejcová, K, Bártů, M, Hojný, J, et al. A comprehensive analysis of the expression, epigenetic and genetic changes of HNF1B and ECI2 in 122 cases of high-grade serous ovarian carcinoma. ONCOL LETT. 2021; 21 (3): 185. doi: 10.3892/ol.2021.12446

[28] Mamelak, AJ, Kowalski, J, Murphy, K, et al. Downregulation of NDUFA1 and other oxidative phosphorylation-related genes is a consistent feature of basal cell carcinoma. EXP DERMATOL. 2005; 14 (5): 336-348. doi: 10.1111/j.0906-6705.2005.00278.x

[29] Guimier, A, Achleitner, MT, Moreau de Bellaing, A, et al. PPA2-associated sudden cardiac death: extending the clinical and allelic spectrum in 20 new families. GENET MED. 2021; 23 (12): 2415-2425. doi: 10.1038/s41436-021-01296-6

[30] Hu, X, Wu, J, Feng, Y, et al. METTL3-stabilized super enhancers-lncRNA SUCLG2-AS1 mediates the formation of a long-range chromatin loop between enhancers and promoters of SOX2 in metastasis and radiosensitivity of nasopharyngeal carcinoma. Clin Transl Med. 2023; 13 (9): e1361. doi: 10.1002/ctm2.1361