

论著·临床研究

肝细胞癌相关的核编码线粒体基因及临床信息的综合预后模型

克德尔亚·艾山江^{1,2}, 傅 怡², 赖冬林^{2,3}, 邬海龙^{2,4#}, 龚 伟^{1,5#}

1. 上海交通大学医学院附属新华医院普外科, 上海 200092; 2. 上海健康医学院协同科研中心, 上海 201318; 3. 南昌大学江西医学院第一附属医院, 南昌 330006; 4. 上海健康医学院药学院, 上海 201318; 5. 上海市胆道疾病研究重点实验室, 上海交通大学医学院胆道疾病研究所, 上海市胆道疾病研究中心, 上海 200092

[摘要] **目的**·建立一个基于线粒体基因和临床信息的肝细胞癌 (hepatocellular carcinoma, HCC) 总生存率 (overall survival, OS) 的预后模型。**方法**·从癌症基因组图谱 (The Cancer Genome Atlas, TCGA) 下载 369 例 HCC 患者和 50 例肝脏正常对照的基因表达谱和临床数据。核编码的线粒体基因 (nuclear encoded mitochondrial gene, NEMG) 从 MitoCarta3.0 数据库获得。使用 “DESeq2” R 包和单变量 Cox 分析选择与 HCC 患者 OS 相关并参与氧化磷酸化、三羧酸循环和细胞凋亡通路的 NEMG [泛素细胞色素 C 还原酶铰链蛋白 (ubiquinol cytochrome C reductase hinge protein, *UQCRRH*)、腺苷三磷酸柠檬酸裂解酶 (ATP citrate lyase, *ACLY*)、磷酸烯醇式丙酮酸羧激酶 2 (phosphoenolpyruvate carboxykinase 2, *PCK2*)、Bcl-2 同源拮抗剂 1 (Bcl-2 homologous antagonist/killer 1, *BAK1*)、Bcl-2 相关 X 蛋白 (Bcl-2-associated X protein, *BAX*) 和 Bcl-2/腺病毒 E1B 相互作用蛋白 3 样 (Bcl-2/adenovirus E1B interacting protein 3-like, *BNIP3L*)]。应用多变量 Cox 回归来确定 HCC OS 的独立危险因素。建立一个基于独立危险因素 (6 个 NEMG 风险特征和 TNM 分期) 的综合预后模型和预后列线图, 计算中位预后评分。以中位预后评分作为分界点, 将 HCC 患者分为高风险组和低风险组。进行 Kaplan-Meier 生存曲线分析, 并进行对数秩检验来评估低风险组和高风险组患者 OS 的差异。使用 “timeROC” 软件包计算受试者操作特征 (receiver operating characteristic, ROC) 曲线下面积 (area under the curve, AUC)。用基因表达数据库 (Gene Expression Omnibus, GEO) 下载 HCC 队列 (GSE14520) 验证综合预后模型对 1、3、5 年 OS 的预测性能。通过实时荧光定量聚合酶链反应 (real-time quantitative polymerase chain reaction, qPCR) 在来自上海交通大学医学院附属新华医院的 34 例 HCC 临床样本中验证 6-NEMG 的相对表达水平。**结果**·ROC 分析结果显示, 与仅 6-NEMG 风险特征 (1、3、5 年 AUC 分别为 0.77、0.66、0.65) 或仅 TNM 分期 (1、3、5 年 AUC 分别为 0.66、0.67、0.63) 相比, 该综合预后模型对 1 年 (AUC, 0.78)、3 年 (AUC, 0.73) 和 5 年 (AUC, 0.69) HCC OS 显示出更好的预测性能。Kaplan-Meier 生存曲线分析结果显示高风险组患者的 OS 明显比低风险组差 ($P=0.001$)。此外, 在 GEO 外部队列中发现该预后模型的预测性能较好 (1、3、5 年 AUC 分别为 0.67、0.66、0.74), 高、低风险组患者的预后差异有统计学意义 ($P=0.001$), 与 TCGA 数据的结果一致。在临床 HCC 队列中, 与癌旁肝脏组织相比, 除 *BNIP3L* 外, 其他 5 个 NEMG 在肿瘤组织的表达水平上调或者下调。相关性分析显示, 在 GSE14520 与临床 HCC 队列中预后评分与 HCC 肿瘤的大小和数量呈正相关。**结论**·构建并验证了一个将 6-NEMG 风险特征与 TNM 分期相结合的 HCC 预后预测模型。该模型可能有助于 HCC 患者的预后预测。

[关键词] 核编码线粒体基因; 肝细胞癌; 癌症基因组图谱 (TCGA); 基因表达数据库 (GEO); 总生存率

[DOI] 10.3969/j.issn.1674-8115.2024.01.001 **[中图分类号]** R657.3; R735.7 **[文献标志码]** A

An integrated prognostic model of nuclear-encoded mitochondrial gene signature and clinical information for hepatocellular carcinoma

Kedeerya Aishanjiang^{1,2}, FU Yi², LAI Donglin^{2,3}, WU Hailong^{2,4#}, GONG Wei^{1,5#}

1. Department of General Surgery, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China;

[基金项目] 国家自然科学基金 (31870905, 82172628, 81974371); 上海市卫生健康委员会基金 (201940352); 上海市科学技术委员会基金 (22ZR1428100); 上海交通大学医学院 “双百人” 项目 (20151001); 上海交通大学医学院附属新华医院临床研究项目 (21XHDB10); 上海市胆道疾病重点实验室研究基金 (17DZ2260200)。

[作者简介] 克德尔亚·艾山江 (1995—), 女, 维吾尔族, 硕士生; 电子信箱: kadirya95@163.com。

[通信作者] 邬海龙, 电子信箱: wuhl@sumhs.edu.cn。龚 伟, 电子信箱: gongwei@xinhumed.com.cn。[#]为共同通信作者。

[Funding Information] National Natural Science Foundation of China (31870905, 82172628, 81974371); Scientific Program of Shanghai Municipal Health Commission (201940352); General Project of Shanghai Municipal Commission of Science and Technology (22ZR1428100); “Two-hundred Talents” Program of Shanghai Jiao Tong University School of Medicine (20151001); Xinhua Hospital Funded Clinical Research (21XHDB10); Shanghai Key Laboratory of Biliary Tract Disease Research Foundation (17DZ2260200)。

[Corresponding Author] WU Hailong, E-mail: wuhl@sumhs.edu.cn. GONG Wei, E-mail: gongwei@xinhumed.com.cn. [#]Co-corresponding authors.



2. Collaborative Innovation Center for Biomedicine, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China; 3. The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang 330006, China; 4. School of Pharmacy, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China; 5. Shanghai Key Laboratory of Biliary Tract Disease Research, Research Institute of Biliary Tract Disease, Shanghai Jiao Tong University School of Medicine, Shanghai Research Center of Biliary Tract Disease, Shanghai 200092, China

[Abstract] **Objective** To establish a prognostic model for the overall survival (OS) of hepatocellular carcinoma (HCC) based on mitochondrial genes and clinical information. **Methods** The gene expression and the clinical data of 369 HCC patients and 50 controls with normal liver were downloaded from The Cancer Genome Atlas (TCGA) database. The nuclear-encoded mitochondrial genes (NEMGs) were obtained from the MitoCarta3.0 database. The "DESeq2" R package and univariate Cox analysis were used to select NEMGs [ubiquinol cytochrome C reductase hinge protein (*UQCRH*), ATP citrate lyase (*ACLY*), phosphoenolpyruvate carboxykinase 2 (*PCK2*), Bcl-2 homologous antagonist/killer1 (*BAK1*), Bcl-2-associated X protein (*BAX*) and Bcl-2/adenovirus E1B interacting protein 3-like (*BNIP3L*)] in HCC that were associated with OS of HCC and participated in dysregulation of oxidative phosphorylation, tricarboxylic acid cycle and cell apoptosis. Multivariate Cox analysis was applied to select independent risk factors for OS of HCC. A comprehensive prognostic model and a prognostic nomogram with 6-NEMG risk characteristics and TNM staging were established. By using the median of prognostic scores as a cut-off, HCC patients were classified into low-risk and high-risk group. Kaplan-Meier survival curve analysis was conducted and log-rank test was performed to evaluate the survival rates between the low-risk and high-risk group. The area under the curve (AUC) values of receiver operating characteristic (ROC) curve were calculated via using the "timeROC" package. The prognostic model for HCC was validated by using the GEO HCC cohort (GSE14520) for 1, 3 and 5 years. Finally, the relative expression level of 6-NEMG was validated in 34 clinical samples of HCC from Xinhua Hospital, Shanghai Jiao Tong University School of Medicine by using real-time quantitative polymerase chain reaction (qPCR) method. **Results** Compared to 6-NEMG risk signature only (AUCs for 1, 3 and 5 years were 0.77, 0.66 and 0.65, respectively) or TNM stage only (AUCs for 1, 3 and 5 years were 0.66, 0.67 and 0.63, respectively), ROC curve analysis showed that this integrated prognostic model displayed better predictive performance for 1-year (AUC, 0.78), 3-year (AUC, 0.73) and 5-year (AUC, 0.69) OS of HCC. The Kaplan-Meier survival curve analysis showed that the OS of HCC patients in the high-risk group was significantly worse than that in the low-risk group ($P=0.001$). In addition, predictive performance of the prognostic model (AUC for 1, 3 and 5 years is 0.67, 0.66 and 0.74, respectively) and prognostic differences between the high-risk and low-risk group ($P=0.001$) were further validated in GEO (GSE14520) external cohort, and these results were consistent with the TCGA data. In addition to *BNIP3L*, dysregulation of five other NEMGs in the clinical HCC cohort was validated. The correlation analysis in GSE14520 and HCC clinical cohort showed a positive correlation between prognosis score and the size and number of tumors. **Conclusion** A new prognostic model that combines 6-NEMG risk characteristics with TNM staging for predicting OS in HCC patients was constructed and validated. This model may help improve the prognosis prediction of HCC patients.

[Key words] nuclear encoded mitochondrial gene; hepatocellular carcinoma; TCGA; GEO; overall survival

在全球范围内,原发性肝癌是最具侵袭性和难以治疗的恶性肿瘤之一^[1]。肝细胞癌 (hepatocellular carcinoma, HCC) 约占所有原发性肝癌病例的90%^[2]。尽管早期和晚期HCC的治疗方法都有所改善^[3-4],但近几十年来,许多国家的发病率和死亡率仍然很高^[5]。由于预后不良,HCC目前是全球第四大常见的癌症相关死亡原因,5年生存率低于21%^[6]。因此,建立一个能够准确预测HCC预后的模型对于改善HCC预后至关重要。虽然许多分类系统,如美国癌症联合委员会 (American Joint Committee on Cancer, AJCC) 的TNM分期 (American Joint Committee on Cancer-Tumor Node Metastasis Staging) 系统、巴塞罗那临床肝癌 (Barcelona Clinic Liver Cancer, BCLC) 分期和意大利肝癌评分 (Cancer of the Liver Italian Program, CLIP) 系统,已

广泛应用于临床评价HCC患者的预后,但这些分类系统主要是基于临床病理特征,没有考虑分子标志物的重要预后作用^[7]。

线粒体是一种双层膜细胞器,在真核细胞中作为动力源,在细胞存活、凋亡、代谢和信号转导中起着至关重要的作用^[8-9]。近99%的线粒体蛋白质是由位于细胞核上的核编码线粒体基因 (nuclear encoded mitochondrial gene, NEMG) 编码的^[10]。这些NEMG参与多种线粒体途径,包括氧化磷酸化 (oxidative phosphorylation, OXPHOS)、细胞凋亡、三羧酸 (tricarboxylic acid, TCA) 循环、 β 氧化、酮体生成等^[11-13],对线粒体稳态至关重要。越来越多的证据表明,线粒体的功能障碍和失调参与癌症的发生和进展。线粒体DNA的存在对癌细胞的生长和致瘤性有着重要作用^[14]。在化学治疗耐药的癌症干细

胞中观察到增强的 OXPHOS^[15-16]。在许多癌症类型中,抑制 OXPHOS 可使癌细胞对化学治疗和靶向治疗重新敏感^[17-18]。此外,TCA 循环中一些酶的显性突变,如异柠檬酸脱氢酶 2 (isocitrate dehydrogenase 2, IDH2)^[19]、琥珀酸脱氢酶 (succinate dehydrogenase, SDH)^[20] 和延胡索酸水合酶 (fumarate hydratase, FH)^[21] 与肿瘤的起始和进展有关。此外,既往研究^[22-23]表明,NEMG 参与 HCC 的发病和进展。尽管 NEMG 在 HCC 的发生发展中很重要,但基于 NEMG 的预后模型研究有限且不够成熟。因此,建立一个基于 NEMG 的预后模型可能对 HCC 的预后有很高的预测价值。

本研究基于生物信息学数据库和 HCC 临床样本,构建并验证了一个基于 NEMG 风险基因和临床特征的 HCC 预后模型,以期改善 HCC 患者的风险分级和生存预测。

1 资料与方法

1.1 数据收集

分别从癌症基因组图谱 (The Cancer Genome Atlas, TCGA) 数据库及基因表达数据库 (Gene Expression Omnibus, GEO) 中下载 419 例 (369 例 HCC 和 50 例非 HCC 样本) TCGA-LIHC (liver hepatocellular carcinoma) 和 247 例 HCC 患者的基因表达和相应的临床信息。在移除没有随访记录和完整临床信息的样本后,把 TCGA 数据库中 339 例 HCC 样本和 50 例正常肝脏样本作为训练队列,GEO 数据集 (GSE14520) 中 219 例 HCC 患者作为验证队列。从 MitoCarta 3.0 数据库 (<https://www.broadinstitute.org/mitocarta>) 下载哺乳动物线粒体蛋白和通路清单,共包含 1 123 个线粒体基因。

1.2 HCC 中差异表达的核编码线粒体基因的分析

根据“DEseq2”R 包筛选差异表达基因 (differentially expressed gene, DEG),筛选标准: $|\log_2 \text{fold change (FC)}| > 0.5$, $Q (P \text{ adjusted}) < 0.05$ 。对 TCGA 数据库中 369 例 HCC 组织与 50 例正常肝脏组织的差异表达 NEMG (differentially expressed nuclear encoded mitochondrial gene, deNEMG) 进行分析,并将 TCGA 数据库中的 deNEMG 与验证队列 GSE14520 数据集中包含的基因取交集,共获得

264 个 deNEMG 以供进一步分析。

1.3 基于参与 OXPHOS、TCA 和凋亡通路的 deNEMG 风险模型构建

采用单变量比例风险回归 (Cox) 方法识别与 HCC 总生存率 (overall survival, OS) 相关的 deNEMG,筛选出参与 3 种重要途径 (OXPHOS、TCA 循环和细胞凋亡) 并与 OS 相关的 deNEMG,来构建基于 NEMG 的风险特征模型。该 6-NEMG 特征模型的风险评分采用线性公式计算: 风险评分 = $\sum \text{coefficient}_i \times \text{gene}_i$ (gene_i , 个体 HCC 患者中 deNEMG 的相对表达水平; coefficient_i , NEMG 的多变量 Cox 回归系数)。

1.4 HCC 的预后模型及列线图的建立及验证

对包括年龄、性别、甲胎蛋白 (α -fetoprotein, AFP) 和 TNM 分期等重要的临床病理特征以及 deNEMG 的风险特征通过“forestplot”和“survival”包进行单因素和多因素 Cox 回归分析,获得独立的预后因素。使用“rms”和“nomogramEx”软件包建立了一个基于独立预后因素的预后模型和预后列线图。根据该预后模型中涉及的独立预后因素的系数,制定公式来计算每个患者的预后评分。以预后评分的中位数作为临界值,将 HCC 患者分为高风险组和低风险组。采用“survival”和“survminer”包进行 Kaplan-Meier 生存曲线分析,并采用 log-rank 检验评估低风险组和高风险组之间的 OS。使用“timeROC”软件包计算受试者操作特征 (receiver operating characteristic, ROC) 曲线下面积 (area under the curve, AUC)。

1.5 临床样本收集

于 2018 年至 2020 年期间在上海交通大学医学院附属新华医院肝移植科收集 34 例行原发性肝切除术或肝移植术后 HCC 患者的癌和癌旁组织。所有病例均经病理确诊为 HCC,术前均未接受放射治疗或化学治疗。

1.6 实时荧光定量聚合酶链反应

用 TRIzol 试剂 (Invitrogen, 美国) 从 34 对肿瘤和癌旁组织中提取总 RNA。使用 PrimeScript RT 试剂盒 (Takara, 日本) 进行反转录。采用 SYBR

PrimeScript RT-PCR 试剂盒（Takara，日本）进行实时荧光定量聚合酶链反应（real-time quantitative

polymerase chain reaction, qPCR）分析。引物序列见表1。采用公式 $2^{-\Delta\Delta C_t}$ 计算基因的相对表达水平。

表1 qPCR引物序列
Tab 1 Primer sequences for qPCR

Primer	Forward sequence (5'→3')	Reverse sequence (5'→3')
<i>UQCRH</i>	GCTCTGTGATGAGCGTGTATCC	GTTGTGTTAAAGAGTTTGTGGGCCAC
<i>ACLY</i>	GCTCTGCCTATGACAGCACCAT	GTCCGATGATGGTCACTCCCTT
<i>PCK2</i>	TAGTGCCTGTGGCAAGACCAAC	GAAGCCGTTCTCAGGGTTGATG
<i>BAK1</i>	TTACCGCCATCAGCAGGAACAG	GGAACCTGAGTCATAGCGTCG
<i>BAX</i>	TCAGGATGCGTCCACCAAGAAG	TGTGTCCACGGCGGCAATCATC
<i>BNIP3L</i>	TGTGGAATGCACACCAGCAGG	CTACTGGACCACTCTGATACCC
18S	GGAGAGGGAGCTGAGAAACG	TTACAGGGCCTCGAAAGAGTCC

Note: *UQCRH*—ubiquinol cytochrome C reductase hinge protein; *ACLY*—ATP citrate lyase; *PCK2*—phosphoenolpyruvate carboxykinase 2; *BAK1*—Bcl-2 homologous antagonist/killer1; *BAX*—Bcl-2-associated X protein; *BNIP3L*—Bcl2/adenovirus E1B interacting protein 3-like; 18S—18S ribosomal RNA.

1.7 统计学分析

所有数据均使用 R 软件包（4.2.0 版）或 GraphPad Prism 软件（9.4.1 版）进行统计分析。定量资料以 $\bar{x} \pm s$ 表示。定性资料采用 n （%）表示。采用“corrplot”包进行预后评分与 HCC 临床信息之间的相关性分析。采用 Fisher 检验比较分类变量（性别、年龄）。采用对数秩检验的 Kaplan-Meier 曲线分析，评价不同风险组 OS 的差异。 $P < 0.05$ 表示差异具有统计学意义。

2 结果

2.1 参与3种重要途径且与预后相关的NEMG筛选

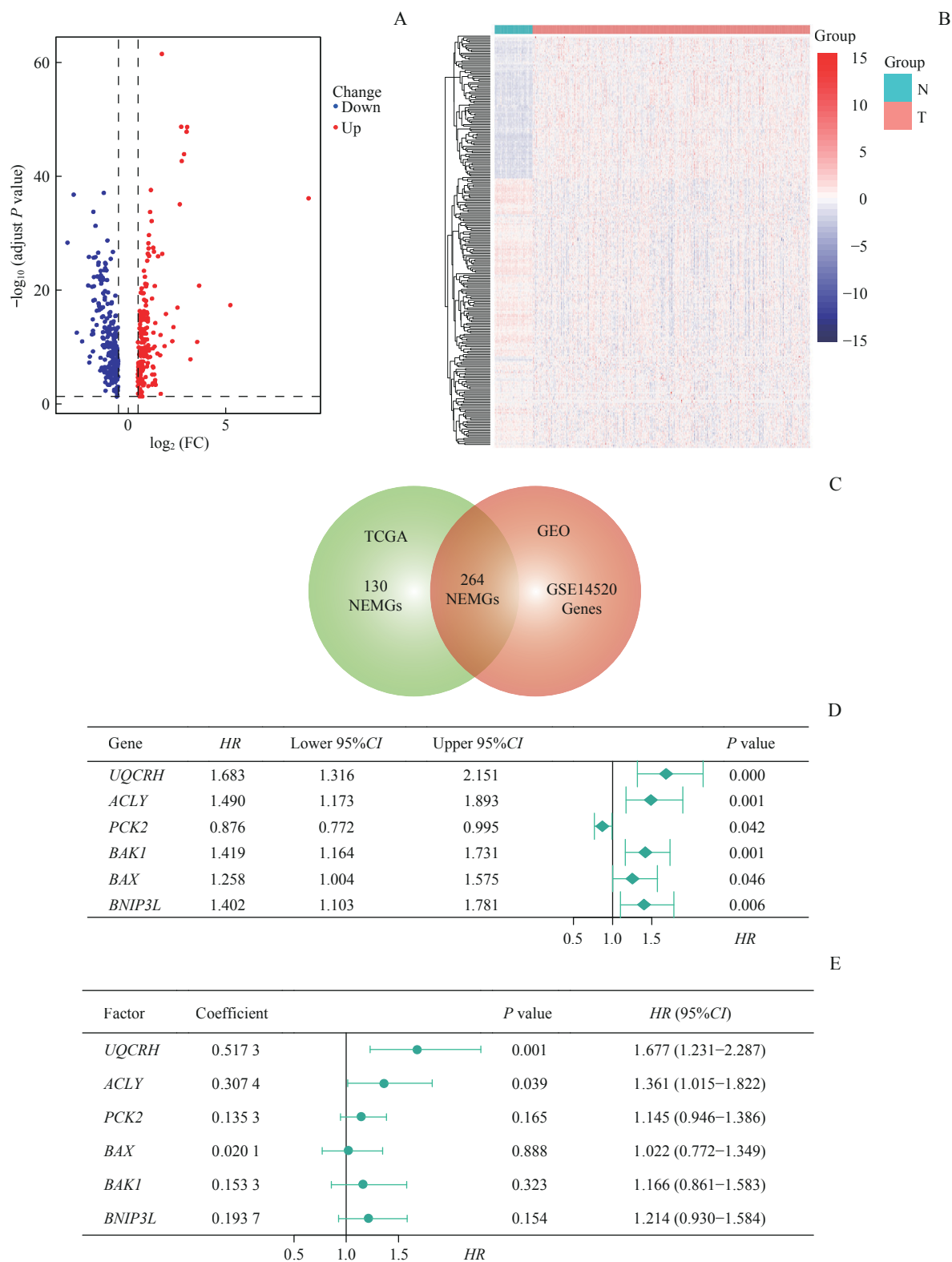
为了鉴定与 HCC 进展相关的 NEMG，使用“DEseq2”软件包在 TCGA-LIHC 队列中鉴定了 369 个 deNEMG ($|\log_2FC| > 0.5$, $P < 0.05$) (图 1A)。在 369 个 deNEMG 中，在 GEO 验证队列 GSE14520 中可检测到 264 个 deNEMG (102 个上调，162 个下调) (图 1B、C)。TCGA-LIHC 队列中 HCC 患者 OS 的单变量 Cox 回归分析显示，95 个 deNEMG 与 OS 显著相关 ($P < 0.05$)。考虑到 OXPHOS、TCA 循环和细胞凋亡途径对 HCC 进展很重要，取 95 个 deNEMG 和参与 3 个通路的 NEMG 的交集。选择 6 个 deNEMG [泛素细胞色素 C 还原酶铰链蛋白 (ubiquinol cytochrome C reductase hinge protein, *UQCRH*)、腺苷三磷酸柠檬酸裂解酶 (ATP citrate lyase, *ACLY*)、磷酸烯醇式丙酮酸羧激酶 2 (phosphoenolpyruvate carboxykinase 2, *PCK2*)、Bcl-2 同源拮抗剂 1 (Bcl-2 homologous

antagonist/killer 1, *BAK1*)、Bcl-2 相关 X 蛋白 (Bcl-2-associated X protein, *BAX*) 和 Bcl-2/腺病毒 E1B 相互作用蛋白 3 样 (Bcl-2/adenovirus E1B interacting protein 3-like, *BNIP3L*)] 来构建基于 NEMG 的风险特征模型 (图 1D)。通过多变量 Cox 回归分析得到 6 个基因的相关系数 (图 1E)。经计算，6-NEMG 特征的风险评分 = $0.517\ 3 \times UQCRH$ 的基因表达水平 + $0.307\ 4 \times ACLY$ 的基因表达水平 + $0.135\ 3 \times PCK2$ 的基因表达水平 + $0.153\ 3 \times BAK1$ 的基因表达水平 + $0.020\ 1 \times BAX$ 的基因表达水平 + $0.193\ 7 \times BNIP3L$ 的基因表达水平。

2.2 基于TCGA队列的HCC预后模型构建

通过对 HCC 预后相关 NEMG 进行单因素和多因素 Cox 分析，发现 TNM 分期和 6-NEMG 特征的风险评分是 HCC OS 的独立预后因素 (图 2A、B)。基于此本研究构建了一个整合了 TNM 分期和 6-NEMG 特征的预后模型。采用如下公式计算该综合预后模型的预后评分：预后评分 = \sum 风险评分 \times 系数 R + TNMi \times 系数 T (风险评分指的是 6-NEMG 风险特征在个体 HCC 患者的风险评分；TNMi 为个体 HCC 患者的 TNM 分期；系数 R 为风险评分的多变量 Cox 回归系数；系数 T 为 TNM 分期的多变量 Cox 回归系数)。经计算，预后评分 = $0.888 \times$ 风险评分 + $0.418 \times$ TNM。

时间依赖的 ROC 曲线显示，该综合预后模型对术后 1 年 (AUC = 0.78)、3 年 (AUC = 0.73) 和 5 年 (AUC = 0.69) 预后表现出良好的预测性能 (图 2C)。利用预后评分的中位数作为临界值，将 TCGA-LIHC

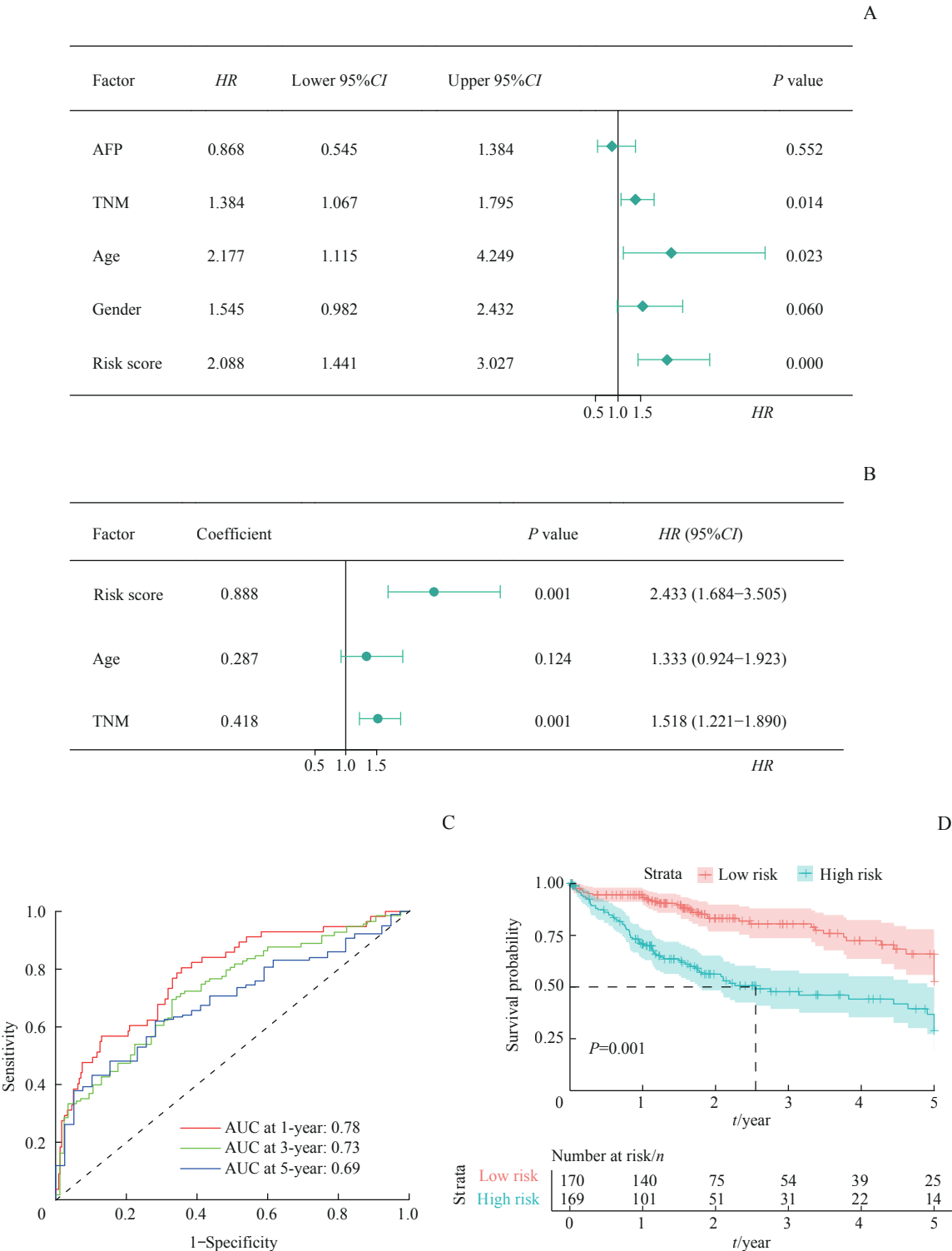


Note: A. DEGs between HCC and adjacent paired normal tissues. Red dots represent significant upregulation and blue dots represent significant downregulation of DEGs in HCC tissues. B. Heatmap of 264 deNEMGs in TCGA cohort. N—adjacent paired normal tissues; T—HCC tumor tissues. C. Venn diagram showing differently expressed deNEMGs between tumor and adjacent tissues. D. Univariate Cox regression analysis of 6 significant prognostic NEMGs associated with OXPHOS, TCA cycle and cell apoptosis. E. Multivariate Cox analysis of 6-NEMG risk signature.

图1 预后相关NEMG的识别和基于6-NEMG的风险特征模型的构建

Fig 1 Identification of prognosis-related NEMGs and construction of a 6-NEMG-based risk signature

队列（表2）中339例临床信息完整的HCC患者分为高风险组（ $n=169$ ）和低风险组（ $n=170$ ）。Kaplan-Meier曲线显示，高风险组患者的OS明显低于低风险组患者（ $P=0.001$ ，图2D）。



Note: A/B. Univariate (A) and multivariate (B) Cox analysis of risk score and clinicopathological features for OS of HCC. C. ROC analysis for the prognostic model at 1-, 3- and 5-year survival in the TCGA-LIHC cohort. D. Kaplan-Meier survival analysis of OS in the high- and low-risk group in the TCGA-LIHC cohort.

图2 6-NEMG 风险特征与 TNM 分期相结合的预后模型的构建
Fig 2 Construction of a prognostic model integrating the 6-NEMG risk signature with TNM stage

表 2 不同风险组患者的基本特征

Tab 2 Baseline characteristics of the patients in different risk groups

Feature	TCGA (n=339)			GEO (n=219)		
	High-risk group	Low-risk group	P value	High-risk group	Low-risk group	P value
Age/n(%)			0.681			0.969
≤50 years	36 (21.30)	39 (22.94)		53 (11.93)	58 (52.73)	
>50 years	133 (78.73)	131 (77.06)		56 (88.07)	52 (47.27)	
Gender/n(%)			0.254			0.156
Male	107 (63.31)	124 (72.94)		96 (88.07)	93 (84.55)	
Female	62 (36.69)	46 (27.06)		13 (11.93)	17 (15.45)	
Stage/n(%)			0.000			0.000
I + II	89 (52.66)	165 (97.06)		65 (59.63)	105 (95.45)	
III + IV	80 (47.34)	5 (2.94)		44 (40.37)	5 (4.55)	
AFP/n(%)			0.031			0.045
≤20 ng·mL ⁻¹ (TCGA) or ≤300 ng·mL ⁻¹ (GEO)	118 (69.82)	111 (65.29)		57 (52.29)	61 (55.45)	
>20 ng·mL ⁻¹ (TCGA) or >300 ng·mL ⁻¹ (GEO)	51 (30.18)	59 (34.71)		52 (47.71)	49 (44.55)	
Event/n(%)						
Survival	92 (54.44)	131 (77.06)		57 (52.29)	78 (70.91)	
Death	77 (45.56)	39 (22.94)		52 (47.71)	32 (29.09)	

2.3 基于TCGA队列的HCC预后预测模型列线图的建立

在基于TCGA队列进行基因筛选以及建立预后预测模型之后，进一步验证该模型在预测HCC患者预后的可行性。ROC曲线分析结果表明，在预测HCC预后方面，预后评分优于仅风险评分或仅TNM分期（图3A~C）。

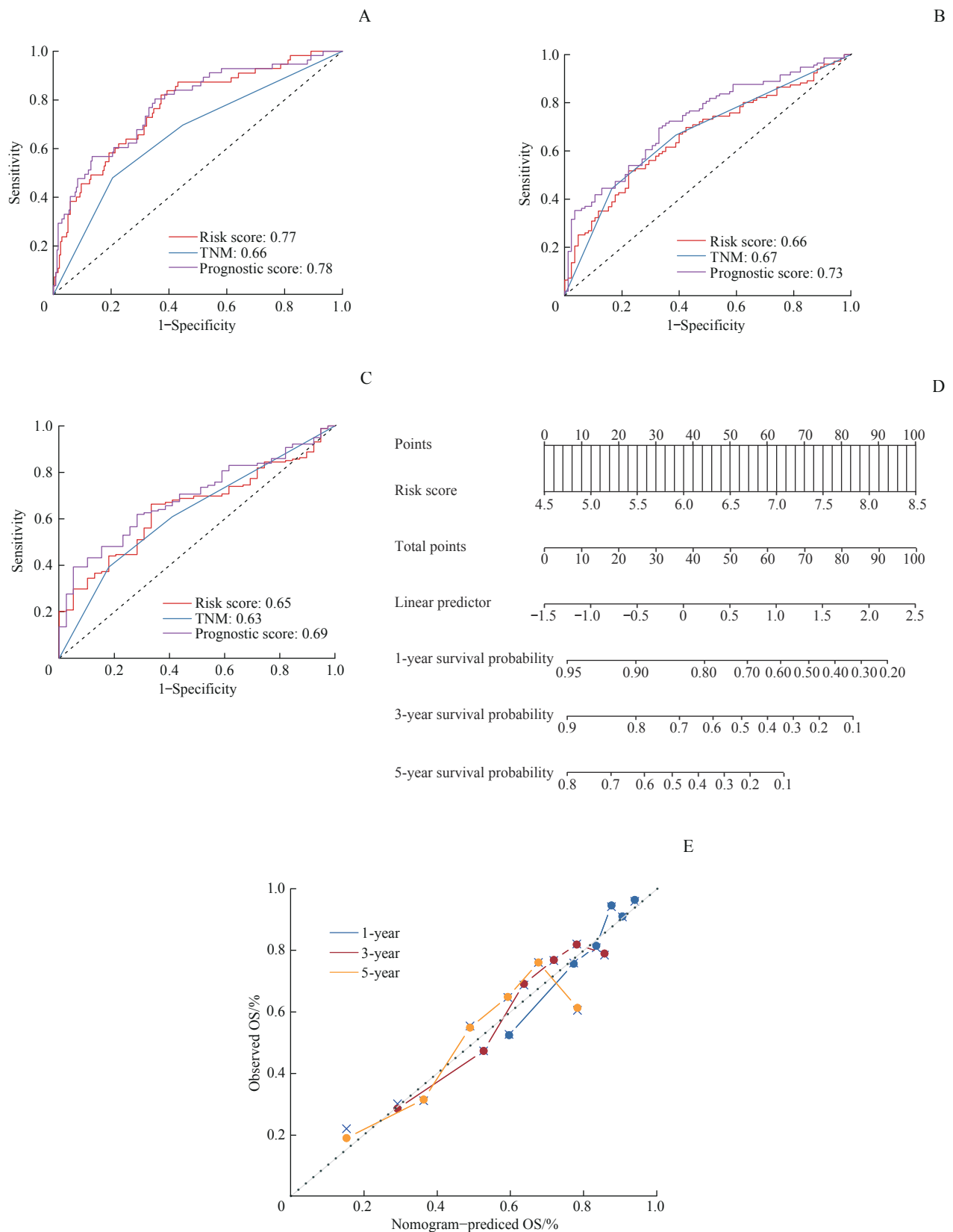
为了验证该预后模型的预测效能，根据预后评分构建了预后列线图（C指数为0.689），见图3D。校准图显示，预测列线图预测的预后与实际OS非常接近（图3E）。这也进一步表明基于NEMG和TNM分期建立的HCC预后模型具有较高的可信度。

2.4 建立的HCC综合预后模型在GEO队列和临床样本中的验证

为了验证预测模型的稳定性，在GEO数据集GSE14520（表2）中对其进行进一步验证。根据预后评分的中位数，GSE14520中的HCC患者被分为高风险组（n=109）和低风险组（n=110）。时间ROC曲线分析证实了该预后模型对GSE14520中HCC OS

良好的预测性能，在1、3和5年OS的AUC值分别为0.67、0.66和0.74（图4A）。Kaplan-Meier曲线还显示，与低风险组患者相比，高风险组患者的预后明显较差（图4B，P=0.001）。

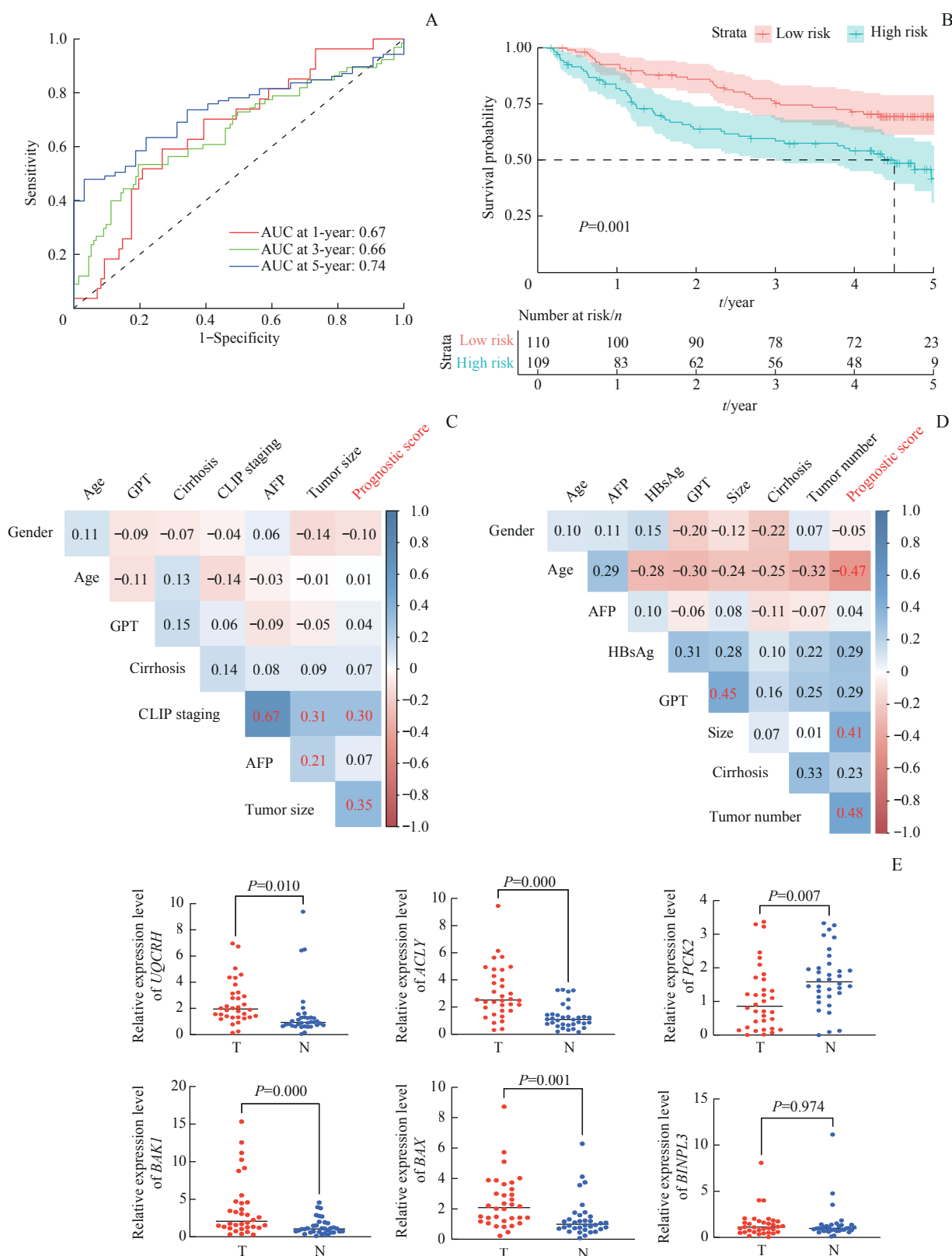
相关性分析显示，预后评分与一些重要的临床病理特征如主要肿瘤大小和CLIP分期呈正相关（P=0.000，P=0.000，图4C）。此外，在34对HCC癌和癌旁组织的队列（表3）中，患者临床信息与预后评分之间的相关性分析也表明，预后评分与肿瘤大小和肿瘤数量呈正相关（图4D），与GEO数据集中的分析结果相一致。随后，在上述的34对HCC肿瘤样本中验证了预后模型中6种NEMGs（UQCRH、ACLY、PCK2、BAK1、BAX和BNIP3L）的相对表达水平。与癌旁组织相比，HCC组织中的UQCRH（P=0.010）、ACLY（P=0.000）、BAX（P=0.001）和BAK1（P=0.000）的表达水平显著上调（图4E），而PCK2的表达水平显著下调（P=0.007，图4E），这与TCGA-LIHC队列中的各基因的表达趋势相一致。但BNIP3L的表达水平在癌和癌旁组织之间差异没有统计学意义（P=0.974，图4E）。



Note: A–C. ROC analysis of the prognostic score, risk score and TNM stage for 1-year OS (A), 3-year OS (B) and 5-year OS (C). D. Nomogram constructed by the prognostic score to predict OS of HCC. E. Calibration curves of the prognostic score-derived nomogram for 1-year, 3-year and 5-year OS of HCC.

图3 基于TCGA队列的HCC预后预测模型列线图建立

Fig 3 Establishment of the prognostic score-derived nomogram based on TCGA cohort



Note: A. ROC analysis of the prognostic model for OS of HCC in GSE14520. B. Kaplan-Meier analysis of OS in HCC patients of GSE14520 which were stratified into high- and low-risk groups by this prognostic model. C. Correlation analysis between clinicopathological features and prognostic score in HCC patients of GSE14520. The correlation coefficient in red indicates $P<0.05$. D. Correlation analysis between clinicopathological features and prognostic score in 34-pair HCC and matched normal tissues. The correlation coefficient in red indicates $P<0.05$. E. Gene expression levels of the 6 NEMGs involved in this prognostic model in 34-pair HCC tumor tissues (T) and matched normal tissues (N).

图4 GEO和临床HCC队列中预后模型的验证

Fig 4 Validation of the prognostic model in GEO and clinical HCC cohort

表3 临床队列中HCC患者的临床病理特征

Tab 3 Clinicopathological features of HCC patients in the clinical cohort

Feature	Total/n(%)
Age	
≤50 year	8 (23.53)
>50 year	26 (76.47)
Gender	
Male	24 (70.59)
Female	10 (29.41)
TNM stage	
I	15 (44.12)
II	5 (14.71)
III	11 (32.35)
IV	3 (8.82)
AFP	
≤20 ng·mL ⁻¹	21 (61.76)
>20 ng·mL ⁻¹	13 (38.24)
Size	
≤5 cm	20 (58.82)
>5 cm	14 (41.18)
Tumor number	
=1	22 (64.71)
>1	12 (35.29)
HBsAg	
-	7 (20.59)
+	27 (79.41)
GPT	
≤40 U·L ⁻¹	25 (73.53)
>40 U·L ⁻¹	9 (26.47)
Cirrhosis	
-	16 (40.06)
+	18 (59.94)

Note: HBsAg—hepatitis B surface antigen; GPT—glutamic pyruvic transaminase.

3 讨论

由于HCC患者的异质性结果以及管理和治疗策略选择的困难性，迫切需要一个有效的HCC患者预后预测指标。临床前研究的证据支持线粒体功能障碍是代谢性肝病和癌症发病机制的关键因素，这进一步表明开发线粒体基因靶向治疗是抑制HCC进展的一种有吸引力的策略^[24]。研究^[25]证明，异常的线粒体能量状态有助于癌症的发生和发展。线粒体TCA循环酶缺陷可产生肿瘤代谢产物，促进肿瘤发生。线粒体功能障碍是由mtDNA突变、TCA循环酶缺陷、电子呼吸链泄漏和随后的氧化应激引起的；致癌和抑

癌信号可以改变细胞代谢途径，破坏氧化还原的平衡，从而导致人类癌症的发生^[26]。NEMG的突变或异常表达可能导致与肿瘤发生相关的异常能量产生。因此建立基于NEMG的HCC患者预后模型对HCC患者的进展及预后至关重要。

本研究着重研究OXPHOS、TCA循环和细胞凋亡这3个通路中的预后相关基因，对95个与HCC预后相关的deNEMG和参与3种有丝分裂通路的NEMG取交集，发现仅6个参与3种有丝分裂通路的deNEMG (*UQCRH*、*ACLY*、*PCK2*、*BAK1*、*BAX*和*BNIP3L*)与HCC的预后密切相关。因此本研究利用这6个基因来构建基于NEMG的风险模型。在这6个NEMG中，*UQCRH*编码一个定位于线粒体和细胞核的铰链蛋白。*UQCRH*作为泛素-细胞色素C还原酶复合物的主要亚基之一，负责复合物Ⅲ的细胞色素C和C1之间的电子转移。事实上，在肺腺癌^[27]和HCC^[28]中已经有关于*UQCRH*表达上调的报道。此外，*UQCRH*的表达与肿瘤大小、分化程度和有无血管侵袭相关，并与HCC患者的不良预后密切相关^[28]。*ACLY*和*PCK2*参与TCA循环，其中*ACLY*形成一种细胞质同源四聚体酶，催化TCA循环中的中间体柠檬酸盐转化为草酰乙酸和乙酰辅酶A^[29]。由于乙酰辅酶A是新脂肪生成和胆固醇生物合成的重要底物之一，因此*ACLY*是一种重要的脂肪生成酶，在结直肠癌、非小细胞肺癌、乳腺癌、膀胱癌和HCC等癌症类型中高表达^[30-32]。同时，在HCC患者中，*ACLY*的高表达与不良预后、肿瘤干性和转移相关^[33]。*PCK2*编码一种线粒体磷酸烯醇式丙酮酸羧激酶，该酶催化TCA循环衍生的草酰乙酸转化为磷酸烯醇式丙酮酸，并伴随GTP水解。*PCK2*在某些癌症类型，如肺癌、前列腺癌、乳腺癌、宫颈癌和睾丸癌中经常上调和激活^[34-35]，可能以此对抗肿瘤微环境中的代谢饥饿状态^[36]。此外，既往研究^[37]表明，*PCK2*的下调与HCC的预后和患癌时间有关，并且*PCK2*的异位过表达可能通过诱导HCC细胞的葡萄糖缺乏而降低细胞活力。*BAX*、*BAK1*和*BNIP3L*是凋亡相关基因，其中*BAX*和*BAK1*是线粒体外膜通透性的决定性关键因素^[38-40]。*BNIP3L*是线粒体吞噬的双重调节因子，而线粒体吞噬是清除受损线粒体的过程^[41]。据报道*BNIP3L*与*BAX*和*BAK*相互作用，导致线粒体外膜通透性增强和凋亡加快^[42]。同时，*BNIP3L*介导的线

粒体吞噬通过糖酵解代谢重编程,促进乙型肝炎病毒X蛋白(hepatitis B virus X protein, HBx)诱导的HCC的肿瘤干性^[42]。

TCGA-LIHC队列的单因素和多因素Cox分析结果表明,6-NEMG风险特征和TNM分期是HCC的独立危险因素。因此,将6-NEMG风险特征与TNM分期相结合,建立了一个新的HCC综合预后模型。该模型在TCGA-LIHC和GSE14520队列中预测HCC预后方面均表现出良好的性能。利用预后评分的中位数作为临界值,将TCGA-LIHC和GSE14520队列患者分为高风险和低风险组,进行了预后差异分析,结果表明在2个队列中高风险组的OS均明显低于低风险组患者。此外,在34对HCC临床样本(癌和癌旁组织)中验证这6个NEMG的表达情况。在HCC肿瘤样本中*UQCRH*、*ACLY*、*BAK1*、*BAX*相对高表达,而*PCK2*相对低表达,这符合在TCGA-LIHC数据库中的分析结果;但*BNIP3L*在癌组织和癌旁组织的表达差异没有统计学意义,可能跟临床样本量有限有关。这些验证进一步表明6个NEMG在HCC的预后预测中起着重要的作用。

本研究用TCGA数据建模,用GEO数据验证的原因在于:TCGA数据库收录的基因表达数据多为通过测序手段获得的,而GEO数据库收录的大部分数据是用芯片技术获得的;前者测序技术挖掘到的基因比后者多,并且临床信息比较全,用能测到更多基因的数据库(TCGA)去训练,可避免模型的重要基因被遗漏。此外,本研究存在以下局限性。首先,TCGA和GEO数据库中有些病例的临床特征不完整,因此仅选择了与肿瘤预后相关并且在2个数据库中均包含的临床特征进行分析。其次,临床HCC队列的样本量不足,临床信息不完整,未能充分验

证该预后模型在临床HCC样本中的预测性能。因此,在将该HCC预后模型扩展到临床应用之前,还需要进行进一步的临床试验。此外,模型中包含的基因的功能也需要通过细胞实验和动物实验模型来进一步验证。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

伦理批准和知情同意/Ethics Approval and Patient Consent

克德尔亚·艾山江和邬海龙设计了该研究;克德尔亚·艾山江、傅怡和赖冬林进行了生物信息学和统计分析;克德尔亚·艾山江完成了验证实验和论文写作;克德尔亚·艾山江、傅怡、赖冬林、龚伟和邬海龙进行了广泛的文献检索和讨论;邬海龙和龚伟修改了稿件。所有作者阅读并同意了最终稿件的提交。

Kedeerya Aishanjiang and WU Hailong designed the project. Kedeerya Aishanjiang, FU Yi and LAI Donglin performed bioinformatics and statistical analysis. Kedeerya Aishanjiang performed the validation experiments and wrote the manuscript. Kedeerya Aishanjiang, FU Yi, LAI Donglin, GONG Wei and WU Hailong performed extensive literature search and discussion. WU Hailong and GONG Wei revised the manuscript. All the authors have read the last version of paper and agreed for the submission.

作者贡献/Authors' Contributions

本研究涉及的所有实验均已通过上海交通大学医学院附属新华医院医学伦理委员会审核批准(文件号XHEC-D-2023-021)。受试对象或其亲属已经签署知情同意书。

All experimental protocols in this study were reviewed and approved by Ethics Committee of Xinhua Hospital, Shanghai Jiao Tong University School of Medicine (Approval Letter No. XHEC-D-2023-021). Consent letters have been signed by the research participants or their relatives.

• Received: 2023-09-18

• Accepted: 2023-12-06

• Published online: 2024-01-28

参·考·文·献

- [1] MILLER K D, FIDLER-BENAOUDIA M, KEEGAN T H, et al. Cancer statistics for adolescents and young adults, 2020[J]. CA Cancer J Clin, 2020, 70(6): 443-459.
- [2] SIEGEL R L, MILLER K D, FUCHS H E, et al. Cancer statistics, 2021[J]. CA A Cancer J Clinicians, 2021, 71(1): 7-33.
- [3] SANGRO B, SAROBE P, HERVÁS-STUBBS S, et al. Advances in immunotherapy for hepatocellular carcinoma[J]. Nat Rev Gastroenterol Hepatol, 2021, 18(8): 525-543.
- [4] RIMASSA L, PERSONENI N, CZAUDERNA C, et al. Systemic treatment of HCC in special populations[J]. J Hepatol, 2021, 74(4): 931-943.
- [5] BERTUCCIO P, TURATI F, CARIOLI G, et al. Global trends and predictions in hepatocellular carcinoma mortality[J]. J Hepatol, 2017, 67(2): 302-309.
- [6] YANG J D, HAINAUT P, GORES G J, et al. A global view of hepatocellular carcinoma: trends, risk, prevention and management[J]. Nat Rev Gastroenterol Hepatol, 2019, 16(10): 589-604.
- [7] ZHU J Y, TANG B F, LI J, et al. Identification and validation of the angiogenic genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma[J]. Aging, 2020, 12(9): 7848-7873.
- [8] DING Z B, ERICKSEN R E, LEE Q Y, et al. Reprogramming of mitochondrial proline metabolism promotes liver tumorigenesis[J]. Amino Acids, 2021, 53(12): 1807-1815.

- [9] MARQUEZ J, FLORES J, KIM A H, et al. Rescue of TCA cycle dysfunction for cancer therapy[J]. *J Clin Med*, 2019, 8(12): 2161.
- [10] SONG B S, MOON J S, TIAN J W, et al. Mitochondrial defects aggravate liver cancer *via* aberrant glycolytic flux and T cell exhaustion[J]. *J Immunother Cancer*, 2022, 10(5): e004337.
- [11] LI M, WANG L, WANG Y J, et al. Mitochondrial fusion *via* OPA1 and MFN1 supports liver tumor cell metabolism and growth[J]. *Cells*, 2020, 9(1): 121.
- [12] LI X J, LI Y M, XU A J, et al. Apoptosis-induced translocation of centromere protein F in its corresponding autoantibody production in hepatocellular carcinoma[J]. *Oncoimmunology*, 2021, 10(1): 1992104.
- [13] CAI Y L, LIN Y X, XIONG X Z, et al. Knockdown expression of MECP, a novel gene of mitochondrial FAS II inhibits growth and colony-formation, promotes apoptosis of hepatocellular carcinoma cells[J]. *Biosci Trends*, 2019, 13(3): 234-244.
- [14] WALLACE D C. Mitochondria and cancer[J]. *Nat Rev Cancer*, 2012, 12(10): 685-698.
- [15] SANCHEZ P, BARNEDA D, HEESCHEN C. Hallmarks of cancer stem cell metabolism[J]. *Br J Cancer*, 2016, 114(12): 1305-1312.
- [16] KUNTZ E M, BAQUERO P, MICHIE A M, et al. Targeting mitochondrial oxidative phosphorylation eradicates therapy-resistant chronic myeloid leukemia stem cells[J]. *Nat Med*, 2017, 23(10): 1234-1240.
- [17] VAZQUEZ F, LIM J H, CHIM H, et al. PGC1 α expression defines a subset of human melanoma tumors with increased mitochondrial capacity and resistance to oxidative stress[J]. *Cancer Cell*, 2013, 23(3): 287-301.
- [18] FARGE T, SALAND E, DE TONI F, et al. Chemotherapy-resistant human acute myeloid leukemia cells are not enriched for leukemic stem cells but require oxidative metabolism[J]. *Cancer Discov*, 2017, 7(7): 716-735.
- [19] ABBAS S, LUGTHART S, KAVELAARS F G, et al. Acquired mutations in the genes encoding IDH1 and IDH2 both are recurrent aberrations in acute myeloid leukemia: prevalence and prognostic value[J]. *Blood*, 2010, 116(12): 2122-2126.
- [20] PASINI B, MCWHINNEY S R, BEI T, et al. Clinical and molecular genetics of patients with the Carney-Stratakis syndrome and germline mutations of the genes coding for the succinate dehydrogenase subunits SDHB, SDHC, and SDHD[J]. *Eur J Hum Genet*, 2008, 16(1): 79-88.
- [21] LEHTONEN H J, BLANCO I, PIULATS J M, et al. Conventional renal cancer in a patient with fumarate hydratase mutation[J]. *Hum Pathol*, 2007, 38(5): 793-796.
- [22] FU J, LIU G X, ZHANG X, et al. TRPM8 promotes hepatocellular carcinoma progression by inducing SNORA55 mediated nuclear-mitochondrial communication[J]. *Cancer Gene Ther*, 2023, 30(5): 738-751.
- [23] HUANG Q C, WU D, ZHAO J, et al. TFAM loss induces nuclear actin assembly upon mDia2 malonylation to promote liver cancer metastasis[J]. *EMBO J*, 2022, 41(11): e110324.
- [24] CHELLA KRISHNAN K, KURT Z, BARRERE-CAIN R, et al. Integration of multi-omics data from mouse diversity panel highlights mitochondrial dysfunction in non-alcoholic fatty liver disease[J]. *Cell Syst*, 2018, 6(1): 103-115. e7.
- [25] WALLACE D C. Mitochondria and cancer[J]. *Nat Rev Cancer*, 2012, 12(10): 685-698.
- [26] LUO Y D, MA J J, LU W Q. The significance of mitochondrial dysfunction in cancer[J]. *Int J Mol Sci*, 2020, 21(16): 5598.
- [27] GAO F, LIU Q C, LI G P, et al. Identification of ubiquinol cytochrome c reductase hinge (UQCRH) as a potential diagnostic biomarker for lung adenocarcinoma[J]. *Open Biol*, 2016, 6(6): 150256.
- [28] PARK E R, KIM S B, LEE J S, et al. The mitochondrial hinge protein, UQCRH, is a novel prognostic factor for hepatocellular carcinoma[J]. *Cancer Med*, 2017, 6(4): 749-760.
- [29] SUN H, WANG F C, HUANG Y Q, et al. Targeted inhibition of ACLY expression to reverse the resistance of sorafenib in hepatocellular carcinoma[J]. *J Cancer*, 2022, 13(3): 951-964.
- [30] GRANCHI C. ATP citrate lyase (ACLY) inhibitors: an anti-cancer strategy at the crossroads of glucose and lipid metabolism[J]. *Eur J Med Chem*, 2018, 157: 1276-1291.
- [31] WEI J, LEIT S, KUAI J, et al. An allosteric mechanism for potent inhibition of human ATP-citrate lyase[J]. *Nature*, 2019, 568(7753): 566-570.
- [32] ICARD P, WU Z R, FOURNEL L, et al. ATP citrate lyase: a central metabolic enzyme in cancer[J]. *Cancer Lett*, 2020, 471: 125-134.
- [33] HAN Q, CHEN C A, YANG W, et al. ATP-citrate lyase regulates stemness and metastasis in hepatocellular carcinoma *via* the Wnt/ β -catenin signaling pathway[J]. *Hepatobiliary Pancreat Dis Int*, 2021, 20(3): 251-261.
- [34] LEITHNER K, HRZENJAK A, TRÖTZMÜLLER M, et al. PCK2 activation mediates an adaptive response to glucose depletion in lung cancer[J]. *Oncogene*, 2015, 34(8): 1044-1050.
- [35] HYROŠŠOVÁ P, ARAGÓ M, MORENO-FELICI J, et al. PEPCK-M recoups tumor cell anabolic potential in a PKC- ζ -dependent manner[J]. *Cancer Metab*, 2021, 9(1): 1.
- [36] BLUEMEL G, PLANQUE M, MADREITER-SOKOLOWSKI C T, et al. PCK2 opposes mitochondrial respiration and maintains the redox balance in starved lung cancer cells[J]. *Free Radic Biol Med*, 2021, 176: 34-45.
- [37] LIU M X, JIN L, SUN S J, et al. Metabolic reprogramming by PCK1 promotes TCA cataplerosis, oxidative stress and apoptosis in liver cancer cells and suppresses hepatocellular carcinoma[J]. *Oncogene*, 2018, 37(12): 1637-1653.
- [38] COSENTINO K, HERTLEIN V, JENNER A, et al. The interplay between BAX and BAK tunes apoptotic pore growth to control mitochondrial-DNA-mediated inflammation[J]. *Mol Cell*, 2022, 82(5): 933-949. e9.
- [39] LUO X, O'NEILL K L, HUANG K. The third model of Bax/Bak activation: a Bcl-2 family feud finally resolved?[J]. *F1000Research*, 2020, 9: 935.
- [40] HAO B B, LI X J, JIA X L, et al. The novel cereblon modulator CC-885 inhibits mitophagy *via* selective degradation of BNIP3L[J]. *Acta Pharmacol Sin*, 2020, 41(9): 1246-1254.
- [41] LI Y, ZHENG W Q, LU Y Y, et al. BNIP3L/NIX-mediated mitophagy: molecular mechanisms and implications for human disease[J]. *Cell Death Dis*, 2021, 13(1): 14.
- [42] CHEN Y Y, WANG W H, CHE L, et al. BNIP3L-dependent mitophagy promotes HBx-induced cancer stemness of hepatocellular carcinoma cells *via* glycolysis metabolism reprogramming[J]. *Cancers*, 2020, 12(3): 655.

[本文编辑] 崔黎明

