

DOI:10.11798/j.issn.1007-1520.202321515

· 论著 ·

# 基于 N6-甲基腺苷相关长链非编码核糖核酸表达的喉鳞状细胞癌预后分析

李念燊, 韩杏倩, 何金阳, 黎国胜, 练森兰, 罗宗华, 尹时华, 侯涛

(广西医科大学第二附属医院耳鼻咽喉科, 广西南宁 530021)

**摘要:** **目的** 探索 N6-甲基腺苷(m6A)相关长链非编码核糖核酸(lncRNA)与喉鳞状细胞癌(LSCC)的预后关系及其临床意义。**方法** 获取癌症基因组图谱(TCGA)数据库 LSCC 的转录组数据和临床数据,共表达分析筛选 m6A 基因相关的 lncRNA,单变量 Cox 分析 m6A 相关 lncRNA 与 LSCC 预后关系、套索回归及交叉验证法迭代分析构建 LSCC 预后模型。采用实时荧光定量聚合酶链式反应(qPCR)验证构建预后模型的 lncRNA 在 LSCC 中的转录水平。通过预后模型计算风险评分,将 LSCC 区分为高、低风险患者,高、低风险患者之间进行基因集富集分析(GSEA)。风险评分与浸润 LSCC 的免疫细胞进行免疫相关性分析。**结果** m6A 相关基因与 lncRNA 的共表达分析筛选出 169 个与 m6A 基因相关的 lncRNA(相关系数  $>0.4$ ,  $P < 0.001$ ),通过单变量 Cox 分析确定了 LSCC 预后相关 lncRNA: ALOX12-AS1 ( $P < 0.05$ )、LINC00528 ( $P < 0.05$ )、STAG3L5P-PVRIG2P-PILRB ( $P < 0.05$ )、MNX1-AS1 ( $P < 0.01$ ) 和 LINC02043 ( $P < 0.05$ )。套索回归、交叉验证迭代分析结果:变量为 4 时模型的均方根误差最小,即 5 个预后相关 lncRNA 仅有 ALOX12-AS1、LINC00528、STAG3L5P-PVRIG2P-PILRB 与 MNX1-AS1 可作为模型变量。预后模型:风险评分 =  $(0.176723096228585 \times \text{MNX1-AS1 表达量}) + (-0.614916717648596 \times \text{ALOX12-AS1 表达量}) + (-0.814201385798827 \times \text{LINC00528 表达量}) + (-0.436537752110547 \times \text{STAG3L5P-PVRIG2P-PILRB 表达量})$ 。qPCR 结果 ALOX12-AS1 ( $P < 0.0001$ )、LINC00528 ( $P < 0.01$ )、STAG3L5P-PVRIG2P-PILRB ( $P < 0.0001$ )、MNX1-AS1 ( $P < 0.0001$ ) 较于癌旁组织,在 LSCC 组织表达水平上调。GSEA 分析结果:高风险患者在细胞-基质粘附(KEGG\_FOCAL\_ADHESION) ( $P < 0.05$ )与细胞外基质受体相互作用(KEGG\_ECM\_RECEPTOR\_INTERACTION) ( $P < 0.05$ )的通路下调,低风险患者在碱基切除修复(KEGG\_SPLICEOSOME) ( $P < 0.05$ )、剪接体(KEGG\_SPLICEOSOME) ( $P < 0.05$ )通路上调。风险评分与免疫浸润细胞相关性分析,风险评分与效应 B 细胞( $R = -0.24$ ,  $P = 0.017$ )、细胞毒性 T 细胞( $R = -0.26$ ,  $P = 0.0091$ )和 T 滤泡辅助细胞( $R = -0.22$ ,  $P = 0.028$ )呈负相关。**结论** 基于 m6A 相关 lncRNA 表达量构建的 LSCC 预后模型可作为预测 LSCC 患者预后的有效工具。

**关键词:** 喉鳞状细胞癌; N6-甲基腺苷; 长链非编码核糖核酸; 基因富集分析; 免疫相关性分析

中图分类号: R739.65

## Prognostic analysis of laryngeal squamous cell carcinoma based on the expression of N6-methyladenosine-related lncRNAs

LI Nianshen, HAN Xingqian, HE Jinyang, LI Guosheng, LIAN Miaolan, LUO Zonghua, YIN Shihua, HOU Tao  
(Department of Otolaryngology, the Second Affiliated Hospital of Guangxi Medical University, Nanning 530021, China)

**Abstract:** **Objective** To explore the relationship between m6A-related lncRNAs and the prognosis of laryngeal squamous cell carcinoma (LSCC) and its clinical significance. **Methods** Transcriptome data from the Cancer Genome Atlas (TCGA) database and clinical data of LSCC were obtained. The lncRNAs associated with the m6A gene were screened by co-expression analysis, a prognostic model of LSCC was constructed by lasso regression analysis and iterative analysis of cross-validated methods, and the relationship between m6A-associated lncRNA and LSCC survival was explored by univariate cox

基金项目:国家自然科学基金(82160213);广西医科大学第二附属医院 NSFC 培育项目(GJFY2019001);广西卫生计生委自筹项目(Z20191075);广西中医药管理局课题(GZZC2020192)。

第一作者简介:李念燊,男,在读硕士研究生。

通信作者:尹时华,Email: shihuayin@gxmu.edu.cn

analysis. Real-time polymerase chain reaction (RT-qPCR) was used to validate the transcript levels of lncRNA in the prognostic model. Risk scores were calculated by the prognostic model to distinguish LSCC into high- and low-risk groups, and gene set enrichment analysis (GSEA) was performed between the two groups. Tumor tissues were detected for immune infiltration, and the correlation between the risk score value and the degree of immune cell infiltration was evaluated by correlation analysis. **Results** A total of 169 lncRNA associated with m6A genes were screened out by co-expression analysis (correlation coefficient  $> 0.4$ ,  $P < 0.001$ ), and univariate cox analysis determined the prognosis-associated lncRNAs in LSCC including ALOX12-AS1 ( $P < 0.05$ ), LINC00528 ( $P < 0.05$ ), STAG3L5P-PVRIG2P-PILRB ( $P < 0.05$ ), MNX1-AS1 ( $P < 0.01$ ), and LINC02043 ( $P < 0.05$ ). Lasso regression and cross-validation iterative analysis revealed that the root-mean-square error of the model was the smallest when the variable was 4, i. e., only ALOX12-AS1, LINC00528, STAG3L5P-PVRIG2P-PILRB and MNX1-AS1 could be used as model variables. In the prognostic model, the risk score =  $(0.176723096228585 \text{MNX1-AS1 expression}) + (-0.614916717648596 \text{ALOX12-AS1 expression}) + (-0.814201385798827 \text{LINC00528 expression}) + (-0.436537752110547 \text{STAG3L5P-PVRIG2P-PILRB expression})$ . RT-qPCR showed that ALOX12-AS1 ( $P < 0.0001$ ), LINC00528 ( $P < 0.01$ ), STAG3L5P-PVRIG2P-PILRB ( $P < 0.0001$ ), and MNX1-AS1 ( $P < 0.0001$ ) were up-regulated in LSCC tissues compared with para-carcinoma tissues. GSEA analysis indicated KEGGFOCALADHESION ( $P < 0.05$ ) and KEGGECMRECEPTORINTERACTION ( $P < 0.05$ ) pathways were down-regulated in the high-risk group, and KEGGSPLICEOSOME pathway ( $P < 0.05$ ) and KEGGSPLICesome pathway ( $P < 0.05$ ) were up-regulated in the low-risk group. Correlation analysis between risk score and immune infiltrating cells disclosed that risk score was negatively correlated with plasma cell ( $R = -0.24$ ,  $P < 0.05$ ), CD8 T cell ( $R = -0.26$ ,  $P < 0.01$ ) and T follicular helper cell ( $R = -0.22$ ,  $P < 0.05$ ). **Conclusion** The prognostic model of LSCC constructed based on m6A-associated lncRNA expression can be used as an effective tool to predict the prognosis of patients with this disease.

**Keywords:** Laryngeal squamous cell carcinoma; N6-Methyladenosine; lncRNA; Prognostic model; Gene enrichment analysis; Immune correlation analysis

喉癌是头颈部最常见的恶性肿瘤之一,占全身恶性肿瘤的1%~5%,其中96%~98%为喉鳞状细胞癌(laryngeal squamous cell carcinoma, LSCC)<sup>[1]</sup>。2020年全球有184 615例喉癌新发病例和99 840例死亡<sup>[2]</sup>。LSCC在世界范围内的新发病例和死亡人数逐年增加,发病率和死亡率均较高,严重危害人类健康。现有的临床病理参数,如淋巴结转移或肿瘤分期,可以识别患者,并根据患者的具体情况采取包括手术、放疗、化疗、免疫治疗、靶向治疗、联合治疗等多种治疗方式<sup>[3]</sup>。虽然在LSCC的诊断和治疗方面取得了一定进展,但尚未获得满意的诊断和治疗效果。多数患者确诊时已属中晚期,总体预后较差。LSCC预后的预测仍是临床上的一大难题。因此,寻找能够预测患者预后的分子标志物和新的治疗靶点意义重大。

长链非编码核糖核酸(long non-coding ribonucleic acid, lncRNA)是指长度超过200 nt的非编码RNA,一般在人类基因组中转录<sup>[4]</sup>。已经证实lncRNA在表观遗传修饰、信使核糖核苷酸(messenger ribonucleic acid, mRNA)转录、剪接、稳定性、翻译等生物学功能中发挥重要作用<sup>[5]</sup>。特别是lncRNA在调节染色质动力学、基因表达、生长、分化和发育中起关键作用<sup>[6]</sup>,多个报道指出lncRNA参

与包括癌症<sup>[7-9]</sup>在内的多种疾病的发生、发展和预后。N6-甲基腺苷(N6-methyladenosine, m6A)是核糖核酸(ribonucleic acid, RNA)中最常见的修饰之一。研究表明m6A是动态可逆的,参与调节RNA转录、输出、剪接、降解和翻译<sup>[10]</sup>。m6A通过甲基转移酶、结合蛋白和去甲基化酶3种同源因子的动态相互作用修饰并作用于RNA,这3种调节因子之间的相互作用被认为与癌症生长和进展有关<sup>[11]</sup>。

目前,利用测序、芯片等大数据进行的生物信息学研究广泛开展,为肿瘤的机制研究和预后标志物的筛选提供了新的思路。本研究基于癌症基因组图谱(the cancer genome atlas, TCGA)数据库,通过生物信息学方法筛选与LSCC预后相关的m6A相关lncRNA,建立预后预测模型,以期对LSCC的诊断、治疗和预后评估提供参考。

## 1 材料和方法

### 1.1 数据整理过程

依据TCGA实用指南的要求,从TCGA数据库下载LSCC转录表达谱数据和匹配的临床信息(表1)<sup>[12]</sup>。截止至2021年3月,TCGA公共数据库包括112例LSCC和12例正常组织标本的表达谱。

依靠注释文件"GRCh38"将转录表达谱数据中的 Ensembl ID 转换为基因名,并提取所有 lncRNA 表达量。

表 1 临床基本数据 (例,%)

临床特征	总样本量 (112)	百分比
年龄 (岁)		
≤65	73	65.2
>65	39	34.8
性别		
女	20	17.9
男	92	82.1
等级		
G1	8	7.1
G2	70	62.5
G3	30	26.8
未知	4	3.6
病理分期		
I 期	2	1.8
II 期	9	8.0
III 期	14	12.5
IV (A-C) 期	72	64.3
未知	15	13.4
T 分期		
T1	7	6.3
T2	12	10.7
T3	26	23.2
T4	54	48.2
未知	13	11.6
M 分期		
M0	40	35.7
M1	1	0.9
未知	71	63.4
N 分期		
N0	39	34.8
N1	12	10.7
N2	40	35.7
N3	2	1.8
未知	19	17.0
生存状态		
死亡	62	55.4
存活	50	44.6

## 1.2 m6A 相关基因与 lncRNA 相关性分析

根据 m6A 相关基因列表 (表 2)<sup>[13-14]</sup> 提取 m6A 相关基因的表达数据。对 m6A 相关基因与 LSCC 的 lncRNA 进行相关性分析,依据相关系数  $>0.4$  和  $P < 0.001$  对结果进行过滤,筛选出 m6A 相关的 lncRNA。

表 2 m6A 相关基因

基因名	基因编码蛋白
<i>METTL14</i>	甲基转移酶
<i>METTL16</i>	甲基转移酶
<i>METTL3</i>	甲基转移酶
<i>RBM15</i>	甲基转移酶
<i>RBM15B</i>	甲基转移酶
<i>VIRMA</i>	甲基转移酶
<i>WTAP</i>	甲基转移酶
<i>ZC3H13</i>	甲基转移酶
<i>KIAA1429</i>	甲基转移酶
<i>METTL15</i>	甲基转移酶
<i>FMR1</i>	结合蛋白
<i>HNRNPA2B1</i>	结合蛋白
<i>HNRNPC</i>	结合蛋白
<i>IGFBP1</i>	结合蛋白
<i>IGFBP2</i>	结合蛋白
<i>IGFBP3</i>	结合蛋白
<i>LRPPRC</i>	结合蛋白
<i>RBMX</i>	结合蛋白
<i>YTHDC1</i>	结合蛋白
<i>YTHDC2</i>	结合蛋白
<i>YTHDF1</i>	结合蛋白
<i>YTHDF2</i>	结合蛋白
<i>YTHDF3</i>	结合蛋白
<i>ALKBH5</i>	去甲基化酶

## 1.3 单因素 Cox 分析筛选预后相关的 m6A 相关 lncRNA

基于癌旁组织与 LSCC 组织 m6A 相关 lncRNA 表达量,结合患者的生存状态、生存时间,对每一个 m6A 相关 lncRNA 进行单因素 Cox 分析,计算结果根据  $<0.05$  进行过滤,筛选出 LSCC 预后相关的 m6A 相关 lncRNA。R 软件中对筛选出来的 lncRNA 在正常组织与肿瘤组织中表达量进行差异分析。

## 1.4 预后模型的构建

将 112 例患者按 1:1 随机分为模型训练组与模型测试组,每组 56 例,基于 5 个预后相关的 m6A 相关 lncRNA 表达量和患者的生存时间、生存状态,利用模型训练组病例进行套索回归及交叉验证方法进行迭代分析,构建 LSCC 预后风险评分模型。

## 1.5 qPCR 验证构建模型 lncRNA 表达

收集 2021 年 8 月—2021 年 12 月在广西医科大学第二附属医院治疗的 LSCC 患者 9 例,9 例 LSCC 组织与癌旁组织 RNA 由 TRIzol 提取,后将其逆转录为互补 DNA。RT-qPCR 采用 StepOnePlus™ 实时荧光

定量 PCR 仪(美国赛默飞公司)进行。使用 TaKaRa 公司 TB Green<sup>®</sup> Premix Ex Taq<sup>™</sup> II 试剂盒进行 qRT-PCR。反应体系总体积为 20  $\mu$ L [10  $\mu$ L TB Green Premix Ex Taq II(2X), 7  $\mu$ L DEPC 水, 1  $\mu$ L cDNA, 0.8  $\mu$ L 上游引物, 0.8  $\mu$ L 下游引物, 0.4  $\mu$ L ROX Reference Dye (50X)]。反应条件为, 98 $^{\circ}$ C 预变性 30s, 1 轮循环。98 $^{\circ}$ C 5s, 60 $^{\circ}$ C 30s 45 轮循环。98 $^{\circ}$ C 15s, 60 $^{\circ}$ C 1s, 98 $^{\circ}$ C 15s, 1 轮循环。内参标准 ACTIN, 每个样品提供 3 个复孔。基因相对表达水平算法为  $2^{-\Delta\Delta Ct}$  法。引物序列见表 3。

表 3 引物序列信息

名称	序列
ALOX12-AS1	F: AAATGCTCACCTTGCCTTCA R: TGCTGTCTCATGTGCGTCTT
LINC00528	F: ATCCTCAAACACGGTGTCTCT R: CCCTGCTGTGGGATTAAGT
STAG3L5P-PVRIG2P-PILRB	F: AAGCAAAGGGCACTCAGAATC R: CACAGGAGGAGGAGGCACAG
MNX1-AS1	F: CTGAAGACCGTGTGTTC R: TTCCCATAGCAGATAAGGTTCC

## 1.6 预后模型的综合分析

风险评分是模型计算得出并用于衡量每个 LSCC 病例预后风险的指标。通过构建的预后模型计算所有 LSCC 患者的风险评分。并以模型训练组风险评分的中位数将模型训练组病例和模型测试组病例分为高风险患者和低风险患者。Kaplan-Meier 生存分析检验高、低风险患者之间生存状况是否存在差异。工作特征(ROC)曲线作为评价模型的预测能力的工具。

对风险评分与患者的临床特征(年龄、性别、病理分期、TMN 分期)进行单因素 Cox 分析和多因素 Cox 分析, 检验风险评分能否作为独立预后因素评估 LSCC 患者预后。

## 1.7 GSEA 基因富集分析

使用 GSEA v4.1.0 对高风险患者、低风险患者的转录表达谱进行差异分析和基因富集分析, 选择基因集“c2. cp. kegg. v7.4. symbols. gmt”进行富集。

## 1.8 免疫相关性分析

通过 R 软件将来自 TCGA 的转录表达谱数据转变为芯片数据格式, 应用“e1071”、“preprocess-Core”, “limma”3 个 R 包对肿瘤样本进行免疫浸润分析, 得到每个样本的免疫细胞含量。风险评分与

每一类免疫细胞进行相关性分析, 计算风险评分与免疫细胞的相关性。

## 1.9 统计学方法

m6A 相关基因与 lncRNA 相关性分析、风险评分与免疫细胞相关性分析采用 pearson 相关性检验; 预后相关的 m6A 相关 lncRNA 采用单因素 Cox 回归分析; 风险评分独立预后因素分析采用单因素 Cox 回归分析和多因素 Cox 回归分析进行两次检验; 由于基因表达不服从正态分布, 基因差异分析使用 Wilcoxon 秩和检验; Kaplan-Meier 曲线采用 Log-rank 检验法进行生存分析; 以  $P < 0.05$  为差异具有统计学意义。模型预测效果可靠性采用 ROC 曲线进行验证曲线下面积(AUC)  $\geq 0.9$  为强有效预测值,  $\geq 0.70$  为有效预测值,  $\geq 0.6$  为可接受预测值<sup>[15]</sup>。qPCR 结果采用  $2^{-\Delta\Delta Ct}$  法计算基因相对表达水平, 配对  $t$  检验对相对表达量进行统计。本实验对高风险患者和低风险患者的差异基因进行 GSEA 分析, 以  $P < 0.05$  为标准筛选富集结果。

## 2 结果

### 2.1 预后相关的 lncRNA

通过相关性检验, 筛选出 169 个与 m6A 修饰基因相关的 lncRNA, 通过网络图展示与 m6A 相关基因对应的 lncRNA。见图 1。

### 2.2 预后相关的 lncRNA

结合 112 例患者生存时间、生存状态, 对这些 m6A 相关 lncRNA 进行单因素 Cox 分析, 5 个与 LSCC 预后相关的 m6A 相关 lncRNA 被筛选出来: ALOX12-AS1 (HR = 0.263,  $P < 0.05$ )、LINC00528 (HR = 0.239,  $P < 0.01$ )、STAG3L5P-PVRIG2P-PILRB (HR = 0.345,  $P < 0.05$ ); ALOX12-AS1、LINC00528 和 STAG3L5P-PVRIG2P-PILRB 的风险比 HR  $< 1$  为 LSCC 的保护因素, MNX1-AS1 (HR = 1.474,  $P < 0.01$ ) 和 LINC02043 (HR = 2.515,  $P < 0.05$ ); MNX1-AS1 和 LINC02043 的 HR  $> 1$  为 LSCC 的高危因素。5 个预后相关的 m6A 相关 lncRNA 在正常组织与在 LSCC 的表达量进行差异分析, ALOX12-AS1 ( $P < 0.05$ )、LINC00528 ( $P < 0.001$ )、STAG3L5P-PVRIG2P-PILRB ( $P < 0.001$ )、MNX1-AS1 ( $P < 0.001$ ) 和 LINC02043 ( $P < 0.001$ ) 在肿瘤组织中均为高表达。见图 2。

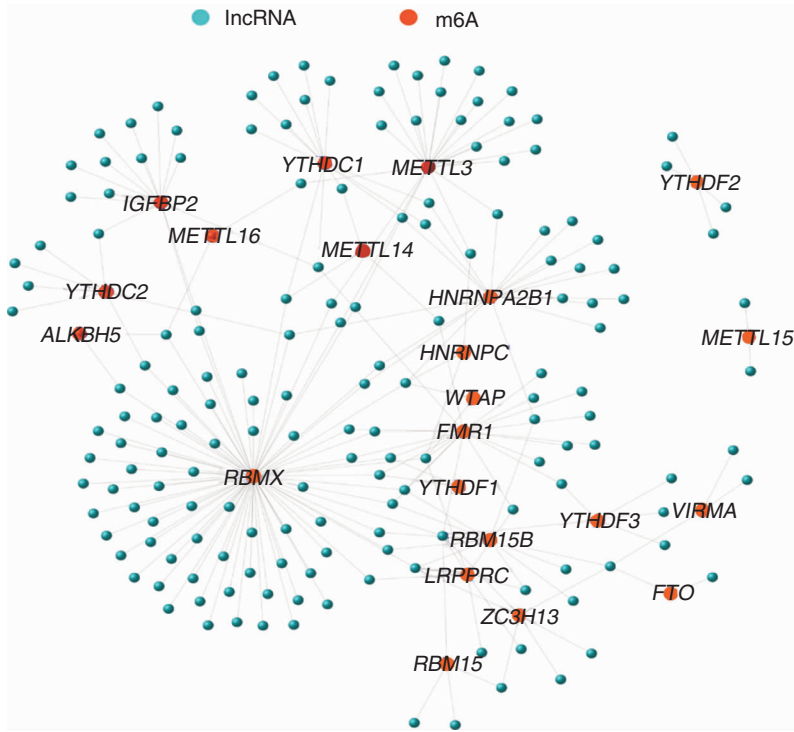


图 1 m6A 相关基因与相关 lncRNA 网络图

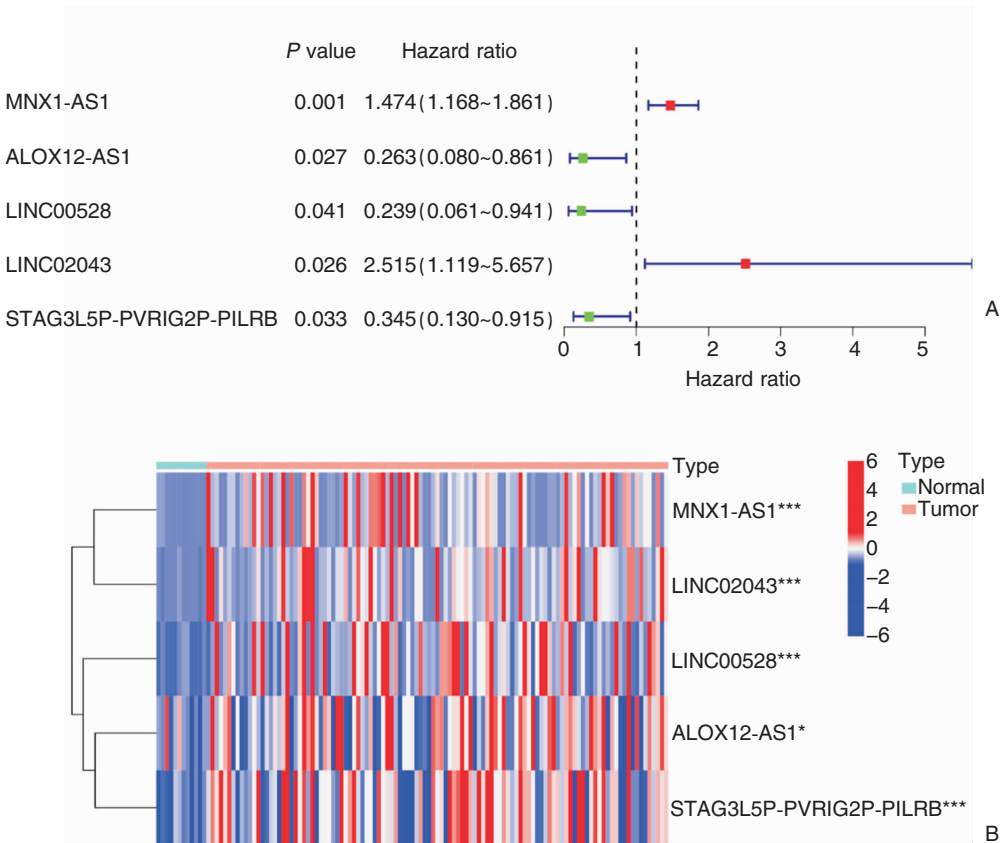


图 2 5 个 LSCC 预后相关的 m6A 相关 lncRNA 表达情况 A:单变量 Cox 回归分析森林图; B:LSCC 预后相关的 m6A 相关 lncRNA 癌旁组织与肿瘤组织表达水平热图 注:低表达用蓝色表示,蓝色越深表示表达量越低,高表达用红色表示,红色越深表示表达量越高( $P < 0.001$  \*\*\*,  $P < 0.05$  \*);LSCC(喉鳞状细胞癌)。下同。

### 2.3 预后模型

为了尽量避免因未纳入重要自变量而出现的模型偏差, 5 个预后相关 lncRNA 均纳入模型构建, 同时为了防止模型构建出现过度拟合的情况, 模型构建运用套索回归构造一个惩罚函数, 压缩变量数目, 以期得到一个较为精炼的模型。套索回归及交叉验证迭代分析结果显示变量个数为 4 时, 模型的均方根误差最小(图 3), 即 ALOX12-AS1、LINC00528、STAG3L5P-PVRIG2P-PILRB 与 MNX1-AS1 成为预后模型的 4 个变量。LSCC 预后预测风险评分模型: 风险评分 = (0.176723096228585 × MNX1-AS1 表达量) + (-0.614916717648596 × ALOX12-AS1 表达量) + (-0.814201385798827 × LINC00528 表达量) + (-0.436537752110547 × STAG3L5P-PVRIG2P-PILRB 表达量)。

### 2.4 qPCR 结果

通过 qPCR 结果验证了 TCGA 转录表达谱数据测算结果可信, ALOX12-AS1 ( $t = 6.077, P < 0.001$ )、LINC00528 ( $t = 8.707, P < 0.0001$ )、STAG3L5P-PVRIG2P-PILRB ( $t = 8.707, P < 0.0001$ ) 和 MNX1-AS1 ( $t = 3.079, P < 0.05$ ) 在 LSCC 中的表达量高于癌旁组织, 结果具有统计学意义。

### 2.5 预后模型的综合分析

预后模型计算 112 例 LSCC 病例的风险评分, 模型训练组的中位评分(风险评分 = 0.783)为截断值, 以模型训练组风险值中位数作为划分高风险患者与低风险患者的临界值, 将模型训练组、模型测试组患者分为高风险组与低风险组(图 4)。

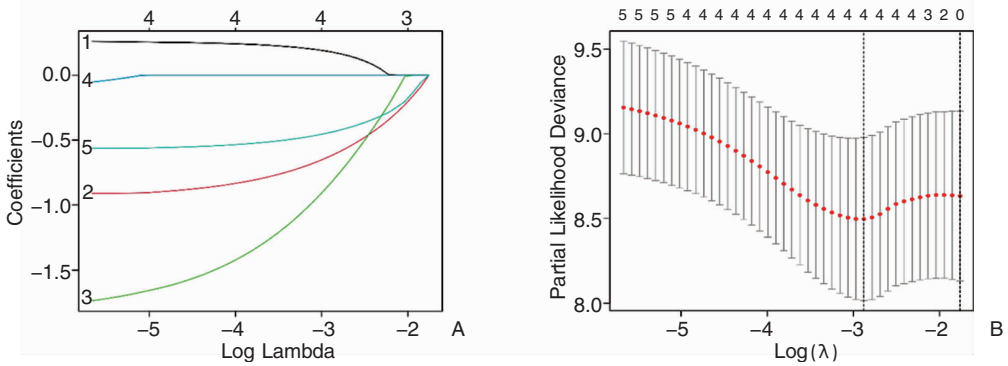


图 3 套索回归筛选变量动态过程图(A)与参数 λ 选择图(B)

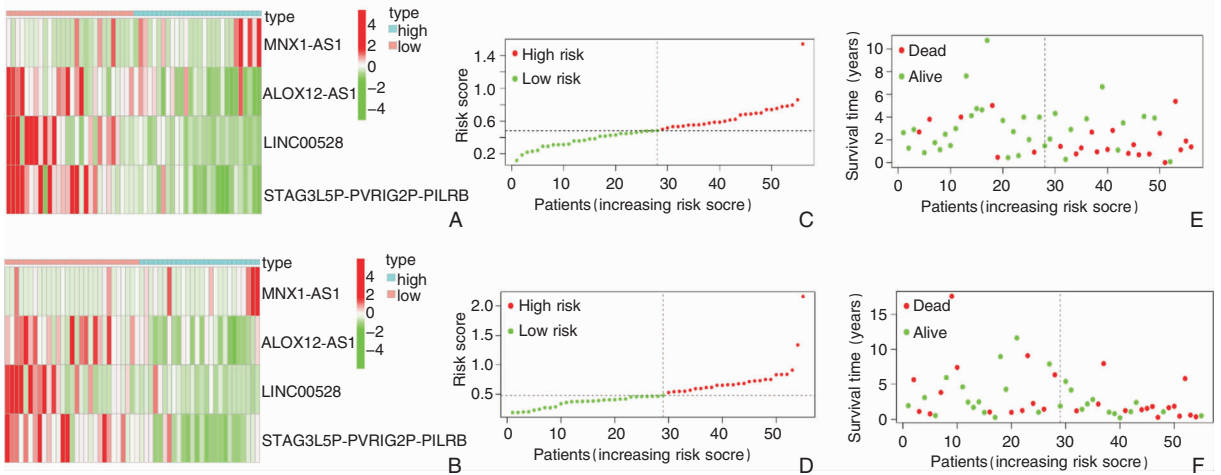


图 4 模型构建 lncRNA 热图, ALOX12-AS1、LINC00528 和 STAG3L5P-PVRIG2P-PILRB, 在低风险患者中高表达, 而 MNX1-AS1 在高风险患者中高表达; 风险评分 - 病例散点图(C、D), 风险评分 - 生存时间分布图(E、F), 散点图的每一个点代表 1 例患者, 红色散点为死亡病例, 绿色散点为仍存活病例, 纵坐标为生存时间, 横坐标为病例数, 低风险患者、高风险患者分别位于虚线的左侧和右侧。图 E 左侧(6 例)红色散点明显少于右侧(16 例), 低风险患者死亡病例明显少于高风险患者, 图 F 左(13 例)、右(14 例)两侧代表死亡病例的红色散点数量相近, 左侧散点位置较高, 低风险患者死亡病例的生存时间分布长于高风险组 注: A、C、E 为模型训练组; B、D、F 为模型测试组。

模型训练组 Kaplan-Meier 曲线,低风险患者的生存率高于高风险患者( $P < 0.01$ );模型测试组 Kaplan-Meier 曲线,低风险患者的生存率高于高风险患者( $P < 0.05$ )。模型训练组与模型测试组病例的 AUC 分别为 0.705、0.720,均大于 0.7,说明模型预测结果有效(图 5)。

模型训练组风险评分:单因素 Cox 独立预后因

素分析( $HR = 41.447, P = 0.017$ )和多因素独立预后分析( $HR = 61.811, P = 0.023$ )。见图 6A、B。模型测试组风险评分:单因素 Cox 独立预后因素分析( $HR = 6.951, P = 0.005$ )、多因素独立预后分析( $HR = 5.093, P = 0.045$ )。见图 6C、D。以上结果说明风险评分可以独立于其他的临床特征作为独立的预后因子。

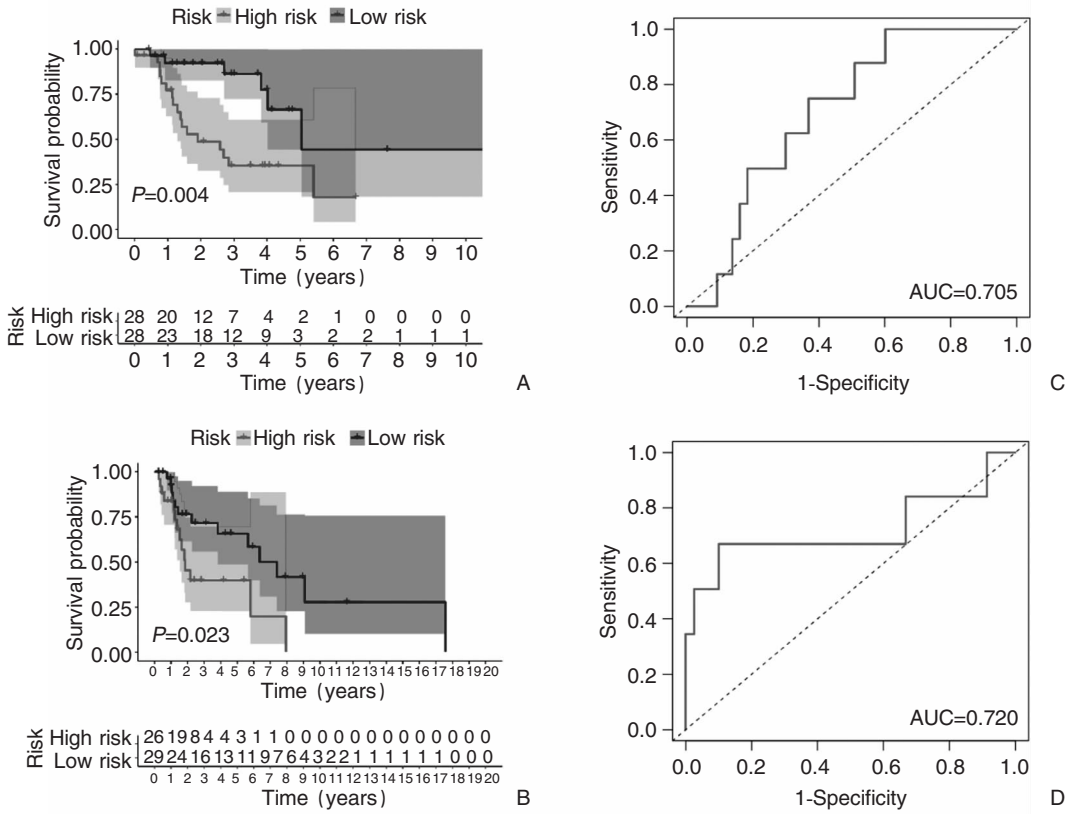


图 5 Kaplan-Meier 生存曲线(A、B),在模型训练组与模型测试组中,低风险患者的生存率均高于高风险患者 A:模型训练组( $P = 0.004$ );B:模型测试组( $P = 0.023$ ),ROC 曲线(C、D); C:模型训练组(AUC = 0.705); D:模型测试组(AUC = 0.720)

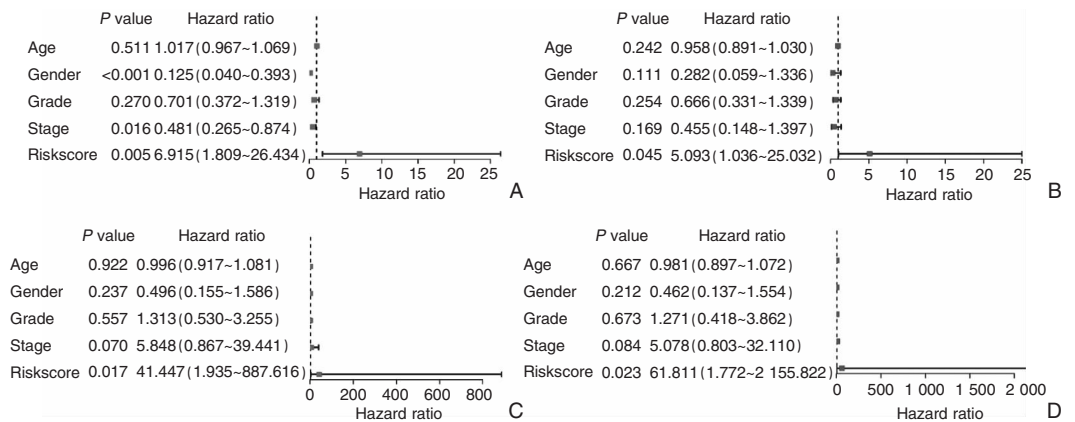


图 6 风险评分与 LSCC 病例临床性状独立预后分析 A、B:模型训练组; C、D:模型测试组

### 2.6 GSEA 富集分析结果

通过 GSEA 富集分析结果显示:高风险患者转录表达谱在细胞 - 基质粘附 (KEGG\_FOCAL\_ADHESION) ( $P < 0.05$ ) 与细胞外基质受体相互作用 (KEGG\_ECM\_RECEPTOR\_INTERACTION) ( $P < 0.05$ ) 通路呈下调趋势。低风险患者转录表达谱在碱基切除修复 (KEGG\_SPLICEOSOME) ( $P < 0.05$ )、剪接体 (KEGG\_SPLICEOSOME) ( $P < 0.05$ ) 通路呈

上调趋势 (图 7)。

### 2.7 免疫相关性

通过免疫浸润分析我们得到了 LSCC 样本中各类免疫细胞的含量 (图 8), 每一类免疫细胞均与风险评分进行相关性检验, 结果显示风险评分与效应 B 细胞 ( $R = -0.24, P = 0.017$ )、细胞毒性 T 细胞 ( $R = -0.26, P = 0.0091$ ) 和 T 滤泡辅助细胞 ( $R = -0.22, P = 0.028$ ) 呈负相关 (图 9)。

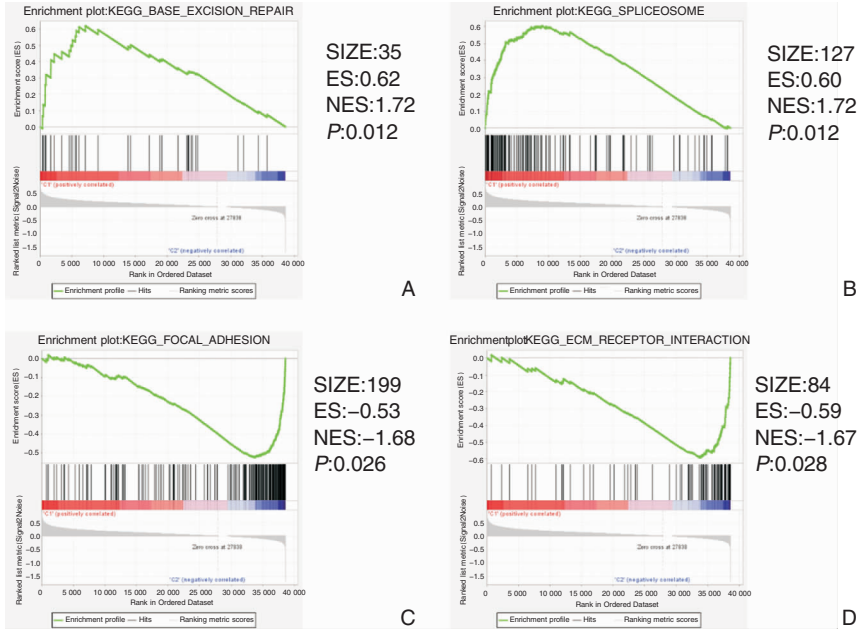


图 7 GSEA 富集分析结果 SIZE:富集到该通路的基因数;ES:富集得分,ES 为正数则该通路呈上调趋势,为负数则为下调趋势;NES:校正后的归一化的 ES 值,其绝对值大于 1 为一条富集标准;A、B:低风险患者富集分析结果;C、D:高风险患者富集结果

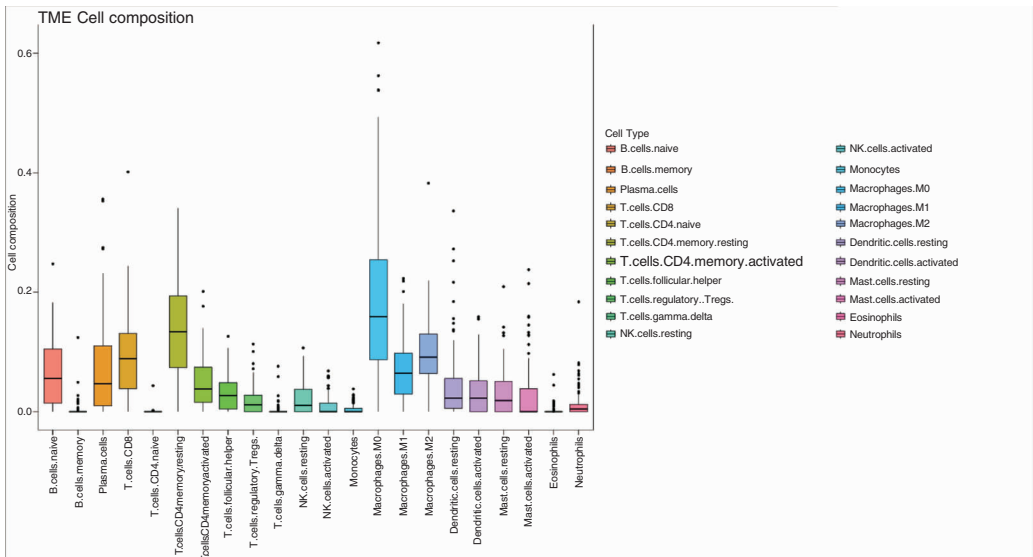


图 8 112 例 LSCC 浸润免疫细胞浸润总体情况



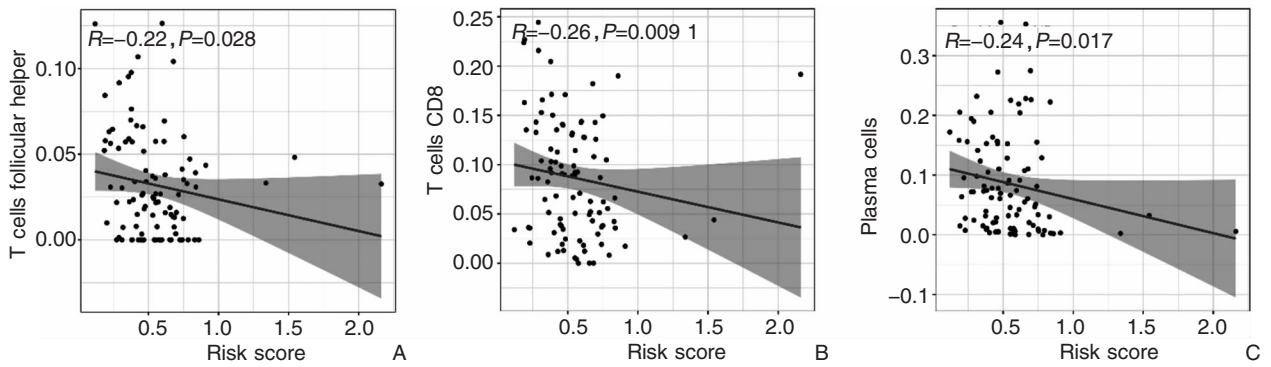


图 9 LSCC 浸润免疫细胞 - 风险评分相关性 A: T 滤泡辅助细胞与风险评分相关性图; B: 细胞毒性 T 细胞与风险评分相关性图; C: 效应 B 细胞与风险评分相关性图

### 3 讨论

m6A 修饰是位于腺嘌呤第 6 位氮原子上的甲基化修饰,它是真核生物 RNA 最常见的一种表观修饰方式,存在包括 mRNA、微小 RNA、环状 RNA 在内的不同类型的 RNA 上<sup>[16]</sup>。m6A 修饰发挥作用依赖甲基转移酶、去甲基化酶和结合蛋白的作用,甲基转移酶能在 RNA 上加入甲基,结合蛋白可识别 m6A 修饰的 RNA,去甲基化酶则调控这一过程的发生与否,m6A 修饰通过这种方式参与了包括 RNA 加工、核输出、翻译、衰变、可变剪切等 RNA 调控过程<sup>[17]</sup>。lncRNAs 同样具有 m6A 修饰,并且参与某些肿瘤的自我更新与分化等过程,调控多种肿瘤的发生、发展和转移。在胶质母细胞瘤干细胞中,m6A 去甲基酶 ALKBH5 通过对 lncRNA FOXM1 新生转录本的 3' UTR 去甲基化来诱导更高水平的 FOXM1 转录,而这促进了胶质母细胞瘤的增殖和致癌性<sup>[18]</sup>。lncRNA MALAT1 是人们最初研究肺癌时发现,是最早发现的癌症相关的 lncRNA 之一,研究证实 MALAT1 参与包括肺癌、肝细胞癌、前列腺癌和胆囊癌等多种人类肿瘤的表达和调控过程<sup>[19]</sup>。而 MALAT1 这种在细胞核中高度表达的 lncRNA,其在细胞亚结构中的定位和活性,以及和蛋白结合的能力受到 m6A 修饰的调控<sup>[20]</sup>。目前的这些研究显示 m6A 修饰的 lncRNA 广泛出现在各种类型肿瘤的调控过程中。

lncRNA 在 LSCC 的发生发展中同样发挥着重要作用,Li 等<sup>[21]</sup>检测了 72 例 LSCC 患者癌组织和癌旁组织中 HOTAIR 的表达水平,发现 lncRNA HOTAIR 在癌组织中表达水平显著高于癌旁组织(16 倍),并且其表达水平与临床分期、肿瘤分化程度、

颈部淋巴结转移等因素有统计学相关性。Chen 等<sup>[22]</sup>对 76 例 LSCC 癌组织和癌旁组织,以及 28 例 LSCC 淋巴结转移组织进行 qPCR 检测,与早期 LSCC 相比,lncRNA AC008440.10 在晚期 LSCC 和存在淋巴转移的 LSCC 患者中的表达水平显著下降。最近,Li 等<sup>[23]</sup>发现通过去甲基酶 ALKBH5 介导 lncRNA KCNQ1OT1 的 m6A 修饰会导致癌基因 HOXA9 上调,诱导 LSCC 的发生。因此,探索更多 LSCC 预后相关的 m6A 修饰 lncRNA,可能为 LSCC 的诊疗提供新思路。

本研究对 112 例来自 TCGA 的 LSCC 病例转录表达谱中的 m6A 相关的基因(甲基转移酶、去甲基化酶和结合蛋白相应基因)与 lncRNA 进行了相关性分析,发现 169 个与 m6A 相关基因相关系数  $> 0.5$ ,  $P < 0.05$  的 lncRNA,这些与 m6A 有着关联的 lncRNA 中不乏与多个 m6A 相关基因有关联的 lncRNA,说明 LSCC 的 lncRNA 中存在广泛的 m6A 修饰,而且这种修饰受到多个 m6A 相关调控因子的参与。结合 LSCC 患者的生存时间、生存状态和这些 m6A 修饰相关的 lncRNA 表达量进行了单因素 Cox 分析,获得了与 LSCC 预后相关联的 5 个 lncRNA: ALOX12-AS1、LINC00528、STAG3L5P-PVRIG2P-PILRB、MNX1-AS1 和 LINC02043。5 个预后相关的 lncRNA 中 ALOX12-AS1、LINC00528、STAG3L5P-PVRIG2P-PILRB 的 HR  $< 1$ ,MNX1-AS1 和 LINC02043 的 HR  $> 1$ ,因此 lncRNA 的 m6A 修饰既可能起促进 LSCC 发生发展,也可能起抑制 LSCC 的作用,这与计锐等<sup>[24]</sup>在 m6A 甲基化修饰与泌尿系统肿瘤关系的回顾中观察到的结果相似。基于 5 个预后相关 lncRNA,通过拉索回归构建惩罚函数,构建 LSCC 预后风险评分模型: 风险评分 =  $(0.176723096228585 \times \text{MNX1-AS1 表达量}) + (-0.614916717648596 \times$

ALOX12-AS1 表达量) + (- 0.814201385798827 × LINC00528 表达量) + (- 0.436537752110547 × STAG3L5P-PVRIG2P-PILRB 表达量)。通过 qPCR 技术,对最终用于构建模型的 4 个 lncRNA 在 LSCC 组织与癌旁组织的表达量进行检验,ALOX12-AS1、LINC00528、STAG3L5P-PVRIG2P-PILRB 和 MNX1-AS1 在 LSCC 中的表达量高于癌旁组织,与依靠 TCGA 数据库 LSCC 转录表达谱数据分析正常组织与 LSCC 组织差异结果一致,证实了 TCGA 转录表达谱数据的可靠性。预后风险评分模型计算 112 例 LSCC 的风险评分,依据风险评分将患者分为了高风险患者和低风险患者,Kaplan-Meier 曲线分析生存率,高风险患者与低风险患者之间的生存率差异具有统计学意义,低风险患者的生存率高于高风险患者,且无论在模型训练组还是在模型测试组,AUC 均大于 0.7。由此可见,通过预后模型计算的风险评分能有效可靠的预测患者的预后。通过风险评分与患者临床特征(年龄、性别、病理等级、分期)的单因素 Cox 分析、多因素 Cox 分析结果可以得知,风险评分可以视为一项独立评估的预后因素。

为了进一步了解 m6A 修饰的 lncRNA 在 LSCC 中可能通过哪些调控通路影响 LSCC 的发生发展,应用 GSEA v4.1.0 对高风险患者和低风险患者进行差异基因富集分析。GSEA 富集分析检测基因集合而不是单个基因的表达变化,因此可以包含细微的表达变化,较传统提取高表达基因或低表达基因进行富集分析,能得到更为理想、可信的富集结果<sup>[25]</sup>。GSEA 富集结果:高风险患者在细胞-基质粘附(KEGG\_FOCAL\_ADHESION)、细胞外基质受体相互作用(KEGG\_ECM\_RECEPTOR\_INTERACTION)两条通路下调。低风险患者则在碱基切除修复(KEGG\_BASE\_EXCISION\_REPAIR)、剪接体(KEGG\_SPLICEOSOME)通路上调。富集分析结果提示 m6A 修饰的 lncRNA 可能参与了 LSCC 碱基切除修复 DNA 过程,以及与远处转移有关的细胞-基质粘附、细胞外基质受体相互作用等调控通路。

肿瘤浸润免疫细胞是指浸润在肿瘤组织中的包括 T 细胞、B 细胞、树突状细胞、NK 细胞和 NKT 细胞等在内的淋巴细胞群,是肿瘤微环境的重要组成部分<sup>[26]</sup>。分析肿瘤浸润免疫细胞在肿瘤组织中的含量,可能成为评估肿瘤预后的重要手段。郝晓龙等<sup>[27]</sup>在回顾了多篇有关 LSCC 与肿瘤浸润免疫细胞的研究后认为,肿瘤浸润 T 淋巴细胞可以作为 LSCC 预后评估指标,不同类别的肿瘤浸润 T 淋巴

细胞在 LSCC 中具有不同的预后价值。风险评分与 LSCC 浸润免疫细胞相关性分析结果显示,风险值与细胞毒性 T 细胞、效应 B 细胞、T 滤泡辅助细胞 3 种浸润免疫细胞呈负相关。从概念上讲,由突变的癌细胞产生的“非自身”新抗原可以触发其自身被 T 细胞消除,这意味着更高的突变负荷可能导致更强的抗肿瘤免疫反应<sup>[26]</sup>。Zhou 和 Keck 等的研究发现,LSCC 中细胞毒性 T 细胞浸润程度较高的患者拥有更好的预后,浸润的密度与肿瘤分期、淋巴结分期和复发之间呈负相关<sup>[28-29]</sup>。效应 B 细胞是免疫系统中释放大量抗体的细胞。T 滤泡辅助细胞对 B 细胞在生发中心的存活和成熟至关重要<sup>[24]</sup>。与风险评分的负相关,可能意味着在肿瘤组织中浆细胞与 T 滤泡辅助细胞协同下释放肿瘤抗体,对 LSCC 细胞产生杀伤作用从而起到抑制肿瘤的作用。结合肿瘤浸润免疫细胞相关性分析,我们可以认为风险评分结合肿瘤组织中的细胞毒性 T 细胞、效应 B 细胞、T 滤泡辅助细胞含量对 LSCC 预后评估有更高的准确性。

综上所述,通过生物信息学的方法,本研通过与 LSCC 预后相关的 m6A 相关联 lncRNA 构建预后风险评分模型,用 Kaplan-Meier 曲线及 ROC 曲线验证了模型的可靠性。并从通路调控及肿瘤浸润免疫细胞两个角度对模型进行了进一步的挖掘分析,一定程度上说明了 m6A 修饰的 lncRNA 在 LSCC 发生发展的潜在作用,但具体的生物功能及机制,有待进一步研究。

#### 参考文献:

- [1] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015[J]. CA Cancer J Clin, 2016, 66(2):115-132.
- [2] Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries [J]. CA Cancer J Clin, 2021, 71(3):209-249.
- [3] Kaidar-Person O, Gil Z, Billan S. Precision medicine in head and neck cancer[J]. Drug Resist Updat, 2018, 40:13-16.
- [4] Xu J, Bai J, Zhang X, et al. A comprehensive overview of lncRNA annotation resources [J]. Brief Bioinform, 2017, 18(2):236-249.
- [5] Pacholewska A, Sung MH. lncRNA expression predicts mRNA abundance[J]. Epigenomics, 2019, 11(10):1121-1128.
- [6] Bhan A, Mandal SS. lncRNA HOTAIR: A master regulator of chromatin dynamics and cancer[J]. Biochim Biophys Acta, 2015, 1856(1):151-164.
- [7] Chen F, Li Z, Deng C, et al. Integration analysis for novel lnc

- cRNA markers predicting tumor recurrence in human colon adenocarcinoma[J]. *J Transl Med*,2019,17(1):299.
- [8] Lan T, Li H, Zhang D, et al. KIAA1429 contributes to liver cancer progression through N6-methyladenosine-dependent post-transcriptional modification of GATA3[J]. *Mol Cancer*,2019,18(1):186.
- [9] Li R, Yang YE, Jin J, et al. Identification of lncRNA biomarkers in lung squamous cell carcinoma using comprehensive analysis of lncRNA mediated ceRNA network[J]. *Artif Cells Nanomed Biotechnol*,2019,47(1):3246–3258.
- [10] Zhang L, Hou C, Chen C, et al. The role of N6-methyladenosine (m6A) modification in the regulation of circRNAs[J]. *Mol Cancer*,2020,19(1):105.
- [11] Panneerross S, Eedunuri VK, Yadav P, et al. Cross-talk among writers, readers, and erasers of m6A regulates cancer growth and progression[J]. *Sci Adv*,2018,4(10):eaar8263.
- [12] Wang Z, Jensen MA, Zenklusen JC. A practical guide to The Cancer Genome Atlas (TCGA) [J]. *Methods Mol Biol*,2016,1418:111–141.
- [13] Edens BM, Vissers C, Su J, et al. FMRP modulates neural differentiation through m6A-dependent mRNA nuclear export[J]. *Cell Rep*,2019,28(4):845–854. e5.
- [14] Tu Z, Wu L, Wang P, et al. N6-methyladenosine-related lncRNAs are potential biomarkers for predicting the overall survival of lower-grade glioma patients [J]. *Front Cell Dev Biol*,2020,8:642.
- [15] Linden A. Measuring diagnostic and predictive accuracy in disease management; an introduction to receiver operating characteristic (ROC) analysis[J]. *J Eval Clin Pract*,2006,12(2):132–139.
- [16] Rasmussen BK, Hansen S, Laursen RJ, et al. Epidemiology of glioma; clinical characteristics, symptoms, and predictors of glioma patients grade I-IV in the the Danish Neuro-Oncology Registry [J]. *J Neurooncol*,2017,135(3):571–579.
- [17] Chen XY, Zhang J, Zhu JS. The role of m6A RNA methylation in human cancer[J]. *Mol Cancer*,2019,18(1):103.
- [18] Zhang S, Zhao BS, Zhou A, et al. m6A Demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program[J]. *Cancer Cell*,2017,31(4):591–606. e6.
- [19] Li CH, Chen Y. Targeting long non-coding RNAs in cancers: progress and prospects[J]. *Int J Biochem Cell Biol*,2013,45(8):1895–1910.
- [20] Zhou KI, Parisien M, Dai Q, et al. N(6)-methyladenosine modification in a long noncoding RNA hairpin predisposes its conformation to protein binding[J]. *J Mol Biol*,2016,428(5 Pt A):822–833.
- [21] Li D, Feng J, Wu T, et al. Long intergenic noncoding RNA HO-TAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma[J]. *Am J Pathol*,2013,182(1):64–70.
- [22] Chen J, Shen Z, Deng H, et al. Long non-coding RNA biomarker for human laryngeal squamous cell carcinoma prognosis [J]. *Gene*,2018,671:96–102.
- [23] Li Y, Yan B, Wang X, et al. ALKBH5-mediated m6A modification of lncRNA KCNQ10T1 triggers the development of LSCC via upregulation of HOXA9[J]. *J Cell Mol Med*,2022,26(2):385–398.
- [24] 计锐,郑君芳. m6A 甲基化修饰与泌尿系统肿瘤关系的研究进展[J]. *中南医学科学杂志*,2022,50(1):141–144,156.
- [25] Subramanian A, Kuehn H, Gould J, et al. GSEA-P: a desktop application for Gene Set Enrichment Analysis[J]. *Bioinformatics*,2007,23(23):3251–3253.
- [26] Jia Q, Wu W, Wang Y, et al. Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer [J]. *Nat Commun*,2018,9(1):5361.
- [27] 郝晓龙,王斌全. 肿瘤浸润 T 淋巴细胞对喉鳞状细胞癌预后评估作用的研究进展[J]. *中国耳鼻咽喉颅底外科杂志*,2021,27(5):534–539.
- [28] Zhou L, Li Y, Gao W, et al. Assessment of tumor-associated immune cells in laryngeal squamous cell carcinoma [J]. *J Cancer Res Clin Oncol*,2019,145(7):1761–1772.
- [29] Keck MK, Zuo Z, Khattri A, et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes[J]. *Clin Cancer Res*,2015,21(4):870–881.

(收稿日期:2021-12-28;网络首发:2022-05-20)

**本文引用格式:**李念燊,韩杏倩,何金阳,等. 基于 N6-甲基腺苷相关长链非编码核糖核酸表达的喉鳞状细胞癌预后分析[J]. *中国耳鼻咽喉颅底外科杂志*,2023,29(1):81–91. DOI:10.11798/j.issn.1007-1520.202321515

**Cite this article as:**LI Nianshen, HAN Xingqian, HE Jinyang, et al. Prognostic analysis of laryngeal squamous cell carcinoma based on the expression of N6-methyladenosine-related lncRNAs[J]. *Chin J Otorhinolaryngol Skull Base Surg*,2023,29(1):81–91. DOI:10.11798/j.issn.1007-1520.202321515