

## • 慢病专题: 癌症 •

# 基于临床肿瘤样本 DNA 甲基化组的机器学习 分类器预测放疗反应<sup>\*</sup>

施小龙<sup>1</sup>, 唐超<sup>1</sup>, 吴颜<sup>1</sup>, 晴俊<sup>2△</sup>

(1. 重庆大学附属肿瘤医院放射肿瘤治疗中心/放射生物中心实验室/重庆市癌症转移与个体化  
诊治转化研究重点实验室, 重庆 400030; 2. 重庆市长寿区人民医院, 重庆 401200)

**[摘要]** 目的 旨在利用 DNA 甲基化特征构建机器学习模型来预测肿瘤患者放射治疗(RT)反应。方法 通过整合分析 10 种癌症类型、843 例患者的全基因组 DNA 甲基化和 RT 疗效数据, 鉴定了与放射的敏感性显著相关的差异甲基化位点(DMSs)。基于这些特征性 DMSs 开发了机器学习分类器模型。结果 基于多个 CpG 位点甲基化信号构建的机器学习分类器, 在区分放射敏感与抗性患者方面(AUC=0.889~1.000)显著优于单一差异甲基化位点(DMS)(AUC=0.594~0.956,  $P<0.001$ )。该模型也可预测患者总生存期。结论 基于 DNA 甲基化组特征的机器学习模型在预测 RT 疗效和预后方面有临床应用价值。

**[关键词]** 放射治疗; DNA 甲基化; 机器学习分类器; 癌症治疗; 放疗敏感性预测

**DOI:** 10.3969/j.issn.1009-5519.2025.11.002      **中图法分类号:** R730.4; TP181

**文章编号:** 1009-5519(2025)11-2506-09

**文献标识码:** A

## Machine learning classifiers based on DNA methylation profiles of clinical tumor samples predict radiotherapy responses<sup>\*</sup>

SHI Xiaolong<sup>1</sup>, TANG Chao<sup>1</sup>, WU Yan<sup>1</sup>, QI Jun<sup>2△</sup>

(1. Radiation Oncology Center/Radiation Biology Center Laboratory/Chongqing Key Laboratory  
of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing  
University Cancer Hospital, Chongqing 400030, China; 2. People Hospital of Changshou  
Chongqing, Chongqing 401200, China)

**[Abstract]** **Objective** To construct a machine learning model using DNA methylation features to predict radiotherapy(RT) response in cancer patients. **Methods** By integrating and analyzing 10 types of cancer and whole-genome DNA methylation data alongside RT efficacy data from 843 patients, differentially methylated sites(DMSs) significantly associated with radio sensitivity were identified. Machine learning classifier models were developed based on these characteristic DMSs. **Results** The machine learning classifier constructed based on methylation signals from multiple CpG sites(AUC=0.889~1.000) was significantly superior to a single DMS(AUC=0.594~0.956,  $P<0.001$ ) in distinguishing between radio-sensitive and radio-resistant patients. The model could also accurately predict the overall survival period of patients. **Conclusion** The machine learning model based on DNA methylome features has clinical application value in predicting the efficacy and prognosis of RT.

**[Key words]** Radiotherapy; DNA methylation; Machine learning classifiers; Cancer treatment;  
Prediction of radio-sensitivity

放射治疗(RT)在多种恶性肿瘤治疗中发挥着关键作用, 既能实现局部控制, 在某些病例中甚至能达到治愈效果<sup>[1]</sup>。然而, 尽管疗效显著, 患者对 RT 的敏感性存在显著的个体差异<sup>[2]</sup>。这种放射敏感性的个体差异促使研究者们致力于寻找能够指导个体化

治疗、优化临床疗效的预测性生物标志物<sup>[2]</sup>。

目前已有多 种方法用于预测放射敏感性, 包括基因表达谱分析和基因组突变检测<sup>[3-4]</sup>。DNA 损伤修复(DDR)基因[如共济失调毛细血管扩张症突变(ATM)和 BRCA1/2]的突变已被用作个体化治疗策

\* 基金项目: 重庆市科卫联合医学科研面上项目(2024MSXM065; 2023GGXM002); 重庆市长寿区科技局项目(2024023CSKJ)。

作者简介: 施小龙(1975—), 博士研究生, 副研究员, 主要从事肿瘤基因组学、肿瘤早筛技术开发、基因检测技术的开发和转化以及生物信息学研究。△ 通信作者, E-mail: qqyc2005@163.com。

网络首发 [https://link.cnki.net/urlid/50.1129.R.20251014.1435.002\(2025-10-15\)](https://link.cnki.net/urlid/50.1129.R.20251014.1435.002(2025-10-15))

略的预测标志物。然而,由于可用突变数量有限,其广泛应用受到制约。基于 RNA 检测的基因表达特征也被开发用于预测 RT 反应,但这些方法存在数据解读复杂、可重复性不足、检测方法缺乏标准化及临床应用受限等缺陷<sup>[3-4]</sup>。

放射敏感性的个体差异部分源于肿瘤细胞基因表达模式的差异。DNA 甲基化作为重要的表观遗传修饰,可通过调控基因表达影响细胞对辐射的反应。研究表明,特定基因启动子区的甲基化状态可作为 RT 反应的预测因子和预后指标<sup>[3]</sup>,这些基因主要涉及 DNA 损伤修复、细胞周期和凋亡等过程。例如,ATM 基因编码的高分子量蛋白激酶在细胞基因组损伤应答中起关键作用。KIM 等<sup>[5]</sup>研究发现 HCT-116 细胞较 LoVo 和 RKO 细胞系表现出更高的放射敏感性,这种差异可能与 ATM 基因启动子区的异常甲基化有关。ROY 等<sup>[6]</sup>在神经胶质瘤细胞系中也观察到类似现象。

与基因表达谱(如 RNA 检测)相比,DNA 甲基化作为 RT 反应预测标志物具有显著优势。首先,甲基化 CpG 位点在基因组中广泛分布,可提供大量候选标志位点;其次,DNA 甲基化标志稳定性高,适用于福尔马林固定石蜡包埋(FFPE)样本检测;此外,基于循环肿瘤 DNA(ctDNA)的液体活检可保留来源肿瘤的表观遗传特征,为无创预测 RT 反应提供可能。

既往研究多采用癌细胞系作为 DNA 甲基化与放射敏感性研究的模型。然而肿瘤组织中的放射反应机制可能比细胞系更为复杂。单个基因的甲基化状态往往不足以准确预测 RT 反应,因此基于临床肿瘤样本 DNA 甲基化状态与 RT 反应分析获得的生物标志物可能更具临床转化价值。高通量 DNA 分析技术(如微阵列和高通量测序)的发展为肿瘤甲基化组研究提供了强大工具。本研究基于 TCGA 数据库,通过对接受 RT 治疗的大规模患者群体进行 DNA 甲基化组综合分析,鉴定出大量差异甲基化位点(DMSs),并采用机器学习算法构建最优 DMSs 组合分类器。这些分类器不仅能显著提高放射敏感性预测效能,还可用于患者生存率预测,为肿瘤精准放疗提供新的决策依据。

## 1 材料与方法

**1.1 样本选择** 从 UCSC XENA-GDC TCGA 数据库(<https://xenabrowser.net/hub/>)下载了 10 种癌症类型的 450K DNA 甲基化数据及临床数据。根据临床资料中的 RT 反应信息,将样本分为 RT 敏感组(S 组,完全/部分缓解)和 RT 抵抗组(R 组,疾病进展)。最终纳入 843 例样本,其中 S 组 586 例,R 组 257 例。

**1.2 DNA 甲基化数据预处理** 采用 R 语言 ChAMP 包<sup>[7]</sup>对原始数据进行质控:(1)剔除低质量样本及探针;(2)使用  $\beta$  值表示甲基化水平(0~1 范围);(3)通过 BMIQ 算法校正技术批次效应;(4)采用  $P$  值过滤

( $P > 0.01$ )剔除信号不可靠位点。

**1.3 基于甲基化组的层次化分析** 按癌种分层后,使用 ChAMP 包进行差异甲基化分析:(1)计算 S 组与 R 组间甲基化水平差异( $t$  检验);(2)设定筛选阈值: $P < 0.05$  且  $|\log 2FC| > 0.1$ ;(3)通过 Benjamini-Hochberg 法校正多重假设检验。可视化分析采用 pheatmap 包绘制热图展示差异甲基化模式;使用 pROC 包生成受试者操作特征(ROC)曲线和 PRROC 包生成精确率-召回率(PPV)曲线评估预测效能<sup>[7-8]</sup>。

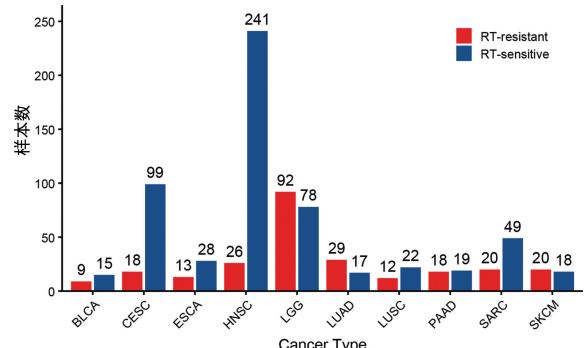
**1.4 机器学习分类器构建** 采用分层抽样将数据按 4:1 比例随机划分为训练集(674 例)与测试集(169 例);通过 LASSO 回归(glmnet 包实现)在训练集上进行特征选择。采用 5 折交叉验证(在训练集内部执行)确定最优正则化参数  $\lambda$ ;选择交叉验证 AUC 最大值对应  $\lambda$  值。逻辑回归模型构建: $Y = W_1 \cdot cg_1 + W_2 \cdot cg_2 + \dots + W_n \cdot cg_n + b$ ,其中  $W$  为权重系数, $cg$  为 DMS 甲基化水平, $b$  为截距项,输出值  $Y > 0.5$  判定为敏感组。

**1.5 预测结果可视化** 使用 ggplot2 绘制预测评分分布图;通过 survival 包进行生存分析。按预测评分中位数分组;使用 Kaplan-Meier 法绘制生存曲线;使用 Log-rank 检验评估组间差异。

**1.6 功能富集分析** 将前 500 个基因注释 DMS 输入 Metascape 平台(<https://metascape.org/gp/index.html>),进行 GO/KEGG 通路富集分析。

## 2 结果

**2.1 患者临床特征分析** 本研究基于 TCGA 公共数据库,对 843 例接受 RT 的癌症患者(S 组 586 例,R 组 257 例)开展甲基化水平与放射敏感性关联分析。样本涵盖 10 种癌症类型,包括膀胱尿路上皮癌(BLCA,24 例)、宫颈鳞癌(CESC,117 例)、食管癌(ESCA,41 例)、头颈鳞癌(HNSC,267 例)、脑低级别胶质瘤(LGG,170 例)、肺腺癌(LUAD,46 例)、肺鳞癌(LUSC,34 例)、胰腺癌(PAAD,37 例)、肉瘤(SARC,69 例)和皮肤黑色素瘤(SKCM,38 例)。图 1 展示了不同癌种样本的分组情况。

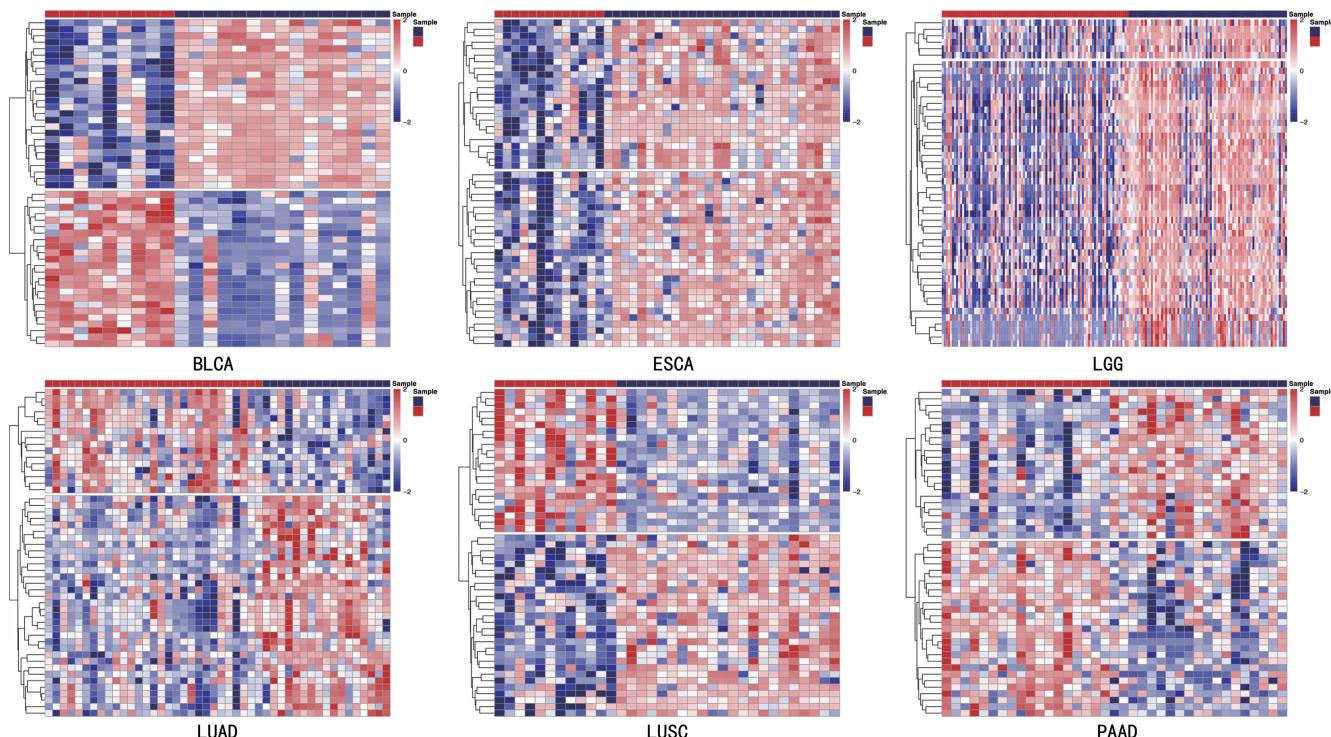


注:根据患者对放疗的临床反应,将癌症样本分为放疗敏感组(蓝色)和放疗抵抗组(红色)。本分析共纳入 TCGA 数据库中 10 种不同类型癌症的 843 例肿瘤样本。

图 1 癌种分布及分组统计图

**2.2 S 组与 R 组差异甲基化位点(DMSs)鉴定** 为鉴定不同癌症类型中与放射敏感性相关的甲基化位点,比较了 S 组和 R 组样本的 DMSs 甲基化水平。通过 p 值筛选,各癌症类型中显著差异的 DMSs 被鉴定。特别值得注意的是,这些 top 50 DMSs 包含 TERT、STAT5A、RAD51L1、PRKDC、PPP1R2P1、MBD3、KDM4B、IL1B、IL12B、BRCA2 和 ARID1A 等基因,这些基因在既往研究中已被证实与甲基化状态相关。无监督聚类分析显示,在 10 种癌症类型中,大多数癌种的两种应答组间甲基化模式存在显著差异。如图 2 所示,在 BLCA、ESCA、LGG、LUAD、LUSC 和 PAAD 肿瘤样本中观察到明显的组间区分能力。以 BLCA 肿瘤样本为例(图 2),在其 top 50 DMSs 中,有 26 个位点在 S 组样本中表现出比 R 组样本更高的甲基化水平趋势,表明这些位点甲基化水平升高可能与放射敏感性存在关联。进一步评估这 10 种癌症类型中 DMSs 对 S 组和 R 组样本的区分效能,对于样本

量相对平衡的癌种(BLCA、ESCA、LGG、LUAD、LUSC、PAAD、SARC、SKCM),采用 ROC 曲线评估其分层的特异性和灵敏度(图 3A~H)。单一 DMSs 的曲线下面积(AUC)值为 0.59~0.96。其中部分 DMSs 展现出较好的区分能力。例如,在 BLCA 中,基因 GJB6 的 cg03568673 位点的 AUC 值达到 0.956。对于存在明显样本不平衡的 CESC 和 HNSC 癌种(灵敏度分别为 84.6% 和 90.3%,一方样本比例低于 20%),额外采用 PRC 进行评估(图 3I、J)。在这两个癌种中表现最佳的单个 DMSs,其 PRC 曲线下面积(AUPRC)值分别高达 0.96 和 0.97。需要指出的是,部分癌种(如 BLCA、LUAD)样本量较少( $n < 50$ ),其单个 DMS 的 ROC AUC 值/精确率-召回率曲线下的面积(AUPRC)值可能存在较大波动。尽管如此,多个位点展现出稳定的高区分效能,表明这些甲基化标记物具有作为放射敏感性预测生物标志物的潜力。



注:横轴代表患者样本,纵轴代表 DMSs 位点。颜色从蓝色到红色表示从低甲基化到高甲基化的变化趋势。样本分类标注于每张热图顶部,其中蓝色代表放疗敏感组,红色代表放疗抵抗组。

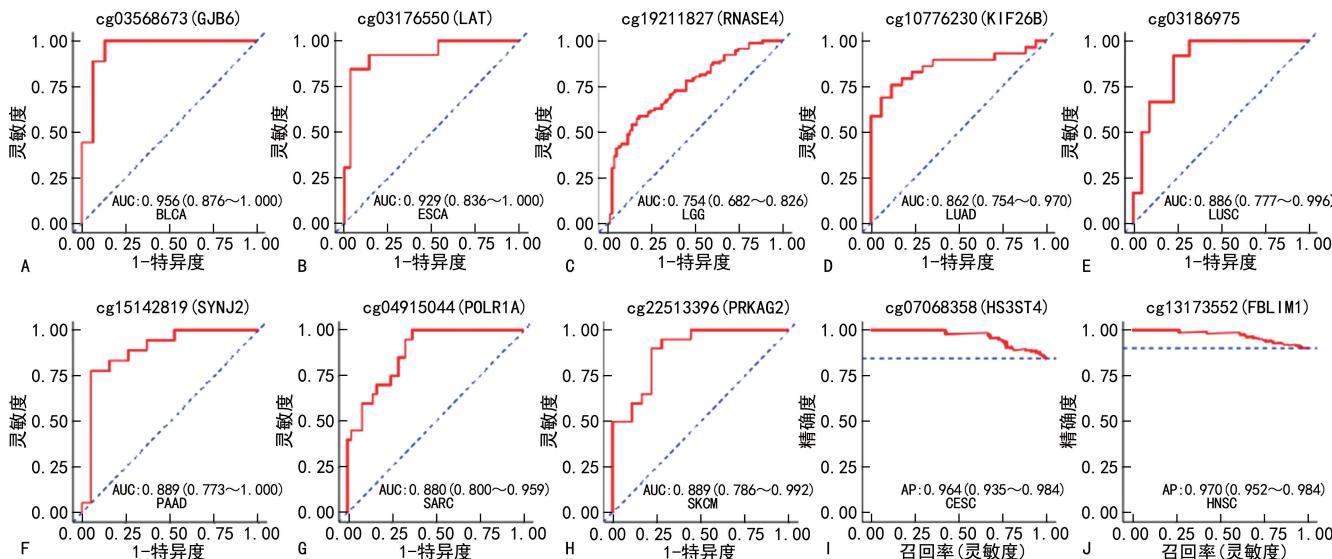
图 2 不同肿瘤类型 DMSs 热图

**2.3 DMSs 功能富集分析** 为探究这些 DMSs 是否与特定基因功能相关,分析了 DMSs 在基因组中的分布特征。结果显示,这些 DMSs 主要分布于启动子区(16%~30%)、基因间区(26%~32%)和基因体区(31%~45%)。这表明大多数 DMSs 定位于基因的启动子区和基因体区域。其中,启动子区作为基因转录起始区域,其 DMSs 的分布可能对基因表达具有重要调控作用;基因间区可能包含远程调控元件;而基因体区覆盖基因的编码区和内含子区,该区域的

DMSs 可能与基因调控相关。通过解析这些区域 DMSs 的功能及调控机制,可深化对肿瘤放射敏感性分子机制的理解。进一步研究中,选择最具代表性的前 500 个 DMSs(top500 DMSs)进行功能富集分析(图 4)。结果显示,这些 DMSs 显著富集的通路与癌症进展密切相关,主要包括:DDR 通路、炎症反应调控通路以及 Wnt 信号通路负调控通路。具体而言:在 LGG 中,发现这些基因显著富集于 Wnt 信号通路负调控和炎症反应调控通路。Wnt 信号通路在肿瘤发

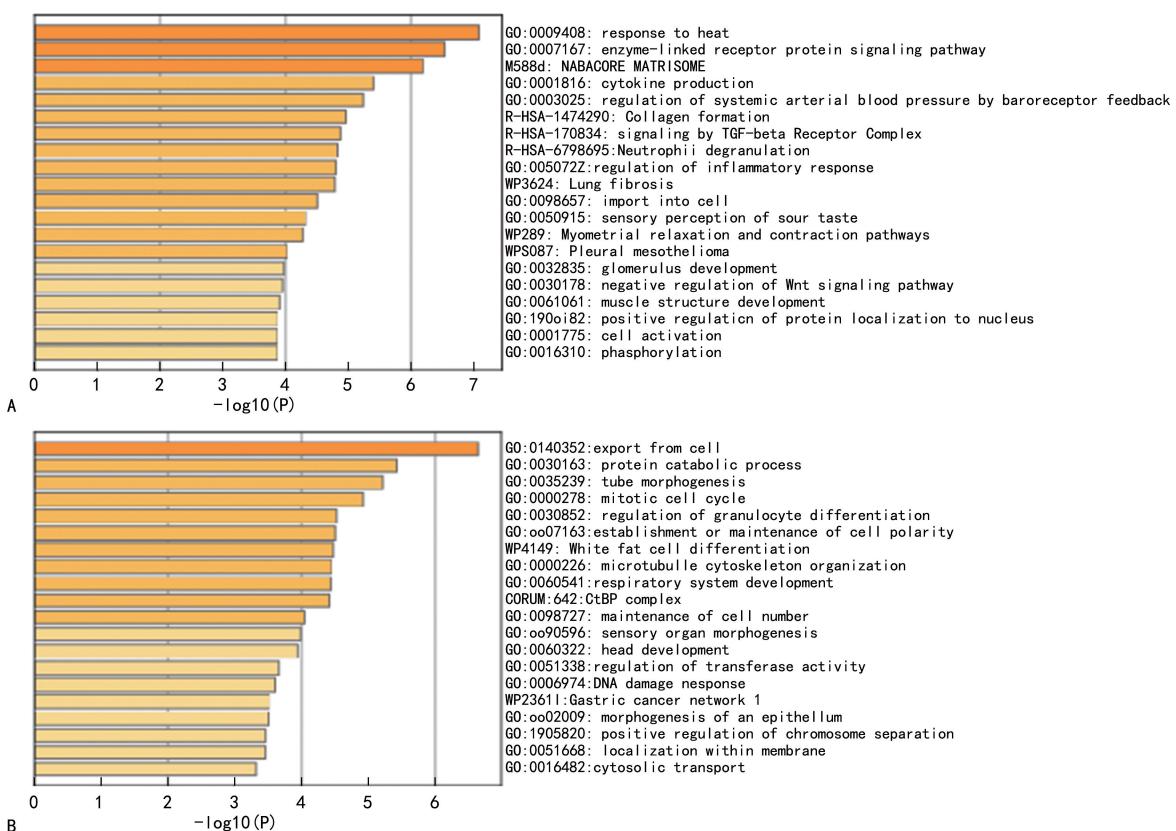
生发展中起关键作用,参与细胞增殖、分化及干细胞自我更新等过程。已有研究证实 Wnt 信号通路与肿瘤放射敏感性存在关联<sup>[9]</sup>。此外,炎症反应在肿瘤侵袭转移过程中也具有重要作用<sup>[10]</sup>。在 LUAD 中,鉴定出 DMSs 相关基因主要与 DDR 通路相关。DDR 是细胞修复 DNA 损伤的关键过程<sup>[11]</sup>。本研究表明,

LUAD 患者肿瘤细胞中 DDR 通路可能存在异常激活或调控,这可能会影响肿瘤的发展进程和治疗敏感性。通过功能富集分析,揭示了 DMSs 相关基因的富集通路与肿瘤细胞对 RT 的应答状态存在显著相关性。这些发现为理解甲基化在肿瘤细胞放射敏感性和抵抗性中的作用提供了重要线索。



注: A~H 样本量相对平衡的 8 种癌症类型(BLCA、ESCA、LGG、LUAD、LUSC、PAAD、SARC、SKCM)中,具有最高 ROC AUC 值的单个 DMS 的 ROC 曲线。I~J 样本量不平衡的 2 种癌症类型(CESC、HNSC)中,具有最高区分效能的单个 DMS 的 PRC 图[以 AUPRC, 即平均精度(AP)评估]。括号中的数字为 95% 的置信区间(CI)。

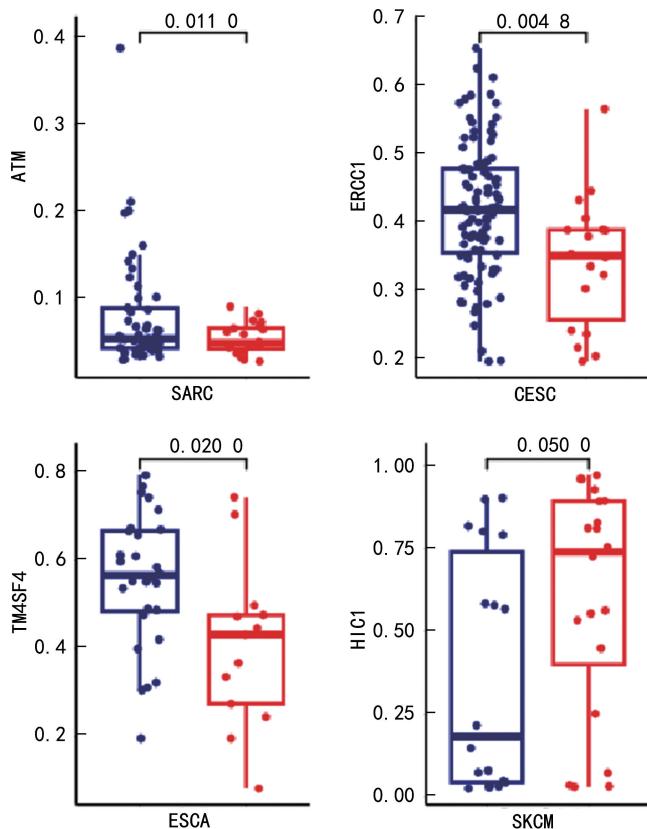
图 3 单个差异甲基化位点(DMSs)对放疗敏感组与放疗抵抗组的区分效能评估



注: A 为 LGG; B 为 LUAD。按 P 值显著性排序。

图 4 DMSs 相关基因功能富集分析

**2.4 已报道 RT 反应相关基因甲基化性能评估** 本研究同时评估了既往研究中多个与 RT 反应预测相关的甲基化标志基因的性能。这些基因包括涉及 DNA 损伤修复的关键因子 ATM<sup>[12]</sup>、NER 相关基因 ERCC1<sup>[13-14]</sup>、DNA 损伤修复检查相关基因 RASSF1A<sup>[15-16]</sup>、促进肿瘤细胞增殖、侵袭、迁移和上皮-间质转化的相关基因 SERPINB5<sup>[17-18]</sup>，参与增殖和血管生成的转录抑制因子 HIC1 基因<sup>[19]</sup>，端粒酶 hTERT 基因<sup>[20]</sup>，参与维持基因组稳定性的 FHIT 基因<sup>[21-22]</sup> 及肿瘤细胞增殖相关的 TM4SF4 基因<sup>[23]</sup>。基于 TCGA 数据库，本研究分析了这些基因甲基化水平与患者 RT 结局的相关性(图 5)。结果显示 ERCC1 基因(CESC,  $P=0.0048$ ) 和 ATM 基因(SARC,  $P=0.0110$ ) 在 S 组样本中启动子区甲基化水平升高；而



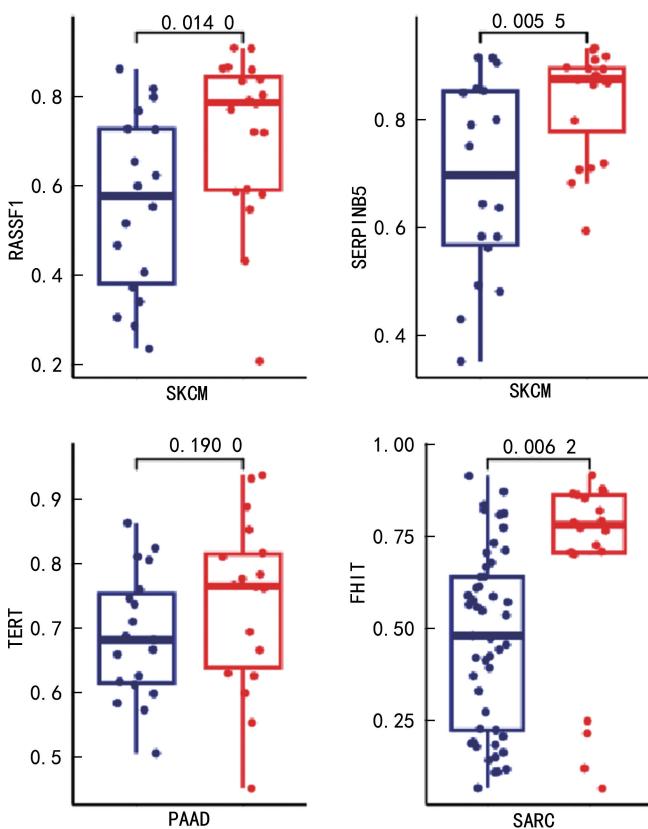
注：采用  $t$  检验比较两组间基因启动子区甲基化水平。蓝色柱状图代表放疗敏感组，红色柱状图代表放疗抵抗组。各组比较均标注相应  $P$  值。

图 5 放疗敏感组与放疗抵抗组间差异基因甲基化水平比较

**2.5 机器学习分类器的构建与验证** 首先通过 ROC 曲线评估单个 DMS 对 S 组和 R 组样本的区分能力，结果显示这些 DMS 的 AUC 值介于 0.594~0.956，已报道的 RT 反应相关基因 DMS 的 AUC 值介于 0.595~0.761。为提高 RT 反应预测的特异性，采用机器学习方法构建了逻辑回归分类器。在构建放射敏感性预测分类器时，通过最优 AUC 值筛选最具预测价值的特征变量进行 LASSO 回归分析(图 7)。将具有非零 LASSO 系数的 DMS 整合成位点

RASSF1 基因(SKCM,  $P=0.0140$ )、FHIT(SARC,  $P=0.0062$ )、HIC1(SKCM,  $P=0.0500$ )、TM4SF5(ESCA,  $P=0.0200$ )、SERPINB5(SKCM,  $P=0.0055$ ) 和 TERT(PAAD,  $P=0.1900$ ) 等基因在 R 组样本中甲基化水平升高(图 5)。这些发现与先前研究结果相符。

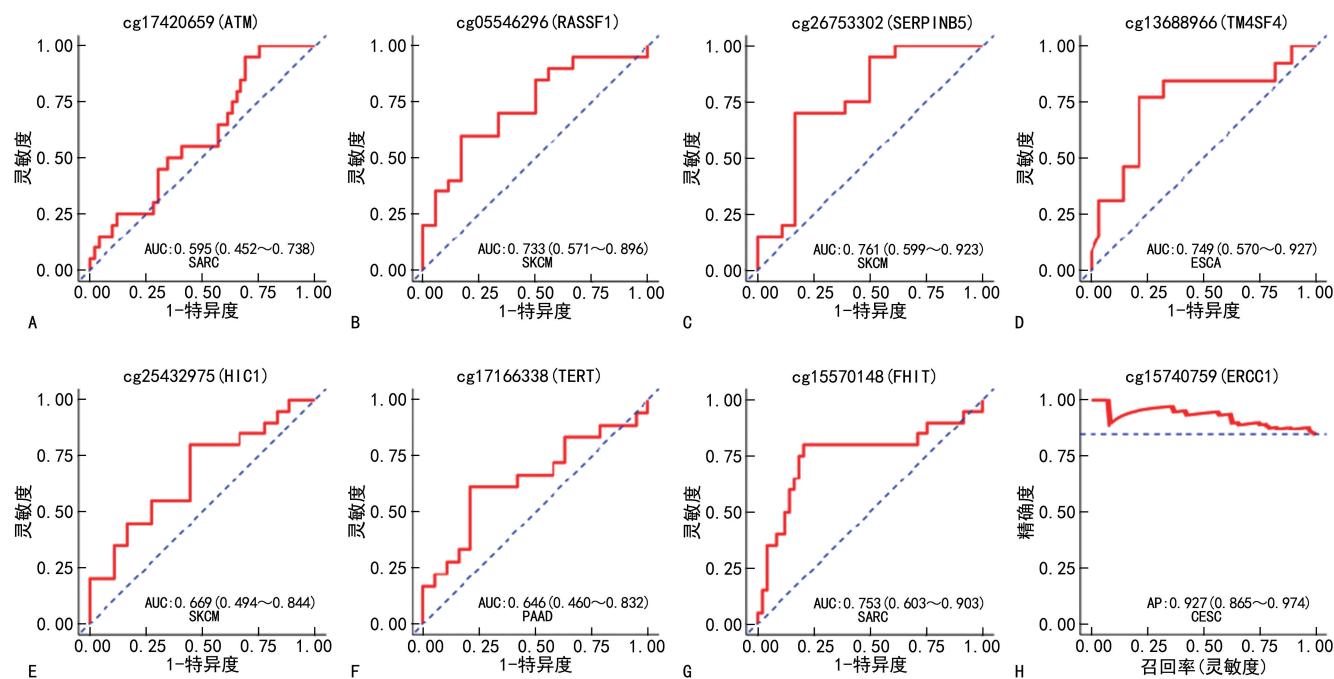
为深入评估这些已报道与 RT 反应相关基因的甲基化位点对 S 与 R 组样本的区分能力，进行了 ROC/PRC 分析。如图 6A~G 所示，在采用 ROC 评估的癌种中，AUC 值为 0.595~0.761；在采用 PRC 评估的 CESC 中(图 6H)，其最佳基因位点的平均精度(AP)值达到 0.927。这些结果表明，尽管存在个别表现优异的基因位点，但多数基因位点对 RT 反应的预测特异性处于中等水平。



集。最终基于材料与方法部分描述的公式计算预测评分(图 8)。如图 8 所示，该分类器能有效区分绝大多数 S 组与 R 组样本。在独立测试集上，通过多维度指标评估模型性能：在样本平衡癌种 ROC 评估体系中，分类器 AUC 值显著超越单点预测(图 9A~D)。例如 LGG 分类器 AUC 达 0.889 (vs. 单点最高 0.754, DeLong 检验  $P = 1.2 \times 10^{-5}$ )。在样本不平衡癌种 PRC 评估体系中(图 9E~F)，AP 值展现卓越稳定性，例如 CESC 分类器测试集实现完美预测(AP

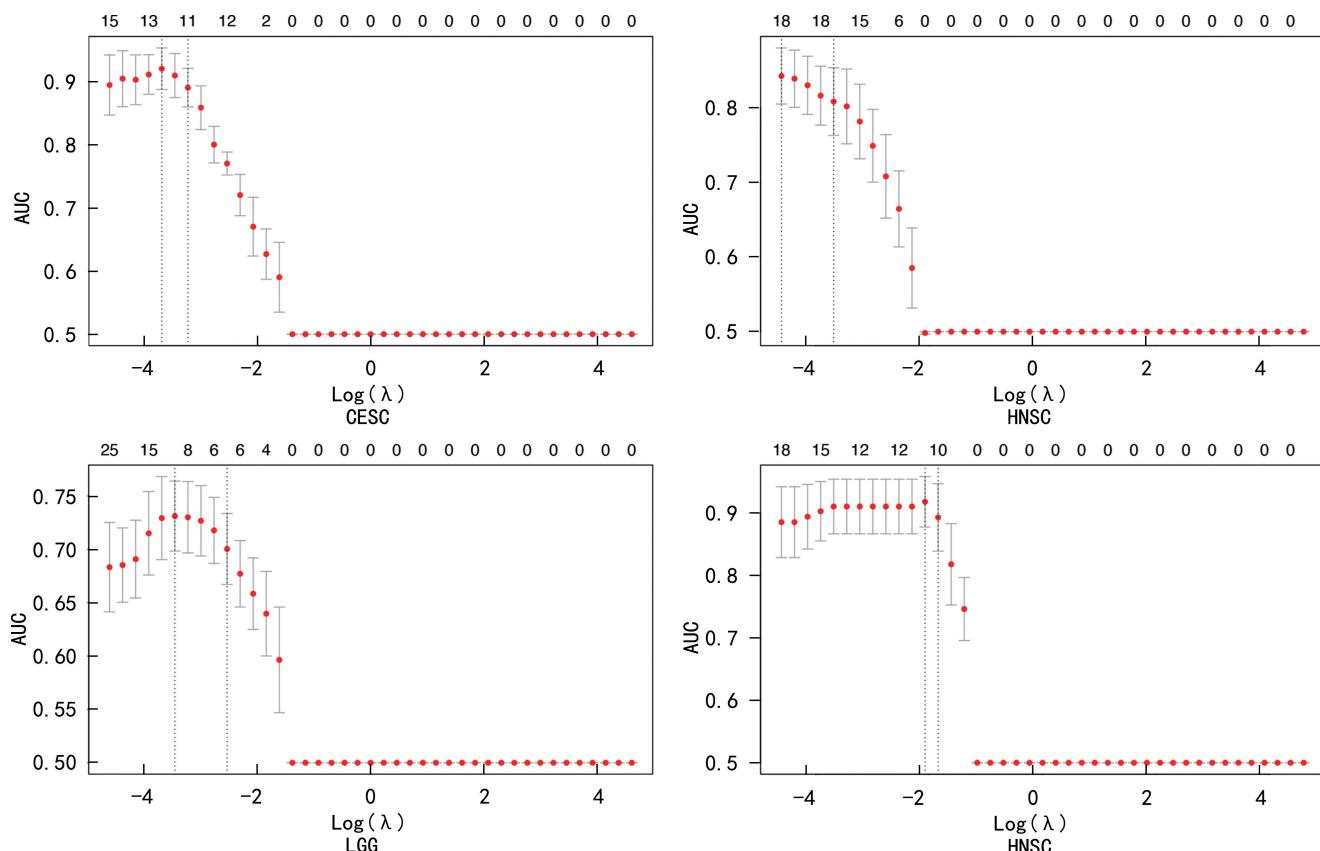
为 1.000), HNSC 分类器(AP 为 0.995)。整合多维甲基化特征构建的分类器,其性能显著超越最优单点标志物,证实多特征整合可突破单标志物的预测瓶

颈。该模型为个体化放疗决策提供了高精度分子工具。



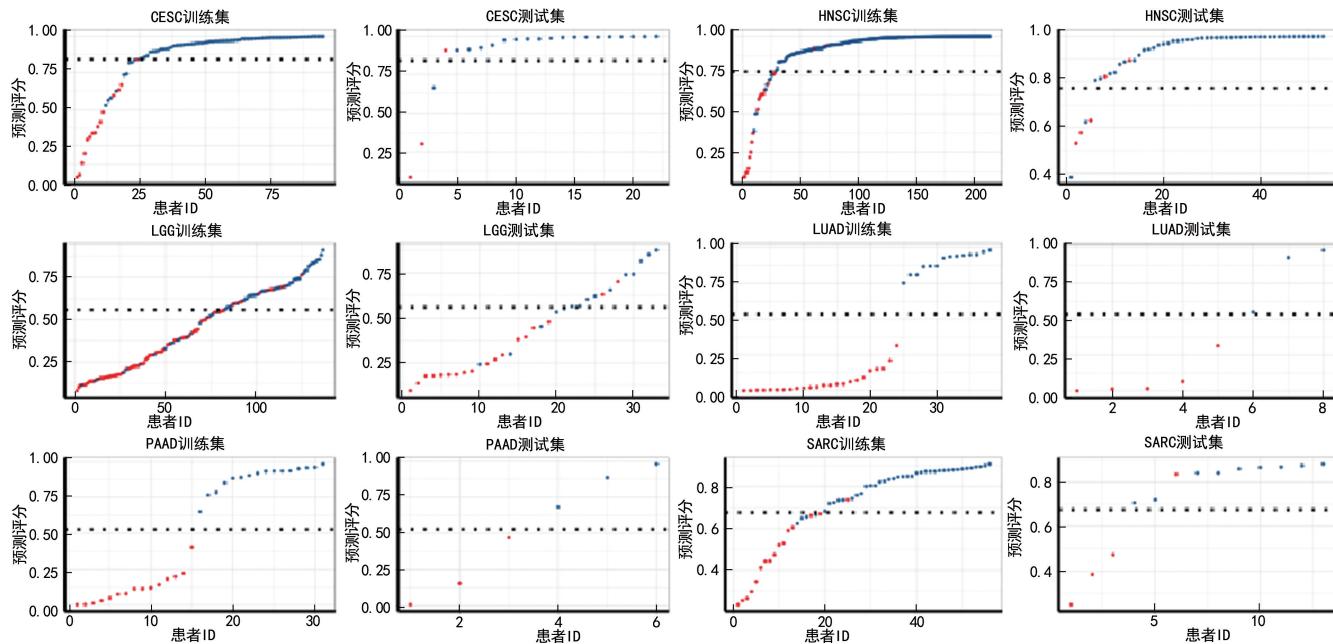
注:展示了在不同癌症类型中,通过 ROC(A~G)或 PRC(H)评估的、已报道与 RT 反应相关的特定基因甲基化位点对 S 组与 R 组的区分能力,多数基因表现出中等程度的特异性。括号中的数字为 95% 的置信区间(CI)。

图 6 评估已报道 RT 反应相关基因甲基化位点的区分效能



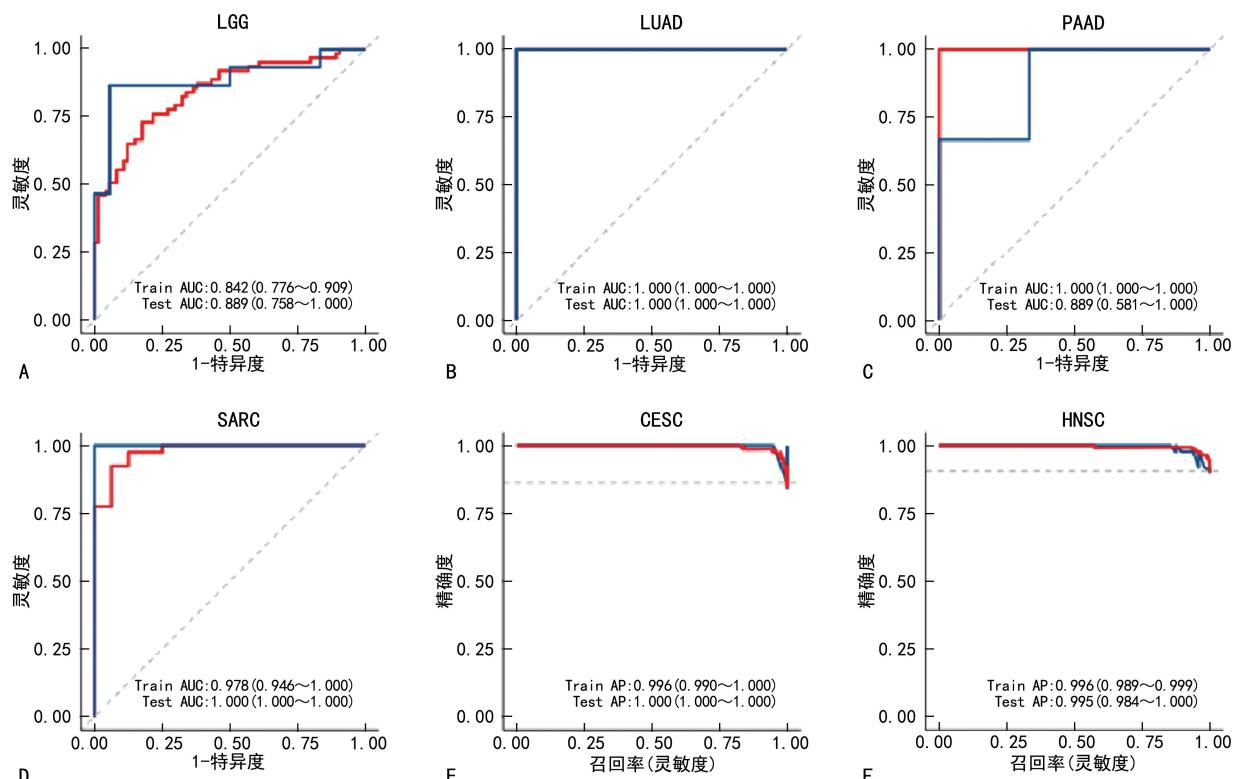
注: AUC 值随  $\log(\lambda)$  的变化趋势,垂直虚线标示最优  $\lambda$  值位置。

图 7 LASSO 回归模型中最优  $\lambda$  值选择曲线



注:红色圆点代表放疗抵抗组(R组)样本,蓝色圆点代表放疗敏感组(S组)样本,虚线表示分类阈值。如图所示,绝大多数R组样本集中在阈值下方,而S组样本主要分布在阈值上方。

图 8 分类器区分 S 组与 R 组样本的性能表现



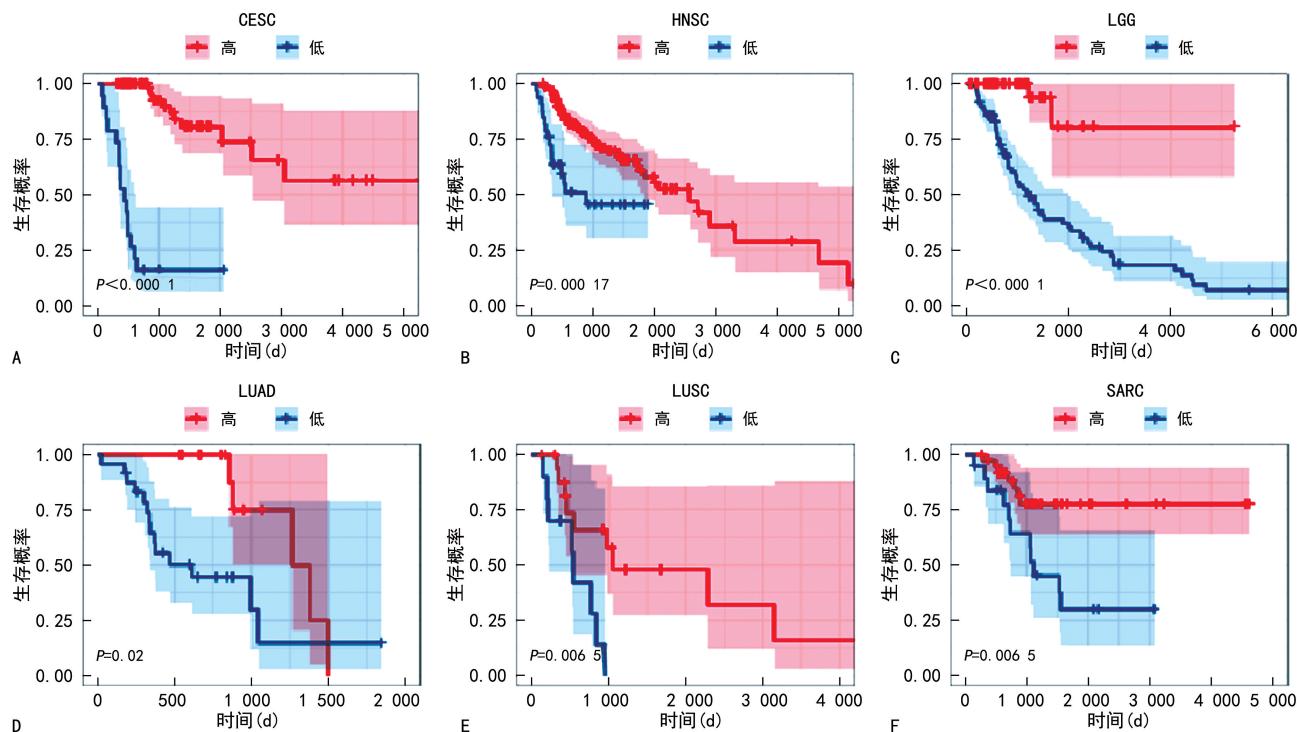
注:A~D 为样本平衡癌种中训练和测试集的 ROC 曲线; E~F 为样本不平衡癌种中训练和测试集的 PRC 曲线; 图中展示该模型的预测性能表现;括号中的数字为 95% 的置信区间(CI)。

图 9 基于多特征整合的逻辑回归分类器预测性能

**2.6 基于机器学习分类器的患者生存预后分层** 为评估分类器的预后价值,以预测评分作为分组标准进行总生存期(OS)分析。根据预测评分将患者分为高分组和低分组,通过 log-rank 检验和 Kaplan-Meier 生存曲线分析预测评分与生存期的相关性。结果显示,在多种癌症类型中,不同预测评分组间的 OS 存在显

著差异。如图 10 所示,高分组的生存优势在以下癌种中差异有统计学意义。例如,宫颈癌(CESC,  $P < 0.0001$ )、头颈鳞癌(HNSC,  $P = 0.0002$ )、脑低级别胶质瘤(LGG,  $P < 0.0001$ )、肺腺癌(LUAD,  $P = 0.0200$ )、肺鳞癌(LUSC,  $P = 0.0065$ )及肉瘤(SARC,  $P = 0.0065$ )。这表明预测评分较高的患者

多数对放疗敏感,且具有更好的生存预后。



注:图中高分/低分组分别对应高/低预测评分患者;较高评分预示放射敏感性倾向,而较低评分提示放射抵抗趋势;经 log-rank 检验证实高预测指数组患者的总生存期显著更优。

图 10 基于预测评分的 OS 曲线分层分析

### 3 讨 论

RT 是癌症治疗的重要手段之一,而肿瘤放射抵抗是导致 RT 失败的主要原因。能够预测患者治疗反应的生物标志物对于制定个体化 RT 策略具有重要价值。临床医生若能预先评估肿瘤放射敏感性,将有助于制定更精准的治疗方案,从而提高疗效、减少不必要的副作用并优化医疗资源配置。

DNA 甲基化作为表观遗传修饰,可在不改变 DNA 序列的情况下调控基因表达,与癌症发生发展密切相关<sup>[24]</sup>。研究表明,DNA 甲基化在放射敏感性调控中起关键作用,其模式改变可能通过影响 DNA 修复、细胞周期调控和凋亡等辐射应答相关通路<sup>[25-26]</sup>,进而调控肿瘤细胞对 RT 的敏感性。KIM 等<sup>[27]</sup>在非小细胞肺癌放射抵抗细胞中发现 747 个高甲基化基因和 344 个低甲基化基因,揭示了敏感与抵抗细胞间显著的甲基化组差异。DNA 修复相关基因(如胶质母细胞瘤中 MGMT 启动子高甲基化和食管癌中 RUNX3 基因高甲基化<sup>[28]</sup>)的甲基化状态可显著影响细胞放射敏感性,表明 DNA 甲基化模式有望成为预测肿瘤放射敏感性的分子标志物。

为系统研究 DNA 甲基化预测 RT 反应的临床价值,分析了 10 种肿瘤的全基因组甲基化数据,鉴定了 S 组与 R 组间的 DMSs。虽然多数 DMSs 仅具有中等预测特异性(AUC 0.6~0.8),但在特定癌种中发现

部分高特异性位点:BLCA 中 GJB6 基因(AUC = 0.956)、ESCA 中 LAT 基因(AUC = 0.929)和 SKCM 中 PRKAG2 基因(AUC = 0.889)的 DMSs 表现出优异预测效能。通过分析 TCGA 临床样本中已报道的 RT 反应相关基因(涉及 DNA 修复、细胞增殖等通路),发现其预测能力仅达中等水平(AUC 0.595~0.761)。这可能源于既往研究多基于肿瘤细胞系,而采用 TCGA 临床样本进行全基因组无偏倚分析更能准确反映肿瘤细胞的放射敏感性特征。

为进一步提升预测性能,采用机器学习算法构建癌种特异性组合模型,通过整合多个 DMSs 的特征实现模型优化。结果显示,测试集中样本平衡的癌种组合模型 AUC 值均超过 0.8,样本不平衡癌种 CESC、HNSC 预测模型 AP > 0.95。这种多特征交互考量策略显著提高了辐射反应预测的准确性。值得注意的是,这些机器学习模型不仅能预测 RT 敏感性,还可评估患者预后。预测评分高分组在 CESC ( $P < 0.001$ ) 和 LGG ( $P < 0.001$ ) 等癌种中均显示出显著生存差异,凸显了模型在生存结局预测中的潜力。DNA 甲基化标志物具有稳定性高的独特优势,且适用于 FFPE 样本和液体活检。与 RNA 标志物相比,DNA 甲基化分析更易于临床转化应用。本研究证实了 DNA 甲基化组模式作为肿瘤放射敏感性预测标志物的潜力,为评估放疗反应提供了稳定、非侵入性的新

方法。

## 参考文献

- [1] LIU B, ZHOU H, TAN L, et al. Exploring treatment options in cancer: tumor treatment strategies[J]. *Sig Transduct Target Ther*, 2024, 9: 175.
- [2] PRICE J M, PRABHAKARAN A, WEST C M L. Predicting tumour radiosensitivity to deliver precision radiotherapy[J]. *Nat Rev Clin Oncol*, 2023, 20: 83-98.
- [3] AVRIL D, FOY J, JEBRANE B, et al. Biomarkers of radioresistance in head and neck squamous cell carcinomas [J]. *Int J Radiat Biol*, 2022, 99: 583-593.
- [4] THIRUTHANEESWARAN N, BIBBY B A S, YANG L, et al. Lost in application: measuring hypoxia for radiotherapy optimisation[J]. *Eur J Cancer*, 2021, 148: 260-276.
- [5] KIM W J, VO Q N, MEENA S, et al. Aberrant methylation of the ATM promoter correlates with increased radiosensitivity in a human colorectal tumor cell line[J]. *Oncogene*, 2002, 21: 3864-3871.
- [6] ROY K, WANG L, MAKRIGIORGOS M, et al. Methylation of the ATM promoter in glioma cells alters ionizing radiation sensitivity[J]. *BBRC*, 2006, 344: 821-826.
- [7] TIAN Y, MORRIS T J, WEBSTER A P, et al. ChAMP: updated methylation analysis pipeline for Illumina Bead-Chips[J]. *Bioinformatics*, 2017, 33: 3982-3984.
- [8] ENGEBRETSEN S, BOHLIN J. Statistical predictions with glmnet[J]. *Clin Epigenetics*, 2019, 11: 123.
- [9] YIN J, DING F, CHENG Z, et al. METTL3-mediated m6A modification of LINC00839 maintains glioma stem cells and radiation resistance by activating Wnt/β-catenin signaling[J]. *Cell Death Dis*, 2023, 14(7): 417.
- [10] DOU Z, GHOSH K, VIZIOLI M G, et al. Cytoplasmic chromatin triggers inflammation in senescence and cancer [J]. *Nature*, 2017, 550(7676): 402-406.
- [11] TUFEGDŽIĆ V A, MITTER R, KELLY G P, et al. Regulation of the RNAPII Pool Is Integral to the DNA Damage Response[J]. *Cell*, 2020, 180(6): 1245-1261. e21.
- [12] HUANG R X, ZHOU P K. DNA damage response signaling pathways and targets for radiotherapy sensitization in cancer[J]. *Signal Transduct Target Ther*, 2020, 5(1): 60.
- [13] YOUSEFZADEH M J, FLORES R R, ZHU Y, et al. An aged immune system drives senescence and ageing of solid organs[J]. *Nature*, 2021, 594: 100-105.
- [14] ZHI G L, HUA Y C, JING J C, et al. Relationship between methylation status of ERCC1 promoter and radiosensitivity in glioma cell lines[J]. *Cell Biol Int*, 2009, 33: 1111-1117.
- [15] DONNINGER H, CLARK J, RINALDO F, et al. The RASSF1A tumor suppressor regulates XPA-mediated DNA repair[J]. *Mol Cell Biol*, 2014, 35: 277-287.
- [16] PABALAN N, KUNJANTARACHOT A, RUANGPR-
- ATHEEP C, et al. Potential of RASSF1A promoter methylation as biomarker for endometrial cancer: a systematic review and meta-analysis[J]. *Gynecol Oncol*, 2017, 146: 603-608.
- [17] HE X, MA Y, HUANG Z, et al. SERPINB5 is a prognostic biomarker and promotes proliferation, metastasis and epithelial-mesenchymal transition (EMT) in lung adenocarcinoma[J]. *Thorac Cancer*, 2023, 14: 2275-2287.
- [18] LUO J L, TAN W, RICONO JM, et al. Nuclear cytokine-activated IKKα controls prostate cancer metastasis by repressing Maspin[J]. *Nature*, 2007, 446: 690-694.
- [19] ZHENG J, XIONG D, SUN X, et al. Signification of hypermethylated in cancer 1 (HIC1) as tumor suppressor gene in tumor progression [J]. *Cancer Microenviron*, 2012, 5: 285-293.
- [20] ZHAI G, LI J, ZHENG J, et al. hTERT promoter methylation promotes small cell lung cancer progression and radiotherapy resistance[J]. *J Radiat Res*, 2020, 61: 674-683.
- [21] PECHERZEWSKA R, NAWROT B. FHIT: tumor suppressor protein involved in induction of apoptosis and cell cycle regulation[J]. *Postepy Biochem*, 2009, 55: 66-75.
- [22] LIN Y, HUNG S K, LEE M S, et al. DNA methylome analysis identifies epigenetic silencing of FHIT as a determining factor for radiosensitivity in oral cancer: an outcome-predicting and treatment-implicating study[J]. *Oncotarget*, 2015, 6: 915-934.
- [23] CHOI S I, KIM S Y, LEE J, et al. TM4SF4 overexpression in radiation-resistant lung carcinoma cells activates IGF1R via elevation of IGF1[J]. *Oncotarget*, 2014, 5: 9823-9837.
- [24] LIU C, TANG H, HU N, et al. Methylomics and cancer: the current state of methylation profiling and marker development for clinical care[J]. *Cancer Cell Int*, 2023, 23: 242.
- [25] BHOOTRA S, JILL N, SHANMUGAM G, et al. DNA methylation and cancer: transcriptional regulation, prognostic, and therapeutic perspective[J]. *Med Oncol*, 2023, 40(2): 71.
- [26] JI J, LI D, ZHAO X, et al. Genome-wide DNA methylation regulation analysis provides novel insights on post-radiation breast cancer[J]. *Sci Rep*, 2025, 15(1): 5641.
- [27] KIM E H, PARK A K, SEUNG M D, et al. Global analysis of CpG methylation reveals epigenetic control of the radiosensitivity in lung cancer cell lines[J]. *Oncogene*, 2010, 29: 4725-4731.
- [28] CHOUHEI S, MIYAGAWA K, FUKUDA K, et al. Frequent silencing of RUNX3 in esophageal squamous cell carcinomas is associated with radioresistance and poor prognosis[J]. *Oncogene*, 2007, 26: 5927-5938.