



doi:10.7659/j.issn.1005-6947.250298
http://dx.doi.org/10.7659/j.issn.1005-6947.250298
China Journal of General Surgery, 2025, 34(11):2368-2379.

· 基础研究 ·

基于凋亡和化疗相关基因的乳腺癌预后模型构建及PCDHB2在肿瘤干性中的作用

杨柳, 代阳阳, 王雅文, 张元元

(中国医科大学附属盛京医院 肿瘤中心, 辽宁 沈阳 110004)

摘要

背景与目的: 乳腺癌是全球女性发病率最高的恶性肿瘤之一, 化疗耐药和复发转移仍是影响患者长期生存的主要难题。细胞凋亡异常被认为是化疗耐药形成的重要机制, 但凋亡和化疗相关基因在乳腺癌预后评估及耐药形成中的协同作用尚缺乏系统性研究。本研究旨在构建基于凋亡和化疗相关基因的多基因预后风险模型, 并进一步探讨关键基因在乳腺癌肿瘤干性及化疗耐药中的潜在作用。

方法: 基于TCGA数据库获取乳腺癌转录组及临床数据, 筛选凋亡和化疗相关差异表达基因。通过单因素Cox回归、LASSO回归及多因素Cox回归构建多基因预后风险模型, 并采用时间依赖性ROC曲线、列线图及GEO外部数据集进行验证。同时分析风险评分与临床特征、免疫细胞浸润及化疗药物敏感性的关系。选取模型中的关键基因, 通过细胞划痕、Transwell及Western blot实验验证其对乳腺癌细胞迁移、侵袭及肿瘤干性相关蛋白表达的影响。

结果: 共筛选出10个与乳腺癌预后显著相关的凋亡和化疗关键基因, 成功构建预后风险模型。该模型在TCGA队列中对3、5年生存率的预测AUC分别为0.705和0.650, 并在GEO外部验证队列中表现出良好的稳定性(3年AUC=0.741)。风险评分为乳腺癌患者的独立预后因素, 并与晚期TNM分期、免疫抑制性细胞浸润增加及NU7441、氟达拉滨等化疗药物耐药显著相关。PCDHB2在高风险患者中高表达, 其沉默可显著抑制MDA-MB-231细胞的迁移和侵袭能力, 并下调干性相关蛋白SOX2的表达。

结论: 本研究构建了基于凋亡和化疗相关基因的多基因预后风险模型, 可用于乳腺癌患者的风险分层和个体化生存预测。PCDHB2可能通过调控肿瘤细胞干性参与化疗耐药过程, 具有潜在的临床转化价值, 为乳腺癌耐药机制研究及精准治疗策略的制定提供了新的理论依据。

关键词

乳腺肿瘤; 多基因风险模型; 免疫微环境; 肿瘤干性; 预后

中图分类号: R737.9

An apoptosis- and chemotherapy-related gene signature for prognosis prediction in breast cancer and the role of PCDHB2 in cancer stemness

YANG Liu, DAI Yangyang, WANG Yawen, ZHANG Yuanyuan

(Department of Oncology, Shengjing Hospital of China Medical University, Shenyang 110004, China)

收稿日期: 2025-05-29; 修订日期: 2025-11-23。

作者简介: 杨柳, 中国医科大学附属盛京医院主管护师, 主要从事乳腺癌防治相关方面的研究。

通信作者: 张元元, Email: sjreml666@163.com

Abstract

Background and Aims: Breast cancer is the most prevalent malignancy among women worldwide, and chemotherapy resistance and tumor recurrence remain major obstacles to long-term survival. Dysregulation of apoptosis is considered a key mechanism underlying chemoresistance; however, the synergistic roles of apoptosis- and chemotherapy-related genes in prognostic stratification and drug resistance have not been fully elucidated. This study aimed to construct a multigene prognostic risk model based on apoptosis- and chemotherapy-related genes and to further investigate the role of the key gene in breast cancer stemness and chemoresistance.

Methods: Transcriptomic and clinical data of breast cancer patients were obtained from The Cancer Genome Atlas (TCGA). Differentially expressed apoptosis- and chemotherapy-related genes were identified, and a multigene prognostic risk model was constructed using univariate Cox regression, LASSO regression, and multivariate Cox regression analyses. The predictive performance of the model was evaluated by time-dependent ROC curves, nomograms, and an external GEO dataset. Associations between the risk score and clinical characteristics, immune cell infiltration, and chemotherapeutic drug sensitivity were further analyzed. Functional assays, including wound healing, Transwell invasion, and Western blot analyses, were performed to validate the biological role of the key gene in breast cancer cells.

Results: A prognostic signature comprising 10 apoptosis- and chemotherapy-related genes was established. The model demonstrated favorable predictive performance in the TCGA cohort, with AUC values of 0.705 and 0.650 for 3- and 5-year overall survival rate, respectively, and was further validated in an external GEO dataset (3-year AUC=0.741). The risk score was identified as an independent prognostic factor and was significantly associated with advanced TNM stage, increased infiltration of immunosuppressive cells, and resistance to chemotherapeutic agents, including NU7441 and fludarabine. PCDHB2 was highly expressed in high-risk patients, and its knockdown markedly inhibited migration and invasion of MDA-MB-231 cells while reducing the expression of the stemness-related protein SOX2.

Conclusion: This study established a robust multigene prognostic model based on apoptosis- and chemotherapy-related genes for risk stratification and personalized survival prediction in breast cancer. PCDHB2 may contribute to chemoresistance by regulating cancer stemness, highlighting its potential as a novel therapeutic target and providing new insights into breast cancer precision therapy.

Key words

Breast Neoplasms; Multigene Risk Model; Immune Microenvironment; Cancer Stemness; Prognosis

CLC number: R737.9

乳腺癌是全球女性发病率最高的恶性肿瘤之一，每年新增病例约 230 万，死亡人数超 68 万^[1-2]。该疾病具有高度异质性，表现为组织学特征、分子分型（如 luminal 型、HER2 阳性型、三阴性型）及临床预后的显著个体差异^[3]。当前乳腺癌治疗以手术切除联合化疗（紫杉类/蒽环类）和靶向治疗为主，但约 30% 患者面临局部复发或远处转移，其中化疗耐药是导致治疗失败的核心因素^[4]。

细胞凋亡作为程序性死亡的核心机制，通过多级基因调控网络维持组织稳态，其异常与肿瘤发生发展密切相关^[5]。研究表明，凋亡相关基因

（如 Bcl-2 家族、caspase 家族）的表达失调可通过以下途径诱导乳腺癌耐药：(1) 靶分子突变（如 TP53 失活突变削弱化疗诱导的 DNA 损伤应答）；(2) 上下游信号蛋白异常（如 survivin 过表达抑制凋亡执行阶段）；(3) 旁路通路激活（如 PI3K/Akt 通路活化拮抗凋亡信号）^[6-7]。然而，目前尚缺乏基于凋亡与化疗相关基因特征的系统性研究，以解析其在乳腺癌预后分层和耐药中的协同作用。

在乳腺癌的发生与治疗抵抗中，凋亡和化疗相关基因的调控失衡是核心机制之一。促凋亡基因（如 BIM、BAX）通过响应化疗药物诱导的 DNA

损伤或微管稳定性改变, 激活线粒体凋亡通路, 但其表达下调或表达沉默(如启动子甲基化)可显著削弱紫杉类药物的疗效^[8]。抗凋亡基因(如 *Bcl-2*、*survivin*) 则通过抑制 caspase 级联反应或稳定线粒体膜电位, 阻断凋亡执行阶段, 导致蒽环类药物耐药。此外, 化疗代谢相关基因(如 *ABCB1* 编码的 P-糖蛋白) 通过主动外排药物降低细胞内浓度^[9], 而代谢重编程基因通过脂质合成重塑细胞膜结构并激活 PI3K/Akt 存活通路, 进一步促进肿瘤细胞在化疗压力下的适应性生存。这些基因的协同作用不仅驱动化疗耐药, 还与乳腺癌分子分型(如三阴性乳腺癌中 *FASN* 高表达) 及不良预后密切相关^[10], 但目前尚未建立基于多基因交互网络的预后预测模型。本研究旨在解析此类基因的协同调控特征, 为乳腺癌精准治疗提供新策略。

本研究基于 TCGA 乳腺癌队列的 RNA 测序数据与临床信息, 筛选出 10 个凋亡和化疗相关关键基因, 构建多基因预后风险模型。该模型可精准预测患者总生存期(overall survival, OS), 并通过体外实验证实 PCDHB2 通过肿瘤细胞干性介导化学治疗抵抗。这一发现为乳腺癌个体化治疗和耐药逆转策略提供了新靶点。

1 资料与方法

1.1 数据来源与处理

从 TCGA 数据库 (<https://portal.gdc.cancer.gov>) 下载 1 222 例乳腺癌组织样本的 RNA-seq 转录组数据及对应的临床信息(包括年龄、种族、TNM 分期等)。纳入标准:(1) 有完整的转录组数据;(2) 有临床随访信息。排除标准:(1) 临床信息缺失(如无生存时间或状态);(2) 低质量样本(基因表达 FPKM<1 在 80% 样本中未表达)。最终 1 083 例患者具有完整临床随访资料, 用于后续预后模型构建(单因素/LASSO/多因素 Cox 回归)。此外, 在风险评分计算后, 进一步排除风险值缺失的样本, 得到 909 例用于生存分析(基于风险评分中位数分组)。数据清洗及预处理均使用 R 语言(v4.1.2) 完成。

1.2 差异表达基因筛选与功能分析

采用 DESeq2 包比较乳腺癌组织与正常组织的基因表达差异, 以 $|\log_2FC| > 1$ 且 $FDR < 0.05$ 为筛选阈值, 从 GeneCards 数据库 (<https://www.genecards.org>) 中获取了 1 416 个与细胞凋亡和化疗相关的基因信

息^[11]。利用 clusterProfiler 包进行 GO 功能注释(生物过程、分子功能、细胞组分) 及 KEGG 通路富集分析, 筛选显著性条目 ($P < 0.05$)。通过 pheatmap 包绘制差异基因热图, ggplot2 包绘制火山图。

1.3 预后风险模型构建

首先, 对于单因素 Cox 回归分析, 将差异基因与患者 OS 进行关联分析, 筛选出与预后显著相关基因 ($P < 0.05$); 随后通过 LASSO 回归分析, 采用 glmnet 包对上述基因进行 10 折交叉验证^[12], 通过最小化似然偏差 (λ .min) 筛选关键基因, 避免模型过拟合; 最后, 采用多因素 Cox 回归分析对基于关键基因构建比例风险模型, 最终确定独立预后基因 ($P < 0.01$), 建立风险评分公式: $RiskScore = \sum_{i=1}^{10} Exp_i \times \beta_i$ 。其中, Exp_i : 模型中凋亡和化疗基因的表达值; β_i : 多因素 Cox 回归分析后凋亡和化疗基因的回归分析系数(coefficient 值)

1.4 模型验证与临床应用评估

首先进行生存分析, 根据风险评分中位数将 909 例患者分为高风险组 ($n=454$) 与低风险组 ($n=455$), 通过 survminer 包绘制 Kaplan-Meier 生存曲线, Log-rank 检验比较组间差异; 随后进行预测效能评估, 使用 timeROC 包计算 3、5、7 年 OS 率的 AUC 值, 评价模型预测准确率; 随后采用列线图方法进行构建与校准^[13], 结合风险评分及临床参数(TNM 分期、年龄) 建立列线图预测模型, 通过校准曲线(500 次 Bootstrap 重采样) 验证预测结果与实际生存率的一致性; 最后进行独立预后分析, 对纳入的 1 083 例患者进行单因素及多因素 Cox 回归, 以期证实风险评分是乳腺癌患者的独立预后因素。

1.5 临床特征与机制探索

采用 Wilcoxon 检验比较不同 TNM 分期、年龄分组的风险评分差异, 通过 ggpubr 包可视化箱线图; 基于 ssGSEA 算法^[14]计算: 14 种免疫细胞浸润评分, Spearman 相关性分析揭示风险评分与免疫微环境的关系; 最后, 对药物敏感性进行预测: 从 GDSC 数据库 (<https://www.cancerrxgene.org>) 获取药物 IC_{50} 值, 比较高低风险组对 NU7441、氟达拉滨等药物的敏感性差异。

1.6 外部数据集验证

从 GEO 数据库(GSE109710 数据集) 获取 134 例乳腺癌患者的表达谱数据, 使用相同模型计算风险评分并进行生存分析, 验证模型泛化能力。

1.7 试剂和仪器

胎牛血清 (海克隆 Hyclone 新西兰胎牛血清); L-15 培养基 (上海源叶生物科技有限公司); RIPA 细胞裂解液 (上海碧云天生物技术有限公司); Transwell 小室 (北京索莱宝生物科技有限公司); riboFECT™ CP Reagent (广州市锐博生物科技有限公司); 青霉素双抗 (上海泽叶生物公司); PCDHB2 小干扰 RNA (si-PCDHB2) 及其阴性对照 (si-NC) (广州市锐博生物科技有限公司); SOX2 单克隆抗体 (美国 Proteintech 公司); OCT4 单克隆抗体 (美国 Proteintech 公司); GAPDH (美国 Santa Cruz 公司)。倒置荧光显微镜 (日本尼康公司); 蛋白免疫印迹电泳仪 (美国 BIO-RAD 公司); 凝胶系统成像仪 (上海天能科技有限公司); 恒温摇床 (常州金坛良友仪器有限公司); 无 CO₂ 细胞培养箱 (美国赛默飞)。

1.8 实验方法

1.8.1 人三阴性乳腺癌 MDA-MB-231 细胞培养

MDA-MB-231 在 L-15 培养基中培养, 将细胞培养至培养皿底部的 80%~90% 的密度后, 将 MDA-MB-231 细胞随机分为两组: si-NC 组和 si-PCDHB2 组。通过转染试剂转染入对应的细胞组 24 h 后, 更换培养基。

1.8.2 Transwell 实验分析 MDA-MB-231 细胞侵袭能力

将转染完成的乳腺癌 MDA-MB-231 细胞按每组 1×10^5 的数量培养于上层 Transwell 室, 该室涂有基质胶, 培养基中不含胎牛血清。含有 5% 胎牛血清的完全培养基被添加到下腔体中作为化学吸引剂。孵化 24 h 后, 用预冷福尔马林固定 MDA-MB-231 细胞, 并用 1% 的结晶紫染色。从每个孔中数出下腔的迁移细胞, 并在显微镜下拍照。

1.8.3 细胞划痕实验分析 MDA-MB-231 细胞迁移能力

细胞转染完成后在 6 孔培养皿的盖面用记号笔勾画 2 条整齐直线并接种 MDA-MB-231 细胞, 在显微镜下使用 200 μ L 移液枪的枪头将标记部位直线内细胞划除, 然后弃掉漂浮的 MDA-MB-231 细胞, 更换新鲜培养基, 并分别在划痕 24 h 后拍照并计算细胞迁移距离, 细胞迁移率=原始距离-现阶段细胞间距离/原始距离。

1.8.4 Western blot 分析蛋白表达水平

收集转染完成的乳腺癌 MDA-MB-231 细胞, RIPA 试剂裂解细胞 30 min, 离心并吸取细胞上清后对总蛋白定量, 通过聚丙烯酰胺凝胶电泳分离不同分子量大小的目

的蛋白, 截取目的蛋白区域的凝胶与聚偏二氟乙烯膜贴合后进行电转。用 5% 的脱脂牛奶封闭, 用 TBST 缓冲液清洗, 然后用一抗孵育, 在 4 $^{\circ}$ C 下过夜。次日用 TBST 清洗 3 次经一抗处理的膜, 并在室温下与二抗孵育 1 h。将聚偏二氟乙烯膜暴露于增强化学发光试剂以观察蛋白表达情况。

1.9 统计学处理

所有分析均通过 R 语言完成。正态分布数据采用 *t* 检验或单因素方差分析, 非正态数据采用 Mann-Whitney *U* 检验或 Kruskal-Wallis 检验。组间连续变量的比较 (药物 IC₅₀ 值) 采用非参数 Mann-Whitney *U* 检验, 数据以中位值表示。生存分析采用 Kaplan-Meier 法, Cox 回归计算风险比 (*HR*)。P<0.05 为差异有统计学意义。

2 结果

2.1 差异表达基因筛选及功能注释

从 TCGA 数据库筛选出 989 个凋亡和化疗相关差异基因, 其中上调基因 493 个, 下调基因 496 个。热图显示差异基因在乳腺癌与正常组织中的表达模式存在显著差异 (图 1A), 火山图进一步验证差异基因分布 (图 1B)。GO 富集分析表明, 差异基因在生物学过程中显著富集于骨化、分枝结构形态发生及白细胞迁移 (P<0.05); 分子功能主要涉及细胞外基质结构成分、生长因子结合及肽酶活性; 细胞组分集中于内质网腔及分泌颗粒腔 (图 1C)。KEGG 通路分析显示, PI3K/Akt 信号通路、细胞因子-细胞因子受体相互作用及人乳头瘤病毒感染通路为关键调控途径 (图 1D)。

2.2 预后风险模型构建及性能验证

单因素 Cox 回归分析: 对 1 083 例患者生存数据与 989 个差异基因进行关联分析, 筛选出 103 个预后相关基因。GO 与 KEGG 富集显示这些基因与免疫调节及细胞周期通路相关 (图 2A-B)。随后进行 LASSO 回归筛选: 通过 10 折交叉验证 ($\lambda_{\min}=0.023$), 从 103 个基因中提取 30 个关键基因 (图 2C-D)。将 LASSO 回归分析中系数不为 0 的基因纳入多因素 Cox 回归分析, 最终筛选出 10 个基因, 以这 10 个基因为基础构建预后风险模型随后, 通过多因素 Cox 回归模型最终构建包含 10 个独立预后基因的评分模型 (图 2E-F), 如上所述, 最终选择 10 个基因建立预后风险模型, 计算公式

如下： $RiskScore = \sum_{i=1}^{10} Exp_i \times \beta_i = 0.402524847 \times (ZNF443) + 0.017657135 \times (ZNF750) + 0.069447928 \times (CD24) + 0.19684665 \times (EPB41L4B) + 0.187101327 \times (PCDHB2) + 0.122269472 \times (IVL) + 0.127091526 \times (MAL2) + 0.130181034 \times (NANOS1) + 0.112130759 \times (PITPNM3) + 0.065384069 \times (PSCA)$ 。根据中位风险值将 909 例患者分为高低风险组，高风险组 OS 率显著低于低风险组 ($HR=0.41, 95\% CI=0.29\sim 0.60, P<0.001$) (图 2G)。随后进行预测效能评估，时间依赖性 ROC 曲线显示，模型对 3、5、7 年 OS 率的 AUC 值分别为 0.705、0.650、0.657 (图 2H)。

2.3 预后模型的可视化与评估

基于风险评分将患者分为高低风险组 (图 3A)，生存状态散点图表明高风险组死亡事件密度显著

增加，风险热图进一步揭示 10 个关键基因 (如 *PCDHB2*、*ZNF443*) 在高风险组中普遍上调。通过列线图预测模型构建，整合风险评分、年龄及 TNM 分期构建列线图模型 (图 3B)，用于个体化生存概率预测。模型一致性指数 (C-index) 为 0.708 (95% $CI=0.683\sim 0.733$)，可以看到绘图得到的线条较好地拟合到了对角线上，证明该模型的预测效果较好，提示其具有较高的预测准确率。例如，总评分为 220 分的患者，其 3、5、7 年 OS 率分别为 80%、76%、50%。校准曲线显示 (图 3C)，模型预测的 3、5 年 OS 率与实际观察值高度吻合 (3 年：平均绝对误差=1.8%；5 年：平均绝对误差=2.3%)，拟合曲线与理想对角线接近 (Bootstrap 重采样 500 次， $P>0.05$)，表明模型预测结果无显著偏倚。

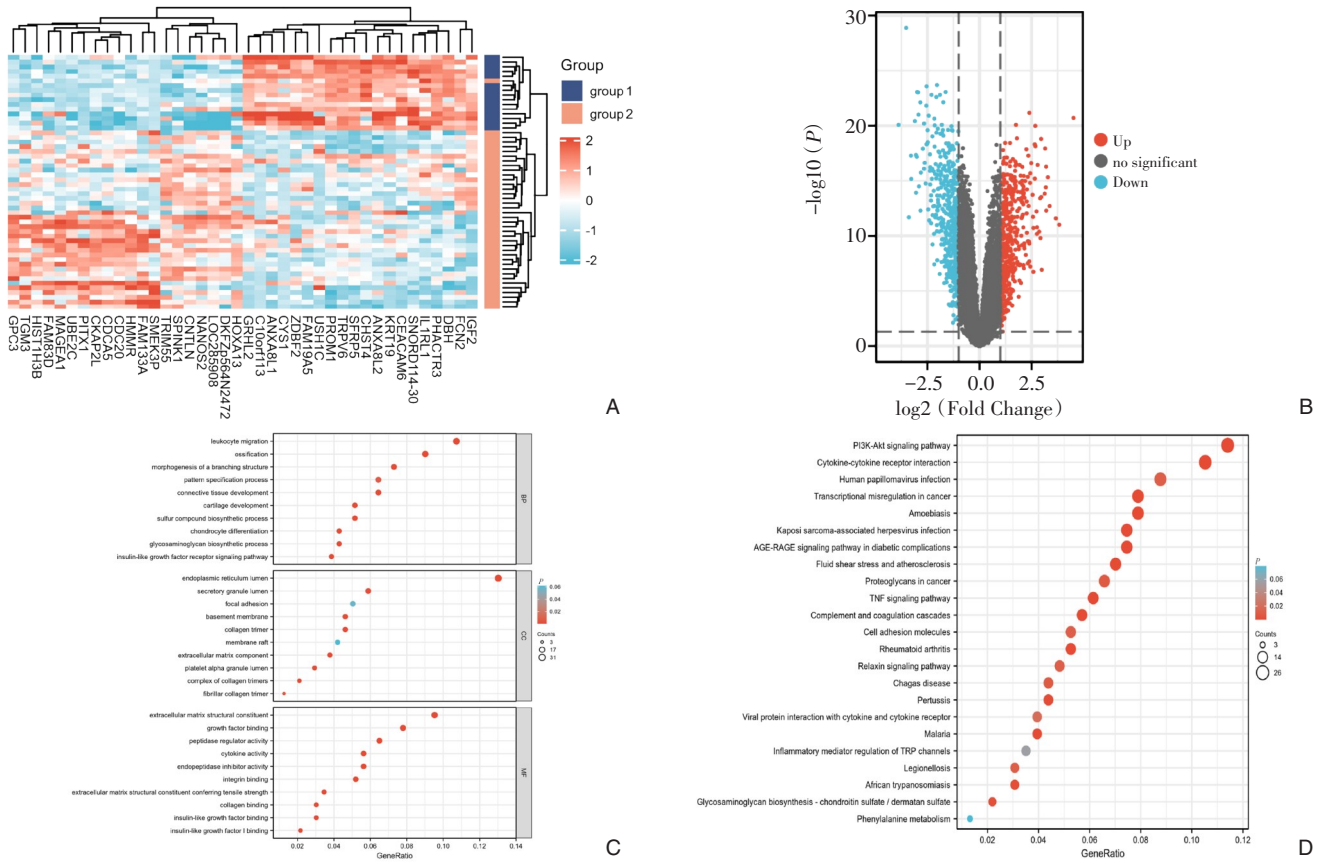


图 1 乳腺癌凋亡和化疗相关基因筛选及功能富集分析 A: 差异基因的表达热图; B: 差异表达火山图; C: GO 富集分析; D: KEGG 通路富集分析

Figure 1 Screening and functional enrichment analysis of apoptosis- and chemotherapy-related genes in breast cancer A: Heatmap of differentially expressed genes; B: Volcano plot of differentially expressed genes; C: GO enrichment analysis; D: KEGG pathway enrichment analysis

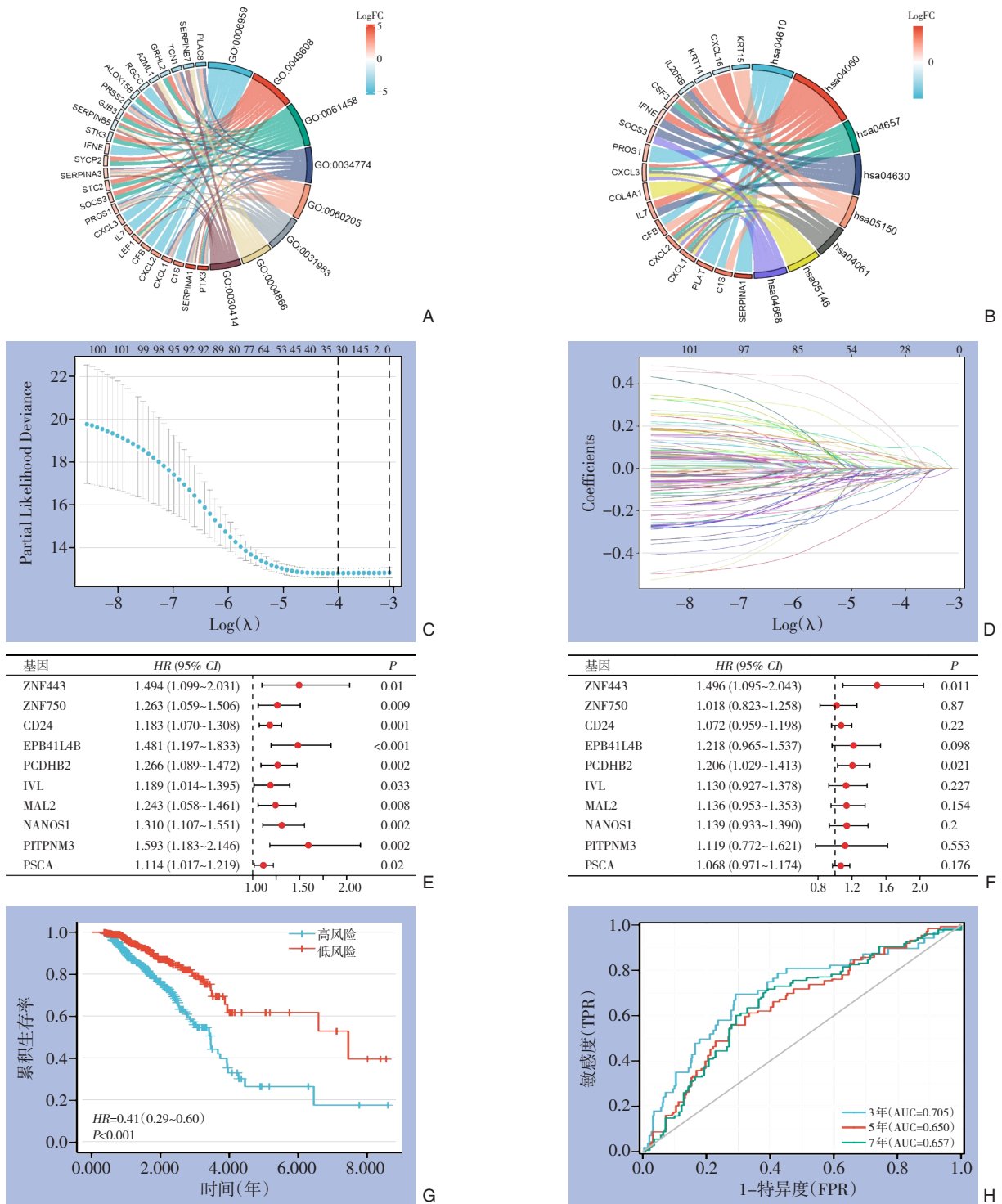


图 2 预后基因筛选及预后模型构建与验证 A: 103 个预后基因的 GO 富集分析; B: 103 个预后基因的 KEGG 富集分析; C: LASSO 系数筛选; D: LASSO 回归变量轨迹图; E: 10 个关键基因的单因素 Cox 回归分析森林图; F: 10 个关键基因的多因素回归分析森林图; G: 高、低风险组的生存曲线; H: 预测模型的 ROC 曲线

Figure 2 Identification of prognostic genes and construction and validation of the prognostic model A: GO enrichment analysis of 103 prognostic genes; B: KEGG enrichment analysis of 103 prognostic genes; C: LASSO coefficient profiles; D: LASSO regression variable trajectory plot; E: Forest plot of univariate Cox regression analysis for 10 key genes; F: Forest plot of multivariate Cox regression analysis for 10 key genes; G: Kaplan-Meier survival curves for high- and low-risk groups; H: Time-dependent ROC curves of the prognostic model

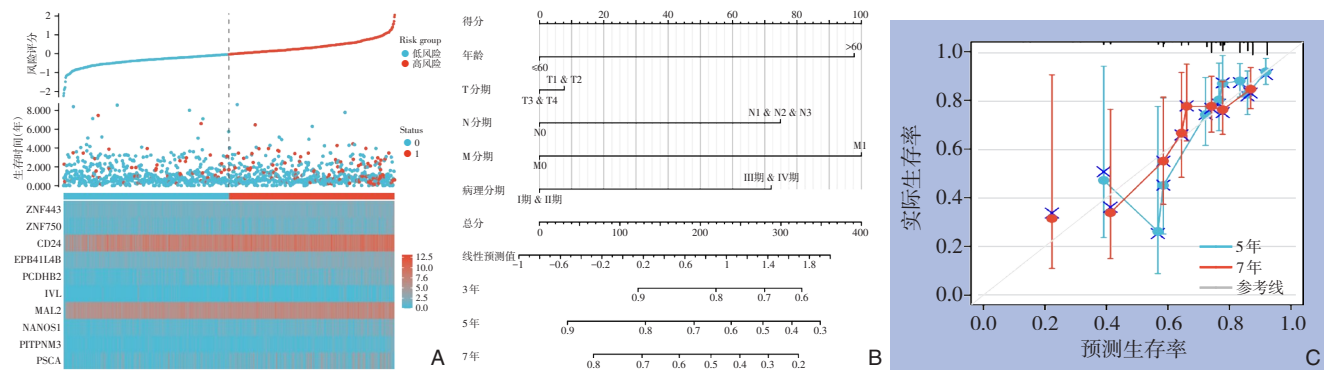


图3 预后模型的可视化分析 A: 风险因子图 (从上至下分别为风险曲线图、生存状态散点图、风险热图); B: 列线图; C: 列线图的校准曲线

Figure 3 Visualization and evaluation of the prognostic model A: Risk factor plot (from top to bottom: risk score distribution, survival status scatter plot, and heatmap of gene expression); B: Nomogram for predicting overall survival; C: Calibration curves of the nomogram

2.4 关键基因与临床特征、免疫细胞及敏感性的关系

首先通过单因素和多因素 Cox 分析结果进行独立预后分析, 多因素 Cox 回归证实, 风险评分是乳腺癌患者的独立预后因素 ($HR=2.31$, $95\% CI=1.86\sim 2.88$, $P<0.001$), 且与年龄 ($P<0.001$)、N 分期 ($P<0.001$) 明显有关 (图 4A-B)。临床相关性方面, PCDHB2 及 PSCA 在病理分期 III 和 IV 期中表达上调 (图 4C-D); 基因 IVL、NANOS1 及 ZNF443 在老年患者中高表达 ($P<0.05$) (图 4E-G), PCDHB2 及 PSCA 分别在 T 分期和 N 分期中表达上调 (图 4H-I)。随后, 分析了预后风险模型的 10 个关键基因与 T 细胞, 中性粒细胞, 树突状细胞, $CD8^+$ T 细胞, B 细胞, 抗原呈递树突状细胞, 肥大细胞, NK 细胞, 浆细胞样树突状细胞, Tgd 细胞, Th1 细胞, 巨噬细胞, Treg 细胞, 效应记忆 T 细胞等 14 种免疫细胞之间的相关性。其中 EPB41L4B、IVL、PSCA、ZNF443 和 MAL2 与 T 细胞, 中性粒细胞和 $CD8^+$ T 细胞密切相关, CD24、PCDHB2 和 ZNF750 与 B 细胞、肥大细胞和 Treg 细胞密切相关 (图 5A-J), 同时绘制了 10 个关键基因与免疫微环境相关性热图 (图 5K), 发现风险评分与 T 细胞 ($r=-0.23$, $P<0.001$)、 $CD8^+$ T 细胞 ($r=-0.18$, $P=0.002$) 浸润呈负相关, 与中性粒细胞 ($r=0.31$, $P<0.001$) 及巨噬细胞 ($r=0.27$, $P<0.001$) 浸润呈正相关。药物敏感性分析显示, 基于 GDSC 数据库的预测模型提示, 高风险组的预测 IC_{50} 标准化评分显著高于低风险组。对于 NU7441, 高风险组的中位评分 (0.39) 显著高于低风险组的中位评

分 (0.31) ($P<0.001$)。对于氟达拉滨, 高风险组的中位评分 (0.38) 同样显著高于低风险组 (0.29) ($P<0.001$)。该结果表明, 基于计算模型预测, 高风险组对这两种药物可能更具耐药性 (图 5L-M)。

2.5 外部数据集验证

基于 GSE109710 数据集 ($n=134$), 风险模型将患者分为高、低风险组 ($n=67/$ 组)。高风险组 OS 率显著降低 ($HR=1.89$, $95\% CI=1.12\sim 3.20$, $P=0.025$) (图 6A), 3、5、7 年 AUC 值分别为 0.741、0.660、0.719 (图 6B), 验证了模型泛化能力。

2.6 细胞实验验证

在本研究构建的乳腺癌凋亡和化疗相关多基因预后模型中, PCDHB2 被确定为 10 个独立预后相关关键基因之一, 其在高风险组患者中显著高表达。进一步分析表明, PCDHB2 的表达水平与多项临床不良特征密切相关: 其在晚期病理分期 (III~IV 期)、肿瘤更大 T 分期 (T3~T4) 和淋巴结转移 N 分期 (N1~N3) 患者中显著上调, 提示其可能参与肿瘤的侵袭与转移过程。此外, 免疫微环境分析显示, PCDHB2 与多种免疫抑制相关细胞 (如 Treg 细胞、B 细胞、肥大细胞) 呈正相关, 进一步暗示其可能通过影响肿瘤免疫逃逸与微环境重塑促进乳腺癌耐药, 同时 PCDHB2 与乳腺癌的关系尚不明确, 因此对 PCDHB2 进行细胞生物学实验验证。细胞划痕实验结果显示, 与 si-NC 组比较, si-PCDHB2 组的 MDA-MB-231 细胞的迁移能力降低 ($P<0.05$) (图 7A); Transwell 实验结果显示, 与 si-NC 组比较, si-PCDHB2 组的 MDA-MB-231 细胞的

侵袭能力降低 ($P<0.05$) (图 7B); Western blot 结果显示, 与 si-NC 组比较, si-PCDHB2 组的 MDA-MB-231

细胞的 SOX2 蛋白表达降低 ($P<0.05$), 而 OCT4 蛋白表达两组无明显差异 ($P>0.05$) (图 7C)。

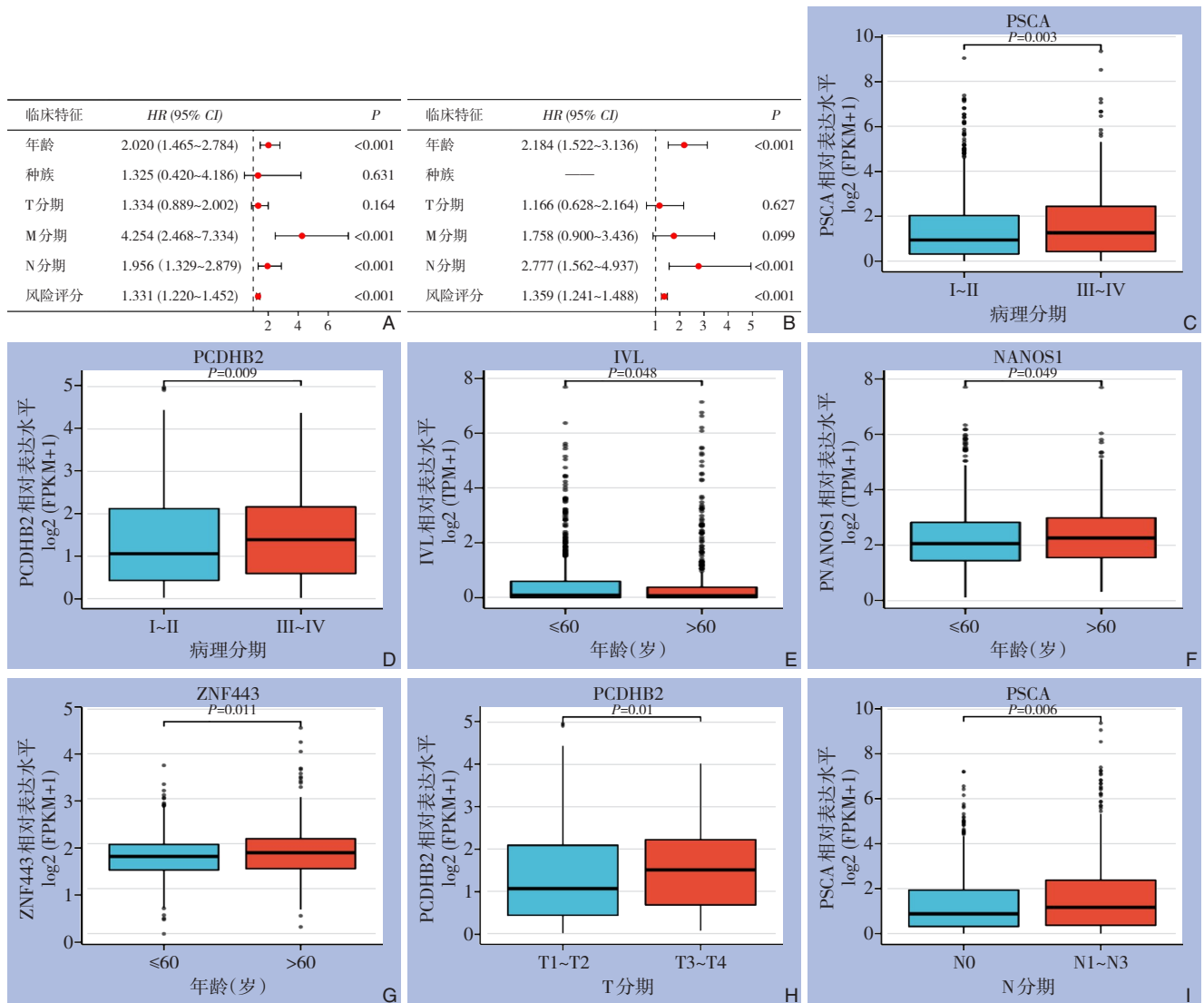


图 4 关键基因与临床特征的关系 A: 单因素 Cox 回归分析森林图; B: 多因素 Cox 回归分析森林图; C-D: PSCA 和 PCDHB2 在不同的病理分期表达差异; E-G: IVL、NANOS1、ZNF443 在不同年龄组的表达差异; H: PCDHB2 在不同 T 分期的表达差异; I: PSCA 在不同 N 分期的表达差异

Figure 4 Associations between key genes and clinical characteristics A: Forest plot of univariate Cox regression analysis; B: Forest plot of multivariate Cox regression analysis; C-D: Expression differences of PSCA and PCDHB2 across pathological stages; E-G: Expression differences of IVL, NANOS1, and ZNF443 across age groups; H: Expression differences of PCDHB2 across T stages; I: Expression differences of PSCA across N stages

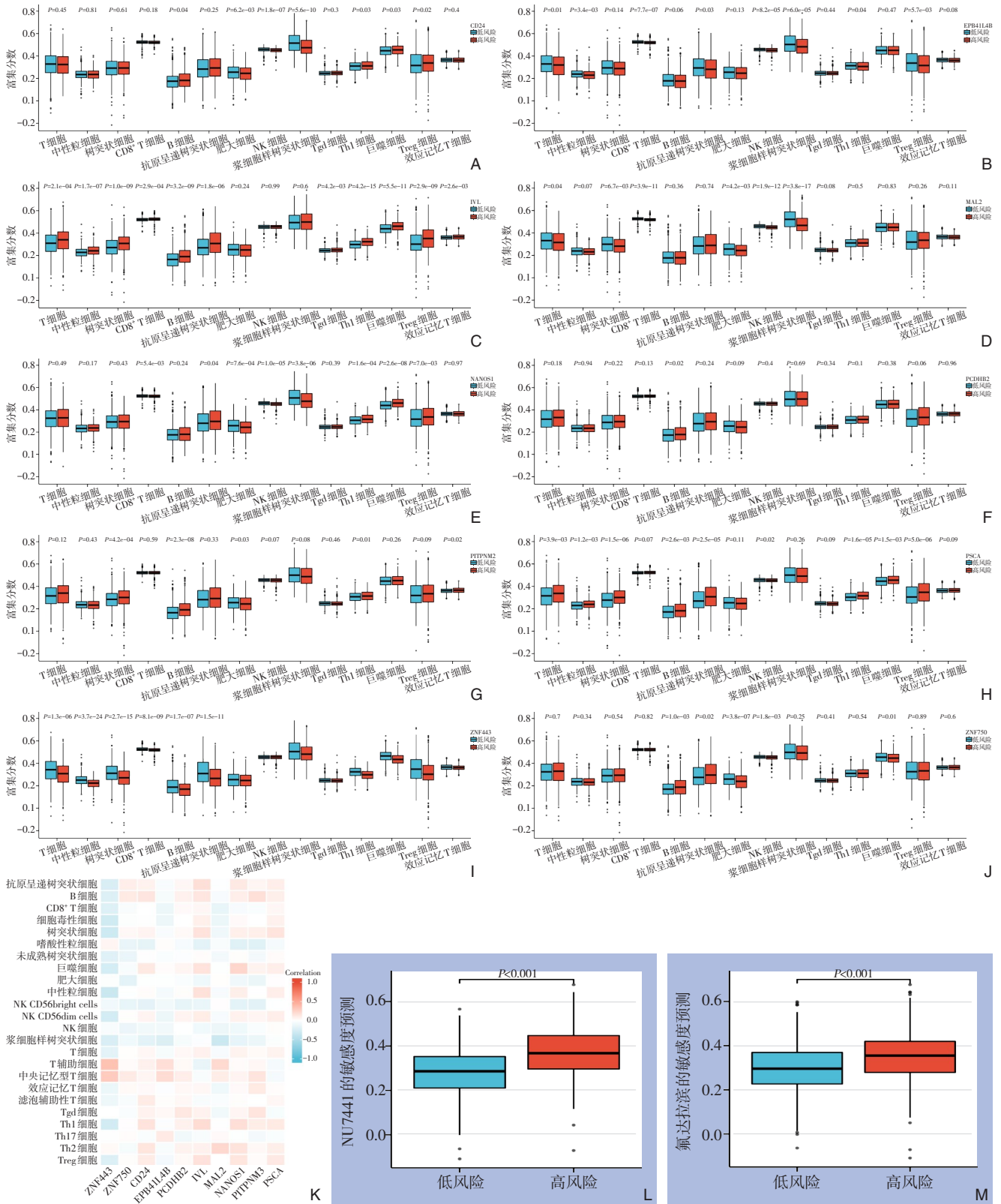


图5 关键基因与免疫细胞及药敏性的关系 A-J: 10个预后相关的关键基因与14种免疫细胞类型丰度的关系; K: 免疫细胞相关性热图; L: 高、低风险组对NU7441的敏感性的差异; M: 高、低风险组对氟达拉滨的敏感性的差异

Figure 5 Relationships between key genes, immune cell infiltration, and drug sensitivity A-J: Correlations between 10 prognostic genes and 14 immune cell types; K: Heatmap of correlations between immune cells and the prognostic model; L: Differences in predicted sensitivity to NU7441 between high- and low-risk groups; M: Differences in predicted sensitivity to fludarabine between high- and low-risk groups

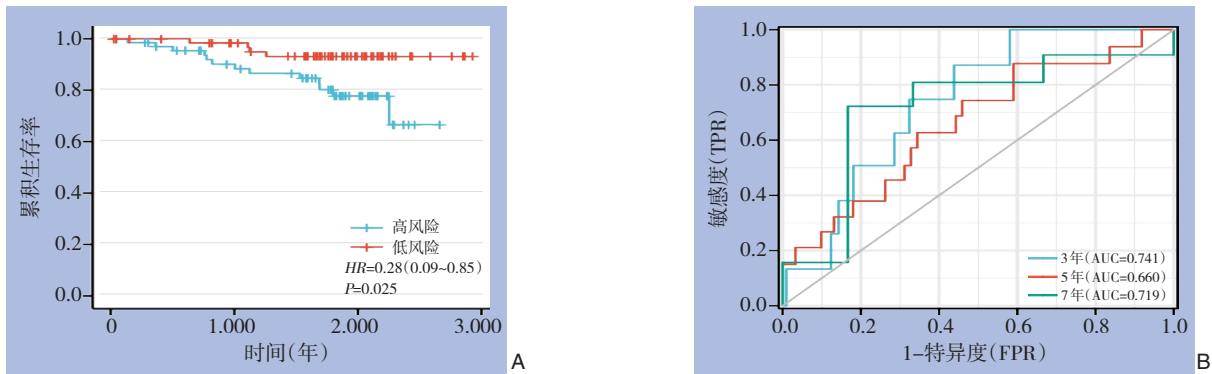


图 6 GEO数据库验证 A: 预后风险模型的Kaplan-Meier曲线; B: 预后风险模型的ROC曲线

Figure 6 External validation of the prognostic model using the GEO dataset A: Kaplan-Meier survival curves based on the prognostic risk model; B: Time-dependent ROC curves of the prognostic model

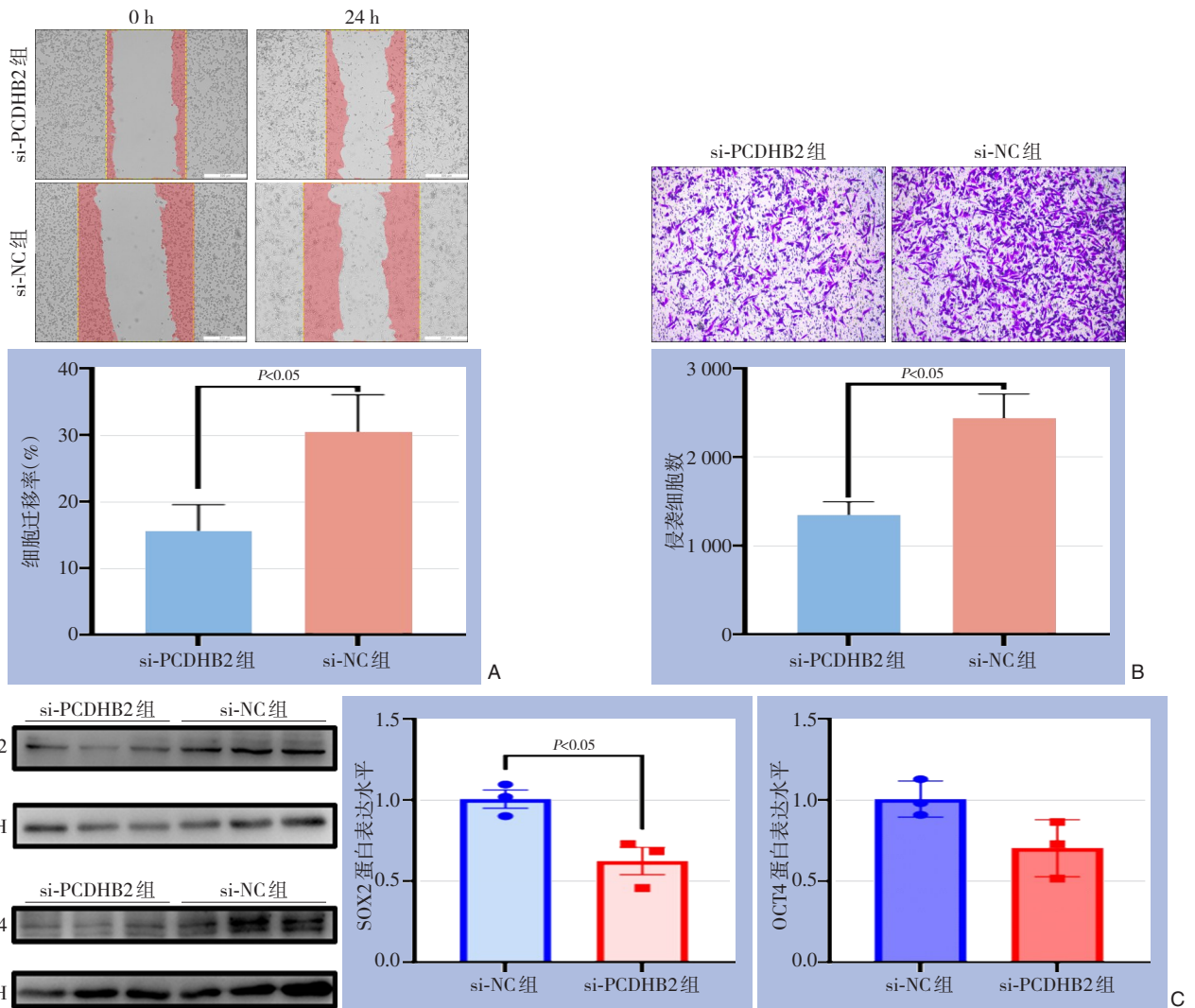


图 7 PCDHB2对乳腺癌MDA-MB-231细胞功能的影响 A: 细胞迁移实验; B: 细胞侵袭实验; C: SOX2与OCT4的蛋白表达检测

Figure 7 Effects of PCDHB2 on the biological behavior of MDA-MB-231 breast cancer cells A: Wound healing assay; B: Transwell invasion assay; C: Western blot analysis of SOX2 and OCT4 protein expression

3 讨论

乳腺癌是一种具有高度异质性的恶性肿瘤，

尽管当前以手术为基础的综合治疗取得了显著进展，但化疗耐药仍是临床治疗失败和复发转移的主要原因之一^[2,15]。本研究围绕凋亡和化疗相关基

因的协同作用,构建了多基因预后风险模型,并通过生物信息学分析与体外实验相结合的方式,系统揭示了其在乳腺癌中的分子机制及临床应用潜力。

首先,通过TCGA数据库的差异表达与生存分析,本研究筛选出10个与凋亡和化疗密切相关的关键基因(如*PCDHB2*、*CD24*、*ZNF443*等),并构建出可用于乳腺癌OS预测的多因素Cox回归风险评分模型。模型在TCGA队列中的生存分析和时间依赖性ROC分析结果显示,模型具有良好的预测能力(3年AUC=0.705,5年AUC=0.650),列线图校准图与实际生存率高度一致(误差<2.5%),提示该模型在乳腺癌患者预后分层和风险评估中具有广泛应用价值。值得关注的是,模型中多个基因(如*PCDHB2*、*PSCA*)在高风险组患者中显著上调,且其表达水平与晚期TNM分期及淋巴结转移密切相关,提示其可能在肿瘤进展过程中发挥驱动作用。

在构建预后风险模型的这10个基因中,已有研究^[16]表明,有的基因与部分癌症有着显著的相关性,如*CD24*、*MAL2*、*PSCA*已有研究证实与乳腺癌显著相关;通过构建*CD24*敲低及过表达三阴性乳腺癌模型,结合体外实验、体内分析及多组学(基因组、脂质组、代谢组)研究发现,*CD24*作为转移性TNBC的肿瘤抑制因子,其表达缺失通过激活PPAR α 介导的脂肪酸 β -氧化通路,诱导线粒体氧化磷酸化并驱动代谢重编程,进而抑制肿瘤进展。通过Meta分析^[17]发现,*CD24*高表达与乳腺癌患者OS率和无病生存率降低及淋巴结侵袭、TNM分期等临床病理因素显著相关,提示*CD24*可作为乳腺癌的有效预后因子。功能实验^[18]表明,敲低*MAL2*可通过抑制 β -catenin/c-Myc信号轴减少乳腺癌细胞迁移侵袭(与上皮-间充质转化相关)、促进细胞凋亡并降低体内转移能力,而 β -catenin激动剂SKL2001可部分逆转上述表型,证实*MAL2*通过调控 β -catenin/c-Myc通路驱动乳腺癌进展。同时,*MAL2*通过直接结合MHC-I复合体及内体相关RAB蛋白促进肿瘤抗原内吞,从而下调抗原呈递并介导免疫逃逸;临床前模型中敲除*MAL2*可显著增强肿瘤浸润CD8⁺T细胞毒性并抑制乳腺癌生长,提示靶向*MAL2*是增强肿瘤免疫治疗应答的潜在策略。笔者的免疫浸润结果也支持这一结论。

此外,本研究还揭示了预后模型与免疫微环境之间的密切关系:高风险评分与CD8⁺T细胞浸

润呈负相关,而与中性粒细胞和巨噬细胞浸润呈正相关。这一免疫特征提示高风险患者可能处于“冷肿瘤”微环境状态,提示其对免疫治疗反应较差,进一步强调了精准免疫干预的重要性。药物敏感性分析亦显示,高风险组对常见化疗药物NU7441和氟达拉滨的敏感性显著下降,提示上述基因的表达变化可能与多药耐药表型密切相关。

已有研究^[19-20]表明,*PCDHB2*在多种实体瘤中表达异常,并可能与细胞迁移、黏附及信号转导相关,但其在乳腺癌中的功能研究尚属空白。目前关于*PCDHB2*在乳腺癌中的分子机制报道极为有限,且其已被申请为乳腺癌骨转移相关专利靶点^[21],具备良好的研究前景和临床转化潜力。基于上述分析,本研究选择*PCDHB2*作为代表性验证基因,采用体外实验探讨其在乳腺癌细胞迁移、侵袭及干性维持等过程中的作用机制。初步实验结果亦证实,沉默*PCDHB2*可显著抑制三阴性乳腺癌细胞(MDA-MB-231)的迁移与侵袭能力,并下调干性相关蛋白SOX2的表达,提示*PCDHB2*可能通过维持肿瘤细胞干性介导化疗耐药。这为其作为乳腺癌潜在治疗靶点提供了坚实基础,这一发现为靶向乳腺癌干性细胞以克服化疗耐药提供了新思路。

尽管本研究构建的多基因风险模型经外部GEO队列验证,表现出良好的泛化能力,但仍存在一定局限性。首先,模型的构建与验证主要基于公开数据库,部分样本信息可能存在不完整或异质性;其次,*PCDHB2*等关键基因的作用机制仍需进一步通过动物模型和临床样本验证;最后,未来应结合单细胞测序技术深入解析不同细胞群在耐药过程中的转录特征。

综上所述,本研究系统揭示了凋亡和化疗相关基因在乳腺癌耐药与预后中的关键作用,构建的多基因预后模型具有良好的预测性能和临床应用潜力,尤其是*PCDHB2*作为潜在治疗靶点,为乳腺癌个体化治疗提供了理论依据和实践指导。后续研究应进一步探索关键基因在免疫微环境重塑与治疗应答中的作用,以推动乳腺癌精准治疗策略的临床转化。

作者贡献声明:杨柳负责研究的设计与实施,并起草了论文初稿,论文返回后的全部修改,并重新分析数据;张元元负责图表制作并参与了论文的起草,论文返

回后的全部修改,并重新制作了图表和部分数据重新分析;代阳阳参与了论文的起草与修改;王雅文负责研究的指导与管理,并对论文进行了修订。

利益冲突:所有作者均声明不存在利益冲突。

参考文献

- [1] Deng T, Zi H, Guo XP, et al. Global, regional, and national burden of breast cancer, 1990–2021, and projections to 2050: a systematic analysis of the global burden of disease study 2021[J]. *Thorac Cancer*, 2025, 16(9):e70052. doi:10.1111/1759-7714.70052.
- [2] Xiong X, Zheng LW, Ding Y, et al. Breast cancer: pathogenesis and treatments[J]. *Sig Transduct Target Ther*, 2025, 10:49. doi:10.1038/s41392-024-02108-4.
- [3] Akram M, Iqbal M, Daniyal M, et al. Awareness and current knowledge of breast cancer[J]. *Biol Res*, 2017, 50(1): 33. doi: 10.1186/s40659-017-0140-9.
- [4] Lee JS, Kim SI, Park HS, et al. The impact of local and regional recurrence on distant metastasis and survival in patients treated with breast conservation therapy[J]. *J Breast Cancer*, 2011, 14(3): 191–197. doi:10.4048/jbc.2011.14.3.191.
- [5] Elmore S. Apoptosis: a review of programmed cell death[J]. *Toxicol Pathol*, 2007, 35(4): 495–516. doi: 10.1080/01926230701320337.
- [6] Neophytou CM, Trougakos IP, Erin N, et al. Apoptosis deregulation and the development of cancer multi-drug resistance[J]. *Cancers (Basel)*, 2021, 13(17):4363. doi:10.3390/cancers13174363.
- [7] Chen X, Duan N, Zhang CG, et al. Survivin and tumorigenesis: molecular mechanisms and therapeutic strategies[J]. *J Cancer*, 2016, 7(3):314–323. doi:10.7150/jca.13332.
- [8] Wang N, Ma T, Yu B. Targeting epigenetic regulators to overcome drug resistance in cancers[J]. *Signal Transduct Target Ther*, 2023, 8(1):69. doi:10.1038/s41392-023-01341-7.
- [9] Tian YC, Lei YR, Wang YN, et al. Mechanism of multidrug resistance to chemotherapy mediated by P-glycoprotein (Review) [J]. *Int J Oncol*, 2023, 63(5): 119. doi: 10.3892/ijo.2023.5567.
- [10] Park JH, Han HS, Lim SD, et al. Fatty acid synthetase expression in triple-negative breast cancer[J]. *J Pathol Transl Med*, 2022, 56(2): 73–80. doi:10.4132/jptm.2021.10.27.
- [11] Zhao K, Xu G, Jin J, et al. Integrated prognostic assessment of apoptosis and chemotherapy related gene in bladder cancer: a prognostic signature[J]. *Discov Oncol*, 2025, 16(1): 718. doi: 10.1007/s12672-025-02581-5.
- [12] Klosa J, Simon N, Westermark PO, et al. Seagull: lasso, group lasso and sparse-group lasso regularization for linear regression models via proximal gradient descent[J]. *BMC Bioinformatics*, 2020, 21(1): 407. doi:10.1186/s12859-020-03725-w.
- [13] Balachandran VP, Gonen M, Joshua Smith J, et al. Nomograms in oncology: more than meets the eye[J]. *Lancet Oncol*, 2015, 16(4): e173–e180. doi:10.1016/S1470-2045(14)71116-7.
- [14] Yan L, Han X, Zhang M, et al. Integrative analysis of TBI data reveals Lgmn as a key player in immune cell-mediated ferroptosis[J]. *BMC Genom*, 2023, 24(1):747. doi:10.1186/s12864-023-09842-z.
- [15] Guo L, Kong D, Liu J, et al. Breast cancer heterogeneity and its implication in personalized precision therapy[J]. *Exp Hematol Oncol*, 2023, 12(1):3. doi:10.1186/s40164-022-00363-1.
- [16] Murthy D, Dutta D, Attri KS, et al. CD24 negativity reprograms mitochondrial metabolism to PPAR α and NF- κ B-driven fatty acid β -oxidation in triple-negative breast cancer[J]. *Cancer Lett*, 2024, 587:216724. doi:10.1016/j.canlet.2024.216724.
- [17] Liu G, Liu GX, Fang Y, et al. Clinicopathological and prognostic value of CD24 expression in breast cancer: a meta-analysis[J]. *Int J Biol Markers*, 2017, 32(2):182–189. doi:10.5301/ijbm.5000254.
- [18] An L, Gong H, Yu X, et al. Downregulation of MAL2 inhibits breast cancer progression through regulating beta-catenin/c-Myc axis[J]. *Cancer Cell Int*, 2023, 21, 23(1):144. doi:10.1186/s12935-023-02993-9.
- [19] Peinado-Serrano J, Quintanal-Villalonga Á, Muñoz-Galvan S, et al. A six-gene prognostic and predictive radiotherapy-based signature for early and locally advanced stages in non-small-cell lung cancer[J]. *Cancers (Basel)*, 2022, 14(9): 2054. doi: 10.3390/cancers14092054.
- [20] Zhang ZY, Zhang X. Identification of m6A-related biomarkers associated with prognosis of colorectal cancer[J]. *Med Sci Monit*, 2021, 27:e932370. doi:10.12659/MSM.932370.
- [21] 任栋, 李宇明, 杨春潇, 等. 用于乳腺癌骨转移诊疗的新靶标 PCDHB2:CN201910569944.3[P]. 2020-01-17.
Ren D, Li YM, Yang CX, et al. A new target for diagnosing and treating of bone metastasis from breast cancer: PCDHB2: CN201910569944.3[P]. 2020-01-17.

(本文编辑 宋涛)

本文引用格式:杨柳,代阳阳,王雅文,等.基于凋亡和化疗相关基因的乳腺癌预后模型构建及PCDHB2在肿瘤干性中的作用[J].中国普通外科杂志,2025,34(11):2368–2379. doi: 10.7659/j.issn.1005-6947.250298

Cite this article as: Yang L, Dai YY, Wang YW, et al. An apoptosis- and chemotherapy-related gene signature for prognosis prediction in breast cancer and the role of PCDHB2 in cancer stemness[J]. *Chin J Gen Surg*, 2025, 34(11): 2368–2379. doi: 10.7659/j.issn.1005-6947.250298