

基于机器学习探究子宫内膜癌免疫微环境与基因表达的关联及其对预后的预测价值

林海宏¹, 郭苑莉², 潘如¹, 雷南香¹, 曾维红¹

1. 梅州市人民医院妇科, 广东 梅州 514130; 2. 广东药科大学附属第一医院妇产科, 广东 广州 510699

【摘要】目的:探究子宫内膜癌(Endometrial Cancer, EC)免疫微环境与基因表达的关联及其对预后的预测价值,通过生物信息学分析与机器学习技术,识别关键的免疫相关基因,构建预后模型,以期为EC的个性化治疗提供新的方向。**方法:**基于TCGA数据库,采用DESeq2、edgeR和limma工具筛选差异表达基因,结合ImmPort数据库筛选免疫相关基因。利用Lasso回归、单变量特征选择、Boruta和随机森林等机器学习算法进行特征基因筛选。通过单因素和多因素Cox回归分析评估基因的预后价值,并构建风险评分模型。采用CIBERSORT算法分析肿瘤免疫浸润,通过免疫组化验证关键基因表达。**结果:**通过3种差异分析结果与免疫相关基因的交集,确定62个差异表达的免疫基因,并使用多种机器学习模型筛选得到25个潜在生物标志物,其中被机器学习筛选为预后相关基因。单因素和多因素Cox回归分析证实,INHBE、SLURP1和TNFSF11基因与EC患者的生存期显著相关。构建的风险评分模型能够有效区分不同预后组别的生存率,且与免疫细胞浸润程度相关。免疫组化分析进一步验证这些基因在肿瘤与正常组织间的表达差异。**结论:**INHBE、SLURP1和TNFSF11是EC免疫微环境中关键的预后生物标志物,其表达水平与免疫细胞浸润及患者生存率密切相关,为EC的精准医疗提供理论基础。

【关键词】子宫内膜癌;机器学习;肿瘤基因组图谱;免疫基因;预后

【中图分类号】R318;R711

【文献标志码】A

【文章编号】1005-202X(2024)12-1568-10

Investigation of the association between endometrial cancer immune microenvironment and gene expression based on machine learning and its predictive value for prognosis

LIN Haihong¹, GUO Yuanli², PAN Ru¹, LEI Nanxiang¹, ZENG Weihong¹

1. Department of Gynaecology, Meizhou People's Hospital, Meizhou 514130, China; 2. Department of Gynaecology, the First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou 510699, China

Abstract: Objective To investigate the association between the immune microenvironment and gene expression in endometrial cancer (EC) and discuss its predictive value for prognosis, and to identify critical immune-related genes through bioinformatics analysis and machine learning techniques and construct a prognostic model for providing new directions for personalized treatment of EC. **Methods** Based on data from the Cancer Genome Atlas (TCGA) program, tools such as DESeq2, edgeR, and limma were utilized to screen for differentially expressed genes. Immune-related genes were selected by integrating data from the Immunology Database and Analysis Portal (ImmPort). Machine learning algorithms including Lasso regression, univariate feature selection, Boruta and random forest were employed to refine the selection of feature genes. Univariate and multivariate Cox regression analyses were conducted to assess the prognostic value of the genes, followed by construction of a risk score model. Additionally, tumor immune infiltration was analyzed using CIBERSORT algorithm, and key gene expressions were validated through immunohistochemistry. **Results** The intersection of 3 difference analysis results and immune-related genes identified 62 differentially expressed immune genes, and 25 potential biomarkers which were selected by a variety of machine learning models were considered as prognosis related genes. Univariate and multivariate Cox regression analyses confirmed that INHBE, SLURP1 and TNFSF11 genes were significantly associated

【收稿日期】2024-09-28

【基金项目】广东省医学科研基金(A2022318);广东省基础与应用基础研究基金(2023A1515220157);梅州市社会发展科学计划(2022C0301003)

【作者简介】林海宏,主治医师,研究方向:妇科肿瘤,E-mail: l_haihong675057@126.com

【通信作者】曾维红,硕士,主任医师,研究方向:妇科肿瘤,E-mail: pasta15914929268@163.com

with survival in EC patients. The constructed risk score model effectively distinguished the survival rate of different prognostic groups, and was related to the degree of immune cell infiltration. Immunohistochemical analysis further verified the differences in the expression of these genes between tumor and normal tissues. **Conclusion** INHBE, SLURP1 and TNFSF11 are key prognostic biomarkers in EC immune microenvironment, and their expression levels are closely associated with immune cell infiltration and patient survival rate, providing theoretical basis for EC precision medicine.

Keywords: endometrial cancer; machine learning; the Cancer Genome Atlas program; immune gene; prognosis

前言

子宫内膜癌(Endometrial Cancer, EC)作为全球女性最常见的妇科恶性肿瘤之一,其发病率和死亡率的逐年攀升已引起国际社会的广泛关注^[1]。随着生物信息学与机器学习技术的飞速进步,这些前沿科技正在深刻改变大家对EC的理解与应对策略,尤其在预后生物标志物的发掘与疾病的精确诊断上展现出了前所未有的潜力。生物信息学方法,如高通量测序技术与生物数据库的深入挖掘,为EC研究提供海量基因表达数据。通过差异表达基因的筛选,研究者们已鉴定出一系列与EC发生、发展紧密相连的关键基因。其中,MYBL2^[2]、KRT7^[3]、PAX8^[4]、HNF-1B^[5]和NAPSA^[6]等基因,因表达水平与EC的临床病理特征高度相关而备受瞩目。MYBL2的高表达被证实与EC患者的总生存期和无病生存期显著相关,暗示其作为预后不良的潜在生物标志物的价值。

在个体基因之外,复合生物标志物,如由多个长链非编码RNA构成的风险评分模型,展现出更强大的预测能力^[7]。基于肿瘤基因组图谱(The Cancer Genome Atlas Program, TCGA)数据库的深入分析,IGFBP4的低表达被识别为EC预后不良的独立风险因素^[8],其机制可能涉及胰岛素样生长因子系统的失衡,进而影响肿瘤细胞的增殖与迁移^[9]。机器学习技术,如支持向量机^[10]、随机森林(Random Forest, RF)^[11]和神经网络^[12],以其卓越的数据处理与模式识别能力,被广泛应用于这些生物标志物的挖掘与验证过程中。RF模型在预测EC病理分级时展示出的高准确率和特异度,以及基于多模型扩散加权成像的EC诊断与预后评估工具的开发,均有力证明机器学习在EC研究领域的巨大潜力^[13]。

EC的分子机制错综复杂,涉及众多基因与信号通路的交互作用。通过生物信息学与机器学习的融合应用,研究者得以从纷繁复杂的基因表达数据中甄别出决定EC进展与预后的关键基因和生物标志物。这些发现不仅深化对EC发病机制的理解,更为

疾病的早期诊断、风险分层与个性化治疗提供坚实的科学依据。未来,随着生物学大数据的不断累积与算法模型的持续优化,机器学习在肿瘤学领域的影响力将持续扩大,有望引领EC诊疗进入一个更加精准、高效的时代。

1 资料与方法

1.1 数据收集

从TCGA中筛选获得EC转录组和患者对应的临床数据(正常组织=35,肿瘤组织=545),其中包含545份EC肿瘤样本和35份癌旁组织的测序数据,并且进行原始数据的读取、清洗、整合,确保数据的质量和一致性。

1.2 差异表达分析与特征基因筛选

在追求精准识别疾病生物标志物的过程中,本研究采取多方位的分析策略,确保结果的稳健性和可靠性。首先,通过综合运用DESeq2、edgeR和limma这3种公认的差异表达分析工具,可以对RNA-seq数据的counts矩阵进行全面而细致的统计检验。设定 $\log_2\text{FoldChange}>2$ 与调整后P值(Padj) <0.01 的严格标准,甄选出在不同病理状态下显著变化的基因,并绘制火山图展示每种软件计算得到的差异基因。通过计算3套分析结果的交集,确定一组共同显著的差异表达基因,构建UPSET图展示各工具筛选出的基因集间的重叠情况。

1.3 免疫相关差异基因筛选

从免疫学数据库和分析门户(ImmPort)数据库(<https://www.immport.org/>)下载2499个免疫基因,与差异分析共同显著的差异表达基因取交集,进一步筛选出差异表达的免疫基因,使用ComplexHeatmap R包绘制差异基因热图,同时标注免疫基因。

1.4 富集分析与功能注释

采用ClusterProfiler包^[14]对免疫相关差异基因做GO与KEGG富集分析,通过ggplot2绘制气泡图和柱形图展现免疫基因集在肿瘤进程中的功能关联,评估其生物学意义。

1.5 机器学习筛选特征基因集

使用 geneSelectR 包进行机器学习筛选,实验采用4种机器学习算法:Lasso回归、单变量特征选择、Boruta和RF,对TCGA数据集进行特征基因筛选。首先,通过VarianceThreshold过滤低方差特征,并应用MinMaxScaler标准化数据。随后,数据经由重复的训练-测试分割,每轮分割都进行模型管道训练及超参数调优,采用网格搜索、随机搜索或贝叶斯优化,通过绘制Barplot展示不同模型的特征基因。通过重叠系数(Overlap Coefficient)、杰卡德相似系数(Jaccard Coefficient)、斯科伦-迪克森系数(Sørensen-Dice Coefficient)评估不同模型之间的异质性,选择合适的模型数据交集作为机器学习筛选到的特征基因集。

1.6 单因素和多因素Cox回归分析

通过机器学习算法筛选出潜在生物标志物后,应用单因素Cox回归分析初步评估每个基因与生存率的关联,选择 P 值小于预设阈值(0.05)的基因进入后续分析。随后,执行多因素Cox回归,调整混杂因素,确认独立预后基因。最后,利用回归系数构建森林图,展示各基因的危险比及其95%置信区间,直观呈现其对生存影响的统计学意义。

1.7 生存分析、构建风险评分模型及验证

为细致解析交集基因在健康对照与病患群体中的表达谱异同,采用easyTCGA包构建表达配对连线图,直观展现目标基因在不同生理状态下的动态变化,揭示其潜在的生物学功能及疾病相关性。随后,构建Kaplan-Meier(KM)生存曲线模型,系统评估特定基因表达量与患者生存结局的定量关系,验证分子标记的临床决策价值。

KM曲线模型依据log-rank检验统计功效,精准衡量高、低表达组间生存期的差异,当观察到显著性分离趋势时,暗示相应基因深度嵌入疾病发展路径,提示其作为诊断指标或治疗靶点的潜力。为进一步增强模型预测性能及可靠性,引入受试者工作特征(ROC)曲线分析,通过计算曲线下面积(AUC),客观评价基因表达水平区分健康与疾病状态的能力,确保所选标志物在临床实践中的实用性与准确性。

1.8 肿瘤免疫浸润分析及预后模型相关性分析

本研究聚焦于肿瘤免疫浸润分析及预后模型的相关性探讨。首先,运用CIBERSORT算法^[15]处理癌症患者的基因表达数据,精确估计22种免疫细胞类型在肿瘤微环境中的相对比例^[16]。其次,根据免疫细胞浸润数据,应用 P 值过滤标准($pFilter=0.05$),排

除非显著结果并整合基因表达与免疫浸润数据,执行Spearman相关性分析^[17],识别基因表达与各免疫细胞类型浸润程度间的潜在联系,对于显著相关的结果,绘制散点图并附加边际密度图,深入解析基因表达与免疫微环境的关系。

1.9 预后模型基因免疫组化

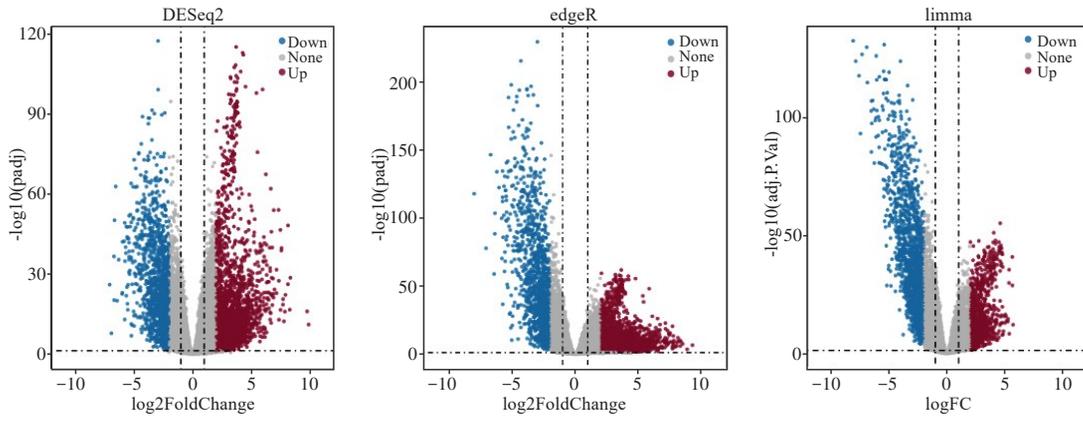
通过人类蛋白图谱(Human Protein Atlas, HPA)数据库,下载TNFSF11(Antibody HPA045142)、INHBE(Antibody CAB034017)和SLURP1(Antibody HPA050967)子宫内膜中肿瘤与正常样本的免疫组化切片图像,对比基因表达在两种状态下的异同,进而为理解基因功能及开发针对性治疗策略提供形态学证据。

2 结果

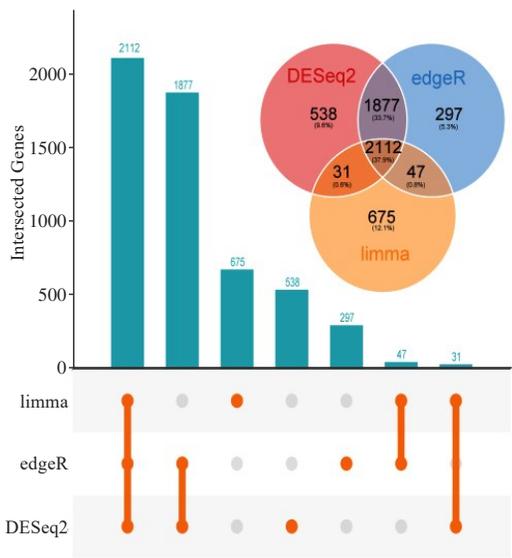
2.1 EC免疫差异基因筛选和GO富集分析、KEGG信号通路分析

从TCGA中获取545例EC和35例正常组织的转录组数据,经过数据清洗和整合,确保数据的一致性和质量。接下来,采用DESeq2、edgeR和limma这3种主流的差异表达分析工具,对数据进行差异分析。结果显示,DESeq2、edgeR和limma分别筛选出4558、4333和2865个差异表达基因(图1a)。通过交叉分析,研究者们找到2112个共同显著的差异表达基因,并绘制UPSET图来展示各个工具筛选出的基因集的重叠情况(图1b)。此外,从免疫学数据库和分析门户(ImmPort)下载2499个免疫基因,与上述差异表达基因进行交集操作,得到62个差异表达的免疫基因。最后,使用ComplexHeatmap R包^[18]绘制差异基因热图展示基因的表达模式(图1c)。

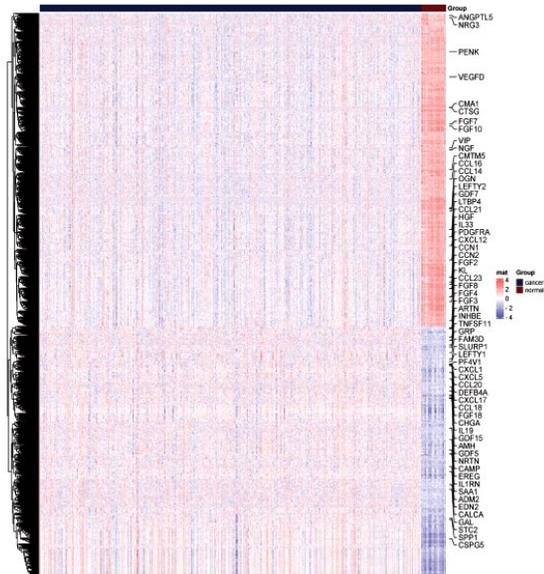
GO富集分析的结果显示(图2a),免疫差异基因主要富集在纤维母细胞生长因子受体信号通路、淋巴细胞趋化、细胞间信号传导、细胞化学趋向性、单核细胞化学趋向性、ERK1和ERK2级联激活的正调节、钙离子信号传导、细胞外基质、细胞因子活性等功能。KEGG通路富集分析的结果显示(图2b),免疫差异基因主要富集在细胞因子-细胞因子受体相互作用、Rap1信号通路、Ras信号通路、MAPK信号通路、PI3K-Akt信号通路、NF-kappa B信号通路、神经活性配体-受体相互作用、细胞因子信号通路、IL-17信号通路等通路。其中,PI3K-Akt信号通路失调的代谢性疾病在报道中有作为EC治疗新策略^[19],如NLRC5通过激活EC中的PI3K/AKT信号通路来促进细胞迁移和侵袭^[20],与筛选基因集功能相互印证。



a: 3种差异分析火山图



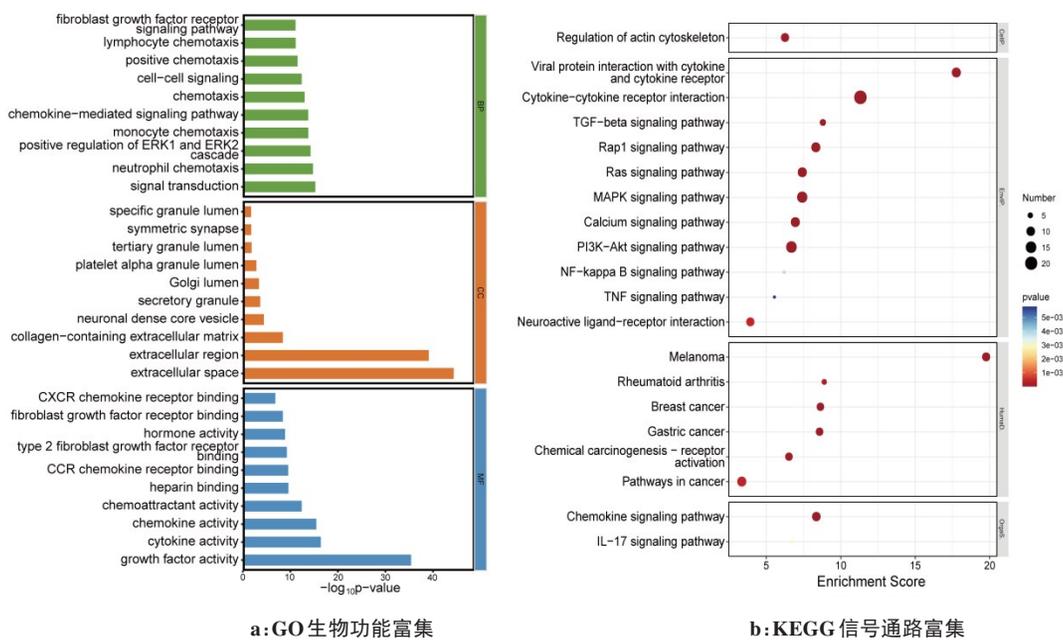
b: 3种差异分析结果交叉分析



c: 差异基因表达热图

图 1 EC 免疫差异基因筛选

Figure 1 Screening of differentially expressed immune genes in EC



a: GO 生物功能富集

b: KEGG 信号通路富集

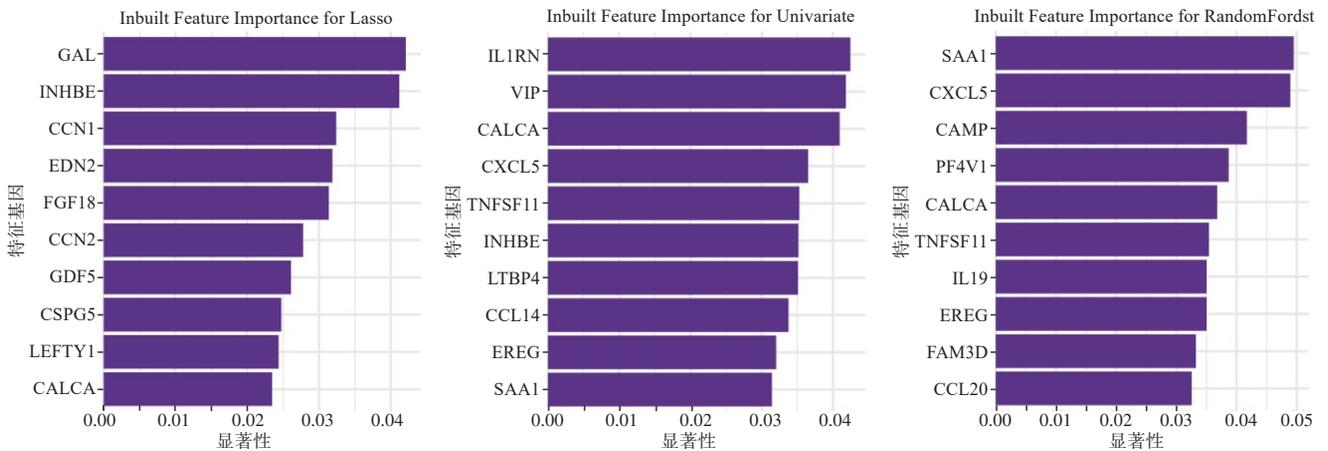
图 2 GO 富集分析、KEGG 信号通路分析

Figure 2 GO enrichment analysis and KEGG pathway analysis

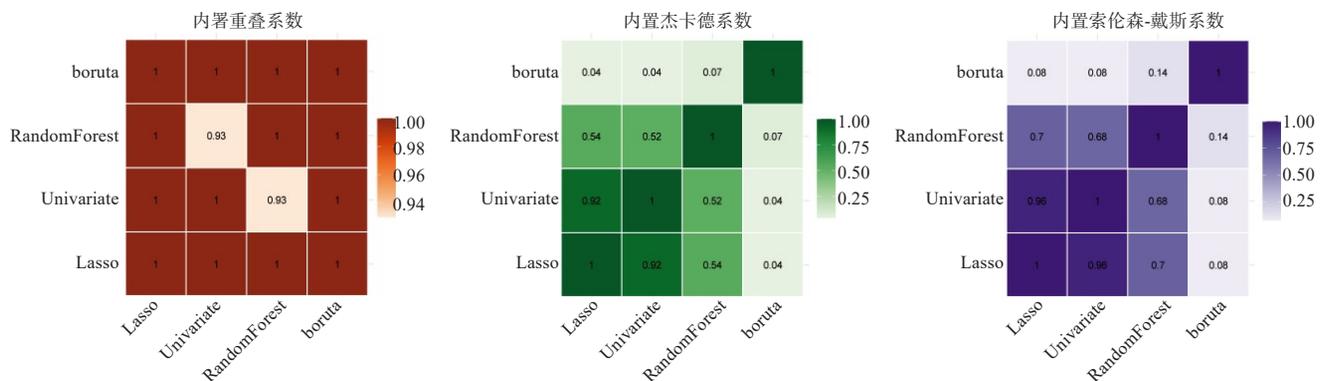
2.2 机器学习筛选特征基因

采用GeneSelectR包中的Lasso^[21]、Univariate^[22]、RandomForest^[23]和boruta^[24]4种机器学习方法以及算法对差异表达基因进行特征选择,图3a显示每个算法内部的重要度指标,并使用重叠系数、杰卡德相似

系数、斯科伦-迪克森系数度量两个集合之间共同元素比例以及集合之间相似度。基于图3b的结果,由于只有boruta算法筛选出一个基因,因此将其排除在外。最终选择由Lasso、Univariate和RandomForest 3种算法的交集作为特征基因,共筛选出25个基因。



a: 机器学习筛选特征基因 top10 条形图



b: 机器学习分析结果交叉分析热图

图3 机器学习筛选特征基因

Figure 3 Machine learning for feature gene selection

2.3 单因素和多因素Cox回归分析

使用单因素Cox回归分析以初步评估机器学习算法筛选出的25个潜在生物标志物与生存率的关系。结果显示,有多个基因的P值小于预设阈值0.05,具体为:INHBE、CALCA、CSPG5、IL1RN、SLURP1、TNFSF11(图4)。这些基因显示出与生存率存在显著关联,暗示它们可能是EC的关键分子标记。

为了进一步考虑混杂因素的影响,以确定具有独立预后的基因,对单因素Cox回归分析结果进行多因素Cox回归分析。结果显示,3个基因的P值小于0.05,分别为INHBE(P=0.01)、SLURP1(P=0.042)和TNFSF11(P=0.007)(图5),这表明它们在患者的生

存期预测中有统计学意义。

2.4 生存分析、构建风险评分模型及验证

为了深入理解交集基因在健康对照和病患群体中的表达谱差异,首先运用easyTCGA包生成表达配对连线图。结果显示,INHBE、SLURP1和TNFSF11在正常组织和肿瘤组织间的表达量存在显著差异(Kruskal-Wallis检验,P值分别为2.7e-05、0.0012和6.9e-07)(图6a)。接着通过构建Kaplan-Meier生存曲线来评估这些基因表达量与患者生存结局的定量关系(图6b)。值得注意的是,高表达组的生存概率普遍低于低表达组,这表明这些基因可能与较差的生存结局有关。为了验证这些分子标记的临床决策价值,进一步进行受试者工作特征(ROC)曲线分析(图

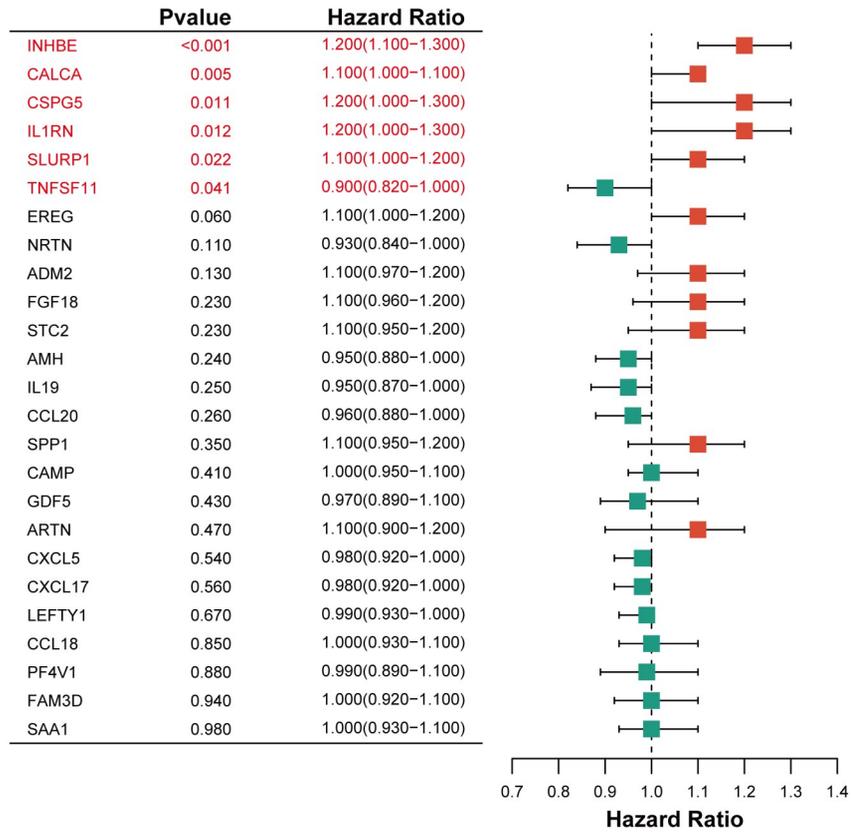


图 4 单因素 Cox 回归分析森林图
Figure 4 Univariate Cox regression analysis forest plot

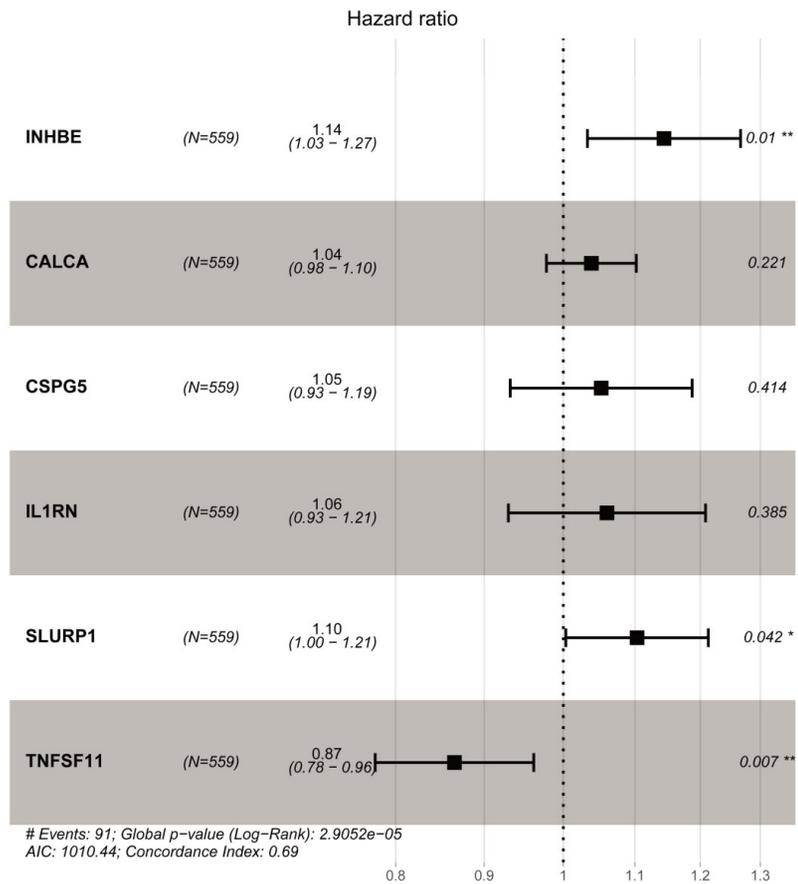


图 5 多因素 Cox 回归分析森林图
Figure 5 Multivariate Cox regression analysis forest plot

6c)。结果显示,这些基因的表达水平能有效地区分健康和疾病状态,INHBE在前5年AUC值均大于0.5,说明在肿瘤早期具有较好的疾病状态判定,SLURP1在1、3、5年的AUC值均大于0.5,10年的

AUC值大于0.7,说明该基因在患病期间具有较好的疾病状态判断,TNFSF11在5和10年的AUC值略低于0.5,相较于其他基因评分略低。

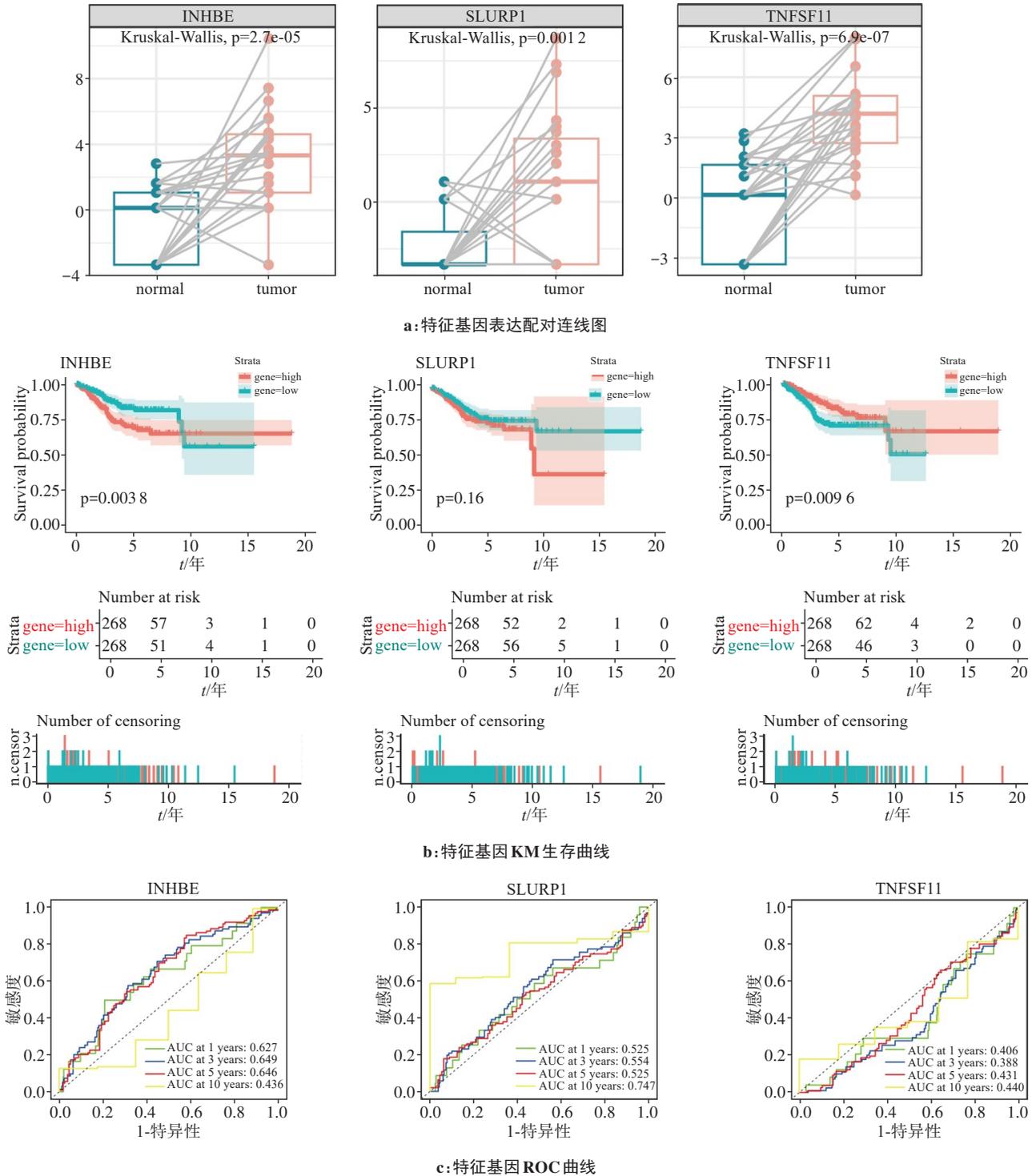


图6 生存分析、构建风险评分模型及验证

Figure 6 Survival analysis, construction of risk score model, and validation

2.5 肿瘤免疫浸润分析及预后模型相关性分析

为了探究肿瘤免疫浸润分析与预后模型的相关

性,使用CIBERSORT算法对癌症患者的基因表达数据进行处理,以准确估算肿瘤微环境中22种免疫细

胞的比例。然后使用 P 值筛选标准 ($pFilter=0.05$) 剔除不显著的数据,并结合基因表达和免疫浸润信息,进行 Spearman 相关性分析,探索基因表达与各类免疫细胞浸润程度之间的关联。结果显示,一些特定的免疫细胞类型如调节性 T 细胞、M0 巨噬细胞、M1 巨噬细胞、CD4 激活 B 细胞、辅助性 T 细胞等在肿瘤组织中的含量较正常组织更高,M2 巨噬细胞、肥大细胞、CD4 记忆 T 细胞等在正常组织中高于肿瘤组织

(图 7a)。此外,研究还发现 INHBE、SLURP1 和 TNFSF11 的表达与 Treg 细胞、浆细胞和静息树突状细胞呈现一定的相关性。具体来说,INHBE 与 Treg 细胞呈正相关 ($R>0.4$),SLURP1 与浆细胞也呈正相关 ($R>0.5$),而 TNFSF11 与静息树突状细胞呈弱正相关 ($R\approx 0.4$) (图 7b)。这些发现表明,这些基因的表达可能与特定免疫细胞类型的浸润有关,从而影响患者的生存率。

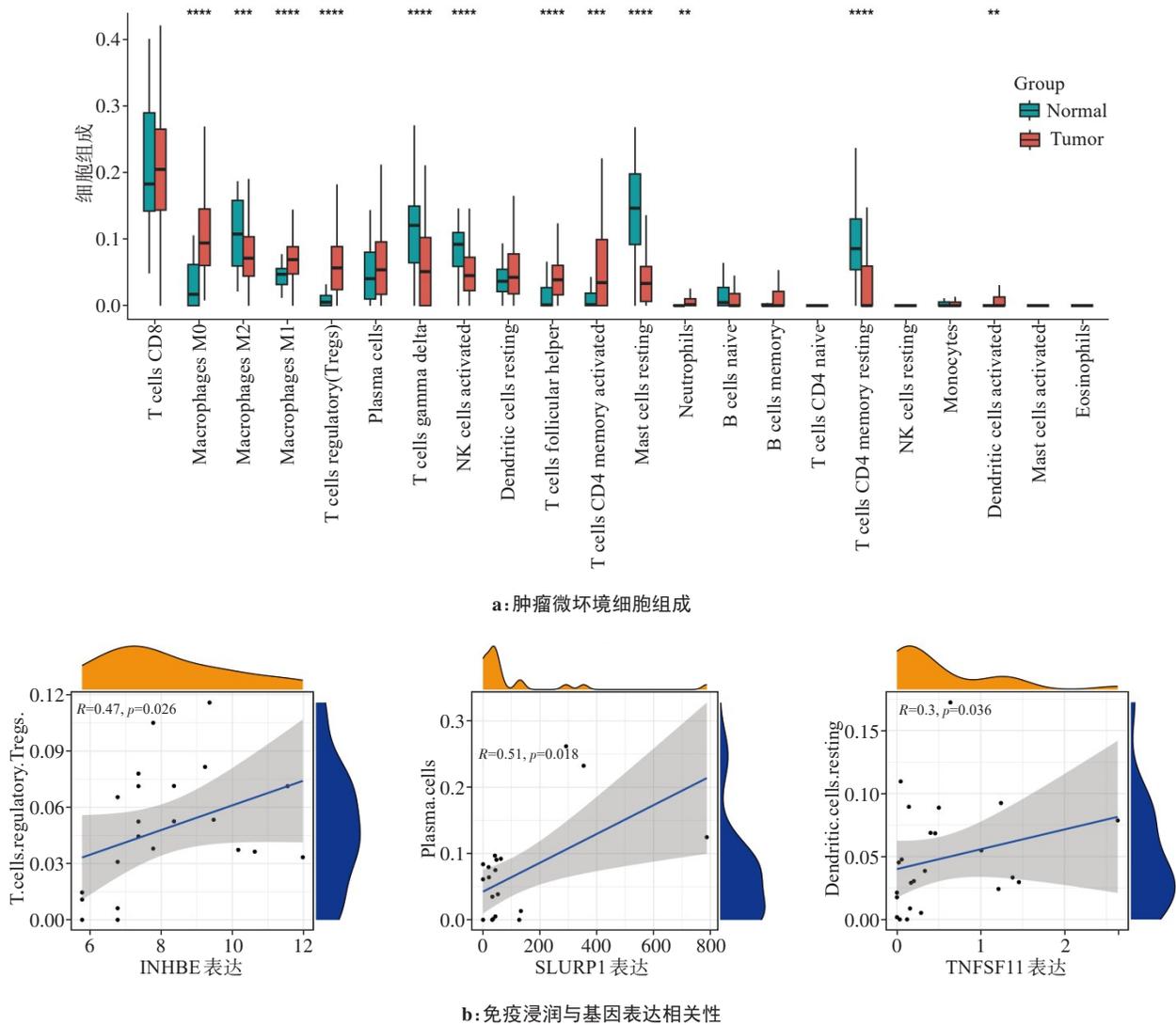


图7 肿瘤免疫浸润分析及预后模型相关性分析

Figure 7 Tumor immune infiltration analysis and prognostic model correlation analysis

2.6 预后模型基因免疫组化

通过 Human Protein Atlas 数据库获取 3 种基因—INHBE、TNFSF11 和 SLURP1 在 EC 及其正常对照中的免疫组化切片图像。通过比较这些图像,可以观察到基因表达在两种状态下的差异:INHBE 在肿瘤样本中局部区域有一定表达,而在正常样本中则表

达较低;相反,TNFSF11 在肿瘤样本中表现出更强的染色,说明其可能在肿瘤发生过程中起着关键作用。同样,SLURP1 也在肿瘤样本中显示出更高的表达(图 8)。这些结果为了了解这些基因的功能及其在疾病进展中的作用提供形态学证据,也为未来制定针对性治疗策略提供参考。

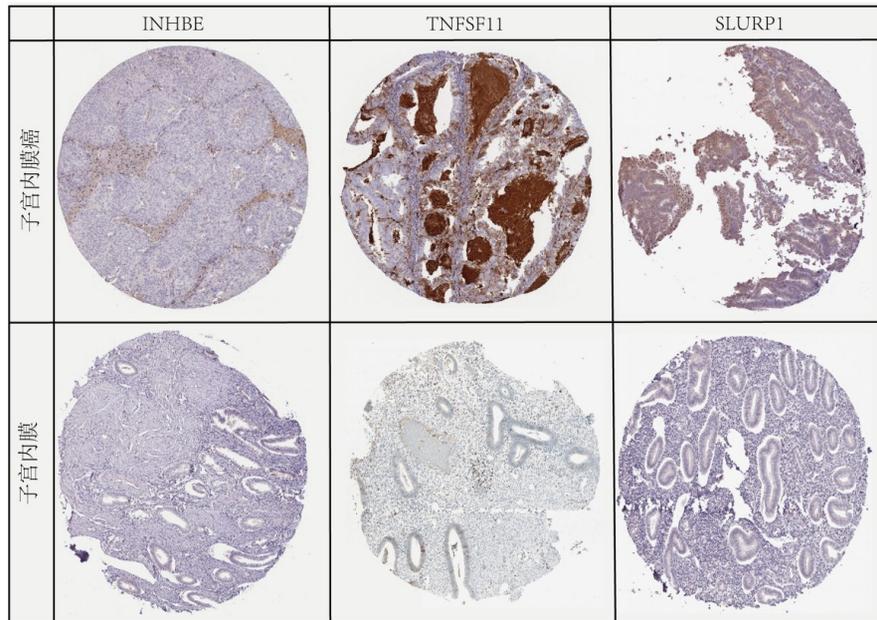


图8 预后模型基因免疫组化图

Figure 8 Immunohistochemistry of prognostic model genes

3 讨论

EC作为妇科恶性肿瘤中的重要成员,近年来其发病率和死亡率的上升趋势引起全球范围内的关注。本研究旨在通过生物信息与机器学习技术的融合,探究EC免疫微环境与基因表达之间的关联及其对预后价值的影响,以期发现新的生物标志物,为EC的早期诊断、风险分层及个性化治疗策略提供理论依据。近年来,EC的研究取得显著进展,尤其是通过高通量测序技术和生物信息学分析,研究者们已经鉴定多个与EC发生发展密切相关的基因^[25],并且关注于EC免疫疗法^[26]。然而,尽管在EC的分子生物学机制和生物标志物发现方面取得突破,但在实际临床应用中,构建稳定可靠的预后基因模型仍面临挑战^[27]。目前,肿瘤样本的获取和临床信息的完整性是构建和验证预后模型的主要瓶颈。高质量的肿瘤样本和详尽的临床随访数据对于建立和验证预后模型至关重要,但这些资源的获取往往受限于伦理考量、样本保存条件以及长期随访的难度。此外,预后模型的验证需要在独立的、多中心的大规模队列中进行,以确保其普遍适用性和可靠性,但这同样受到资金、资源分配和国际合作等方面的限制。

本研究创新性综合运用多种差异表达分析工具(DESeq2, edgeR 和 limma)与 ImmPort 数据库筛选免疫相关差异基因,再结合 GO 和 KEGG 富集分析,深入解析免疫基因在EC进程中的功能和信号通路。此外,通过机器学习筛选特征基因集,利用 Lasso 回

归、单变量特征选择、随机森林等多种算法,有效克服单一算法的局限性,提高基因筛选的准确性和可靠性。单因素和多因素 Cox 回归分析揭示 INHBE、SLURP1 和 TNFSF11 等基因与 EC 患者生存期的显著相关性,这些差异表达的基因可能成为 EC 预后的关键生物标志物。

尽管本研究通过综合分析方法获得有价值的结果,但仍存在一些局限性。首先,虽然 TCGA 数据库提供丰富的基因表达数据,但样本的地域、种族和年龄分布可能影响结果的普遍适用性,且子宫内膜正常样本数相对肿瘤样本较少,也会造成差异分析的偏差。其次,机器学习算法的选择和参数调优过程较为复杂,不同的算法可能会产生不同的特征基因列表,需要进一步验证其稳定性和可重复性。最后,本研究基于回顾性数据分析,缺乏前瞻性临床试验验证,未来研究应结合更多独立队列和临床试验数据,以提高结果的可靠性和临床转化价值。

随着生物医学大数据的持续积累和算法模型的不不断优化,机器学习在 EC 研究领域展现出巨大的潜力。未来研究方向应包括但不限于以下几个方面:(1)跨组学数据分析:整合基因组学^[28]、转录组学、蛋白质组学^[29]和代谢组学^[30]等多组学数据,构建综合性的 EC 分子图谱;(2)深度学习应用:利用深度学习技术,如卷积神经网络^[31],从影像学资料中提取更深层次的特征,辅助 EC 的诊断和预后评估;(3)个体化医疗:基于分子分型和基因表达特征,为 EC 患者制

定个性化的治疗方案,实现精准医疗的目标;(4)免疫治疗靶点开发:探索与免疫细胞浸润相关的基因,为开发新型免疫疗法提供靶点,改善EC的治疗效果。

综上所述,本研究通过生物信息学和机器学习技术,揭示EC免疫微环境与基因表达的复杂关联,鉴定多个潜在的预后生物标志物。这些发现不仅加深了对EC发病机制的理解,而且为EC的早期诊断、预后评估和个体化治疗提供新的视角和策略。未来,随着研究的深入和技术的进步,期待看到更多精准有效的EC诊疗方案的诞生。

【参考文献】

- [1] Di Donato V, Giannini A, Bogani G. Recent advances in endometrial cancer management[J]. *J Clin Med*, 2023, 12(6): 2241.
- [2] Le LL, Luo J, Wu HF, et al. Overexpression of MYBL2 predicts poor prognosis and promotes oncogenesis in endometrial carcinoma[J]. *Eur J Histochem*, 2021, 65(2): 3226.
- [3] Sulaiman R, De P, Aske JC, et al. Identification and morphological characterization of features of circulating cancer-associated macrophage-like cells (CAMLs) in endometrial cancers[J]. *Cancers (Basel)*, 2022, 14(19): 4577.
- [4] Gorbokov N, Baltruschat S, Lennartz M, et al. PAX8 expression in cancerous and non-neoplastic tissue: a tissue microarray study on more than 17 000 tumors from 149 different tumor entities[J]. *Virchows Archiv*, 2024, 485(3): 491-507.
- [5] Minamikawa T, Yachida N, Takahashi K, et al. Endometrial cancer with and without endometriosis: clinicopathological differences [J]. *Cancers (Basel)*, 2023, 15(23): 5635.
- [6] Ji JX, Cochrane DR, Tessier-Cloutier B, et al. Use of immunohistochemical markers (HNF-1 β , napsin a, ER, CTH, and ASS1) to distinguish endometrial clear cell carcinoma from its morphologic mimics including Arias-Stella reaction[J]. *Int J Gynecol Pathol*, 2020, 39(4): 344-353.
- [7] Liu HY, Wan JH, Feng QL, et al. Long non-coding RNA SOS1-IT1 promotes endometrial cancer progression by regulating hypoxia signaling pathway[J]. *J Cell Commun Signal*, 2022, 16(2): 253-270.
- [8] Ruan P, Wang SJ, Yang CY, et al. m6A mRNA methylation regulates the ERK/NF- κ B/AKT signaling pathway through the PAPPA/IGFBP4 axis to promote proliferation and tumor formation in endometrial cancer[J]. *Cell Biol Toxicol*, 2023, 39(4): 1611-1626.
- [9] Ray I, Möller-Levet CS, Michael A, et al. Circulating adipocytokines and insulin like-growth factors and their modulation in obesity-associated endometrial cancer[J]. *Cancers (Basel)*, 2024, 16(3): 531.
- [10] Xu SH, Zhang Y, Yang Y, et al. A m6A regulators-related classifier for prognosis and tumor microenvironment characterization in hepatocellular carcinoma[J]. *Front Immunol*, 2024, 15: 1374465.
- [11] Wang YX, Li C, He JX, et al. Multi-omics analysis and experimental validation of the value of monocyte-associated features in prostate cancer prognosis and immunotherapy[J]. *Front Immunol*, 2024, 15: 1426474.
- [12] Li M, Cai YH, Zhang MZ, et al. NNBGWO-BRCA marker: neural network and binary grey wolf optimization based breast cancer biomarker discovery framework using multi-omics dataset[J]. *Comput Methods Programs Biomed*, 2024, 254: 108291.
- [13] De Perrot T, Sadjo Zoua C, Glessgen CG, et al. Diffusion-weighted MRI in the genitourinary system[J]. *J Clin Med*, 2022, 11(7): 1921.
- [14] Wu TZ, Hu EQ, Xu SB, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data[J]. *Innovation (Camb)*, 2021, 2(3): 100141.
- [15] Guan M, Jiao Y, Zhou LL. Immune infiltration analysis with the CIBERSORT method in lung cancer[J]. *Dis Markers*, 2022, 2022: 3186427.
- [16] Lu KN, Wei S, Wang ZY, et al. Identification of novel biomarkers in Hunner's interstitial cystitis using the CIBERSORT, an algorithm based on machine learning[J]. *BMC Urol*, 2021, 21(1): 109.
- [17] Chen JJ, Jiang XL, Gao X, et al. Ferroptosis-related genes as diagnostic markers for major depressive disorder and their correlations with immune infiltration[J]. *Front Med (Lausanne)*, 2023, 10: 1215180.
- [18] Gu ZG, Hübschmann D. Make interactive complex heatmaps in R[J]. *Bioinformatics*, 2022, 38(5): 1460-1462.
- [19] Kyo S, Nakayama K. Endometrial cancer as a metabolic disease with dysregulated PI3K signaling: shedding light on novel therapeutic strategies[J]. *Int J Mol Sci*, 2020, 21(17): 6073.
- [20] Fan YJ, Dong Z, Shi YC, et al. NLRC5 promotes cell migration and invasion by activating the PI3K/AKT signaling pathway in endometrial cancer[J]. *J Int Med Res*, 2020, 48(5): 300060520925352.
- [21] Yu SH, Cai JH, Chen DL, et al. LASSO and bioinformatics analysis in the identification of key genes for prognostic genes of gynecologic cancer[J]. *J Pers Med*, 2021, 11(11): 1177.
- [22] Cea García J, Rodríguez Jiménez I, Ríos-Pena L, et al. Incidence and univariate models for lymphatic drainage disorders following management for cervical cancer[J]. *J Obstet Gynaecol Res*, 2021, 47(1): 343-351.
- [23] Zhou QM, Liu CX, Zhou JP, et al. Machine Learning-Based radiological features and diagnostic predictive model of xanthogranulomatous cholecystitis[J]. *Front Oncol*, 2022, 12: 792077.
- [24] Zhu J, Luo JJ, Ma Y. Screening of serum exosome markers for colorectal cancer based on Boruta and multi-cluster feature selection algorithms[J]. *Mol Cell Toxicol*, 2024, 20(2): 343-351.
- [25] Xiong WQ, Jin J, Liu Y. METTL3-dependent m6A modification facilitates decreased endometrial decidualization *via* attenuation of MET in endometriosis[J]. *Reproduction*, 2024, 168(3): e230336.
- [26] Mahdi H, Chelariu-Raicu A, Slomovitz BM. Immunotherapy in endometrial cancer[J]. *Int J Gynecol Cancer*, 2023, 33(3): 351-357.
- [27] Cao WY, Ma XY, Fischer JV, et al. Immunotherapy in endometrial cancer: rationale, practice and perspectives[J]. *Biomark Res*, 2021, 9(1): 49.
- [28] Miller EM, Patterson NE, Gressel GM, et al. Utility of a custom designed next generation DNA sequencing gene panel to molecularly classify endometrial cancers according to the cancer genome atlas subgroups[J]. *BMC Med Genomics*, 2020, 13(1): 179.
- [29] Serambeque B, Mestre C, Hundarova K, et al. Proteomic profile of endometrial cancer: a scoping review[J]. *Biology (Basel)*, 2024, 13(8): 584.
- [30] Raffone A, Troisi J, Boccia D, et al. Metabolomics in endometrial cancer diagnosis: a systematic review[J]. *Acta Obstet Gynecol Scand*, 2020, 99(9): 1135-1146.
- [31] Urushibara A, Saida T, Mori K, et al. The efficacy of deep learning models in the diagnosis of endometrial cancer using MRI: a comparison with radiologists[J]. *BMC Med Imaging*, 2022, 22(1): 80.

(编辑:陈丽霞)