

· 临床研究 ·

## m5C 甲基化相关基因预测肾透明细胞癌的预后

陈晓超,卜恒涛,梁超,李杰\*

南京医科大学第一附属医院泌尿外科,江苏 南京 210029

**[摘要]** 目的:利用肿瘤基因组图谱(the Cancer Genome Atlas,TCGA)数据库中的数据,构建 m5C 甲基化相关基因组成的生存预后模型,分析与肾透明细胞癌生存有关的独立预后影响因素以预测肾透明细胞癌患者的预后。方法:从 TCGA 数据库下载肾透明细胞癌患者资料,从已发表的文献中获取 m5C 甲基化相关基因并分析其在肿瘤组和对照组间的表达差异。对这些基因进行共识聚类分析以揭示 m5C 甲基化相关基因与肾透明细胞癌预后之间的关系。通过单变量 Cox 分析和 Lasso-Cox 回归分析构建出生存预后模型并分析与肾透明细胞癌相关的预后影响因素,GO 和 KEGG 富集分析进一步探索生物功能和潜在的信号通路。通过 qRT-PCR 验证 m5C 甲基化相关基因在肾透明细胞癌细胞系及组织和正常肾细胞系及组织之间的差异表达。结果:m5C 甲基化相关基因在肾透明细胞癌肿瘤组和正常组之间存在差异性表达。共识聚类分析结果显示亚群 1 的肾透明细胞癌患者预后优于亚群 2。构建的生存预后模型将肾透明细胞癌分为高风险组和低风险组,单变量和多变量 Cox 分析显示分级和分期可能是肾透明细胞癌的独立预后因素,而 GO 和 KEGG 分析显示 m5C-RNA 甲基化修饰可影响肾透明细胞癌的进展。此外,qRT-PCR 实验也证实 m5C 甲基化相关基因在肾透明细胞癌肿瘤组和正常组之间表达存在差异。结论:本文构建的 m5C 甲基化相关基因的生存预后模型可以预测肾透明细胞癌患者的预后。

**[关键词]** 肾透明细胞癌;m5C-RNA 甲基化;生存预后模型;TCGA

**[中图分类号]** R737.11

**[文献标志码]** A

**[文章编号]** 1007-4368(2022)07-983-11

**doi:**10.7655/NYDXBNS20220711

## m5C methylation-related genes predict prognosis in renal clear cell carcinoma

CHEN Xiaochao, BU Hengtao, LIANG Chao, LI Jie\*

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China

**[Abstract]** **Objective:** This study aims to use the data from the Cancer Genome Atlas (TCGA) database to construct a survival prognostic model composed of m5C methylation-related genes, and to analyze the independent prognostic factors related to the survival of patients with renal clear cell carcinoma (RCCC) to predict the prognosis of RCCC. **Methods:** The data of patients with RCCC were downloaded from the TCGA database, m5C methylation-related genes were obtained from the published literature, and their differential expression between the tumor group and the control group was analyzed. Consensus clustering analysis was then performed to reveal the relationship between m5C methylation-related genes and prognosis of RCCC. A survival prognostic model was constructed by univariate Cox analysis and Lasso-Cox regression analysis to analyze the prognostic factors associated with RCCC. Finally, GO and KEGG enrichment analysis was performed to further explore biological functions and potential signaling pathways. In addition, we also performed qRT-PCR experiments to measure the differential expression of m5C methylation-related genes between cancer cells and normal cell lines or between RCCC and adjacent normal tissue. **Results:** We found that m5C methylation-related genes were differentially expressed between the tumor group and the normal group in RCCC. The results of consensus clustering analysis showed that the prognosis of patients with RCCC in cluster 1 was better than that in cluster 2. The constructed survival prognostic model divided RCCC into high-risk group and low-risk group. Univariate and multivariate Cox analysis showed that grade and stage may be independent prognostic factors for RCCC. GO and KEGG analysis showed that m5C-RNA modification helps to regulate the progression of RCCC. In addition, qRT-PCR experiments also confirmed that m5C methylation-related genes are differentially expressed between the tumor group and the normal group in RCCC. **Conclusion:** The survival prognostic model of m5C methylation-

**[基金项目]** 国家自然科学基金(82002718,81672532)

\*通信作者(Corresponding author), E-mail: drc\_lijie@hotmail.com

related genes constructed by this research can predict the prognosis of patients with RCCC.

[Key words] renal clear cell carcinoma; m5C-RNA methylation; survival prognostic model; TCGA

[J Nanjing Med Univ, 2022, 42(07):983-993]

肾细胞癌是世界上第三常见的泌尿系恶性肿瘤,主要包括肾透明细胞癌、乳头状肾细胞癌、嫌色细胞癌、集合管癌和未分类肾细胞癌。其中肾透明细胞癌约占80%~90%<sup>[1]</sup>。尽管目前开发了许多靶向药物和免疫抑制药物,但肾根治性切除术仍然是金标准治疗方法<sup>[2]</sup>。此外,肾透明细胞癌的特点是复发率和转移率高<sup>[3]</sup>,这对患者的健康和生活质量造成严重影响。因此,寻找肾透明细胞癌中的相关基因以构建预测模型至关重要,同时这些基因也可以作为未来分子靶向治疗的潜在靶点。

最新研究表明, RNA 修饰是一种新兴的基因调控机制,其主要包括 m6A、m5C 和 m1A, m6A 和 m5C 修饰技术是 160 多种化学修饰中最主要和最具代表性的两种 RNA 转录后修饰类型<sup>[4-5]</sup>。其中, m5C-RNA 甲基化修饰主要依赖于甲基转移酶、去甲基化酶和结合蛋白, m5C-RNA 甲基化修饰发挥多种生物学功能。如调节 mRNA 的转运,增加 RNA 的稳定性,调节蛋白质翻译,维持 RNA 正常结构等,近年来许多研究表明, m5C-RNA 甲基化修饰参与调节多种肿瘤的发生发展、侵袭和转移。如在肺癌中, m5C 甲基化相关基因 NSUN3、NSUN4 在肿瘤组织中的表达高于正常肺组织,且与临床病理特征及生存率有关,此外,还与 6 种主要免疫细胞的浸润有关,从而调节肺癌的肿瘤免疫微环境<sup>[6]</sup>。在膀胱尿路上皮癌中, m5C 甲基化转移酶 NSUN2 表达较正常组织高,从而通过增强癌基因肝癌衍生生长因子(HDGF)mRNA 的稳定性来促进疾病的发生发展。而对头颈鳞状细胞癌的研究显示, m5C RNA 甲基化修饰不仅有助于促进疾病的进展, NSUN5、DNMT1 和 DNMT3A 的表达模式还有助于预测疾病的预后。因此,寻找 m5C 甲基化相关基因对肿瘤患者的早期诊断和预后评估具有重要意义。

本研究基于肿瘤基因组图谱(the Cancer Genome Atlas, TCGA)数据库下载的肾透明细胞癌患者的数据,通过生物信息学和统计学分析,构建了由 m5C 甲基化相关基因组成的生存预后模型,分析了与肾透明细胞癌生存有关的独立预后影响因素,同时用此模型将肾透明细胞癌患者分为高低风险组,通过

GO 和 KEGG 分析对高低风险组中的差异基因进行了分析,进一步揭示这些差异基因的生物学功能和潜在的信号通路。

## 1 资料和方法

### 1.1 资料

从 TCGA 数据库中下载了 572 个肾透明细胞癌样本和 72 个正常样本的转录组数据用于进一步分析,临床数据也从 TCGA 数据库下载(<https://www.cbioportal.org/>)<sup>[9]</sup>。在 cBioportal 数据库中研究了选定的 4 种基因与拷贝数变异之间的关系(<https://www.cbioportal.org/>)<sup>[10]</sup>。在 GEPIA2 数据库找到这 4 种基因的不同器官表达特征<sup>[11]</sup>。肾透明细胞癌和癌旁正常组织样本来自 2008 年 4 月—2019 年 4 月南京医科大学第一附属医院接受根治性肾切除术的患者,本研究经过了南京医科大学第一附属医院伦理委员会授权(伦理批准号:2021-SR-430)。所有患者都签署了协议,允许他们的组织样本和相应临床信息用于进一步的研究。

### 1.2 方法

#### 1.2.1 m5C-RNA 甲基化相关基因分析

从已发表的文献中找到 13 个 m5C 甲基化相关基因,其中包括 11 个甲基转移酶(NOP2、NSUN2、NSUN3、NSUN4、NSUN5、NSUN6、NSUN7、DNMT1、DNMT3A、DNMT3B、TRDMT1)、1 个去甲基化酶(TET2)和 1 个结合蛋白(ALYREF)。然后系统地提取并分析了 13 种调节基因的表达谱,以及临床病理参数。

#### 1.2.2 m5C 甲基化相关基因差异表达的数据处理

提取 572 例肿瘤组织和 72 例正常肾组织中 m5C 甲基化相关基因的表达矩阵和临床资料。使用 limma 软件包的 R 版本(4.1.0)确定肿瘤和对照组之间差异表达的 m5C 甲基化相关基因。以  $P < 0.05$  和  $|\log_2FC| > 1$  作为所有检验的显著性阈值。此外,用热图和小提琴图展示肿瘤组和对照组表达的差异。

#### 1.2.3 m5C 甲基化相关基因的共识聚类分析

用 R(4.1.0)的“ConsensusClusterPlus”软件包对

筛选出来的13个m5C甲基化相关基因进行共识聚类分析。此外,还对不同的聚类进行生存分析,以确定肾透明细胞癌样本中的最佳聚类。然后进行了Kaplan-Meier生存分析,以揭示不同集群之间总体生存率的差异。比较临床病理参数,包括性别、分级、年龄、TNM分期和不同聚类之间的关系。

#### 1.2.4 蛋白质相互作用及功能分析

为了进一步筛选出中心基因,使用STRING数据库检索相互作用基因以分析差异表达的m5C甲基化相关基因,构建蛋白质-蛋白质相互作用(protein-protein interaction, PPI)网络(<http://string-db.org/>)<sup>[12]</sup>。

#### 1.2.5 构建预测模型

首先,计算危险比(hazard ratio, HR,危险组个体和非危险组个体的比值)和95%置信区间(confidence interval, CI),通过单变量Cox回归分析识别适当的候选基因。其次,使用Lasso-Cox回归分析,挑选出合适的能够预测肾透明细胞癌的预后基因,风险值计算公式为:风险值= $\sum_{i=1}^n \text{coef} \times x_i$ (其中, $n$ 代表mRNA的数目,coef是系数, $x_i$ 是基因的表达量)。接下来,根据风险值中位数,将TCGA的肾透明细胞癌队列分为高风险组和低风险组,再用Kaplan-Meier生存分析/风险预测模型估计候选风险基因的

预后价值。

#### 1.2.6 GO和KEGG富集分析

使用R软件包“clusterProfiler”、“enrichplot”和“ggplot2”对这13个选定基因进行GO和KEGG富集分析,以确定分子功能(molecular function, MF)、细胞成分(cellular components, CC)和生物过程(biological process, BP)并寻找潜在的信号通路。

#### 1.2.7 细胞培养

肾癌细胞系(786-O、A498)和人肾小管上皮细胞系(HK-2)均购自中国科学院(上海)细胞库。其中,786-O和HK-2细胞在RPMI 1640中培养,A498细胞在DMEM/F12中培养。所有细胞系均在37℃、含5%CO<sub>2</sub>的培养箱中培养。

#### 1.2.8 总RNA的提取和qRT-PCR

使用TRIzol试剂(美国赛默飞公司)从培养的细胞和组织样本中提取总RNA,并将其反转录为cDNA。然后使用StepOne Plus实时PCR系统和SYBR Premix Ex Taq试剂进行qRT-PCR实验,目的基因引物见表1。

#### 1.3 统计学方法

统计分析使用R软件(4.1.0)进行。此外,Perl编程语言(版本5.30.2)用于数据处理。Kaplan-Meier生存曲线分析用于分析总体生存率(overall survival, OS)。单变量和多变量Cox回归分析用于评估预后意义。 $P < 0.05$ 为差异有统计学意义。

表1 实时定量PCR的引物序列

Table 1 The sequences of primers for RT-qPCR

目的基因	正向引物(5'→3')	反向引物(5'→3')
GAPDH	GAAGGTGAAGGTCGGAGTC	GAAGATGGTGATGGGATTTTC
NSUN5	AGTGGCGGAGACTGTAAGG	GGCAATGGAATGGCTAAA
NSUN6	AATAATGGCGAGGGTCTT	GTTGGGCTAATGCTCTGT
DNMT3B	CTTCTTGTTCACCCACCTT	CCCACTGCCTCTTTACCC
TET2	TCTGAGGGTAAGGTGGGC	GAGGTGGGAGGTTGGGAT

## 2 结果

### 2.1 m5C甲基化相关基因在肾透明细胞癌中的表达情况

为了明确肾透明细胞癌中m5C甲基化相关基因的基本生物学功能的表达,首先从TCGA数据库中下载并提取了有效的基因表达数据,如热图所示(图1A)。在572例肿瘤患者中,13个m5C甲基化相关基因中有12个在肿瘤和正常组织中有差异性表

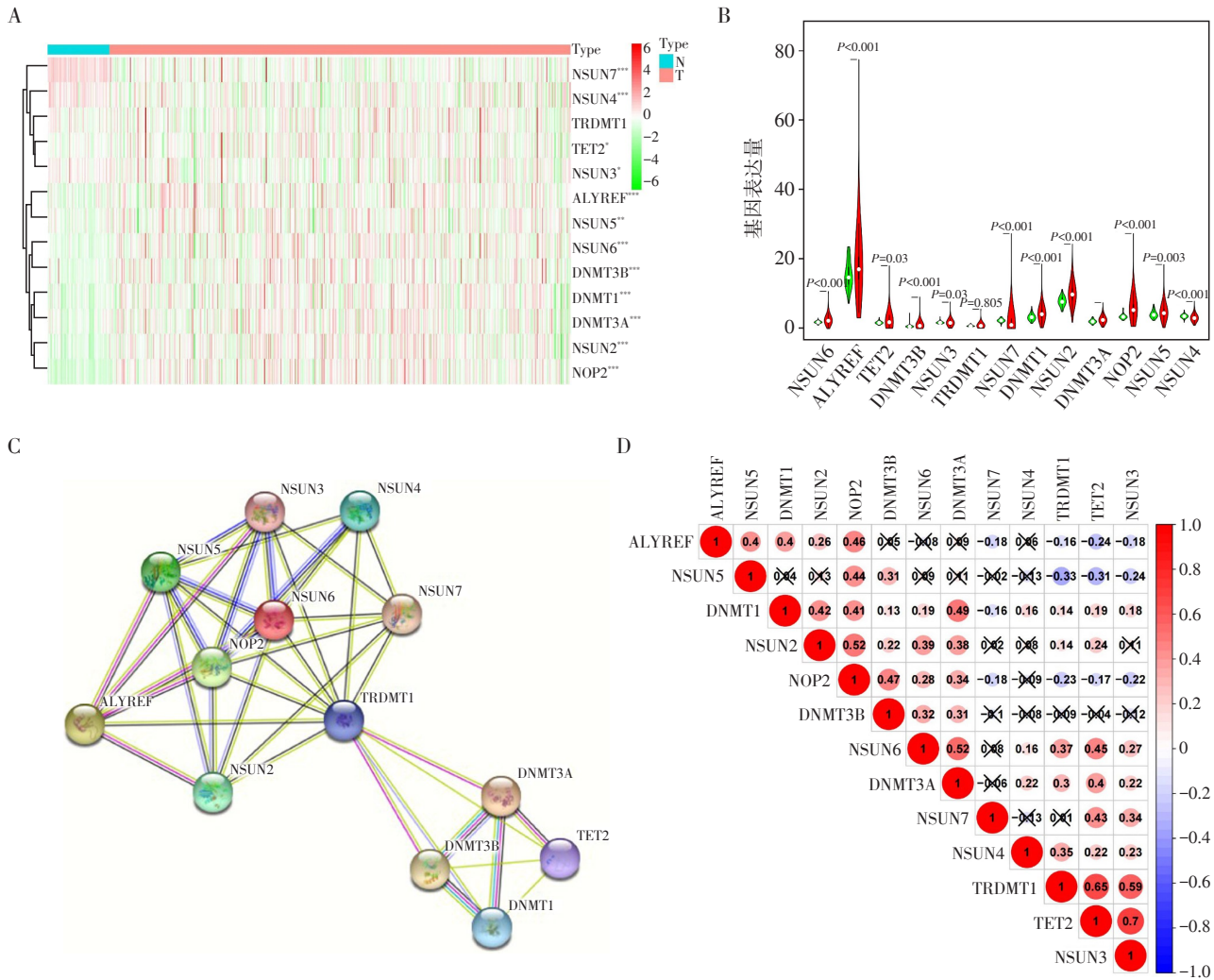
达( $P < 0.05$ )。其中包括NOP2、NSUN2、NSUN3、NSUN4、NSUN5、NSUN6、NSUN7、DNMT1、DNMT3A、TET2、DNMT3B以及ALYREF。然而,与正常组织相比,肾透明细胞癌组织中TRDMT1的表达没有明显差异(图1B)。在这12个基因中,除了NSUN4和NSUN7外,其余基因在肿瘤组织中表达较正常组织明显上调( $P < 0.05$ )。说明m5C甲基化相关基因在肾透明细胞癌组织和相应的正常组织中的表达有明显差异。



### 2.2 肾透明细胞癌中 m5C 甲基化相关基因之间的相互作用

为了研究主要 m5C 甲基化相关基因之间的关联,我们构建了 PPI 网络(图 1C),以显示 13 个 m5C 甲基化相关基因之间的相互作用。TRDMT1 似乎是相互作用网络的中心基因,与其余大多数 m5C 甲基化相关基因相关。由于 PPI 网络没有提供相关性的详

细信息,对肾透明细胞癌进行了进一步的相关性分析(图 1D)。TET2 和 NSUN3 之间存在密切的相关性。除 NSUN4 外,NOP2 与其他 11 个 m5C 甲基化相关基因相关。除 NSUN7、NSUN4、NSUN3 外,NUSN2 与其他 m5C 甲基化相关基因呈正相关。DNMT3A、DNMT3B、NSUN4、NSUN6 亦如此。最后,TRDMT1 和 TET2 与 NSUN5 负相关性最高,分别为-0.33 和-0.31。



A: 热图显示每个临床样本中 13 种 m5C 甲基化相关基因的表达水平; N: 正常组; T: 肿瘤组; 两组比较, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; B: 小提琴图显示 13 种 m5C 甲基化相关基因在肿瘤组织和正常对照组之间存在显著差异; C: 13 个差异表达的 m5C 甲基化相关基因的 PPI 网络; D: Pearson 分析确定 TCGA 肾透明细胞癌中 13 个 m5C 甲基化相关基因之间的相关性。

图 1 m5C 甲基化相关基因在肾透明细胞癌中的表达和相互作用

Figure 1 Expression and interaction of m5C methylation-related genes in renal clear cell carcinoma

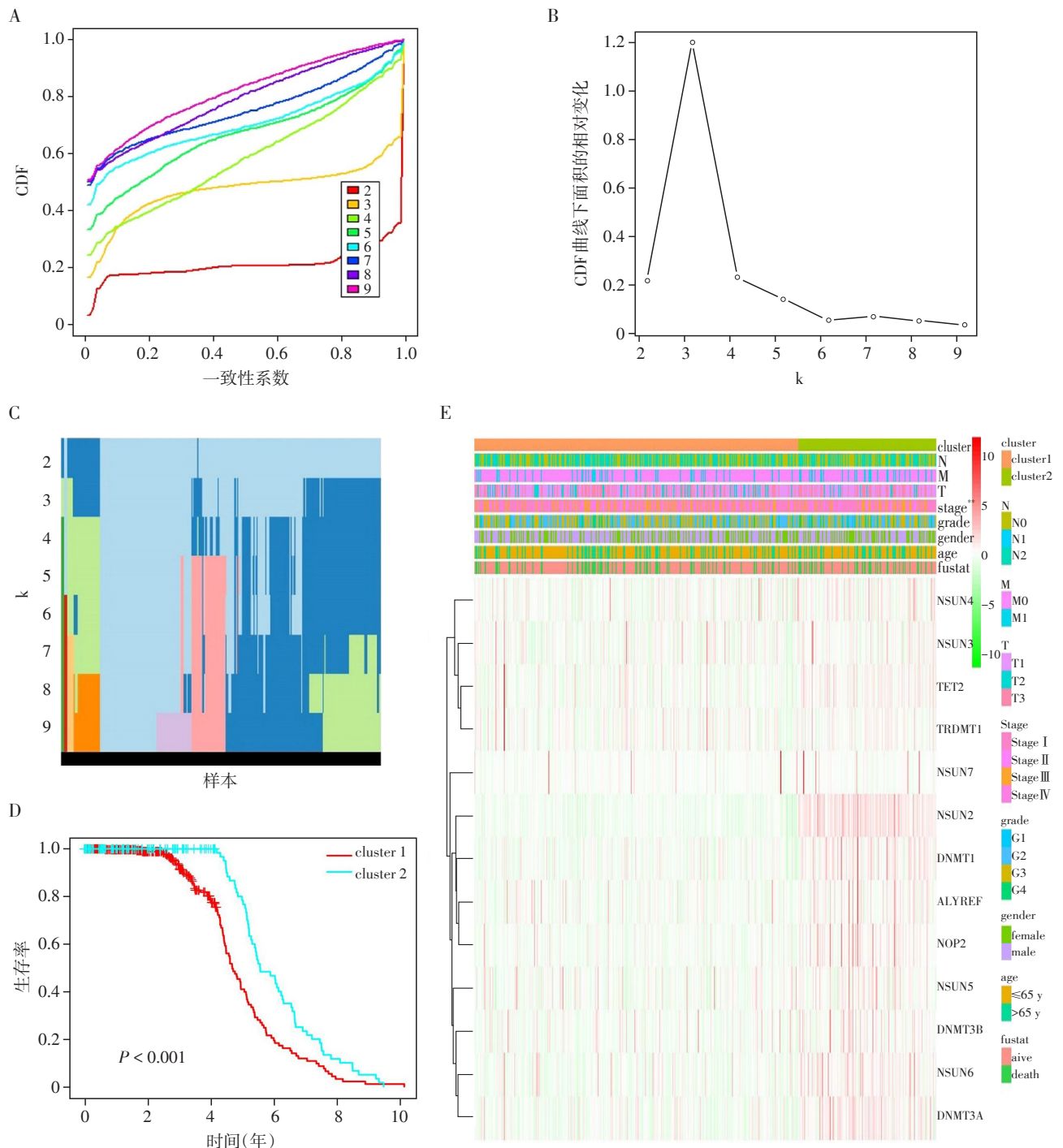
### 2.3 鉴定具有不同临床结果和特征的两组肾透明细胞癌样本

使用 TCGA 数据库中 572 份肾透明细胞癌样本的转录组数据进行共识聚类分析。从上述差异表达的 m5C 甲基化相关基因中,使用 13 个基因(NOP2、

NSUN2、NSUN3、NSUN4、NSUN5、NSUN6、NSUN7、TRDMT1、DNMT1、DNMT3A、DNMT3B、TET2 和 ALYREF)进行进一步研究。基于 13 个基因表达的相似性分析,结合共识聚类累积分布函数(CDF)和 CDF 曲线下区域的相对变化(图 2A、B),当聚类稳定性数

据集中k在2~10之间变化时, k=2被认为是合适的聚类数(图2C)。然后, 根据生存参数将肾透明细胞癌样本分为两个亚群, 可见亚群1中肾透明细胞癌患者的总生存期(overall survival, OS)较亚群2缩短( $P < 0.001$ , 图2D)。此外, 为了更好地预测这两个亚群肾

透明细胞癌的临床病理特征, 应用热图来显示分期的显著差异( $P < 0.05$ )。而其他特征, 如TNM分类、性别、年龄和临床分级等无显著差异(图2E)。从而, 我们可以得出结论, 13个m5C甲基化相关基因的表达特征与肾透明细胞癌患者的分期有关。



A, B: 共识聚类累积分布函数(CDF)和CDF曲线下区域的相对变化, 从k=2到k=9; C: 肾透明细胞癌的样本追踪图; D: 通过Kaplan-Meier方法对两个亚群进行生存分析; E: 根据临床病理参数分组的患者中13个m5C甲基化相关基因的表达热图; 两组比较,  $P < 0.01$ 。

图2 共识聚类分析显示肾透明细胞癌两个亚群的预后差异

Figure 2 Consensus clustering analysis revealed prognostic differences between two subgroups of clear cell renal cell carcinoma

## 2.4 生存预后模型的构建

为了研究m5C甲基化相关基因在肾透明细胞癌预后中的作用,对13个m5C甲基化相关基因表达谱进行了单变量Cox回归分析。结果显示,有4个基因(NSUN5、NSUN6、DNMT3B、TET2)可以用来预测肾透明细胞癌的预后(图3A)。在这4个基因中,TET2是HR < 1的保护基因,而NSUN5、NSUN6、DNMT3B是HR > 1的风险基因。然后将筛选出的基因应用Lasso-Cox回归建立生存预后模型,根据最低标准生成了单个候选基因的系数(图3B)。随后,TCGA数据库中每个肾透明细胞癌患者的风险值计算: $0.258\ 493 \times \text{NSUN6 表达量} + 0.058\ 203 \times \text{NSUN5 表达量} + (-0.502\ 707) \times \text{TET2 表达量} + 0.070\ 293 \times \text{DNMT3B 表达量}$ (图3C)。之后,根据风险值中位数界限,肾透明细胞癌患者被分成高风险组和低风险组,高风险组患者生存率显著低于低风险组(图3D)。这些结果表明,这4个m5C甲基化相关基因可以作为肾透明细胞癌的预测基因。

## 2.5 生存预后模型与肾透明细胞癌临床特征的相关性

热图显示,TET2在低风险组中高表达,而NSUN5、NSUN6和DNMT3B在高风险组中高表达。此外,T、M分期和分级存在显著差异(图3E,  $P < 0.05$ )。进一步运用单变量Cox回归分析确定预后因素。结果显示:年龄( $P < 0.001$ , HR=1.03, 95% CI: 1.017~1.044),分级( $P < 0.001$ , HR=2.302, 95% CI: 1.858~2.851),分期( $P < 0.001$ , HR=1.926, 95% CI: 1.678~2.209),T( $P < 0.001$ , HR=1.987, 95% CI: 1.674~2.357),N( $P < 0.001$ , HR=1.856, 95% CI: 1.528~2.255),M( $P < 0.001$ , HR=4.388, 95% CI: 3.178~6.06)为单变量分析的独立预后因素(图3F)。再对这些参数进行多变量Cox回归分析,有价值的因素包括年龄( $P < 0.01$ , HR=1.034, 95% CI: 1.018~1.049),分级( $P < 0.001$ , HR=1.549, 95% CI: 1.218~1.968),分期( $P=0.027$ , HR=1.689, 95% CI: 1.062~2.689,图3G)。

## 2.6 4个模型基因在泛癌中的表达及其在肾透明细胞癌中的遗传变异

为了明确上述4个基因在人类相应肿瘤中的表达差异,在泛癌样本中对这4个基因的转录水平进行比较,结果表明,TET2、NSUN5、NSUN6、DNMT3B在人类肿瘤中几乎均是过表达(图4A~D)。此外,在cBioportal数据库中研究了这4种基因与拷贝数变异之间的关系,结果显示:NSUN5、NSUN6、TET2、

DNMT3B的表达量和拷贝数变异之间的Pearson系数分别为0.36、0.13、0.30、0.34(图5),我们认为TET2、NSUN5、DNMT3B可能是由于其拷贝数的变异而对肾透明细胞癌的发生和进展产生影响。

## 2.7 4个模型基因对肾透明细胞癌预后产生影响的可能途径

通过limma软件包筛选出336个差异表达基因,其中281个上调( $P < 0.05$ ,  $\log_2\text{FC} > 1$ ),55个下调( $P < 0.05$ ,  $\log_2\text{FC} < -1$ ),使用GOplot和ggplot2软件包进行了GO和KEGG富集分析,以说明差异表达基因的功能。KEGG分析结果显示:在高风险组中,这些差异表达基因的功能主要与补体和凝血级联相关( $P < 0.05$ ),而在低风险组中,则主要与癌症中的蛋白聚糖相关( $P < 0.05$ ,图6A、B)。GO分析结果显示:在高风险组中,这些差异表达基因在生物过程方面主要与体液免疫反应和水解酶活性的负调控相关( $P < 0.05$ ),在细胞成分方面主要与血液微粒相关( $P < 0.05$ ),在分子功能方面主要与酶抑制剂活性相关( $P < 0.05$ );而在低风险组中,这些差异表达基因在生物过程、细胞成分和分子功能方面分别与对类固醇激素的反应、含胶原蛋白的细胞外基质和碳水化合物结合相关( $P < 0.05$ ,图6C、D)。这些结果为我们寻找m5C甲基化相关基因在肾透明细胞癌中的潜在功能提供了新的途径。

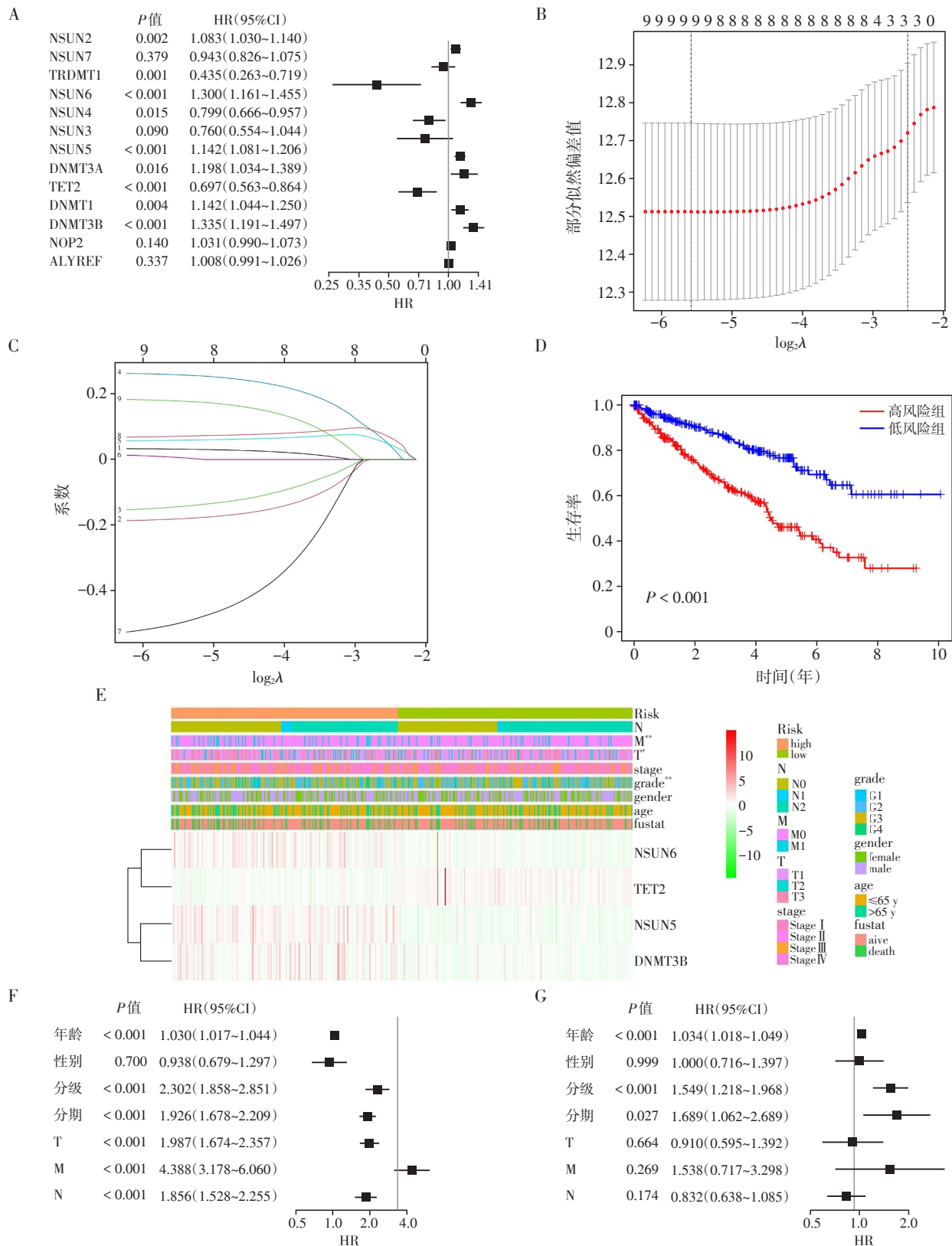
## 2.8 4种m5C甲基化相关基因在肾透明细胞癌细胞系及组织样本中的表达

为了进一步验证4个选定的模型基因在肾透明细胞癌和组织样本中的mRNA表达情况,本研究在肾透明细胞癌细胞系和组织样本中进行了qRT-PCR实验。结果显示:与人肾小管上皮细胞系(HK-2)和癌旁正常组织相比,肾癌细胞系(786-O、A498)和肾透明细胞癌组织样本中的NSUN5、NSUN6、TET2、DNMT3B表达水平显著上调(图7)。

## 3 讨论

本研究首先分析了这13个m5C甲基化相关基因在肾透明细胞癌中的表达情况,结果显示,NOP2、NSUN2、NSUN3、NSUN4、NSUN5、NSUN6、NSUN7、DNMT3A、DNMT3B、TET2和ALYREF在肾透明细胞癌组及正常组中差异表达。其中,除了NSUN4和NSUN7外,其余基因在肿瘤组中均高表达。接下来,根据m5C甲基化相关基因将肾透明细胞癌患者样本分为两个亚群,分析显示,亚群1中肾透明细胞癌患者总体生存率较亚群2低,这预示着m5C甲基

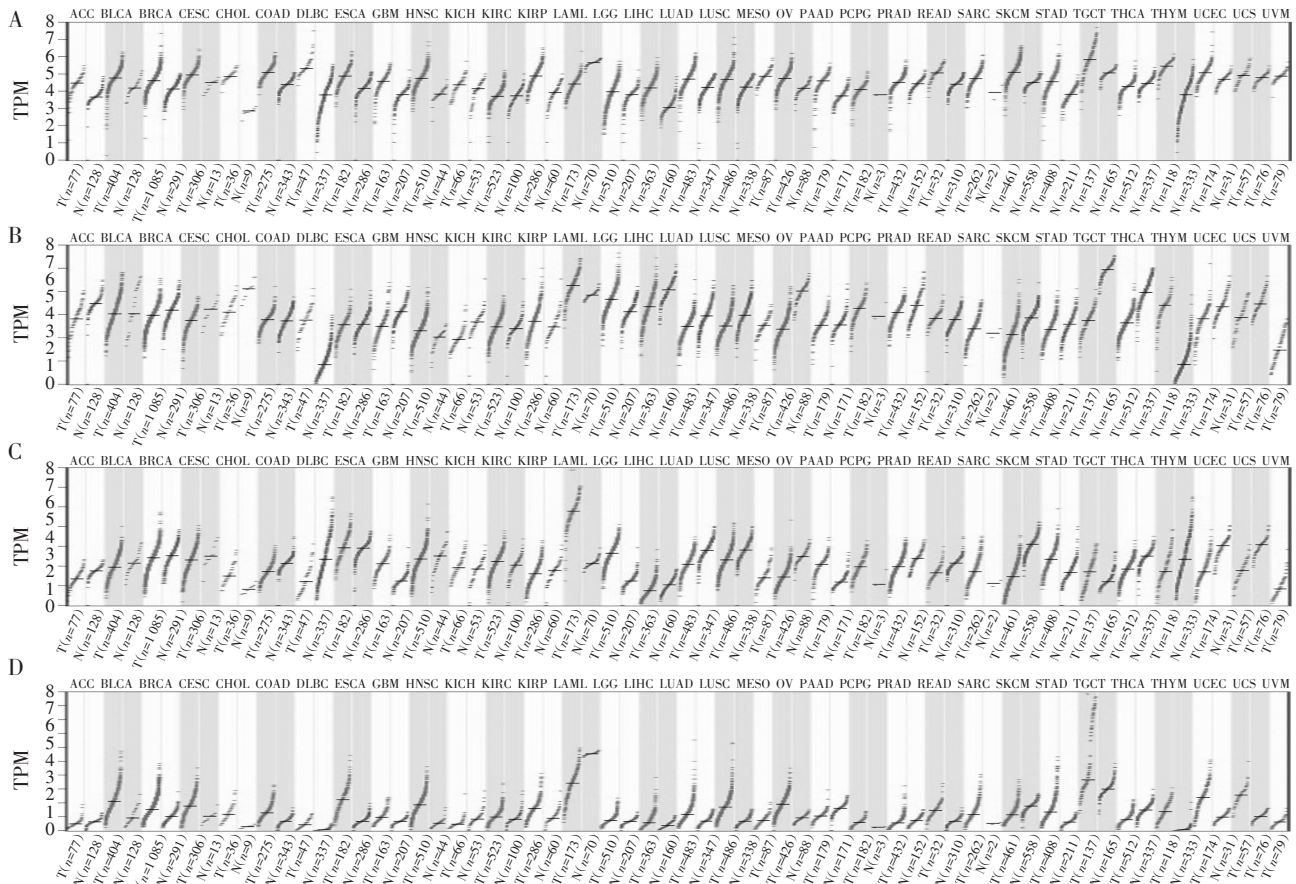




A: 单变量 Cox 回归分析结果显示了 13 个 m5C 甲基化相关基因的 P 值和风险比 (HR) 以及置信区间 (CI); B, C: Lasso-Cox 回归分析显示了 4 个选定的 m5C 甲基化相关基因 (NSUN5, NSUN6, DNMT3B, TET2) 的系数; D: Kaplan-Meier 总体生存分析将 TCGA 肾透明细胞癌队列的分为高风险组和低风险组, 然后进行风险评分; E: 高风险组和低风险组在 T、M 分期和分级方面存在差异,  $P < 0.01$ ; F, G: 单变量和多变量 Cox 回归分析揭示了可能的独立预后因素。

图 3 4 种 m5C 甲基化相关基因所构建的生存预后模型

Figure 3 A survival prognostic model constructed by four m5C methylation-related genes



NSUN5(A)、NSUN6(B)、TET2(C)和DNMT3B(D)在泛癌和正常组织中的表达值。T:肿瘤组;N:正常组;TPM:每百万条reads的转录本。

图4 4个模型基因在泛癌中的表达及其在肾透明细胞癌中的遗传变异

Figure 4 Pan-cancer expression of four model genes and their genetic variation in renal clear cell carcinoma

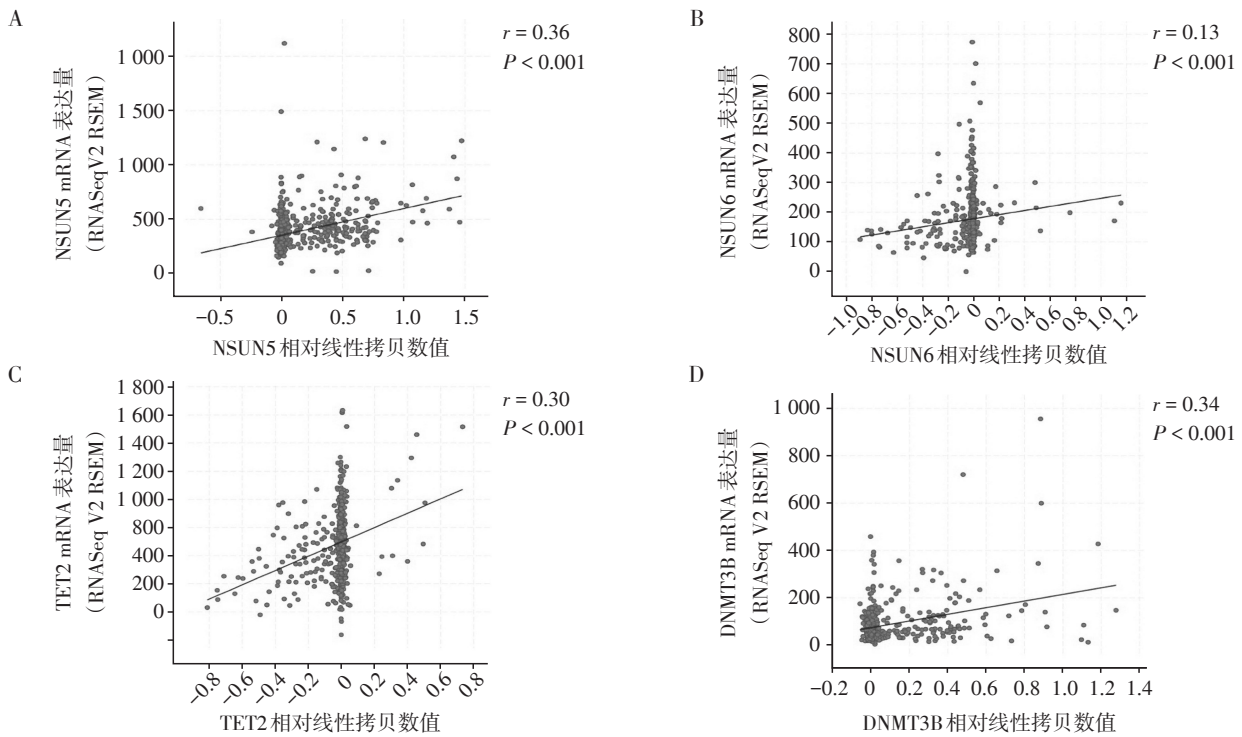
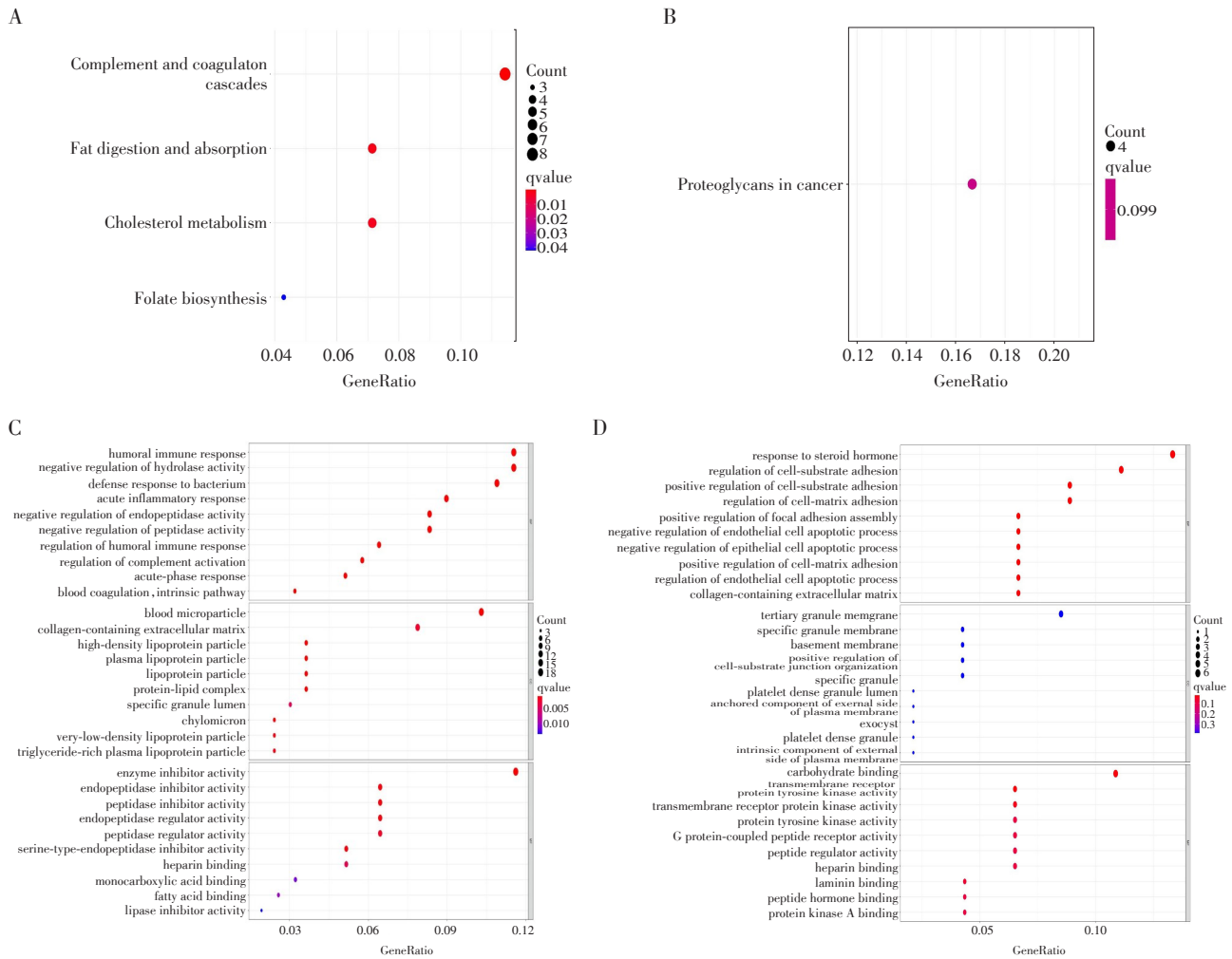


图5 NSUN5(A)、NSUN6(B)、TET2(C)和DNMT3B(D)与拷贝数变异之间的关系

Figure 5 Association between NSUN5(A), NSUN6(B), TET2(C) and DNMT3B(D) and copy number variation





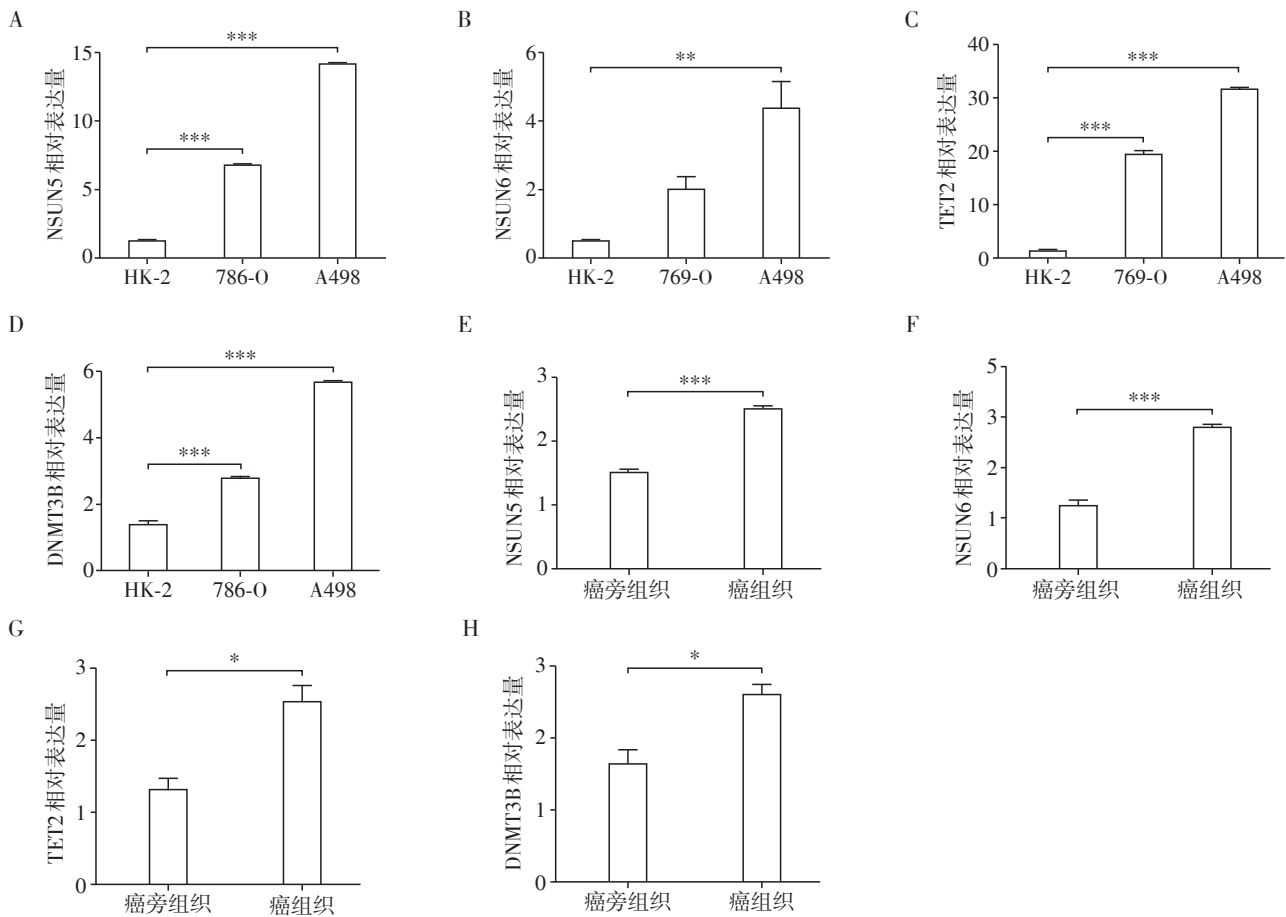
A、B:KEGG分析结果,气泡图显示了高低风险组中排名靠前的信号通路;C、D:GO分析结果,气泡图显示了高低风险组中排名靠前的信号通路。

图6 GO和KEGG富集分析结果  
Figure 6 GO and KEGG enrichment analysis results

化相关基因可能与肾透明细胞癌的预后相关。通过 Lasso-Cox 回归分析建立了由 NSUN5、NSUN6、DNMT3B 和 TET2 4 个基因组成的生存预后模型,根据每个肾透明细胞癌患者风险值将样本分为高风险组和低风险组,观察发现,高风险组患者生存率显著低于低风险组。这一结果提示,这 4 个基因可以作为肾透明细胞癌的预测因子。接着先制作热图评估了 4 个 m5C 甲基化相关基因的临床参数,然后通过单变量和多变量 Cox 回归分析确立了分级和分期是肾透明细胞癌的独立预后因素。此外,GEPIA2 数据库分析发现这 4 种基因在肿瘤组中的表达均上调,而在 cBioportal 数据库中发现这 4 种基因的表达量与拷贝数变异之间有关,这一结果显示, TET2、NSUN5、DNMT3B 可能是由于其拷贝数的变异而对肾透明细胞癌的发生和进展产生影响。

最后,GO 和 KEGG 富集分析验证这些 m5c 甲基化相关基因的主要生物学功能。总体而言,本研究结果表明 m5C 甲基化相关基因对肾透明细胞癌具有相对稳定的预测能力。

目前,NSUN5、NSUN6、DNMT3B 和 TET2 在肿瘤中的作用已被广泛研究。根据报道,NSUN5 可以通过修饰真核 rRNA 来维持小鼠的蛋白质合成和正常的生长<sup>[13]</sup>。此外,NSUN5 在结直肠癌肿瘤组织和细胞中表达上调,并通过促进细胞增殖,在体外触发细胞周期停滞进而促进肿瘤的生长<sup>[14]</sup>;NSUN6 可以通过调节 CDK10 参与有丝分裂纺锤体的组装进而影响胰腺癌细胞的增殖<sup>[15]</sup>,此外,NSUN6 可以通过 NELFB 和 RPS6KB2 相互作用来介导胶质母细胞瘤对替莫唑胺的反应<sup>[16]</sup>;另外有研究显示, DNMT3B 能够与前列腺癌细胞中的转录抑制因子结合控制



qRT-PCR 验证 m5C 甲基化相关基因在细胞系以及组织样本中的表达差异。A~D: NSUN5(A)、NSUN6(B)、TET2(C)和 DNMT3B(D)在肾透明细胞癌细胞系中显著上调;E~H: NSUN5(E)、NSUN6(F)、TET2(G)和 DNMT3B(H)在肿瘤组织样本中显著上调。两组比较,  $P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  ( $n=4$ )。

图7 肾癌细胞系及组织样本中4个m5C甲基化相关基因的验证

Figure 7 Validation of four m5C methylation-related genes in renal cancer cell lines and tissue samples

RAD9 的表达,进而调节前列腺癌细胞的进展<sup>[17]</sup>, DNMT3B 还可以通过表观遗传学抑制 miR-34a 促进膀胱癌的迁移和侵袭<sup>[18]</sup>。TET2 对调节正常造血至关重要,并且可以作为肿瘤抑制因子以维持造血细胞的稳态,TET2 的缺失可能会导致骨髓恶性肿瘤的发生<sup>[19]</sup>,此外,有研究表明,TET2 可以抑制细胞的增殖和转移,诱导乳腺癌细胞的凋亡<sup>[20]</sup>。

总而言之,本研究构建了一个新的 m5C 相关基因的风险预后模型用于预测肾透明细胞癌患者风险程度和生存预后;根据该模型的风险值中位值对患者进行分组,生存分析结果显示:高风险组患者的分级和分期更高,这证明了所构建模型的可靠性。通过生物信息学技术对大规模的肾透明细胞癌患者数据进行了分析,并在细胞系和组织中进行验证,相较于以往的研究适用更广泛、可靠性更高,本研究结果对肾透明细胞癌患者的风险程度以及

预后的预测具有一定意义,同时,本研究构建的模型基因可以作为肾透明细胞癌个体化治疗的治疗靶点。但是,本研究仍然存在着一些局限性,本研究数据主要来自 TCGA 公共数据库,缺乏其他的数据库来源的数据,并且在使用模型之前需要进一步的多中心前瞻性临床研究进行验证。

[参考文献]

[1] SUN Z, JING C, XIAO C, et al. Prognostic risk signature based on the expression of three m6A RNA methylation regulatory genes in kidney renal papillary cell carcinoma [J]. Aging, 2020, 12(21): 22078-22094

[2] CROCEROSSA F, CARBONARA U, CANTIello F, et al. Robot-assisted radical nephrectomy: a systematic review and meta-analysis of comparative studies [J]. Eur Urol, 2021, 80(4): 428-439

[3] RUAN B L, FENG X Z, CHEN X Y, et al. Identification

- of a set of genes improving survival prediction in kidney renal clear cell carcinoma through integrative reanalysis of transcriptomic data [J]. *Dis Markers*, 2020, 2020: 8824717
- [4] RONG D W, SUN G S, WU F, et al. Epigenetics: roles and therapeutic implications of non-coding RNA modifications in human cancers [J]. *Mol Ther Nucleic Acids*, 2021, 25:67-82
- [5] WILKINSON E, CUI Y H, HE Y Y. Roles of RNA modifications in diverse cellular functions [J]. *Front Cell Dev Biol*, 2022, 10:828683
- [6] PAN J F, HUANG Z D, XU Y Q. m5C RNA methylation regulators predict prognosis and regulate the immune microenvironment in lung squamous cell carcinoma [J]. *Front Oncol*, 2021, 11:657466
- [7] CHEN X, LI A, SUN B F, et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs [J]. *Nat Cell Biol*, 2019, 21(8):978-990
- [8] HAN Z, YANG B, WANG Y, et al. Identification of expression patterns and potential prognostic significance of m(5)C-related regulators in head and neck squamous cell carcinoma [J]. *Front Oncol*, 2021, 11:592107
- [9] MOUNIR M, LUCCHETTA M, SILVA T C, et al. New functionalities in the TCGA biolinks package for the study and integration of cancer data from GDC and GTEX [J]. *PLoS Comput Biol*, 2019, 15(3):e1006701
- [10] ZHANG X K, WANG Y Y, A G R, et al. Pan-cancer analysis of PARP1 alterations as biomarkers in the prediction of immunotherapeutic effects and the association of its expression levels and immunotherapy signatures [J]. *Front Immunol*, 2021, 12:721030
- [11] TANG Z F, KANG B X, LI C W, et al. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis [J]. *Nucleic Acids Res*, 2019, 47(W1):W556-W560
- [12] SZKLARCZYK D, GABLE A L, LYON D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets [J]. *Nucleic Acids Res*, 2019, 47(D1):D607-D613
- [13] HEISSENBERGER C, LIENDL L, NAGELREITER F, et al. Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth [J]. *Nucleic Acids Res*, 2019, 47(22):11807-11825
- [14] JIANG Z, LI S, HAN M J, et al. High expression of NSUN5 promotes cell proliferation via cell cycle regulation in colorectal cancer [J]. *Am J Transl Res*, 2020, 12(7):3858-3870
- [15] YANG R, LIANG X, WANG H, et al. The RNA methyltransferase NSUN6 suppresses pancreatic cancer development by regulating cell proliferation [J]. *EBioMedicine*, 2021, 63:103195
- [16] AWAH C U, WINTER J, MAZDOOM C M, et al. NSUN6, an RNA methyltransferase of 5-mC controls glioblastoma response to temozolomide (TMZ) via NELFB and RPS6KB2 interaction [J]. *Cancer Biol Ther*, 2021, 22(10/11/12):587-597
- [17] ZHU A P, HOPKINS K M, FRIEDMAN R A, et al. DNMT1 and DNMT3B regulate tumorigenicity of human prostate cancer cells by controlling RAD9 expression through targeted methylation [J]. *Carcinogenesis*, 2021, 42(2):220-231
- [18] XU K, CHEN B, LI B, et al. DNMT3B silencing suppresses migration and invasion by epigenetically promoting miR-34a in bladder cancer [J]. *Aging*, 2020, 12(23):23668-23683
- [19] PADDA J, KHALID K, YADAV J, et al. JAK2 and TET2 mutation in polycythemia vera [J]. *Cureus*, 2021, 13(9):e17854
- [20] PENG B, LI C, HE L, et al. miR-660-5p promotes breast cancer progression through down-regulating TET2 and activating PI3K/AKT/mTOR signaling [J]. *Braz J Med Biol Res*, 2020, 53(12):e9740

[收稿日期] 2022-04-18

(责任编辑:蒋莉)