

生物信息分析验证基于lncRNA标签预测HPV阴性头颈部鳞状细胞癌患者总生存率

黄宇¹, 杨博², 李伟忠¹

1. 南方医科大学南方医院口腔颌面外科, 广东 广州 510000; 2. 南方医科大学南方医院口腔综合科, 广东 广州 510000

【摘要】目的:采用生物信息分析方法研究影响HPV阴性头颈部鳞状细胞癌(HNSCC)患者总生存率的相关因素,为HNSCC患者的预后预测及治疗提供依据。**方法:**从TCGA数据库获取HPV阴性HNSCC患者的病变组织和癌旁组织的原始RNA测序数据以及相应病人的临床信息, Survival R工具包对所有差异表达的长链非编码RNA(lncRNA)进行单变量Cox回归分析,使用JAVA程序,利用MSigDB C2经典通路基因集集合进行基因集富集分析,采用LASSO方法的Cox回归对数据降维和模型构建,计算每个患者的风险评分。**结果:**研究发现153个lncRNA与HPV阴性HNSCC患者的总生存率显著相关,时间依赖性分析表明,13个lncRNA特征在训练集中具有良好的预测性能,多变量Cox回归和分层分析表明,基于13-lncRNA标签的风险评分可作为HPV阴性HNSCC患者的独立预后因素。**结论:**13-lncRNA标签可能是HPV阴性HNSCC预后的新型独立生物标志物。

【关键词】头颈部鳞状细胞癌;长链非编码RNA;人乳头瘤病毒;生存率;生物信息分析

【中图分类号】R318;R739.8

【文献标志码】A

【文章编号】1005-202X(2025)02-0261-08

Bioinformatic analysis validates lncRNA signature as a predictor of overall survival in patients with HPV-negative head and neck squamous cell carcinoma

HUANG Yu¹, YANG Bo², LI Weizhong¹

1. Department of Oral and Maxillofacial Surgery, Nanfang Hospital, Southern Medical University, Guangzhou 510000, China; 2. Department of Stomatology, Nanfang Hospital, Southern Medical University, Guangzhou 510000, China

Abstract: Objective To investigate the factors affecting the overall survival in patients with HPV-negative head and neck squamous cell carcinoma (HNSCC) by bioinformatics analysis, thereby providing evidence for the prognosis prediction and therapy in HNSCC. **Methods** The raw RNA sequencing data of the primary tumor and adjacent tissue of HPV-negative HNSCC patients were obtained from the TCGA database; and the corresponding clinical information were collected. Survival R toolkit was used to perform univariate Cox regression analysis on all differentially expressed lncRNA; a Java program was used to carry out gene set enrichment analysis using the MSigDB C2 classic pathway gene set collection; and LASSO Cox regression was used for data dimensionality reduction and model building, and the risk score for each patient was calculated. **Results** There were 153 lncRNA which were significantly associated with overall survival in HPV-negative HNSCC patients. Time-dependent analysis showed that 13 lncRNA had good predictive performance in the training set. Multivariate Cox regression and stratified analysis revealed that the risk score based on the 13-lncRNA signature could be served as an independent prognostic factor for HPV-negative HNSCC patients. **Conclusion** The 13-lncRNA signature may be a novel independent biomarker for the prognosis of HPV-negative HNSCC.

Keywords: head and neck squamous cell carcinoma; long non-coding RNA; human papillomavirus; survival; bioinformatics analysis

【收稿日期】2024-10-26

【基金项目】广东省自然科学基金(2015A030313309)

【作者简介】黄宇, 硕士, 研究方向: 口腔颌面外科学, E-mail: huangyukq@smu.edu.cn

【通信作者】李伟忠, 教授, 研究方向: 口腔颌面外科学, E-mail: liwz@smu.edu.cn

前言

头颈部鳞状细胞癌(Head and Neck Squamous Cell Carcinoma, HNSCC)是常见恶性肿瘤,约70%为HPV阴性^[1],研究发现HNSCC患者中HPV阴性与HPV阳性患者相比不但预后差,其治疗反应和生存率也比HPV阳性患者差^[2-3],因而促使人们探求用于

风险分层的预后标记物及新的治疗靶点。长链非编码 RNA (long non-coding RNA, lncRNA) 是一类长度超过 200 个核苷酸的非蛋白质编码 RNA, 广泛分布于基因组中, 在转录、转录后和表观遗传水平上发挥重要作用, 并参与转录调控、细胞生长和肿瘤发生等各种生物学过程和途径^[4-11], 这提示它们作为癌症预后的新型独立生物标志物的潜在作用。本研究通过生物信息学分析的方法, 对肿瘤基因组图谱 (The Cancer Genome Atlas, TCGA) 数据库中 HPV 阴性 HNSCC 患者的 lncRNA 表达谱及相应的临床信息进行分析, 鉴定与患者生存显著相关的 lncRNA, 构建一个基于 13 个 lncRNA (13-lncRNA) 标签, 为分析 HPV 阴性的 HNSCC 患者预后提供依据。

1 方法

1.1 数据来源与预处理

从 TCGA 数据中获取 359 个 HPV 阴性 HNSCC 患者的病变组织和 36 个相邻正常组织的原始 RNA 测序 (RNA-seq) 以及相应病人的临床信息。通过 RNASeqV2 后处理流程从 RNA 测序原始数据生成 lncRNA 表达谱。使用 edgeR 的 R/Bioconductor 工具包筛选 359 个 HPV 阴性的 HNSCC 组织和 36 个相邻正常组织之间的差异表达 lncRNA (Differentially Expressed lncRNAs, DELs)^[12], $|\log_2$ 倍数变化 ≥ 1 且 $P < 0.05$ 被设定为阈值, 经过 edgeR 归一化后, lncRNA 的表达谱进一步转化为 \log_2 (归一化值为 1) 变换, 用于下一步操作, 本研究遵守 TCGA 出版指南和数据访问政策, 无需当地伦理委员会的批准。

1.2 与预后相关的关键 lncRNA 的鉴别和筛选

使用 Survival R 工具包对所有 DELs 进行单变量 Cox 回归分析^[13], $P < 0.05$ 的 lncRNA 被定义为靶向 lncRNA。本研究采用稳健的基于可能性的生存建模方法, 使用 rbsurv R 包鉴别影响 HPV 阴性 HNSCC 患者预后的关键 lncRNA^[14]。(1) 随机分配一个包含 $N \cdot (1-P)$ 个样本的训练集, 或一个包含 $N \cdot P$ 个样本的测试集 ($P=1/3$)。将一个 lncRNA 拟合到样本训练集, 并获得了该 lncRNA 的参数估计值, 然后用参数估计和样本测试集评估对数可能值。对每个 lncRNA 重复此评估。(2) 上述步骤独立重复 10 次, 从而获得每个 lncRNA 的 10 个对数可能值。选择对数可能值平均值最大的 lncRNA, 即 $g(1)$ 。(3) 通过评估每两个 lncRNA 模型来寻找下一个最佳 lncRNA, 并选择具有最大平均对数可能值的最佳模型。(4) 继续此前 lncRNA 选择程序, 从而得到一系列预测模型, 计算所有候选预测模型的赤池信息量准则 (Akaike Information Criteria, AIC), 并选择具有最低 AIC 的最优模型。

1.3 基因集合富集度分析

使用 JAVA 程序, 利用 MSigDB C2 经典通路基因集集合 (包含 1 329 个基因集^[15]) 进行基因集合富集度分析 (Gene Set Enrichment Analysis, GSEA), 具有正富集得分 (或负富集得分) 且 $P < 0.05$ 的基因集被视为显著富集的基因集, 其中大多数基因上调, 并伴有高风险或低风险得分。

1.4 统计学分析

采用 LASSO 方法的 Cox 回归对数据降维和模型构建, 通过 10 倍交叉验证确定参数 λ 的最优值, 筛选出具有相关系数的预后 lncRNA, 随后根据训练数据集中每个预后 lncRNA 的表达水平及其相关系数计算每个患者的风险评分^[16]。根据风险评分公式, 以风险评分中值为界值, 将 HPV 阴性 HNSCC 患者分为高风险组和低风险组, 通过训练集中的 Kaplan-Meier 估计量和对数秩检验评估低风险组和高风险组之间的差异。随后, 为了进一步确定 lncRNA 特征是否独立于其他临床变量, 进行了单变量、多变量 Cox 回归和分层分析。绘制 5 年内时间依赖性受试者工作特征 (Receiver Operating Characteristic, ROC) 曲线, 比较基于 13 个 lncRNA (13-lncRNA) 风险评分的生存预测的灵敏度和特异性, $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 训练集中与患者总体生存率相关的 lncRNA 差异表达

对 359 例 HPV 阴性 HNSCC 组织和 36 例癌旁正常组织样品的 14461 个 lncRNA 的原始 3 级 RNAseq 数据进行 DELs 分析, edgeR 鉴定出 2 281 个 DELs, 其中 762 个 lncRNA 上调表达, 1 519 个 lncRNA 下调表达 (图 1)。进一步分析 DELs 与 HPV 阴性 HNSCC 患者总体生存率之间的相关性, 将共 359 个样本进一步随机分为训练集和测试集 (表 1), 训练集中的单变量 Cox 回归分析确定了 153 个 lncRNA ($P < 0.05$) 为靶向 lncRNA, 随后对 153 个目标 lncRNA 进行稳健的可能性生存分析, 筛选出 19 个与预后相关的 lncRNA。

2.2 训练集中的 13-lncRNA 特征和患者生存率

应用 LASSO Cox 回归模型对训练集中 19 个与预后相关 lncRNA 进行分析, 根据对数 (λ) 序列生成系数剖面图, 在使用 10 倍交叉验证选定的值处绘制垂直线, 最小化 λ 方法得到 13 个 lncRNA (图 2); 最后根据 13 个 lncRNA 的表达水平和 LASSO Cox 回归系数计算训练集中每个患者的风险评分: 风险评分 = $(0.121\ 4 \times \text{RP11-366H4.1 表达值}) + (-0.137\ 5 \times \text{RP11-70C1.1 表达值}) + (0.118\ 9 \times \text{RP11-367J11.3 表达值}) + (0.315\ 5 \times \text{RP3-329E20.2 表达值}) + (-0.146\ 4 \times \text{RP11-567M16.1 表达值}) + (0.146\ 6 \times \text{KDM4A-AS1 表达值}) +$

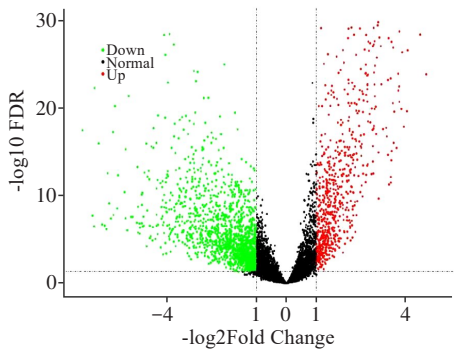


图1 火山图展示 2281 个差异表达 lncRNA

Figure 1 Volcano plot showing 2281 differentially expressed lncRNA

绿色代表上调表达,红色代表下调表达

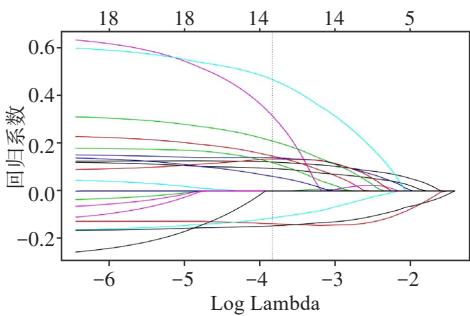


图2 LASSO 回归获得 13 个 lncRNA 的回归系数

Figure 2 LASSO regression obtained the regression coefficients of 13 lncRNA

竖线表示 10 倍交叉验证获得的最小λ值

表 1 训练集和测试集基本资料[例(%)]
Table 1 Basic information of training set and testing set [cases (%)]

| 指标 | TCGA 训练集 | TCGA 测试集 | P 值 |
|--------|------------|------------|-------|
| 年龄/岁 | | | 1.000 |
| <60 | 73(40.56) | 73(40.78) | |
| ≥60 | 107(59.44) | 106(59.22) | |
| 性别 | | | |
| 男 | 127(70.56) | 127(70.95) | |
| 女 | 53(29.44) | 52(29.05) | |
| 吸烟历史 | | | 1.000 |
| 是 | 136(79.53) | 142(79.33) | |
| 否 | 35(20.47) | 37(20.67) | |
| 饮酒历史指标 | | | 0.673 |
| 是 | 109(63.01) | 117(65.73) | |
| 否 | 63(36.99) | 61(34.27) | |
| 肿瘤病理分期 | | | 0.676 |
| I | 8(5.16) | 13(7.69) | |
| II | 29(18.71) | 26(15.38) | |
| III | 28(18.06) | 34(20.12) | |
| IV | 90(58.06) | 96(56.8) | |
| 评估 | | | 0.935 |
| 高分化组 | 26(14.77) | 28(16.09) | |
| 中分化组 | 112(63.64) | 108(62.07) | |
| 低分化组 | 38(21.59) | 38(21.84) | |

(0.139 1×RP11-467L19.16 表达值)+(0.113 3×OVAAL 表达值)+(0.094 3×RP11-25I15.3 表达值)+(0.133 7×DLEU7-AS1 表达值)+(0.209 6×RP11-401O9.3 表达值)+(0.061 3×LINC01232 表达值)+(0.467 3×RP11-250B2.6 表达值)。

计算训练集中每个 HPV 阴性 HNSCC 患者基于 lncRNA 的风险评分,并使用中位风险评分作为截止值,

将其分为高风险组($n=90$)和低风险组($n=90$),结果显示高风险组患者的预后比低风险组患者更差。为了评估 13-lncRNA 特征的预后能力,进行了时间依赖性 ROC 曲线分析,13-lncRNA 特征预后模型的曲线下面积(Area Under Curve, AUC)在预测训练集 3 年和 5 年总生存率时分别为 0.82 和 0.77(图 3)。训练集中 13-lncRNA 表达风险评分、总体生存率及表达谱分布如图 4 所示,高风险患者的死亡率高于低风险患者。对于具有高风险的患者, lncRNA (RP11-70C1.1、RP11-567M16.1 和 OVAAL)的表达谱显著下调,其余 10 个 lncRNA(RP11-467L19.16、RP11-11-401O9.3、RP11-366H4.1、RP11-367J11.3、RP3-329E20.2、KDM4A-AS1、RP11-25I15.3、DLEU7-AS1、LINC01232 和 RP11-250B2.6)表达上调。

2.3 进一步验证 lncRNA 表达特征在测试集中的生存预测

为了进一步评估 13-lncRNA 特征的预测能力,本研究在测试集中验证了 13-lncRNA 特征,利用从训练数据集得出的风险评分公式,测试集中每位患者都使用相同的阈值分为高风险组和低风险组,两组总体生存率存在显著差异($HR=2.36$, 95%CI: 1.44~3.87, $P<0.001$) (图 5),同时,测试集中 13-lncRNA 特征预后模型在预测 3 年、5 年总生存率时 AUC 分别为 0.68 和 0.72(图 3b)。

2.4 lncRNA 特征对生存预测的独立性与其他临床变量

单变量和多变量 Cox 回归分析研究 13-lncRNA 特征的预测能力是否独立于性别、年龄、肿瘤分级和 AJCC 肿瘤分期等临床病理因素,结果表明在调整临床因素后,13-lncRNA 特征与生存预后紧密相关(表 2)。根据肿瘤分期进行分层分析,将患者分为早期(AJCC I 期和 II 期)和晚期(AJCC III 期和 IV 期),无论处于哪个阶段,13-lncRNA 特征都可以区分高、

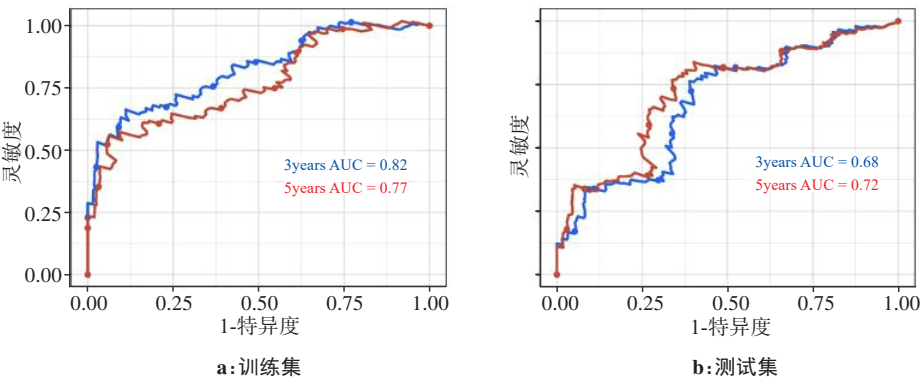


图 3 13 个 lncRNA 特征预后模型预测 HPV 阴性 HNSCC 患者 3 年和 5 年生存率的 ROC 曲线分析

Figure 3 ROC curve analyses of 13 lncRNA signature based prognostic models for predicting 3- and 5-year survival rates in HPV-negative HNSCC patients

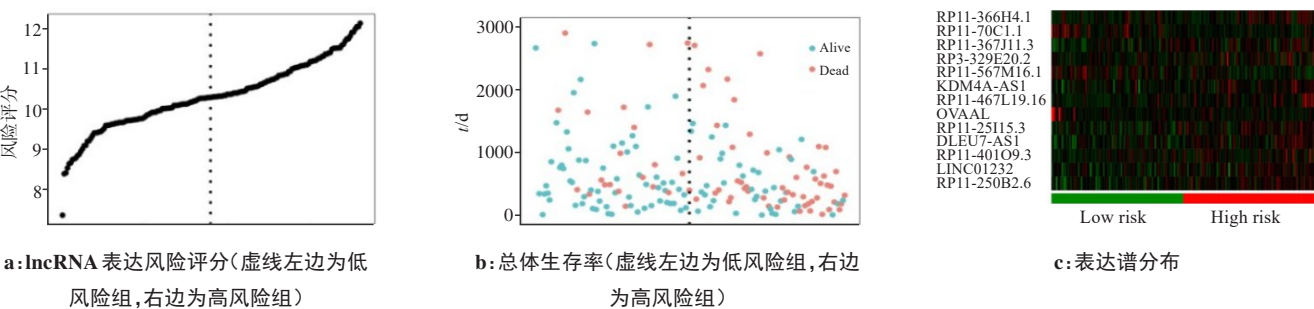


图 4 训练集中 13-lncRNA 表达风险评分、总体生存率及表达谱分布

Figure 4 Risk scores of 13 lncRNA expressions, overall survival rate and expression profile distribution in the training set

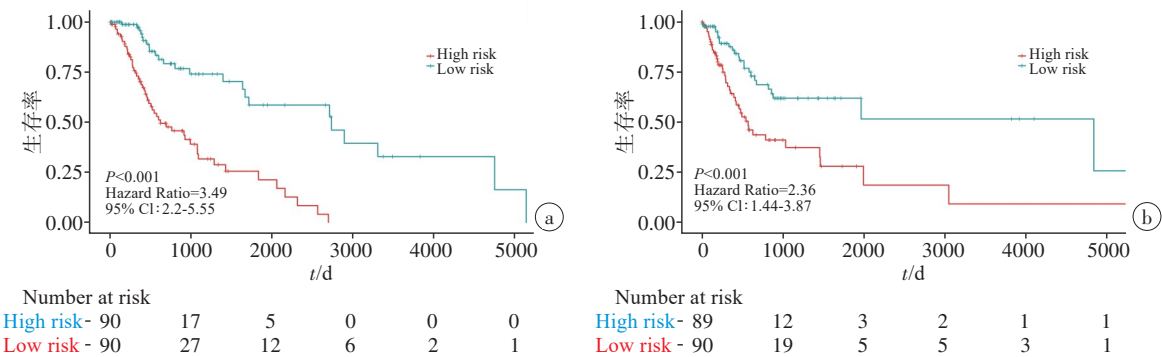


图 5 Kaplan-Meier 曲线分析训练集(a)和测试集(b)中 HPV 阴性 HNSCC 高、低风险患者生存曲线

Figure 5 Kaplan-Meier curve analyses of the survival curves of high- and low-risk HPV-negative HNSCC patients in the training set (a) and the test set (b)

风险危患者,显示全组肿瘤早期高风险患者的预后明显差于低危患者(HR=3.86,95%CI1.67~8.92, $P=0.003$),与全组肿瘤晚期结果一致(HR=2.50,95%CI1.67~3.37, $P<0.001$)(图6)。当进行肿瘤分级(图7)和不同解剖位置(图8)的分层分析时,也得到了类似的结果。这些结果表明,基于13个基因标记的风险评分可能是 HPV 阴性 HNSCC 患者生存的独立预测因素。本研究还进行了多指标的时间相关 ROC 曲线

分析,比较整个数据集中 13-lncRNA 特征风险评分、肿瘤分级和肿瘤分期之间的生存预测敏感性和特异性,计算并比较这 3 个预后因素的 AUC。结果显示风险评分的 AUC 为 0.71,明显大于肿瘤分级(AUC=0.56, $P<0.01$),亦是大于肿瘤分期(AUC=0.60, $P<0.01$)。结果表明风险评分比肿瘤分级或肿瘤分期具有更好的预测能力(图9)。

表 2 单变量和多变量 Cox 回归分析临床参数在 HPV 阴性 HNSCC 患者的预后价值

Table 2 Prognostic significance of univariate and multivariate Cox regression analyses of clinical parameters for HPV-negative HNSCC patients

| 指标 | 单变量模型 | | | 多变量模型 | | |
|-------------|-------|-----------|--------|-------|-----------|--------|
| | 风险比 | 95%CI | P 值 | 风险比 | 95%CI | P 值 |
| 训练集 (n=180) | | | | | | |
| 年龄 | 1.54 | 0.95~2.51 | 0.082 | 1.29 | 0.29~0.86 | 0.389 |
| 性别 | 0.75 | 0.47~1.21 | 0.243 | 0.69 | 0.29~1.28 | 0.202 |
| AJCC 病理肿瘤分期 | 2.26 | 1.16~4.41 | 0.017 | 1.94 | 0.37~1.80 | 0.072 |
| 分级 | 2.23 | 1.09~4.55 | 0.028 | 1.70 | 0.40~1.31 | 0.189 |
| 风险评分 | 4.28 | 2.48~7.40 | <0.001 | 3.82 | 0.31~4.28 | <0.001 |
| 测试集 (n=179) | | | | | | |
| 年龄 | 1.23 | 0.73~2.08 | 0.436 | 1.68 | 0.29~1.77 | 0.077 |
| 性别 | 0.87 | 0.51~1.47 | 0.593 | 0.94 | 0.31~0.19 | 0.847 |
| AJCC 病理肿瘤分期 | 1.55 | 0.78~3.06 | 0.211 | 1.90 | 0.38~1.68 | 0.093 |
| 分级 | 1.59 | 0.68~3.73 | 0.287 | 0.99 | 0.46~0.02 | 0.987 |
| 风险评分 | 2.59 | 1.55~4.32 | <0.001 | 3.23 | 0.29~4.01 | <0.001 |

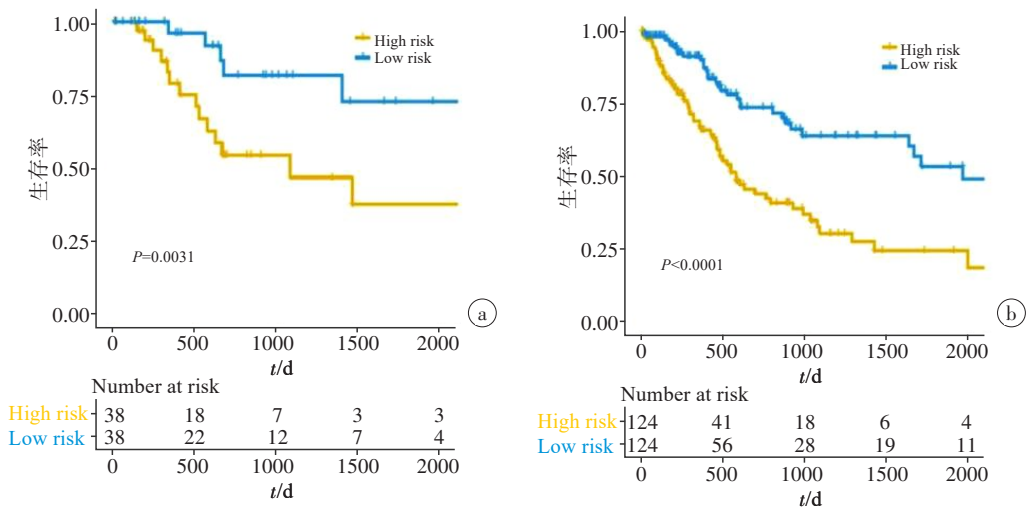


图 6 Kaplan-Meier 曲线分析早期(a)和晚期(b)HPV 阴性 HNSCC 患者高、低风险的生存差异

Figure 6 Kaplan-Meier survival analyses of the survival differences of high- and low-risk HPV-negative HNSCC patients in early (a) and late (b) stages

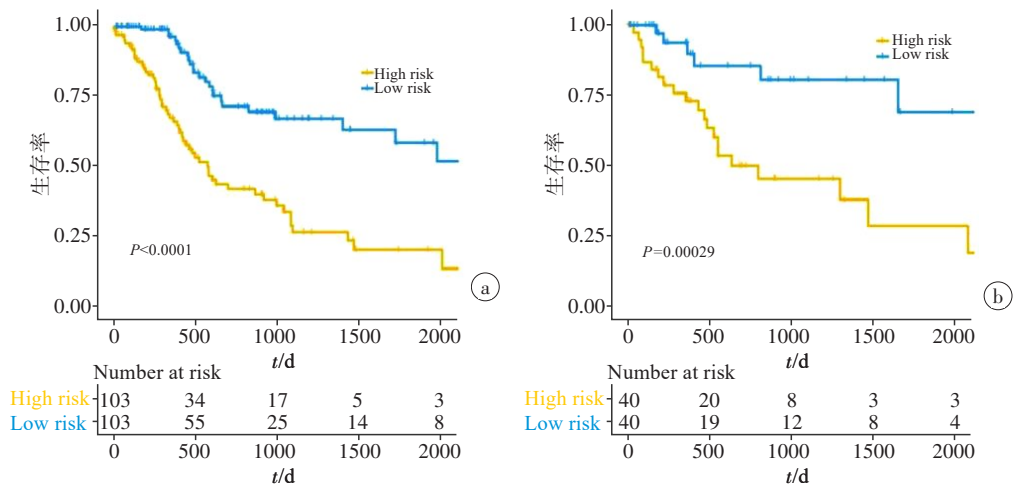


图 7 Kaplan-Meier 曲线分析高、中分化(a)和低分化(b)HPV 阴性 HNSCC 高、低风险患者的生存差异

Figure 7 Kaplan-Meier survival analyses of high- and low-risk HPV-negative HNSCC patients with high- (a) and low-grade (b) differentiations

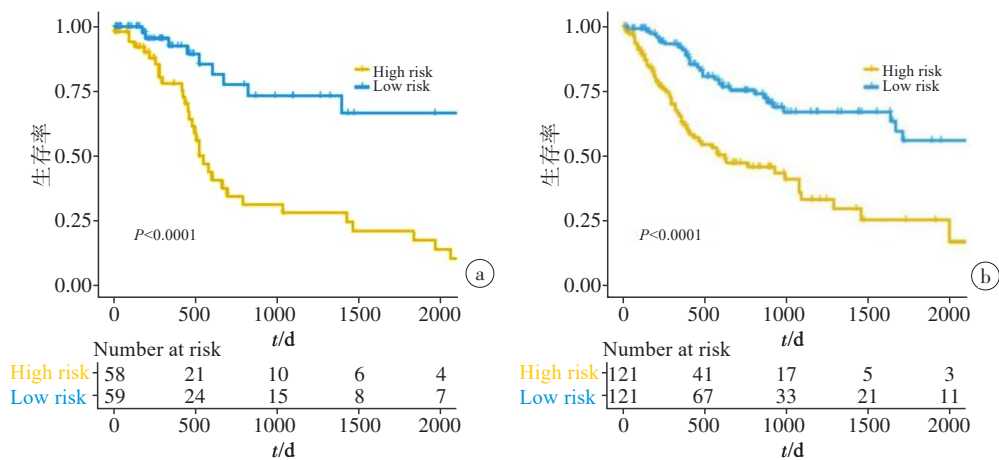


图 8 Kaplan-Meier 曲线分析口腔癌(a)和非口腔癌(b)HPV 阴性 HNSCC 高、低风险患者的生存差异
Figure 8 Kaplan-Meier survival analyses of the survival differences of HPV-negative HNSCC patients with high and low risks of oral cancer (a) and non-oral (b) cancer

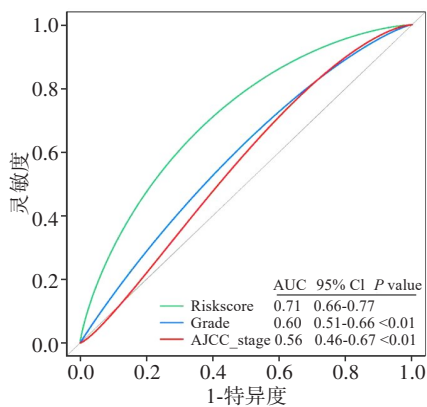


图 9 13-lncRNA 风险评分、肿瘤分期、分化程度预后模型 ROC 曲线
Figure 9 ROC curve analyses of 13-lncRNA risk score, tumor stage and differentiation degree prognostic models

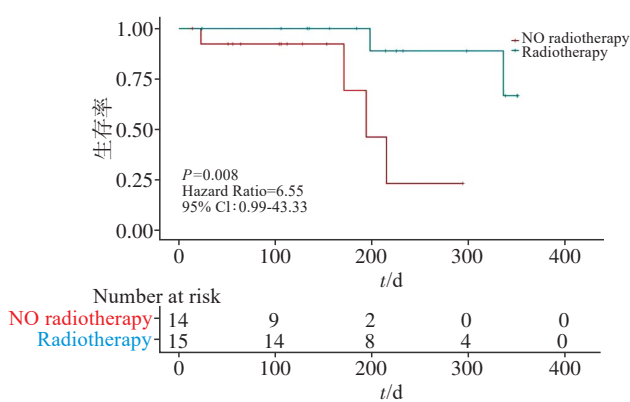


图 10 Kaplan-Meier 生存分析低风险 HPV 阴性 HNSCC 患者是否接受放疗的生存差异
Figure 10 Kaplan-Meier survival analyses of low-risk HPV-negative HNSCC patients receiving radiotherapy or not

2.5 13-lncRNA 标签与低风险人群放射治疗的疗效相关性

本研究分析了 13 个 lncRNA 特征是否可用于预测 HPV 阴性 HNSCC 患者放疗的疗效,低风险评分患者同时进行放疗有利增强疗效 (HR=6.55, 95%CI: 0.99~43.33, $P=0.008$),见图 10。而高风险患者同时进行放疗,其疗效增强无统计学意义 (HR=3.08, 95%CI: 0.62~7.02, $P=0.183$)。

2.6 高风险和低风险人群中不同表达基因的功能评估

GSEA 用于识别相关的生物过程和信号通路本研究比较了 TCGA 数据库中按 13-lncRNA 标签分类的 HPV 阴性 HNSCC 患者在高风险组和低风险组的基因表达谱,挑选出表达差异显著 [错误发现率 (False Discovery Rate, FDR) <0.05] 的基因集进行 GSEA^[17],低风险组高表达的差异基因主要富集在花

生四烯酸代谢、Fc epsilon RI 信号通路(图 11),相反,高风险组高表达的差异基因主要与蛋白酶体和 p53 信号通路相关(图 12)。

3 讨论

研究证实 lncRNA 与恶性肿瘤的发生、发展及转移密切相关,其异常表达可能成为包括头颈部恶性肿瘤在内的肿瘤诊断和预后判断的独立生物标志物^[18-24],但 lncRNA 在 HPV 阴性 HNSCC 中的综合预后价值的研究报道不多。

本研究采用基于生物信息分析的方法,研究 lncRNA 标签预测 HPV 阴性 HNSCC 患者总体生存率。研究发现 HPV 阴性 HNSCC 中 2 281 个 lncRNA 与癌旁正常组织相比表达差异显著,提示 lncRNA 在肿瘤发生中可能发挥关键作用。在训练集中对 359

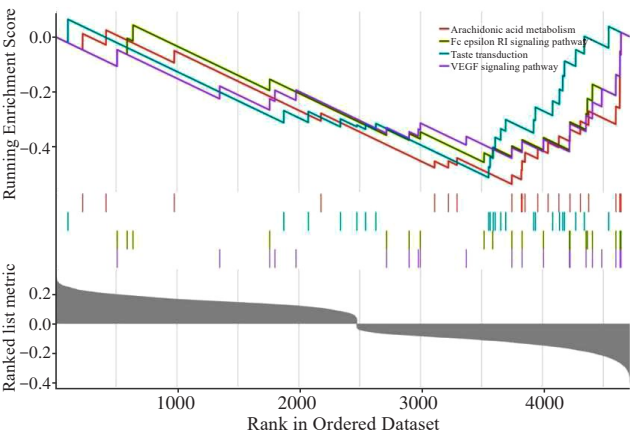


图 11 GSEA 分析低风险组中富集的京都基因与基因组百科全书(KEGG)通路

Figure 11 GSEA analysis of KEGG pathways enriched in low-risk patients

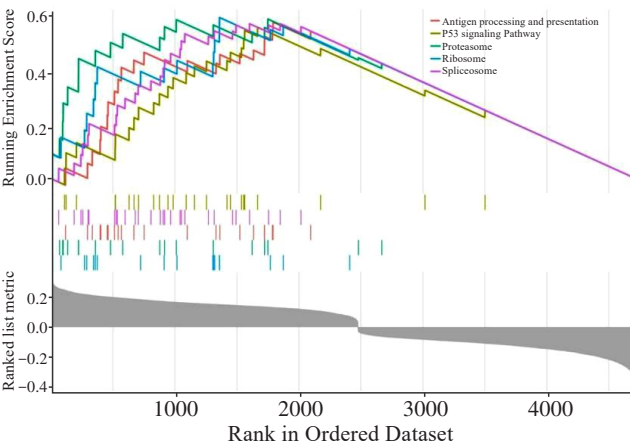


图 12 GSEA 分析高风险组中富集的 KEGG 通路

Figure 12 GSEA analysis of KEGG pathways enriched in high-risk patients

例患者的lncRNA 表达谱进行不同的lncRNA 表达与临床结果之间的相关性进行综合分析,发现 153 个lncRNA 与 HPV 阴性 HNSCC 患者的总生存率显著相关。通过整合训练集中 13 个预后 lncRNA 的表达模式及其对 LASSO Cox 回归分析的回归系数,构建了一个 lncRNA 风险评分模型,13 个 lncRNA 特征在训练集和测试集中均表现出显著的预后能力,可将患者分为高风险组和低风险组,且总生存期存在显著差异,且 13-lncRNA 特征与其他传统临床变量如年龄、性别、分级和分期等无关。

目前,虽然 TNM 分期和肿瘤分级已被广泛用于辅助预测 HNSCC 患者的生存率^[25],但与患者的临床预后判断仍有较大差异,因此,探索分子标签有助于协助临床诊断。在分层分析中,风险评分对不同分期、不同细胞分化程度的头颈恶性肿瘤患者均具有

预后判断价值。风险评分可将 TNM 分期、肿瘤分级和解剖位置不同的患者分为高风险组和低风险组,比肿瘤分期或分级具有更好的预后价值。笔者发现,与不进行放疗相比,放疗有助于提高低风险患者的生存期,但对提高高风险组患者的生存期则无显著差异,提示高风险患者需要其它治疗措施来防止肿瘤复发或转移。对 TCGA 数据集中的基因表达数据进行 GSEA 分析,分析了按 13-lncRNA 标签分层的高风险组和低风险组之间的差异,发现高风险人群中基因的过度表达主要集中在蛋白酶体、p53 信号通路,这可能导致 HPV 阴性 HNSCC 的发生和发展,相反,低风险组的基因过度表达聚集在花生四烯酸代谢、Fc epsilon RI 信号通路,其可能涉及炎症相关通路。

综上所述,本研究确定 13 个与 HPV 阴性 HNSCC 患者总生存率预后显著相关的 lncRNA 生物标志物,尽管 13-lncRNA 特征具有显著的应用前景,但也有其局限性。首先,13-lncRNA 特征作为预后分析独立因素需要在更大的队列试验中进一步验证;其次,应进一步研究获得关于 lncRNA 潜在机制的试验数据;最后,该标签是否有助于预测 HPV 阴性 HNSCC 患者在手术切除后辅助治疗的获益,需要患者有明确的术后辅助治疗信息。此外,13-lncRNA 特征的预后能力与其他临床特征无关,并且可以稳定地预测相同 TNM 分期、分级和位置的 HPV 阴性 HNSCC 患者的生存率,然而,需要在更大的队列试验中进行进一步的验证研究,以评估 13-lncRNA 特征的预后能力。

【参考文献】

[1] Kreimer AR, Clifford GM, Boyle P, et al. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review[J]. Cancer Epidemiol Biomarkers Prev, 2005, 14(2): 467-475.

[2] Gorphe P, Classe M, Ammari S, et al. Patterns of disease events and causes of death in patients with HPV-positive versus HPV-negative oropharyngeal carcinoma[J]. Radiother Oncol, 2022, 168: 40-45.

[3] Jiang BY, Elkashif A, Coulter JA, et al. Immunotherapy for HPV negative head and neck squamous cell carcinoma [J]. Biochim Biophys Acta Rev Cancer, 2024, 1879(5): 189138.

[4] Yi Q, Feng JG, Lan WW, et al. CircRNA and lncRNA-encoded peptide in diseases, an update review[J]. Mol Cancer, 2024, 23(1): 214.

[5] Ao YQ, Gao J, Jiang JH, et al. Comprehensive landscape and future perspective of long noncoding RNAs in non-small cell lung cancer: it takes a village[J]. Mol Ther, 2023, 31(12): 3389-3413.

[6] Guo XB, Hua Z, Li C, et al. Biological significance of long non-coding RNA FTX expression in human colorectal cancer[J]. Int J Clin Exp Med, 2015, 8(9): 15591-15600.

[7] Shan TD, Xu JH, Yu T, et al. Knockdown of linc-POU3F3 suppresses the proliferation, apoptosis, and migration resistance of colorectal cancer[J]. Oncotarget, 2016, 7(1): 961-975.

[8] Shibutani M, Maeda K, Nagahara H, et al. Prognostic significance of the preoperative serum C-reactive protein level in patients with stage IV colorectal cancer[J]. Surg Today, 2015, 45(3): 315-321.

[9] Chen LF, Zhang CX, Gui QF, et al. Ultra-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry-based metabolic profiling of human serum prior to and

- following radical resection of colorectal carcinoma[J]. *Mol Med Rep*, 2015, 12(5): 6879-6886.
- [10] Park EJ, Baik SH, Kang J, et al. The impact of postoperative complications on long-term oncologic outcomes after laparoscopic low anterior resection for rectal cancer[J]. *Medicine (Baltimore)*, 2016, 95(14): e3271.
- [11] Corvaisier M, Bauzone M, Corfiotti F, et al. Regulation of cellular quiescence by YAP/TAZ and cyclin E1 in colon cancer cells: implication in chemoresistance and cancer relapse[J]. *Oncotarget*, 2016, 7(35): 56699-56712.
- [12] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data[J]. *Bioinformatics*, 2010, 26(1): 139-140.
- [13] Valle J, Wasan H, Palmer DH, et al. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer[J]. *N Engl J Med*, 2010, 362(14): 1273-1281.
- [14] Cho HJ, Yu AM, Kim S, et al. Robust likelihood-based survival modeling with microarray data[J]. *J Stat Softw*, 2008, 29(1): 1-16.
- [15] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles[J]. *Proc Natl Acad Sci U S A*, 2005, 102(43): 15545-15550.
- [16] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent[J]. *J Stat Softw*, 2010, 33(1): 1-22.
- [17] Yu GC, Wang LG, Han YY, et al. clusterProfiler: an R package for comparing biological themes among gene clusters[J]. *OMICS*, 2012, 16(5): 284-287.
- [18] Sur S, Davray D, Basu S, et al. Novel insights on oral squamous cell carcinoma management using long non-coding RNAs[J]. *Oncol Res*, 2024, 32(10): 1589-1612.
- [19] Peña-Flores JA, Bermúdez M, Ramos-Payán R, et al. Emerging role of lncRNAs in drug resistance mechanisms in head and neck squamous cell carcinoma[J]. *Front Oncol*, 2022, 12: 965628.
- [20] Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis[J]. *Nature*, 2010, 464(7291): 1071-1076.
- [21] Khaitan D, Dinger ME, Mazar J, et al. The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion[J]. *Cancer Res*, 2011, 71(11): 3852-3862.
- [22] Bo CW, Li N, Li XL, et al. Long noncoding RNA uc.338 promotes cell proliferation through association with BMI1 in hepatocellular carcinoma[J]. *Hum Cell*, 2016, 29(4): 141-147.
- [23] Zhang XQ, Sun S, Lam KF, et al. A long non-coding RNA signature in glioblastoma multiforme predicts survival[J]. *Neurobiol Dis*, 2013, 58: 123-131.
- [24] Li L, Gu M, You B, et al. Long non-coding RNA ROR promotes proliferation, migration and chemoresistance of nasopharyngeal carcinoma[J]. *Cancer Sci*, 2016, 107(9): 1215-1222.
- [25] Ghosh S, Shah PA, Johnson FM. Novel systemic treatment modalities including immunotherapy and molecular targeted therapy for recurrent and metastatic head and neck squamous cell carcinoma[J]. *Int J Mol Sci*, 2022, 23(14): 7889.

(编辑:薛泽玲)