



· 论 著 ·

利用单细胞测序和转录组测序建立结直肠癌免疫细胞的9基因预后模型

唐楠¹, 黄慧霞^{2, 3, 4}, 刘晓健¹

1. 复旦大学附属肿瘤医院肿瘤内科, 复旦大学上海医学院肿瘤学系, 上海 200032;
2. 复旦大学附属肿瘤医院病理科, 复旦大学上海医学院肿瘤学系, 上海 200032;
3. 复旦大学附属肿瘤医院肿瘤研究所, 上海 200032;
4. 复旦大学病理学研究所, 上海 200032

[摘要] **背景与目的:** 结直肠癌 (colorectal carcinoma, CRC) 是一种常见的恶性肿瘤, 其发病率和死亡率仅次于胃癌和食管癌, 在消化系统恶性肿瘤中居第二位。越来越多的证据表明, 免疫细胞在CRC的发生、发展中发挥着重要作用。本研究旨在构建一个与免疫细胞相关的基因的预后模型, 以预测CRC患者的预后并进行精准管理。**方法:** 从基因表达综合 (Gene Expression Omnibus, GEO) 数据库和癌症基因组图谱 (The Cancer Genome Atlas, TCGA) 数据库下载结直肠癌的单细胞RNA测序 (single-cell RNA sequencing, scRNA-seq) 和普通转录组测序 (RNA sequencing, RNA-seq) 数据以及临床信息, 提取免疫细胞亚型的差异基因, 通过Cox回归和LASSO回归分析在TCGA数据中筛选出与预后相关的基因, 同时使用GSE39582和GSE41258进行外部验证。基于预后模型进行化疗药物敏感性分析、免疫治疗效果分析、风险评分相关的通路分析以及临床相关性分析。采用实时荧光定量聚合酶链反应 (real-time fluorescence quantitative polymerase chain reaction, RTFQ-PCR) 和免疫组织化学检测验证模型基因在10例经术后病理学检查的新鲜冷冻结直肠癌样本和细胞系中的表达水平, 样本采集均已获得复旦大学附属肿瘤医院伦理委员会的批准 (伦理编号: 050432-4-2108*)。**结果:** 本研究根据scRNA-seq数据集 (GSE161277) 定义了16个细胞群, 并通过R包cellDex将这些群标记为不同的细胞类型。然后对免疫细胞亚型进行差异分析, 共获得了374个差异表达基因。通过单因素Cox和LASSO回归分析, 本研究在CRC中构建了一个9基因风险预后模型。该风险模型表现出对预后的可靠预测效果, 对预测抗肿瘤药物敏感性、免疫治疗效果、潜在分子机制及临床特征等方面起重要作用, 高风险分数的患者从免疫疗法中受益概率较低。**结论:** 我们基于CRC肿瘤微环境中免疫细胞的异质性构建了一个包含9个基因的风险预后模型, 并据此预测了CRC患者的生存和治疗效果。

[关键词] 结直肠癌; 预后模型; 肿瘤免疫与免疫治疗; 药物筛选; scRNA-seq

中图分类号: R735.3+4 文献标志码: A DOI: 10.19401/j.cnki.1007-3639.2024.06.003

Integrated single-cell sequencing and transcriptome sequencing to reveal a 9-gene prognostic signature of immune cells in colorectal cancer TANG Nan¹, HUANG Huixia^{2, 3, 4}, LIU Xiaojian¹ (1. Department of Medical Oncology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China; 2. Department of Pathology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China; 3. Cancer Institute, Fudan University Shanghai Cancer Center, Shanghai 200032, China; 4. Institute of Pathology, Fudan University, Shanghai 200032, China)

Correspondence to: LIU Xiaojian E-mail: liuxiaojian3000@hotmail.com

[Abstract] **Background and purpose:** Colorectal carcinoma (CRC) is a common malignant tumor with incidence and mortality rates second only to gastric cancer and esophageal cancer among digestive system malignancies. Increasing evidence suggests that immune cells play a significant role in the occurrence and development of CRC. The aim of this study was to construct a

第一作者: 唐楠 (ORCID: 0000-0002-4771-5324), 硕士研究生; 并列第一作者: 黄慧霞 (ORCID: 0000-0002-6967-6872), 硕士研究生。

通信作者: 刘晓健 (ORCID: 0000-0002-8884-4448), 主任医师, 教授, E-mail: liuxiaojian3000@hotmail.com。

prognostic model related to immune cell-associated genes to predict the prognosis of CRC patients and enable precise management. **Methods:** Single-cell RNA sequencing (scRNA-seq) and bulk RNA sequencing (RNA-seq) data, along with clinical information for colorectal cancer, were downloaded from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. Differential genes of immune cell subtypes were extracted, and genes related to prognosis were screened in TCGA data using Cox regression and LASSO regression, with external validation using GSE39582 and GSE41258. The prognostic model was used to analyze chemotherapy drug sensitivity, assess immunotherapy efficacy, analyze pathways related to risk scores, and perform clinical correlation analysis. Finally, the expression levels of model genes were validated in 10 fresh frozen CRC tissues with post-operative pathological examination and cell lines using real-time fluorescence quantitative polymerase chain reaction (RTFQ-PCR) and immunohistochemistry. All samples were approved by the Fudan University Shanghai Cancer Center Ethics Committee (No. 050432-4-2108*). **Results:** We defined 16 cell clusters based on the scRNA-seq dataset (GSE161277) and labeled these clusters as different cell types using the R package cellDex. Differential analysis of immune cell subtypes yielded 374 differentially expressed genes. Using univariate Cox and LASSO analyses, we constructed a 9-gene risk prognostic model in CRC. This risk model exhibited reliable prognostic prediction performance and played an important role in predicting anti-tumor drug sensitivity, immunotherapy efficacy, potential molecular mechanisms, and clinical characteristics. Patients with high-risk scores had a lower probability of benefiting from immunotherapy. **Conclusion:** We constructed a 9-gene risk prognostic model based on the heterogeneity of immune cells in the CRC tumor microenvironment, which predicted the survival and treatment outcomes of CRC patients.

[**Keywords**] Colorectal carcinoma; Prognostic model; Cancer immunity and immunotherapy; Drug screening; scRNA-seq

结直肠癌 (colorectal carcinoma, CRC) 是全球第四大致命癌症, 每年约导致90万人死亡。5年和10年的生存率分别为65%和58%, 男性的患病率和死亡率比女性高25%^[1]。CRC的广泛筛查是减少CRC发病率的重要措施之一^[2-3]。近年来, 肿瘤研究越来越关注肿瘤微环境 (tumor microenvironment, TME), 因其不仅涉及肿瘤本身, 也影响局部恶性组织中的整个微环境, 对于上皮-间质转化 (epithelial-mesenchymal transition, EMT) 至关重要^[4]。

近年来, 各种高通量测序技术得到了长足的发展, 相比过去其可靠性、准确性和重复性都有了显著提高。其中, 最广泛使用的是普通转录组测序 (RNA sequencing, RNA-seq), RNA-seq可以同时分析样本内所有细胞的基因表达水平, 但无法分析单个细胞水平即恶性肿瘤组织中的异质性。单细胞RNA测序 (single cell RNA sequencing, scRNA-seq) 可在单个细胞分辨率的水平分析基因表达继而揭示TME内的异质性。但与RNA-seq相比, scRNA-seq成本较高, 数据量较少, 不足以用于构建可靠的预后模型。RNA-seq和scRNA-seq可从不同角度揭示TME在CRC进展和预后预测中的关键作用^[5-8]。因此, 利用现有的研究资料对RNA-seq和scRNA-seq的结果进

行联合分析将获得较为可靠的结果。

本研究从基因表达综合数据库 (Gene Expression Omnibus, GEO) 中获取了CRC的scRNA-seq数据集, 结合癌症基因组图谱 (The Cancer Genome Atlas, TCGA) 数据库的表达数据。基于scRNA-seq和RNA-seq数据的整合, 我们筛选出了与CRC中免疫细胞亚群相关的关键基因, 并据此建立了一个9基因风险预后预测模型, 可望为CRC患者的诊断和治疗提供新的参考和策略。

1 资料和方法

1.1 数据获取

本研究的整体工作流程见图1。从NCBI GEO公共数据库下载GSE161277的单细胞数据文件, 下载具有完整单细胞表达谱的12例样本数据进行单细胞分析。从NCBI GEO公共数据库下载GSE39582的Series Matrix File数据文件, 注释平台为GPL570, 共提取到562例带有生存信息的完整表达谱。下载GSE41258的Series Matrix File数据文件, 注释平台为GPL96, 共提取到149例带有生存信息的完整表达谱。本研究下载了处理过的COADREAD原始mRNA表达数据, 对应样本包括正常组 ($n=51$) 和肿瘤组 ($n=650$)。

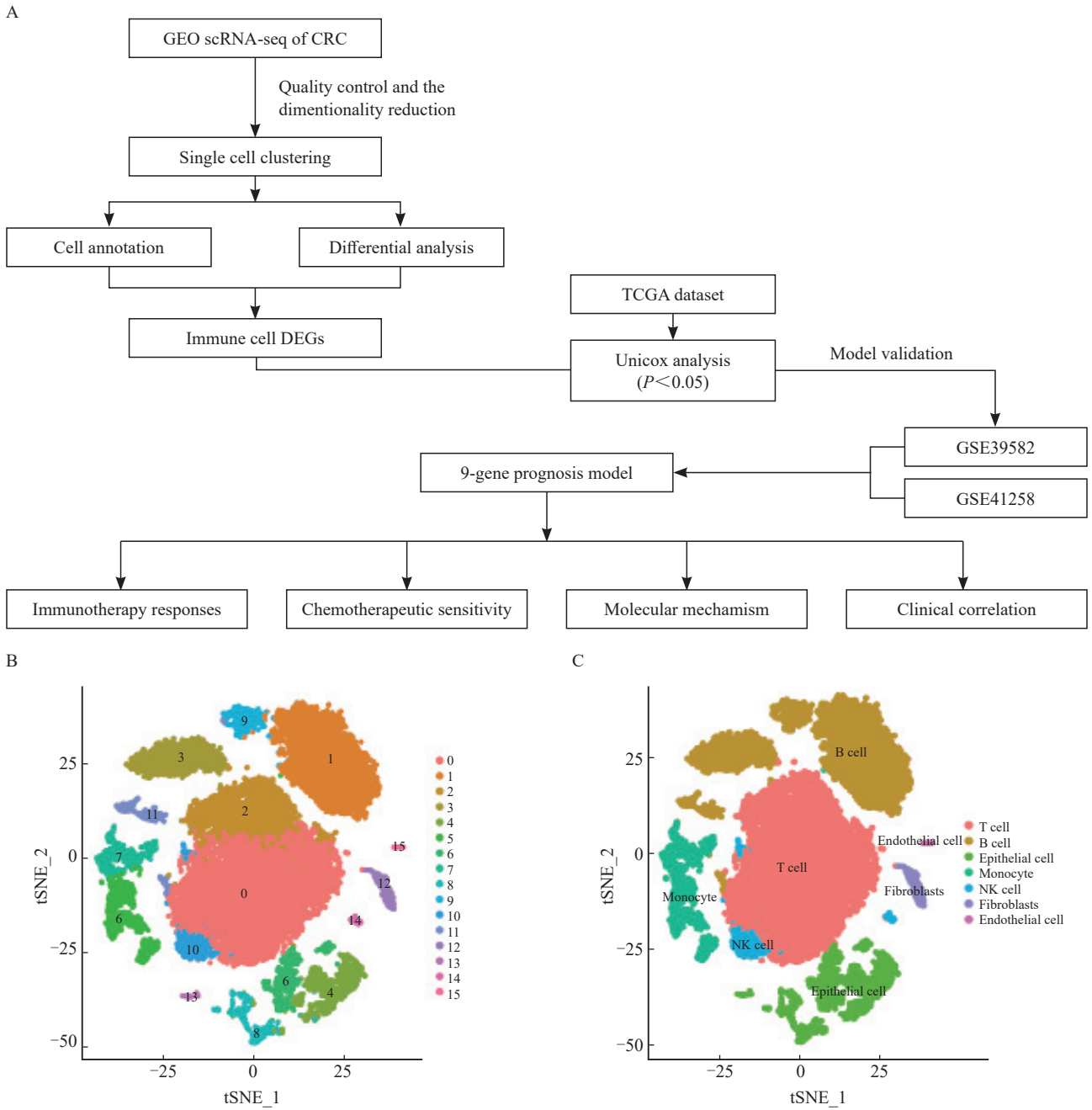


图1 鉴定不同的细胞群

Fig. 1 Identification of different cell clusters

A: Flowchart of this study. B: Identification of 16 cell clusters by the t-SNE algorithm. C: Further detailed annotation of different cell clusters.

1.2 单细胞分析

使用Seurat软件包读取表达谱并筛除低表达基因 (nFeature_RNA > 50 & percent.mt < 5)。依次对数据进行标准化、均一化、PCA分析处理, 并使用ElbowPlot来观察最佳的pc数 (n=15)。通过tSNE分析获得每个聚类群的位置关系, 聚类群由cellDex软件包进行注释。最后,

我们筛选了免疫细胞亚群中的差异表达基因, 筛选条件为logFCfilter > 1 & adjPvalFilter < 0.05。

1.3 模型构建与预后分析

选择与免疫细胞亚群相关的差异基因, 并采用LASSO回归进一步构建与预后相关的模型。在纳入每个特定基因的表达值后, 为每例患者构建了一个风险评分公式, 并在LASSO回归分析中用

其估计的回归系数加权。根据风险评分公式，以风险评分中位值为分界点，将患者分为低风险组和高风险组。通过Kaplan-Meier分析评估两组之间患者的生存差异，并使用log-rank统计方法进行比较。使用LASSO回归分析和分层分析来测试风险评分在预测患者预后中的作用。

1.4 肿瘤免疫功能障碍和排斥 (tumor immune dysfunction and exclusion, TIDE)

采用TIDE方法模拟肿瘤免疫逃避机制(包括T淋巴细胞功能障碍和T淋巴细胞排斥)。该软件可以在<http://tide.dfci.harvard.edu>上免费获得。在本研究中，我们使用TIDE来评估患者对免疫治疗的反应。TIDE得分越高，表明免疫逃逸的可能性越高，患者免疫治疗的效果越差。

1.5 化疗敏感性评估

基于抗癌药物敏感性基因组学数据库(Genomics of Drug Sensibility in Cancer, GDSC)的药物敏感性数据，采用“pRRophetic”包(<https://github.com/paulgeeleher/pRRophetic/>)预测不同药物在高、低危组的半数抑制浓度(IC₅₀)。药物的IC₅₀值越低，说明该药物治疗癌症的效果越好。

1.6 基因集差异分析 (gene set variation analysis, GSVA)

在本研究中，从MSigDB(<https://www.gseamsigdb.org/gsea/msigdb/index.jsp>)数据库下载基因集合，使用GSVA算法对每个基因集进行全面评分，以评估不同样本潜在的生物学功能变化。

1.7 基因集富集分析 (gene set enrichment analysis, GSEA)

GSEA使用预定义的基因集将关键基因按照在两种样本中的差异表达程度进行排序，然后检查预设的基因集是否在排序表的顶部或底部富集。在本研究中，使用GSEA比较高表达组和低表达组之间信号转导通路的差异，并探索两组患者的核心基因的分子机制，设置置换次数为1 000，置换类型为phenotype。

1.8 细胞培养

人结肠癌细胞系(DLD1、RKO、

SW620、SW480、SW1116、HCT116、HCT15、HCT8和LOVO)和人正常结肠上皮细胞系NCM460来自美国典型培养物保藏中心(American Type Culture Collection, ATCC)，在含有10%胎牛血清(fetal bovine serum, FBS)和1%的青霉素/链霉素的DMEM培养基中，在37℃、CO₂体积分数为5%条件下培养。所有使用的细胞系均无支原体污染。

1.9 临床样本收集

本研究收集10例在复旦大学附属肿瘤医院诊治的CRC患者经手术切除的CRC组织，均经过术后病理学检查明确诊断，包括肿瘤和癌旁组织。样本采集获得了复旦大学附属肿瘤医院伦理委员会的批准(伦理编号:050432-4-2108*)。每位参与本研究的患者均签署了书面知情同意书。

1.10 RNA提取与实时荧光定量聚合酶链反应 (real-time fluorescence quantitative polymerase chain reaction, RTFQ-PCR)

使用TRIzol试剂(美国Invitrogen公司)提取结肠癌细胞系和患者样本的总RNA，然后进行反转录(ABScript II RT Mix, 美国Abclonal公司)并存储在-20℃条件下。使用TB Green Premix Ex Taq II(日本TaKaRa公司)在Roche LightCycler® 480系统中进行RTFQ-PCR检测，重复3次。数据由内部参考基因 β -actin校正，所用引物序列见表1。

1.11 免疫组织化学检测

石蜡包埋组织切片在65℃加热2 h，在3个二甲苯槽中进行30 min的去石蜡处理和30 min乙醇溶液的再水化处理。然后，组织切片在EDTA缓冲液中进行15 min的高温高压抗原修复。内源性过氧化物酶活性被灭活并且组织被封闭后，切片将在室温下与抗TKT抗体(1:50; Cat#A6314)、抗TIMP1抗体(1:1 000; Cat#A22903)和抗G0S2(1:100; Cat#A9970)一抗共同温育2 h，然后与生物素结合的二抗温育45 min，上述试剂均购自武汉爱博泰克生物科技有限公司。最后用3,3'-二氨基联苯(3,3'-diaminobenzidine, DAB)染色和伊红反染，最终的免疫组织化学检测评分由两位有经验

的病理学研究生独立计算, 根据染色强度和染色面积进行IRS评分。

表1 RTFQ-PCR所用引物列表

Tab. 1 The primer list used in RTFQ-PCR

Primer	Sequence (5'-3')
CD36-1-F	GGCTGTGACCGGAACTGTG
CD36-1-R	AGGTCTCCAACCTGGCATTAGAA
CTSD-F	ATTCAGGGCGAGTACATGATCC
CTSD-R	CGACACCTTGAGCGTGTAG
G0S2-F	TCGGCCTGATGGAGACTGT
G0S2-R	TTCTGGAGAGCCTGTCGCT
TIMP1-F	AGAGTGTCTGCGGATACTTCC
TIMP1-R	CCAACAGTGTAGGTCTTGGTG
HSPA1A-F	AGCTGGAGCAGGTGTGTAAC
HSPA1A-R	TACCTCTCAATGGTGGGGC
VSIG4-1-F	GGGGCACCTAACAGTGGAC
VSIG4-1-R	GTCTGAGCCACGTTGTACCAG
CLEC4A-F	GACTGTGCTAGAATGGAGGCT
CLEC4A-R	GTCGCTGACCTTCTGGATCTG
TKT-1-F	TCATCGAGTGCTACATTGCTG
TKT-1-R	GCCATGCGAATCTGGTCAAAG
ASAH1-1-F	AGATGTCATGTGGATAGGGTTCC
ASAH1-1-R	GGGGCCAATATCTTGGTCTTG
β -actin-F	TTGTTACAGGAAGTCCCTTGCC
β -actin-R	ATGCTATCACCTCCCCTGTGTG

F: Forward; R: Reverse.

1.12 统计学处理

生存曲线通过Kaplan-Meier法生成, 并通过log-rank进行比较。采用Cox比例风险模型进行单因素分析。所有统计分析均采用R语言和SPSS 22.0软件进行。所有统计检验均为双侧。 $P < 0.05$ 为差异有统计学意义。

2 结 果

2.1 鉴定不同的免疫细胞亚群和差异表达基因

我们对从GSE161277数据集中提取的所有基因进行了标度, 通过PCA降维分析发现样本之间

的批次效应并不明显, 并通过ElbowPlot观察到最佳的主成分数目为15。最后通过tSNE分析获得了16个亚群(图1B)。我们将16个聚类注释为7个细胞类别(图1C), 通过logFCfilter > 1 & adjPvalFilter < 0.05的筛选条件对其中的免疫细胞亚型进行差异分析, 包括T淋巴细胞、B淋巴细胞、单核细胞和自然杀伤(natural killer, NK)细胞, 共获得了374个差异表达基因。

2.2 基于重要亚型的差异基因建立免疫细胞相关的预后模型并进行外部验证

我们从374个差异表达基因中筛选出17个预后相关基因($P < 0.05$), 然后进一步通过LASSO回归筛选CRC中的特征基因(图2A~C), 得到了最佳风险评分值用于后续分析[风险评分 = $ASAH1 \times (-0.372\ 380\ 553) + TKT \times (-0.299\ 171\ 578) + CLEC4A \times (-0.251\ 595\ 951) + VSIG4 \times 0.064\ 583\ 754 + HSPA1A \times 0.086\ 799\ 323 + TIMP1 \times 0.098\ 931\ 341 + G0S2 \times 0.120\ 197\ 581 + CTSD \times 0.146\ 006\ 52 + CD36 \times 0.240\ 287\ 725$]。根据风险评分中位值, 将患者分为高风险组和低风险组。Kaplan-Meier曲线分析表明, 在训练集和验证集中, 高风险组的总生存率(overall survival, OS)明显低于低风险组(图3A、B)。

在构建预后模型后, 本研究下载了GEO数据库中(GSE39582、GSE41258)带有生存资料的患者数据, 根据模型预测了GEO数据库中患者的临床分型, 通过Kaplan-Meier分析评估了两组之间患者的生存差异, 并探讨了预后模型的稳定性。结果显示, 在GEO外部验证集中, 高风险组的OS显著低于低风险组(图3C、D)。

RTFQ-PCR检测结果显示, 这些基因在混合CRC细胞系(HCT116、HCT15、CACO2、HCT8、SW620、SW480、SW1116、RKO和DLD1)和健康细胞系NCM460中的表达水平有差异(图4A、B), 说明这些基因在CRC的进展中发挥着重要作用。此外, 本研究利用免疫组织化学检测进一步验证模型基因中上调基因在癌和癌旁组织中的蛋白表达情况, 结果显示, TKT、TIMP1和G0S2在肿瘤组织中蛋白水平显著

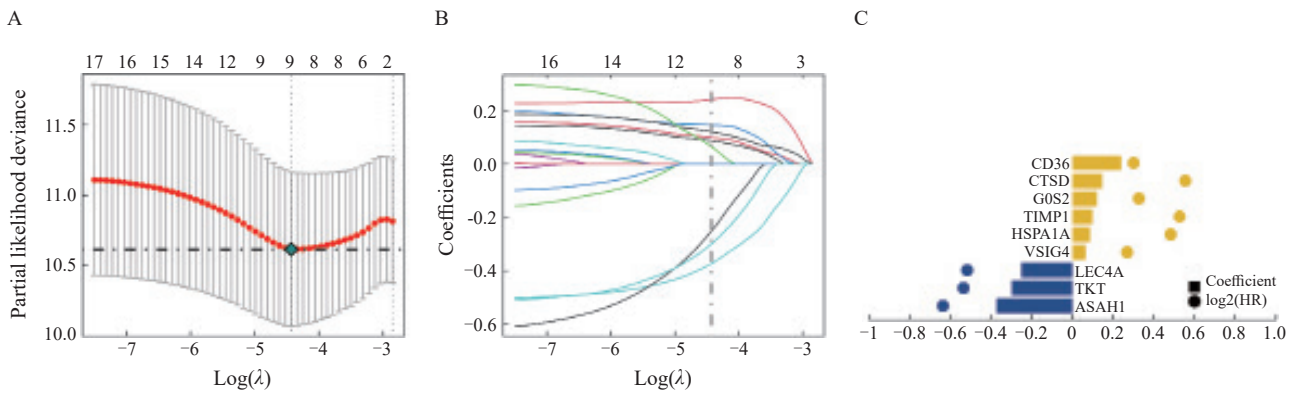


图2 LASSO回归分析构建预后模型

Fig. 2 Construction of prognostic model through LASSO regression analysis

A: The partial likelihood deviance for the LASSO Cox regression analysis. B: LASSO coefficient profile plots of each independent variable. C: Coefficient of LASSO genes.

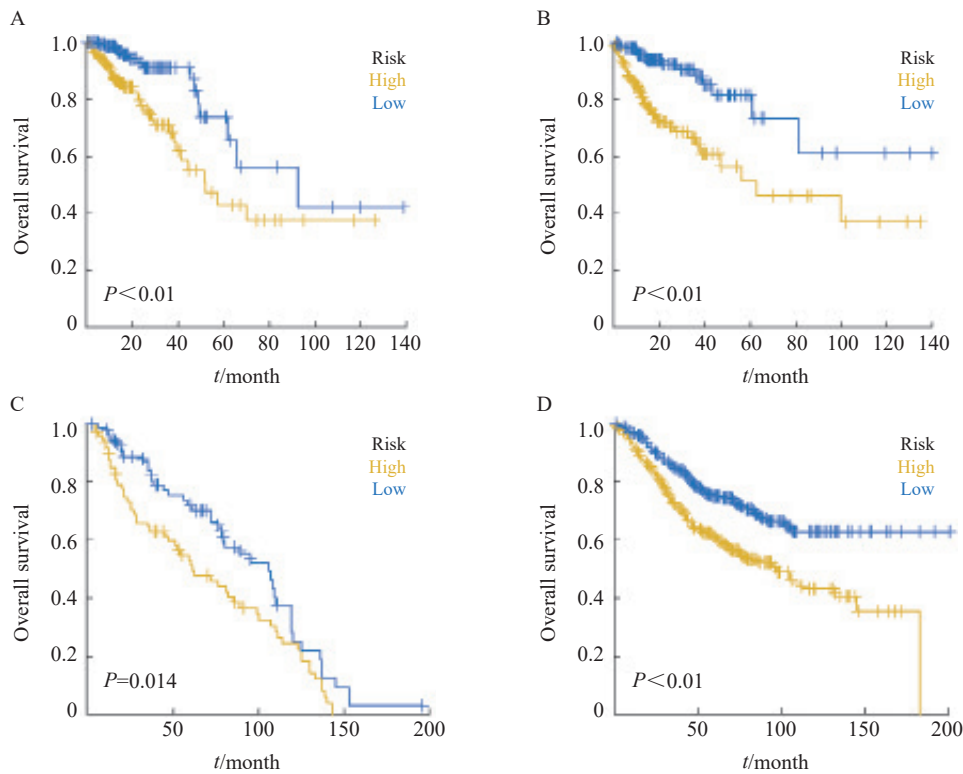


图3 训练集和验证集中高低风险组的预后分析

Fig. 3 Prognostic analysis of the high/low-risk group in the training set and test set

A-B: Kaplan–Meier survival curves of the model in the training set and testing set. C-D: Overall survival of CRC patients in the high- and low-risk groups in the test set (GSE39582, GSE41258).

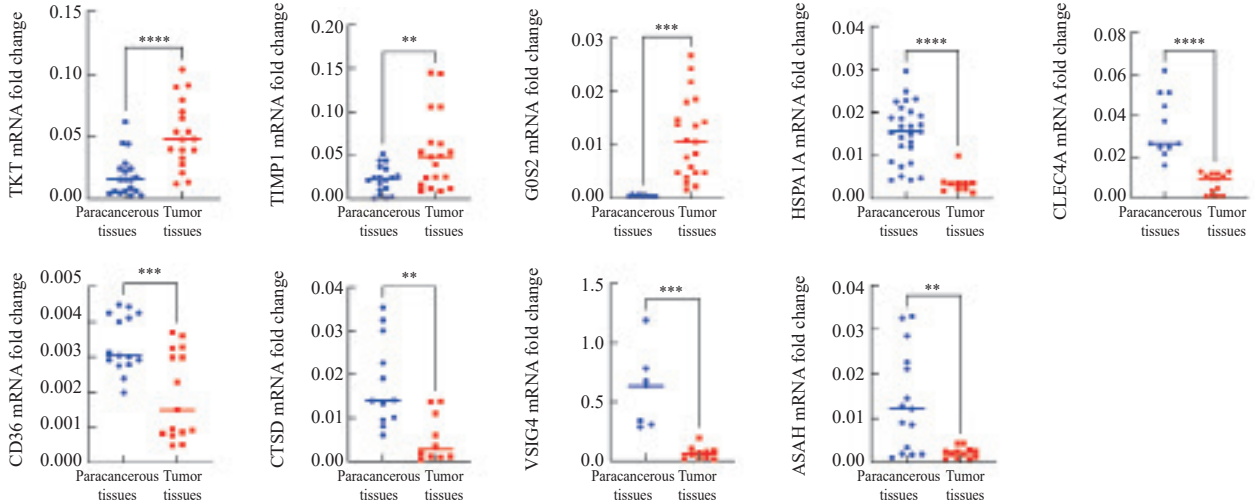
升高 (图4C), 与RTFQ-PCR的结果一致, 提示 TKT、TIMP1和G0S2可能作为原癌基因影响CRC的发生、发展, 这对于指导CRC的诊断治疗及机制研究有一定的参考价值。所有的证据表明, 我们构建了一个能够用于预后评估的可靠的9基因 [N-酰基鞘氨醇酰胺水解酶1 (N-acylsphingosine

amidohydrolase 1, ASAHI)、转酮醇酶 (transketolase, TKT)、C型凝集素结构域家族4成员A (C-type lectin domain family 4 member A, CLEC4A)、含V组和免疫球蛋白结构域4 (V-set and immunoglobulin domain containing 4, VSIG4)、热激蛋白家族A成员1A (heat shock

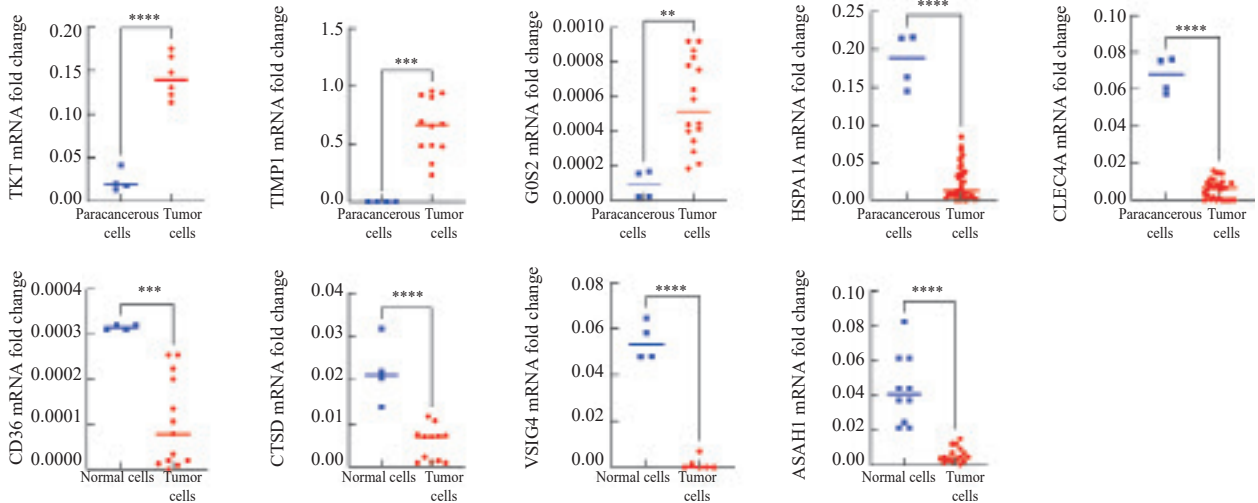
protein family A member 1A, HSPA1A)、TIMP 金属肽酶抑制因子1 (TIMP metalloproteinase inhibitor 1, TIMP1)、G₀/G₁ 开关基因2 (G₀/G₁

switch 2, G0S2)、组织蛋白酶D (cathepsin D, CTSD) 和CD36] 预后模型。这些基因都被报道参与肿瘤的生长、转移和侵袭等过程。

A



B



C

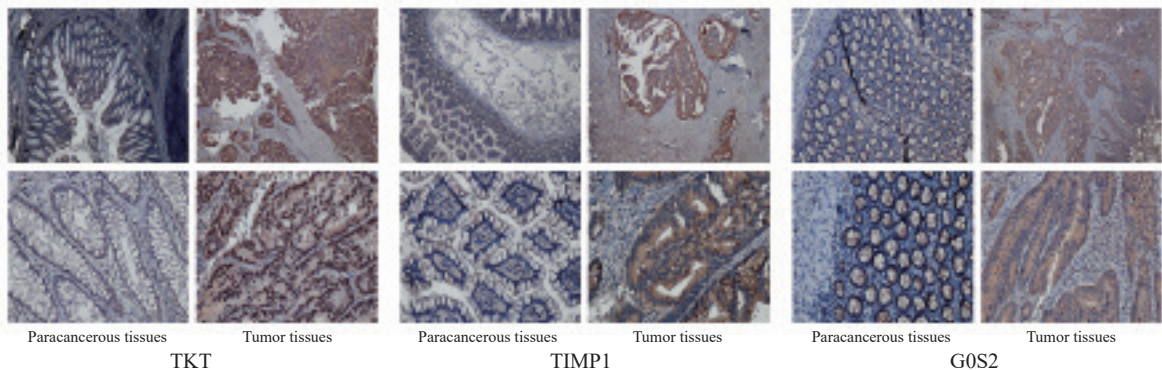


图4 检测结直肠癌组织和细胞中模型基因的表达情况

Fig. 4 Measurement of signature gene expression in CRC tissues and cells

A: RNA expression of TKT, TIMP1, G0S2, HSPA1A, CLEC4A, CD36, CTSD, VSIG4 and ASAH1 in tumor and paracancerous tissues. B: RNA expression of TKT, TIMP1, G0S2, HSPA1A, CLEC4A, CD36, CTSD, VSIG4 and ASAH1 in the mixed CRC cell lines (HCT116, HCT15, CACO2, HCT8, SW620, SW480, SW1116, RKO and DLD1) and the normal colon cell line NCM460. C: Representative images of immunohistochemical staining of TKT, TIMP1 and G0S2 in tumor and paracancerous tissues ($\times 20$, DAB; $\times 40$). *: $P < 0.05$, compared with each other; **: $P < 0.01$, compared with each other; ***: $P < 0.001$, compared with each other; ****: $P < 0.0001$, compared with each other.

2.3 免疫治疗效果和化疗敏感性预测

本研究分析了高风险组和低风险组患者对免疫治疗的反应情况，结果显示，高风险组对免疫治疗的反应显著低于低风险组（图5A）。高风险

组的TIDE得分较高，免疫逃逸的可能性更高，而且风险得分与TIDE得分呈正相关，相关系数为0.273，说明风险越高则免疫治疗效果越差（图5B、C）。

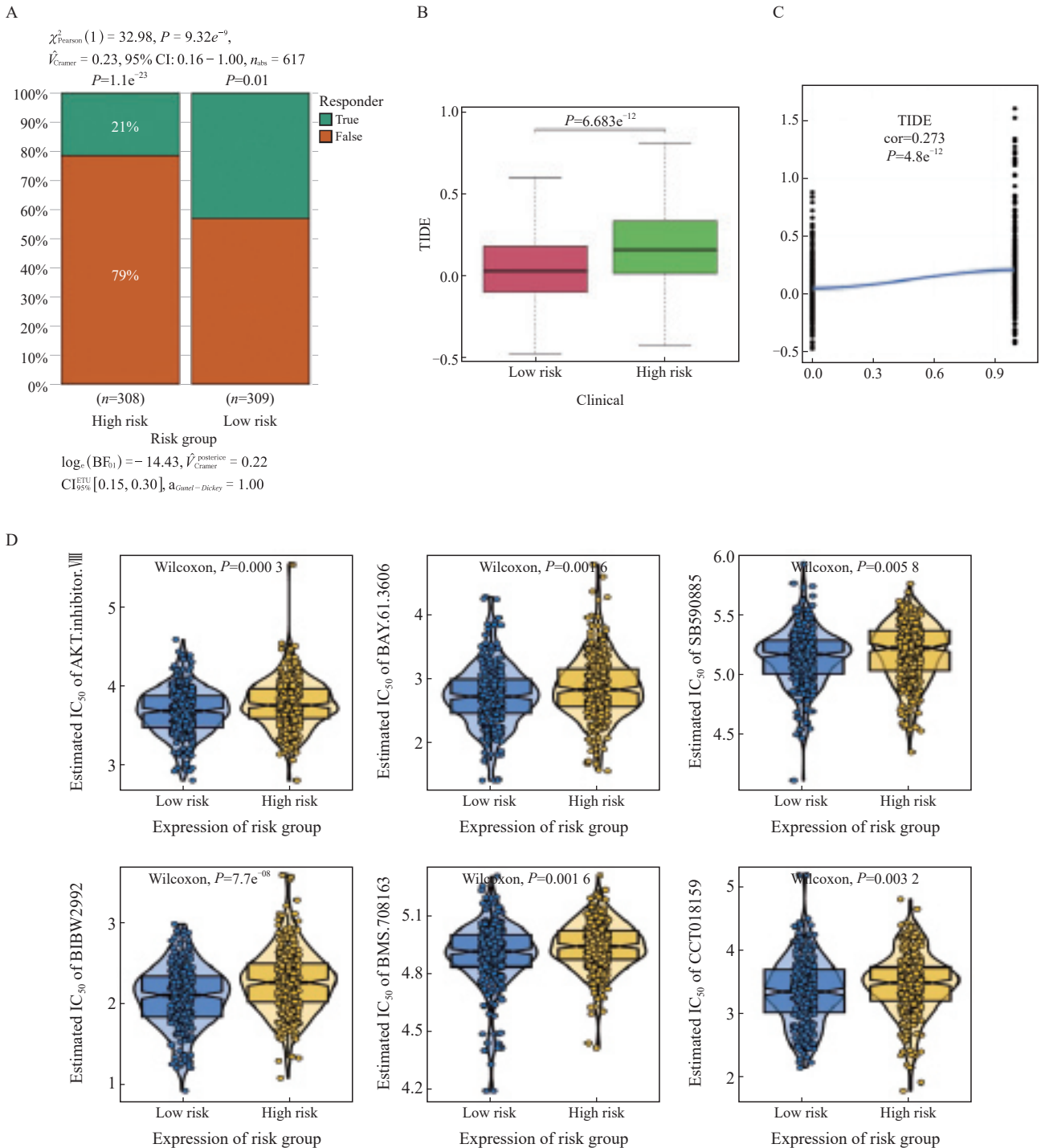


图5 免疫治疗效果和化疗敏感性预测

Fig. 5 Immunotherapy response and chemotherapy sensitivity prediction

A: Histogram of the response to immunotherapy in the high- and low-risk groups. B: Boxplot of TIDE scores in the high- and low-risk groups. C: Scatterplot of correlations between TIDE scores and risk scores. D: The IC_{50} values of 6 anti-cancer drugs of 2 risk groups of CRC patients.

手术联合化疗对早期CRC的效果明显，基于抗癌药物敏感性基因组学数据库（genomics of drug sensibility in cancer, GDSC）的药物敏感性数据，我们预测了每个肿瘤样本的化疗敏感性，并进一步探讨了风险评分与常用化疗药物敏感性的关系。研究表明，风险评分与患者对AKT抑制剂.V III、BAY.61.3606、SB590885、BIBW2992、BMS.708163、CCT018159等化疗药物的敏感性显著相关（图5D）。

2.4 风险评分影响肿瘤进展的潜在分子机制

我们探究了高低风险相关模型中涉及的具体信号转导通路，以探讨风险评分影响肿瘤进展的潜在分子机制。GSVA结果显示，两组患者的差异信号转导通路主要富集在Myogenesis、Epithelial-mesenchymal transition、Apical junction等信号转导通路中（图6A）。GSEA结果显示，富集的通路包括Cell adhesion molecules CAMS、ECM receptor interaction、Focal adhesion以及 Neuroactive ligand receptor interaction（图6B）。

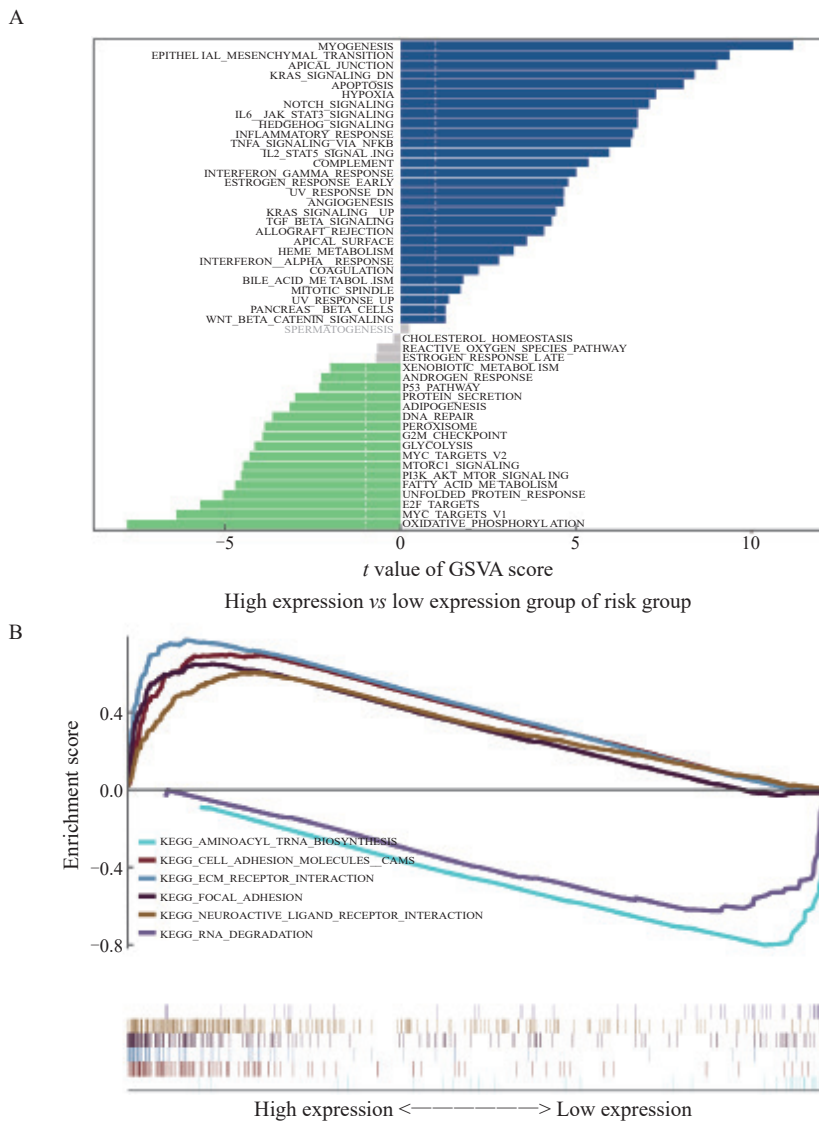


图6 探究影响肿瘤进展的潜在分子机制

Fig. 6 Exploration of the underlying molecular mechanisms affecting tumor progression

A: The results of GSVA showed the differential pathways between the high and low groups. B: The GSEA results showed the enriched pathways between the high- and low-expression groups.

2.5 发病风险与多个临床指标的相关性

为了探究临床特征与风险评分之间的关联，我们根据临床指标值的大小将样本对应的风险评分分成不同组，并以箱线图的形式展示了每个临床指标的分组结果。通过秩和检验，本研究发现风险评分的分布在Fustat分期、T分期、N分期和M分期等临床指标中的组间差异有统计学意义 ($P < 0.05$)，在晚期CRC中，Fustat分期、T分期、N分期和M分期的风险评分较高

(图7A~E)。这提示建模分析得到的风险评分对CRC样本的分型具有良好的适用性。为了进一步评估该模型在预测患者预后方面的临床应用价值，我们利用TCGA-CRC中的数据进行了单变量和多变量分析，结果显示，风险评分与预后相关(图7F)。此外，多变量分析进一步确认了风险评分是CRC患者的独立预后因子(图7G)。总之，这些结果表明，9基因模型具有优异的预后预测效能。

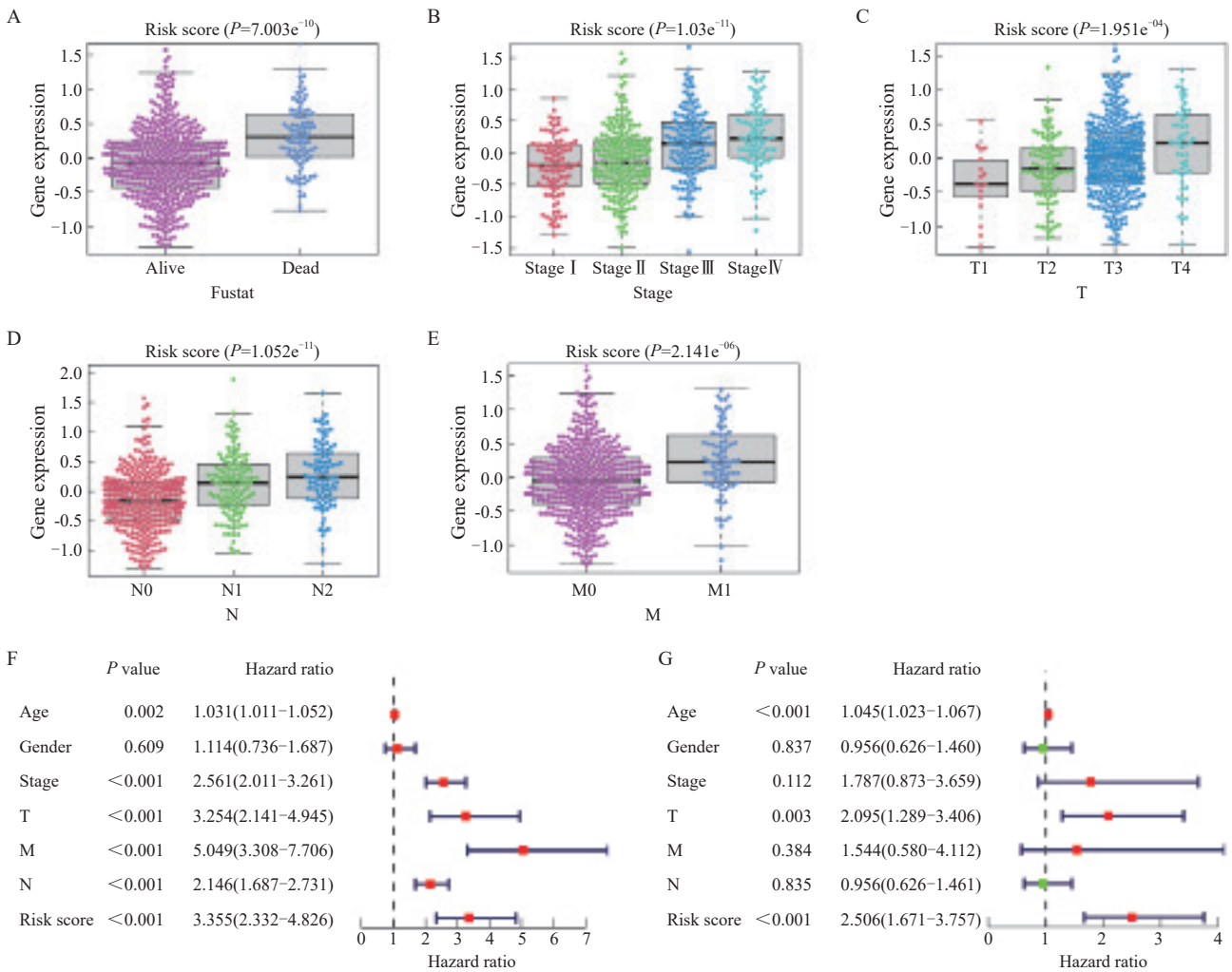


图7 模型的临床预测价值

Fig. 7 The clinical predictive values of the model

A-E: The differences in risk score by Fustat (A), stage (B), T stage (C), N stage (D), and M stage (E) of CRC. F-G: Univariate (F) and multivariate (G) Cox regression analysis in TCGA-CRC.

3 讨论

CRC是世界第三常见的肿瘤，分别占有

男性和女性癌症死亡数的9%和8%^[9]。CRC患者的预后受多种因素影响，其中重要的因素之一为TME。TME是由肿瘤细胞、周围免疫和炎症细胞、肿瘤相关成纤维细胞、肿瘤间质组织、微

血管以及各种细胞因子和趋化因子共同组成的整体^[10]。通过分泌信号分子,肿瘤影响TME,促进血管生成和免疫耐受,而TME中的免疫细胞则往往促进肿瘤生长^[11]。已有大量研究^[12-14]发现TME中的各种细胞直接参与肿瘤的发生、发展,进而影响患者预后。

高通量测序技术的发展使人们对CRC的发生、发展有了更加深入的了解^[15]。有研究表明,通过对TME特征进行分析,可以预测患者对化疗方案的响应情况,进而预测患者预后^[16-18]。scRNA-seq可以在单细胞水平上揭示肿瘤组织中细胞类型与表达谱,从而识别可能影响预后的相关标志基因。

我们的工作分析了scRNA-seq数据并联合TCGA数据库中的基因表达数据^[19],构建了一个含有9个差异表达基因的预后预测模型。ASAHI是参与鞘脂代谢的酶,有研究报道其参与调控乳腺癌的发展,可作为潜在的治疗靶点^[20-22]。在CRC中,研究发现carmofur作为一种新的AHAS1抑制剂具有作为化疗增敏剂的潜在价值^[23-25]。TKT作为参与葡萄糖代谢的酶,与多种肿瘤如胰腺癌、乳腺癌、宫颈癌和CRC的发生、发展相关^[26-29]。此外,在CRC中,TKT可通过与sirtuin 5协作对抗DNA损伤,从而促进结肠癌细胞存活^[30]。CLEC4A是一种C型凝集素受体,在小鼠树突状细胞中特异性表达,并具有促肿瘤效应^[31]。VSIG4是一种巨噬细胞上的补体激活受体,VSIG4⁺巨噬细胞和VSIG4的细胞外结构域抑制了补体途径和T淋巴细胞的激活。它还促进诱导调节性T淋巴细胞(induced regulatory T cells, iTreg)的分化以介导免疫抑制,促进肿瘤发生、发展^[32]。因此,VSIG4可作为癌症治疗中的免疫检查点靶点^[33]。HSPA1A是热激蛋白70(heat-shock protein 70, HSP70)的同源蛋白,是一种分子伴侣,可维持蛋白质稳态^[34]。HSPA1A在癌症中具有促进作用,如在ARID2缺陷的肺癌中作为其下游效应器^[35]。CRC肿瘤移植模型中H3K14ac修饰升高可促进Hsp1a1的转录和分泌,促进细胞毒性T淋巴细胞的免疫活性,表现出抗肿瘤活性^[36]。TIMP1是内源性基

质金属蛋白酶抑制蛋白,其在肿瘤中的具体作用存在争议。Guccini等^[37]认为,TIMP1是前列腺癌中决定衰老效应的分子开关,TIMP1敲除通过激活基质金属蛋白酶重编程肿瘤细胞衰老相关的分泌性表型,促进肿瘤转移。而Tian等^[38]则发现在胰腺导管腺癌(pancreatic ductal adenocarcinoma, PDAC)中,肿瘤细胞通过TIMP1与其周边的Schwann细胞互作,实现周围神经侵袭。Schoeps等^[39]鉴定了TIMP1在PDAC微环境中对中性粒细胞细胞外诱捕网形成的作用,并证实其可促进PDAC的转移。TIMP1在不同肿瘤中的作用差异可能是由其激活的不同效应细胞和下游信号转导通路导致,在CRC中,Ma等^[40]报道抑制TIMP1/FAK/Akt信号转导通路可抑制右侧CRC的增殖和侵袭。G0S2是一种主要定位于线粒体的蛋白质,具有调节能量利用、脂肪代谢和免疫功能的作用^[41-43]。G0S2被认为具有抑制肿瘤的作用,并且在各种肿瘤中存在高度甲基化修饰^[44-46]。CTSD是与细胞自噬相关的酶,可作为三阴性乳腺癌免疫治疗的靶点^[47]。此外,CTSD/ERK被发现与非小细胞肺癌(non-small cell lung cancer, NSCLC)的进展相关,抑制CTSD/ERK信号转导通路可以有效地抑制NSCLC细胞的增殖和转移^[48]。有相关预后模型分析工作认为CTSD可能是CRC的治疗靶点^[49]或早期筛查的标志物^[50]。此外,在periplocin处理的CRC细胞中,Wang等^[51]发现CTSD的表达下降引起溶酶体活性增加,发挥抗肿瘤作用,提示CTSD可能与CRC的进展相关。

本研究结果显示,G0S2、TIMP1和TKT在CRC细胞系和肿瘤组织中均高度表达,而HSPA1A、CLEC4A、CD36、CTSD、VSIG4和ASAHI在肿瘤组织和细胞系中下调。值得注意的是,有研究^[25]分析了ASAHI在结肠癌中促进肿瘤发展的机制,并指出其具有成为肿瘤靶点的潜力,这也印证了我们所建模型的可靠性。CLEC4A则从未被预测或证明与CRC相关。综上所述,本模型具有良好的预测CRC患者预后的潜力,针对模型内基因的进一步研究也将加深我们对CRC的理解,并提供新的治疗策略。

通过整合scRNA-seq和大量RNA-seq数据,本研究成功地构建了包括9个基因的免疫细胞预后预测模型,可用于评估CRC患者的生存和治疗反应。

利益冲突声明: 所有作者均声明不存在利益冲突。

作者贡献声明:

唐楠: 提出概念、基本框架、提出研究方向、设计论文框架; 设计研究思路、设计研究方案; 起草论文、修订论文、审核论文; 数据分析; 绘制图表; 黄慧霞: 进行实验验证、参与撰写论文; 刘晓健: 提出概念、基本框架、提出研究方向、设计论文框架; 参与论文修订、论文最终版本修订。

[参 考 文 献]

- [1] DEKKER E, TANIS P J, VLEUGELS J L A, et al. Colorectal cancer [J] . Lancet, 2019, 394(10207): 1467–1480.
- [2] EDWARDS B K, NOONE A M, MARIOTTO A B, et al. Annual Report to the Nation on the status of cancer, 1975–2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer [J] . Cancer, 2014, 120(9): 1290–1314.
- [3] PATEL S G, KARLITZ J J, YEN T, et al. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection [J] . Lancet Gastroenterol Hepatol, 2022, 7(3): 262–274.
- [4] HAN L, WANG S Y, WEI C, et al. Tumour microenvironment: a non-negligible driver for epithelial-mesenchymal transition in colorectal cancer [J] . Expert Rev Mol Med, 2021, 23: e16.
- [5] LUO X J, ZHAO Q, LIU J, et al. Novel genetic and epigenetic biomarkers of prognostic and predictive significance in stage II/III colorectal cancer [J] . Mol Ther, 2021, 29(2): 587–596.
- [6] CHEN B, SCURRAH C R, MCKINLEY E T, et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps [J] . Cell, 2021, 184(26): 6262–6280.e26.
- [7] GUO W, ZHANG C Y, WANG X, et al. Resolving the difference between left-sided and right-sided colorectal cancer by single-cell sequencing [J] . JCI Insight, 2022, 7(1): e152616.
- [8] NORKIN M, ORDÓÑEZ-MORÁN P, HUELSKEN J. High-content, targeted RNA-seq screening in organoids for drug discovery in colorectal cancer [J] . Cell Rep, 2021, 35(3): 109026.
- [9] SIEGEL R L, MILLER K D, WAGLE N S, et al. Cancer statistics, 2023 [J] . CA Cancer J Clin, 2023, 73(1): 17–48.
- [10] JIN M Z, JIN W L. The updated landscape of tumor microenvironment and drug repurposing [J] . Sig Transduct Target Ther, 2020, 5: 166.
- [11] PITT J M, MARABELLE A, EGGERMONT A, et al. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy [J] . Ann Oncol, 2016, 27(8): 1482–1492.
- [12] JACKSTADT R, VAN HOOFF S R, LEACH J D, et al. Epithelial NOTCH signaling rewires the tumor microenvironment of colorectal cancer to drive poor-prognosis subtypes and metastasis [J] . Cancer Cell, 2019, 36(3): 319–336.e7.
- [13] TIAN S B, CHU Y N, HU J, et al. Tumour-associated neutrophils secrete AGR2 to promote colorectal cancer metastasis via its receptor CD98hc - xCT [J] . Gut, 2022, 71(12): 2489–2501.
- [14] YIN Y, LIU B X, CAO Y L, et al. Colorectal cancer-derived small extracellular vesicles promote tumor immune evasion by upregulating PD-L1 expression in tumor-associated macrophages (adv. sci. 9/2022) [J] . Adv Sci, 2022, 9(9).
- [15] DEL VECCHIO F, MASTROIACO V, MARCO A D, et al. Next-generation sequencing: recent applications to the analysis of colorectal cancer [J] . J Transl Med, 2017, 15(1): 246.
- [16] LIANG L L, YU J, LI J, et al. Integration of scRNA-seq and bulk RNA-seq to analyse the heterogeneity of ovarian cancer immune cells and establish a molecular risk model [J] . Front Oncol, 2021, 11: 711020.
- [17] QU X D, ZHAO X Y, LIN K X, et al. M2-like tumor-associated macrophage-related biomarkers to construct a novel prognostic signature, reveal the immune landscape, and screen drugs in hepatocellular carcinoma [J] . Front Immunol, 2022, 13: 994019.
- [18] LU J, CHEN Y F, ZHANG X Q, et al. A novel prognostic model based on single-cell RNA sequencing data for hepatocellular carcinoma [J] . Cancer Cell Int, 2022, 22(1): 38.
- [19] ZHENG X B, SONG J N, YU C E, et al. Single-cell transcriptomic profiling unravels the adenoma-initiation role of protein tyrosine kinases during colorectal tumorigenesis [J] . Signal Transduct Target Ther, 2022, 7(1): 60.
- [20] VIJAYAN Y, LANKADASARI M B, HARIKUMAR K B. Acid ceramidase: a novel therapeutic target in cancer [J] . Curr Top Med Chem, 2019, 19(17): 1512–1520.
- [21] LUCKI N C, SEWER M B. Genistein stimulates MCF-7 breast cancer cell growth by inducing acid ceramidase (ASAHI) gene expression [J] . J Biol Chem, 2011, 286(22): 19399–19409.
- [22] LI Y H, LIU H T, XU J, et al. The value of detection of S100A8 and ASAH1 in predicting the chemotherapy response for breast cancer patients [J] . Hum Pathol, 2018, 74: 156–163.
- [23] REALINI N, SOLORZANO C, PAGLIUCA C, et al. Discovery of highly potent acid ceramidase inhibitors with in vitro tumor chemosensitizing activity [J] . Sci Rep, 2013, 3: 1035.
- [24] MACHALA M, PROCHÁZKOVÁ J, HOFMANOVÁ J, et al.

- Colon cancer and perturbations of the sphingolipid metabolism [J]. *Int J Mol Sci*, 2019, 20(23): 6051.
- [25] VIJAYAN Y, JAMES S, VISWANATHAN A, et al. Targeting acid ceramidase enhances antitumor immune response in colorectal cancer [J]. *J Adv Res*, 2023: S2090–S1232(23)00403–4.
- [26] LI M L, ZHAO X, YONG H M, et al. Transketolase promotes colorectal cancer metastasis through regulating AKT phosphorylation [J]. *Cell Death Dis*, 2022, 13(2): 99.
- [27] YANG H, WU X L, WU K H, et al. MicroRNA-497 regulates cisplatin chemosensitivity of cervical cancer by targeting transketolase [J]. *Am J Cancer Res*, 2016, 6(11): 2690–2699.
- [28] SHUKLA S K, PUROHIT V, MEHLA K, et al. MUC1 and HIF-1 α signaling crosstalk induces anabolic glucose metabolism to impart gemcitabine resistance to pancreatic cancer [J]. *Cancer Cell*, 2017, 32(3): 392.
- [29] DASGUPTA S, RAJAPAKSHE K, ZHU B K, et al. Metabolic enzyme PFKFB4 activates transcriptional coactivator SRC-3 to drive breast cancer [J]. *Nature*, 2018, 556(7700): 249–254.
- [30] WANG H L, CHEN Y, WANG Y Q, et al. Sirtuin5 protects colorectal cancer from DNA damage by keeping nucleotide availability [J]. *Nat Commun*, 2022, 13(1): 6121.
- [31] UTO T, FUKAYA T, MITOMA S, et al. Clec4A4 acts as a negative immune checkpoint regulator to suppress antitumor immunity [J]. *Cancer Immunol Res*, 2023, 11(9): 1266–1279.
- [32] LIU B, CHENG L, GAO H H, et al. The biology of VSIG4: implications for the treatment of immune-mediated inflammatory diseases and cancer [J]. *Cancer Lett*, 2023, 553: 215996.
- [33] CHOW A, SCHAD S, GREEN M D, et al. Tim-4⁺ cavity-resident macrophages impair anti-tumor CD8⁺ T cell immunity [J]. *Cancer Cell*, 2021, 39(7): 973–988.e9.
- [34] WU S, PEI Q, NI W, et al. HSPA1A protects cells from thermal stress by impeding ESCRT-0-mediated autophagic flux in epidermal thermoresistance [J]. *J Invest Dermatol*, 2021, 141(1): 48–58.e3.
- [35] WANG X, WANG Y T, FANG Z Y, et al. Targeting HSPA1A in ARID2-deficient lung adenocarcinoma [J]. *Natl Sci Rev*, 2021, 8(10): nwab014.
- [36] JIANG Y, XU Y J, ZHENG C, et al. Acetyltransferase from *Akkermansia muciniphilab* blunts colorectal tumorigenesis by reprogramming tumour microenvironment [J]. *Gut*, 2023, 72(7): 1308–1318.
- [37] GUCCINI I, REVANDKAR A, D'AMBROSIO M, et al. Senescence reprogramming by TIMP1 deficiency promotes prostate cancer metastasis [J]. *Cancer Cell*, 2021, 39(1): 68–82.e9.
- [38] TIAN Z F, OU G S, SU M X, et al. TIMP1 derived from pancreatic cancer cells stimulates Schwann cells and promotes the occurrence of perineural invasion [J]. *Cancer Lett*, 2022, 546: 215863.
- [39] SCHOEPS B, ECKFELD C, PROKOPCHUK O, et al. TIMP1 triggers neutrophil extracellular trap formation in pancreatic cancer [J]. *Cancer Res*, 2021, 81(13): 3568–3579.
- [40] MA B B, UEDA H, OKAMOTO K, et al. TIMP1 promotes cell proliferation and invasion capability of right-sided colon cancers via the FAK/Akt signaling pathway [J]. *Cancer Sci*, 2022, 113(12): 4244–4257.
- [41] HECKMANN B L, ZHANG X D, XIE X T, et al. The G₀/G₁ switch gene 2 (GOS2): regulating metabolism and beyond [J]. *Biochim Biophys Acta*, 2013, 1831(2): 276–281.
- [42] NIELSEN T S, MØLLER N. Adipose triglyceride lipase and G₀/G₁ switch gene 2: approaching proof of concept [J]. *Diabetes*, 2014, 63(3): 847–849.
- [43] KIOKA H, KATO H, FUJIKAWA M, et al. Evaluation of intramitochondrial ATP levels identifies G₀/G₁ switch gene 2 as a positive regulator of oxidative phosphorylation [J]. *Proc Natl Acad Sci U S A*, 2014, 111(1): 273–278.
- [44] CHANG X F, MONITTO C L, DEMOKAN S, et al. Identification of hypermethylated genes associated with cisplatin resistance in human cancers [J]. *Cancer Res*, 2010, 70(7): 2870–2879.
- [45] BARREAU O, ASSIÉ G, WILMOT-ROUSSEL H, et al. Identification of a CpG island methylator phenotype in adrenocortical carcinomas [J]. *J Clin Endocrinol Metab*, 2013, 98(1): E174–E184.
- [46] KUSAKABE M, KUTOMI T, WATANABE K, et al. Identification of GOS2 as a gene frequently methylated in squamous lung cancer by combination of in silico and experimental approaches [J]. *Int J Cancer*, 2010, 126(8): 1895–1902.
- [47] ASHRAF Y, MANSOURI H, LAURENT-MATHA V, et al. Immunotherapy of triple-negative breast cancer with cathepsin D-targeting antibodies [J]. *J Immunother Cancer*, 2019, 7(1): 29.
- [48] LIU C M, SHEN H T, LIN Y A, et al. Antiproliferative and antimetastatic effects of praeurptorin C on human non-small cell lung cancer through inactivating ERK/CTSD signalling pathways [J]. *Molecules*, 2020, 25(7): 1625.
- [49] XU J S, DAI S Q, YUAN Y, et al. A prognostic model for colon cancer patients based on eight signature autophagy genes [J]. *Front Cell Dev Biol*, 2020, 8: 602174.
- [50] XIE L Q, ZHAO C, CAI S J, et al. Novel proteomic strategy reveal combined alpha1 antitrypsin and cathepsin D as biomarkers for colorectal cancer early screening [J]. *J Proteome Res*, 2010, 9(9): 4701–4709.
- [51] WANG K, FU S Y, DONG L X, et al. Periplocin suppresses the growth of colorectal cancer cells by triggering LGALS3 (galectin 3)-mediated lysophagy [J]. *Autophagy*, 2023, 19(12): 3132–3150.

(收稿日期: 2024-02-06 修回日期: 2024-06-05)