

网络出版时间:2024-05-06 13:21:08 网络出版地址:https://link.cnki.net/urlid/34.1065.R.20240429.1131.007

# 基于免疫基因相关的结直肠癌预后模型的构建分析及药物筛选

郑伟<sup>1</sup>, 赵佳佳<sup>2</sup>, 程祥<sup>1</sup>, 谭泓鑫<sup>1</sup>, 黄琦<sup>1</sup>

**摘要** 目的 寻找新的生物标志物预测结直肠癌(CRC)患者的预后。方法 利用单因素 Cox 回归分析和最小绝对收缩和选择算子(LASSO)回归分析,从癌症基因组图谱(TCGA)数据库中建立了一个免疫相关基因的结直肠癌预后模型。通过利用表达数据估算恶性肿瘤组织中的基质细胞和免疫细胞(ESTIMATE)软件包以及通过估算 RNA 转录本的相对子集识别细胞类型(CIBERSORT)软件包比较了高低风险组患者的免疫浸润特征。此外,还分析了不同风险组患者的免疫检查点的表达水平。通过癌症药物敏感性基因组学(GDSC)数据库比较了两个风险组患者对化疗药物的敏感性。结果 结果显示,基于免疫基因构建的预后模型可以较好地预测 CRC 患者的总体生存期(OS),结果显示 1 年、3 年和 5 年 OS 的曲线下与坐标轴面积(AUC)值分别为 0.764(95% CI:0.751~0.793)、0.773(95% CI:0.761~0.779)和 0.760(95% CI:0.742~0.774)。低风险组患者的免疫检查点表达水平更高和免疫细胞更丰富如 T 细胞( $P<0.001$ )、树突状细胞( $P<0.001$ )、巨噬细胞( $P<0.001$ )、中性粒细胞( $P<0.001$ )。高风险组患者可能对于阿昔替尼、伊马替尼、甲氨蝶呤、帕唑帕尼、雷帕霉素、舒尼替尼和他普西加金等一些化疗药物更敏感。结论 基于 19 个免疫基因的预后模型能够有效的预测 CRC 患者的预后。不同患者中免疫微环境中的免疫细胞数量和活跃程度可能是影响其对免疫检查点抑制剂和化疗药物响应的重要因素。

**关键词** 结直肠癌;免疫检查点抑制剂;肿瘤微环境;预后模型;免疫相关基因

**中图分类号** R 735

**文献标志码** A **文章编号** 1000-1492(2024)05-0789-08  
doi:10.19405/j.cnki.issn1000-1492.2024.05.007

结直肠癌(colorectal cancer, CRC)的发病率和病死率在男性和女性中均位居前三位<sup>[1]</sup>。CRC 患者往往预后较差,且确诊时已是中晚期,限制了治疗效果<sup>[2]</sup>。免疫疗法在 CRC 中发挥着重要的作用。

免疫检查点抑制剂(immune checkpoint inhibitors, ICIs)是较为成熟的免疫治疗方法之一<sup>[3-4]</sup>。但是,ICIs 仅对部分患者展现出良好的治疗效果<sup>[5]</sup>。

由于肿瘤微环境的复杂性,不同患者对于免疫治疗的应答情况不尽相同,实际上仅不到三分之一的患者对免疫治疗有效果<sup>[5-6]</sup>。因此,构建有助于全面了解患者的免疫细胞浸润水平的预后模型,对于提高患者对免疫治疗及化疗等治疗方式的有效性具有重要价值。但目前,仍然缺少全面了解结直肠癌患者的免疫细胞浸润水平的预后模型。该研究通过单因素 Cox 回归分析和最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)回归分析在结直肠癌患者中构建了免疫基因相关的预后模型,并分析了不同风险组患者的免疫细胞浸润状态、免疫检查点的表达以及对化疗的药物敏感性。为 CRC 患者的免疫细胞浸润状态水平与预后之间的联系提供了重要的参考。

## 1 材料与方法

**1.1 数据来源** 患者的基因测序数据和临床信息来自癌症基因组图谱项目(the cancer genome atlas program, TCGA)。免疫相关基因来自免疫学数据库和分析门户网站(immunology database and analysis portal, ImmPort)数据库。免疫相关的预后基因免疫组化结果来自人类蛋白质组数据库(human protein atlas, HPA)。药物敏感性数据基于癌症药物敏感性基因组学(genomics of drug sensitivity in cancer, GDSC)数据库。

**1.2 预后模型的建立和评估** 利用单因素 Cox 回归分析得到与患者的预后相关的差异表达基因(differently expressed genes, DEGs),将 DEGs 与免疫基因取交集。使用 LASSO 回归分析对 LI 进行范数约束,模型的复杂程度通过 Lambda( $\lambda$ )控制所得,利用最大似然比法估算出 19 个基因的 95% 置信区间(CI)、危险比(HR)和系数( $\beta$ )。风险评分的计算方式为 19 个基因的表达量乘以相应的系数值并求和。根据风险评分的中位值划分高低风险组的。通过与时间依赖的接收者操作特征曲线(time-dependent

2023-12-28 接收

基金项目:国家自然科学基金(编号:82003037);安徽省转化医学研究院科研基金(编号:2022zhyx-C79)

作者单位:<sup>1</sup> 安徽医科大学第二附属医院肿瘤科,合肥 230601

<sup>2</sup> 安徽医科大学第二附属医院神经外科,合肥 230601

作者简介:郑伟,男,硕士研究生;

黄琦,女,博士,主治医师,硕士生导师,责任作者, E-mail: huangqi214@163.com

receiver operating characteristic curves, *t*-ROC) 和卡普兰-梅耶尔生存曲线(Kaplan-Meier, KM)评估了模型的准确性。

**1.3 GO 和 KEGG 分析** 将两组患者之间的差异基因进行了基因本体分析论(gene ontology, GO)和京都基因与基因组百科全书(kyoto encyclopedia of genes and genomes, KEGG)分析。两组患者之间差异基因筛选条件是 $|\log_2^{FC}| > 1$  和  $P < 0.05$ 。

**1.4 免疫浸润特征分析** 通过利用表达数据估算恶性肿瘤组织中的基质细胞和免疫细胞(stimulation of stromal and immune cells in malignant tumor tissues using expression data, ESTIMATE)软件包以及估算 RNA 转录本的相对子集识别细胞类型(cell-type identification by estimating relative subsets of RNA transcripts, CIBERSORT)软件包,评估了两组患者的肿瘤免疫浸润特征。

**1.5 免疫检查点抑制剂基因的表达水平和药物敏感性分析** 研究了不同风险组患者的免疫检查点基因的表达水平。利用 GDSC 数据库和 pRRophetic 软件包分析患者对化疗药物的敏感性。

**1.6 统计学处理** 使用 R 4.1.2 软件完成了统计分析。使用 Wilcoxon 秩和检验绘制箱形图。使用斯皮尔曼系数(Spearman's coefficient)进行相关性分析。所有双侧假设检验均要求  $P < 0.05$ 。

## 2 结果

### 2.1 免疫相关预后基因的筛选和预后模型的建立

471 例肿瘤样本和 41 例正常组织样本中共发现了 9 135 个 DEGs(图 1A)。单因素 Cox 回归分析得到了 1 346 个与预后相关的 DEGs。最后,将 Immport 数据库中得到的 1 793 个免疫基因与二者取交集的得到 26 个免疫相关预后基因(图 1B)。使用 1 000 次交叉验证其准确性(图 1C)。为避免模型的该风险模型的过度拟合, LASSO 回归分析后得到 19 个基因(图 1D),包括 ULBP2、PMCH、CXCL8、UCN、MC1R、CD1A、NTF4、LHB、TNFRSF19、CD1B、PLXNA3、IL17A、INHBB、HAMP、NMB、OXT、CXCL1、STC2 和 FGF19 等。如图 1E 所示,验证该 19 个免疫相关的预后基因在高低风险组患者风险评分。19 个免疫相关预后基因与 CRC 患者总体生存期的相关性如表 1。

**2.2 19 个免疫相关预后基因的特征分析** 19 个免疫相关的预后基因在肿瘤组织和正常组织中的表达情况(表 1)。除 TNFRSF19 外,这些基因在肿瘤组织中的高表达差异均有统计学意义。通过 KM 生存分析,研究表明 PMCH、CXCL8 和 CXCL1 等的高表达对 CRC 患者的总生存有利(表 2)。此外,该实验还从 HPA 数据库中验证了肿瘤组织和癌旁组织中

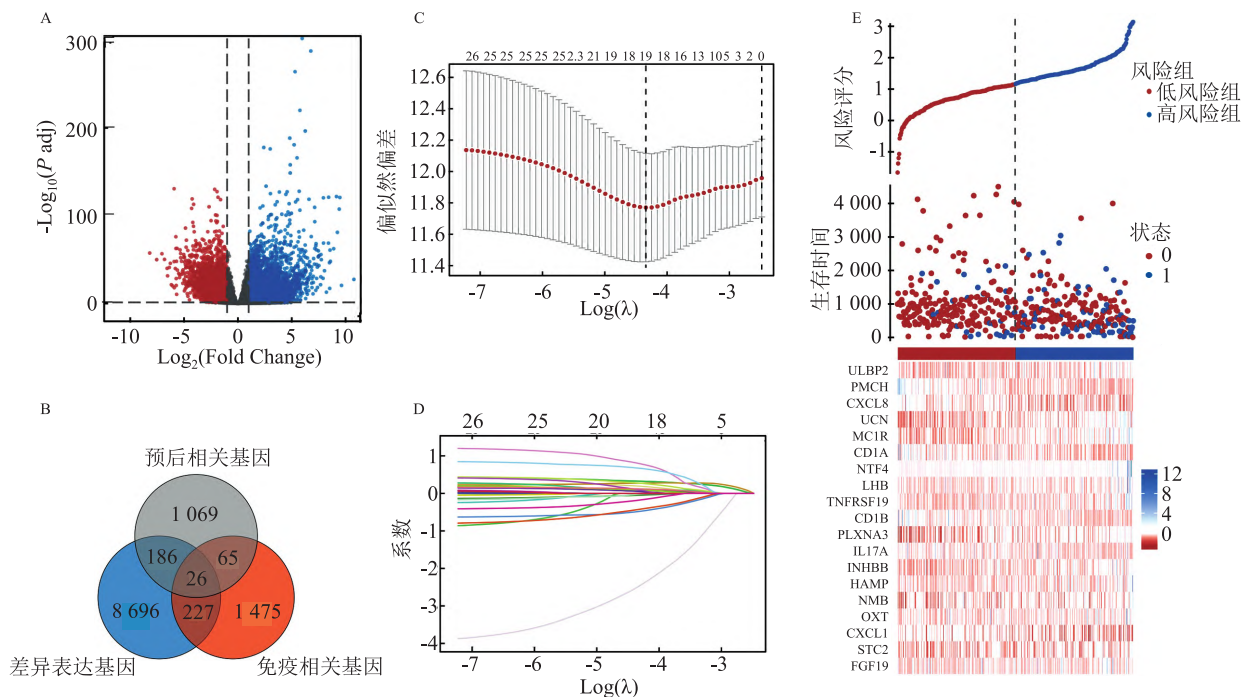


图1 免疫相关预后风险模型的建立

A: 9 135 个 DEGs 的火山图; B: 预后相关基因、免疫相关基因和差异表达基因三者的韦恩图; C: LASSO 模型的十次交叉验证; D: LASSO 系数曲线; E: 19 个免疫相关预后基因在高低风险组结肠患者中的风险因子图,包括风险评分、生存状态和热图

表1 相关基因在正常组织和结肠癌组织中的相对表达水平( $\bar{x} \pm s$ )

| 基因名      | 正常组织                  | 结肠癌组织                 | P 值    |
|----------|-----------------------|-----------------------|--------|
| INHBB    | 0.662 1 $\pm$ 0.334 8 | 1.500 0 $\pm$ 1.012 0 | <0.001 |
| HAMP     | 0.065 8 $\pm$ 0.104 3 | 0.246 1 $\pm$ 0.213 0 | <0.001 |
| NMB      | 1.446 0 $\pm$ 0.332 3 | 2.887 0 $\pm$ 0.596 4 | <0.001 |
| OXT      | 0.077 8 $\pm$ 0.076 6 | 0.179 9 $\pm$ 0.329 5 | <0.05  |
| CXCL1    | 2.730 0 $\pm$ 1.162 0 | 5.047 0 $\pm$ 1.325 0 | <0.001 |
| STC2     | 0.385 2 $\pm$ 0.342 8 | 1.865 0 $\pm$ 1.073 0 | <0.001 |
| FGF19    | 0.067 3 $\pm$ 0.363 4 | 0.682 3 $\pm$ 1.011 0 | <0.001 |
| ULBP2    | 0.112 5 $\pm$ 0.065 8 | 1.160 0 $\pm$ 0.968 5 | <0.001 |
| PMCH     | 0.019 3 $\pm$ 0.026 2 | 0.056 0 $\pm$ 0.063 7 | <0.01  |
| CXCL8    | 1.547 0 $\pm$ 1.306 0 | 4.386 0 $\pm$ 1.817 0 | <0.001 |
| UCN      | 0.291 3 $\pm$ 0.146 2 | 0.926 6 $\pm$ 0.442 9 | <0.001 |
| MC1R     | 0.319 3 $\pm$ 0.128 3 | 0.799 6 $\pm$ 0.582 4 | <0.001 |
| CD1A     | 0.197 4 $\pm$ 0.164 6 | 0.389 9 $\pm$ 0.327 6 | <0.001 |
| NTF4     | 0.006 8 $\pm$ 0.008 2 | 0.019 7 $\pm$ 0.028 4 | <0.01  |
| LHB      | 0.058 0 $\pm$ 0.068 3 | 0.172 6 $\pm$ 0.148 9 | <0.001 |
| TNFRSF19 | 0.547 7 $\pm$ 0.154 7 | 0.898 1 $\pm$ 0.843 0 |        |
| CD1B     | 0.143 1 $\pm$ 0.134 0 | 0.314 4 $\pm$ 0.326 1 | <0.01  |
| PLXNA3   | 1.701 0 $\pm$ 0.366 3 | 2.707 0 $\pm$ 0.599 1 | <0.001 |
| IL17A    | 0.051 3 $\pm$ 0.098 4 | 0.156 5 $\pm$ 0.203 9 | <0.01  |

表2 相关基因在结肠癌患者中的 KM 预后分析

| 基因名      | HR(95% CI)         | P 值    |
|----------|--------------------|--------|
| INHBB    | 1.61 (1.08 ~ 2.41) | 0.019  |
| HAMP     | 1.57 (1.06 ~ 2.32) | 0.024  |
| NMB      | 1.54 (1.04 ~ 2.28) | 0.033  |
| OXT      | 1.51 (1.02 ~ 2.24) | 0.038  |
| CXCL1    | 0.67 (0.45 ~ 0.99) | 0.044  |
| STC2     | 1.49 (1.01 ~ 2.20) | 0.045  |
| FGF19    | 1.48 (1.00 ~ 2.20) | 0.050  |
| ULBP2    | 1.91 (1.28 ~ 2.84) | 0.001  |
| PMCH     | 0.52 (0.35 ~ 0.78) | 0.002  |
| CXCL8    | 0.54 (0.38 ~ 0.80) | 0.003  |
| UCN      | 1.78 (1.20 ~ 2.85) | 0.004  |
| MC1R     | 2.28 (1.61 ~ 3.23) | <0.001 |
| CD1A     | 0.55 (0.38 ~ 0.78) | <0.001 |
| NTF4     | 1.71 (1.15 ~ 2.54) | 0.008  |
| LHB      | 1.36 (0.96 ~ 1.93) | 0.083  |
| TNFRSF19 | 1.65 (1.11 ~ 2.48) | 0.013  |
| CD1B     | 0.61 (0.40 ~ 0.91) | 0.015  |
| PLXNA3   | 1.62 (1.09 ~ 2.40) | 0.018  |
| IL17A    | 0.73 (0.51 ~ 1.04) | 0.079  |

表达的 19 个免疫相关的预后基因。结果显示, CXCL8、UCN、CD1B、IL-17A、NMB、STC2 和 FGF19 在肿瘤组织中的表达量明显较高。但 HPA 数据库中没有任何其他基因的表达情况(图 2D)。在 CRC 中表达的这 19 个免疫相关预后基因相互独立(图 2C)。利用 Spearman 相关分析, 分析了 19 个免疫相关的预后基因与免疫浸润的关系, 包括 ESTIMATE 和 ssGSEA, 结果显示 CXCL8、CD1A、CD1B、HAMP 和 OXT

与免疫浸润有较强的相关性(图 2A、2B)。

**2.3 评估风险模型的准确性** 通过前文所提到的方法即通过单因素 Cox 回归分析和 LASSO 回归分析筛选出与患者预后相关的 19 个免疫基因。接着计算出每例患者的风险评分。对每例患者的得分进行排序, 按照中位值(1.158)将患者分为高风险组和低风险组。在 TCGA 数据集中, 高风险组患者有更高的病死率倾向(图 3A)。与高风险评分的患者相比, 低风险评分的患者生存状态更好( $P < 0.001$ )。图 1E 显示了两个评分组的风险因素图包含风险评分、生存状态和免疫相关预后基因的热图。然后, 该实验应用时间依赖性 ROC 曲线预测患者的 OS, 结果显示 1 年、3 年和 5 年 OS 的 AUC 值分别为 0.764 (95% CI: 0.751 ~ 0.793)、0.773 (95% CI: 0.761 ~ 0.779) 和 0.760 (95% CI: 0.742 ~ 0.774) (图 3B)。

**2.4 GO 和 KEGG 分析两组的 DEGs** 使用 DESeq2 R 软件包鉴定了两组的 DEGs(图 4A)。在 GO 和 KEGG 分析中, DEGs 主要参与单核细胞和淋巴细胞分化以及 T 细胞活化等(图 4B)。结果显示, 丝裂原活化蛋白激酶(mitogen-activated protein kinases, MAPK)、磷脂酰肌醇-3-激酶/丝苏氨酸蛋白激酶(phosphatidylinositol 3-kinase/serine-threonine kinase, PI3K/AKT)和环磷腺苷(cyclic adenosine monophosphate, cAMP)通路可能与 CRC 患者的预后有关(图 4C)。

**2.5 免疫浸润特征分析** 为了更好地了解两组患者的免疫功能, 利用 ESTIMATE 和 CIBERSORT 来测量两组风险患者的免疫细胞浸润水平。根据 ESTIMATE 分析结果显示低风险组患者的基质、免疫和 ESTIMATE 得分更高, 肿瘤纯度更低。此外, CIBERSORT 分析显示, 低危患者含有较高比例的活化免疫细胞, 包括记忆 CD4<sup>+</sup>T 细胞、嗜酸性粒细胞、中性粒细胞、树突状细胞(dendritic cells, DC)和自然杀伤细胞(natural killer, NK)。除此之外, 还有几种细胞亚型在低风险组患者中被激活, 包括 B 细胞、多种 T 细胞、巨噬细胞。调节性 T 细胞(regulatory T cells, Tregs)则在低风险患者中表达较低。这些结果表明, 与高危组相比, 低危组患者的免疫细胞浸润程度更高, 尤其是巨噬细胞、DC 细胞、CD8<sup>+</sup>T 细胞和中性粒细胞(图 5)。

**2.6 评估免疫检查点抑制剂基因的表达水平和化疗药物敏感性** 为了验证两组 CRC 患者对 ICIs 和某些化疗药物的敏感性, 该实验检测了两组患者某

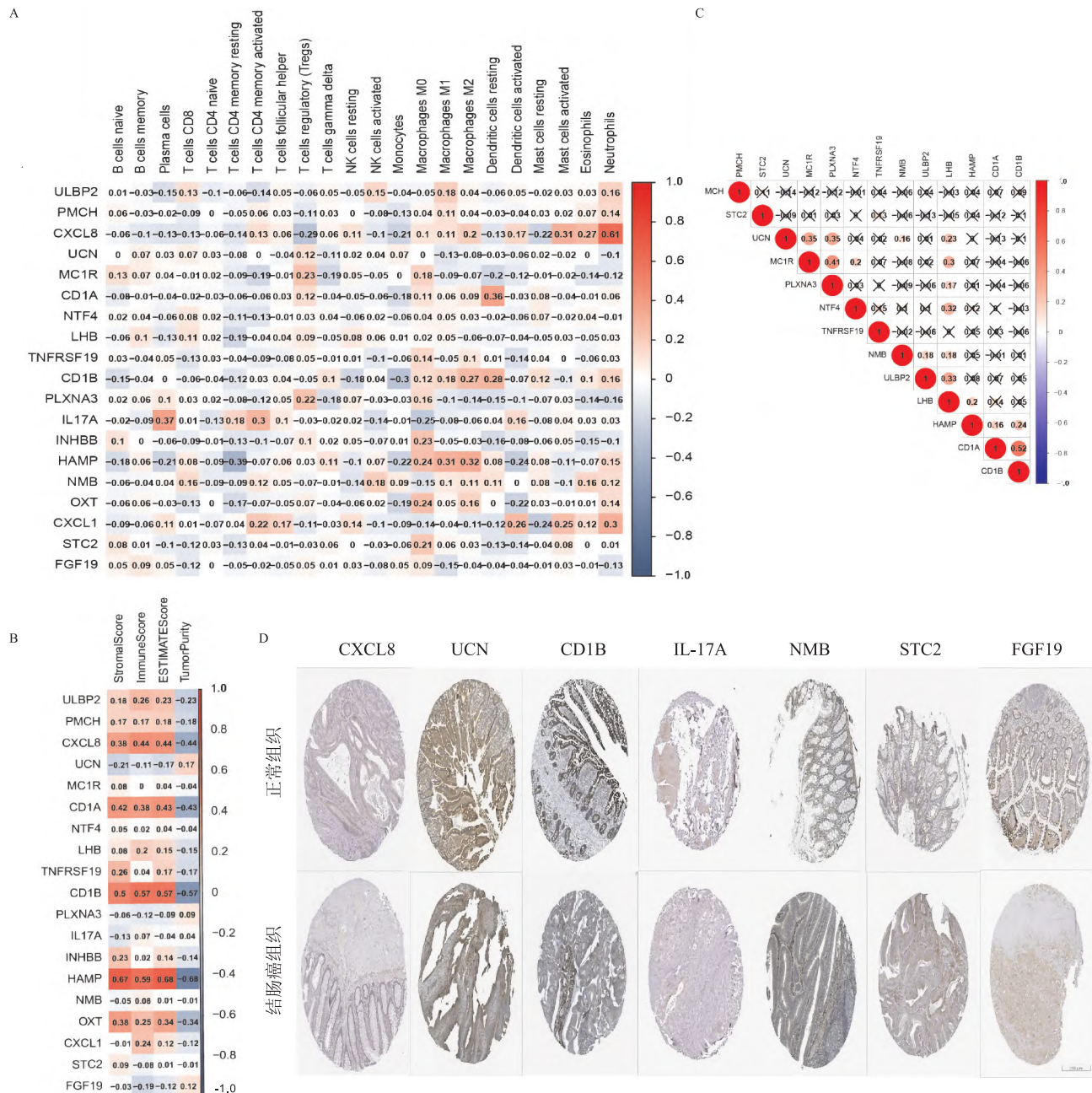


图2 19个免疫相关预后基因的特征分析

A: ssGSEA 分析; B: ESTIMATE 分析; C: 基因表达的相关性分析; D: 部分免疫相关预后基因在肿瘤组织和正常组织中的表达情况 二氨基联苯胺法染色  $\times 4$

些免疫检查点的表达和某些化疗药物的  $IC_{50}$ 。该分析显示低风险组的 CRC 患者的免疫检查点基因表达水平较高, 如程序性细胞死亡 1 (programmed cell death protein 1, PD1) 相关靶点[程序性细胞死亡 1 配体(programmed cell death 1 ligand 1, PDL1)、程序性细胞死亡 2 配体(programmed cell death 1 ligand 1, PDL2) 和程序性细胞死亡 3 配体(programmed cell death 1 ligand 1, PDL3)]、细胞毒性 T 淋巴细胞抗原 4 相关(cytotoxic t-lymphocyte associated pro-

tein 4, CTLA4) 以及 CD80 和 CD86、活化 T 细胞激动剂[肿瘤坏死因子受体超家族 9 (tumor necrosis factor receptor superfamily 9, 4-1BB) 和诱导刺激协同分子(inducible synergistic co-stimulation molecules, ICOS)]、T 细胞免疫球蛋白-3(T cell immunoglobulin domain and mucin domain-3, TIM3) 和 T 细胞免疫球蛋白 ITMI 结构域(T cell immunoreceptor with Ig and ITIM domains, TIGIT) (图 6)。此外, 高风险组患者对化疗药物(阿昔替尼、伊马替尼、甲氨蝶呤、帕唑

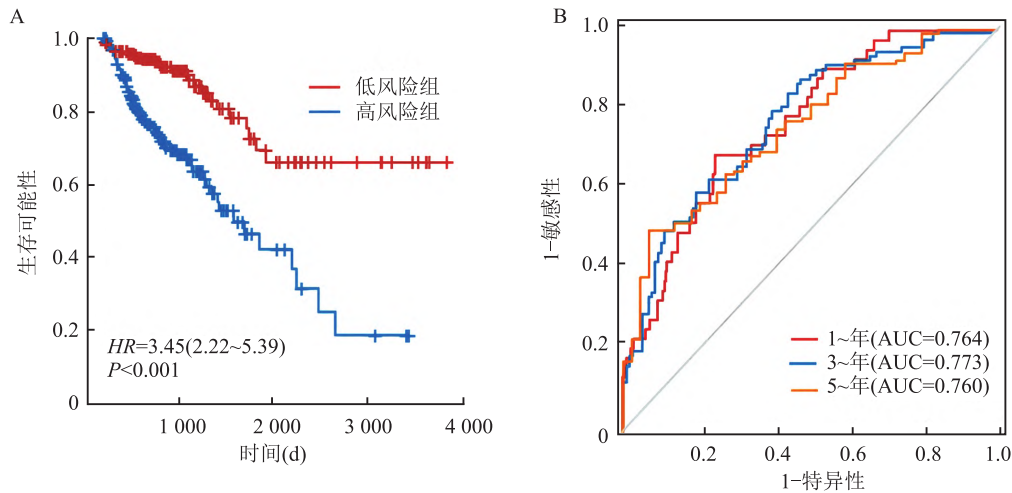


图3 评估 TCGA 数据库的模型的准确性

A:KM 曲线;B:随时间变化的 ROC 曲线预测风险模型的 1 年、3 年和 5 年生存率

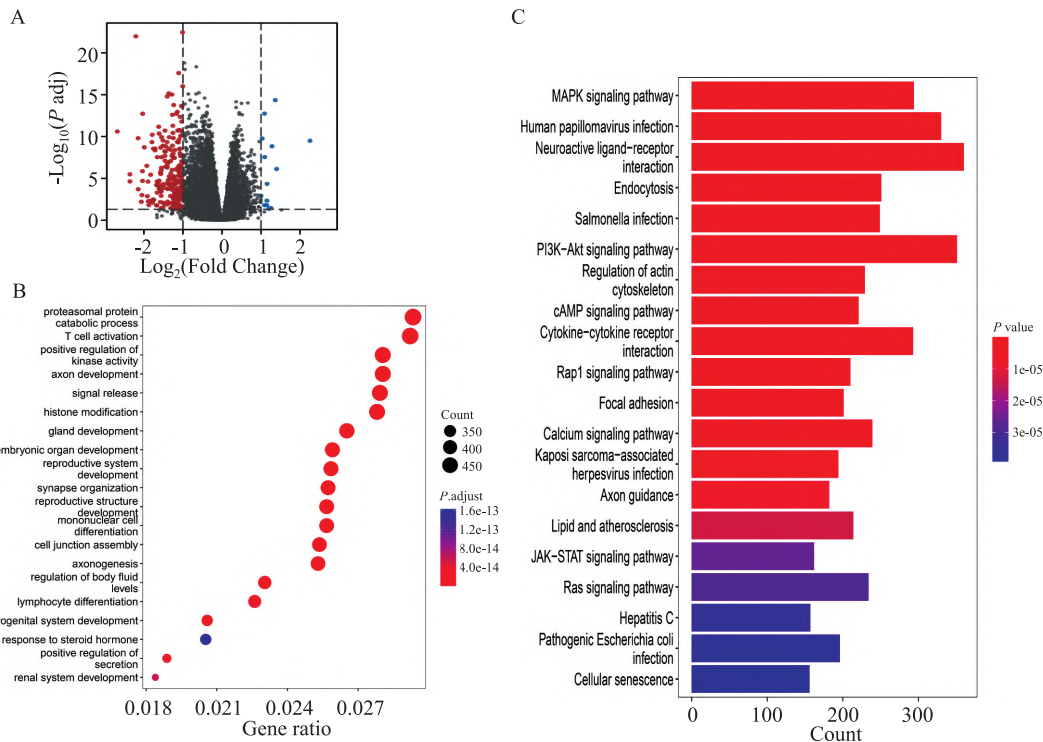


图4 不同风险组 CRC 患者中的差异基因涉及的相关通路

A:两个风险组 CRC 患者不同表达基因的火山图;B:GO 分析;C:KEGG 分析

帕尼、雷帕霉素、舒尼替尼和他普西加金)可能更敏感(图7)。这些结果为不同风险组别的患者提供了更准确的治疗方案。

### 3 讨论

肿瘤微环境是影响肿瘤进展和治疗的重要因素<sup>[7]</sup>。尽管人们对 CRC 的发病机制有了一定的了解,但仍缺乏能准确预测 CRC 预后的标志物。目

前,已有研究<sup>[8-9]</sup>通过免疫基因水平,建立预后模型,预测患者的预后状况。Zhu et al<sup>[8]</sup>通过食管癌相关的 6 个免疫基因的表达水平建立的预后模型可以有效的预测多种上消化道肿瘤患者的预后水平。Zhang et al<sup>[9]</sup>也通过免疫基因的表达水平成功构建了能够有效预测胶质瘤患者的预后水平。但是仍然缺少基于免疫基因表达水平来预测结直肠癌患者的预后模型。该研究基于 19 个免疫相关预后基因构

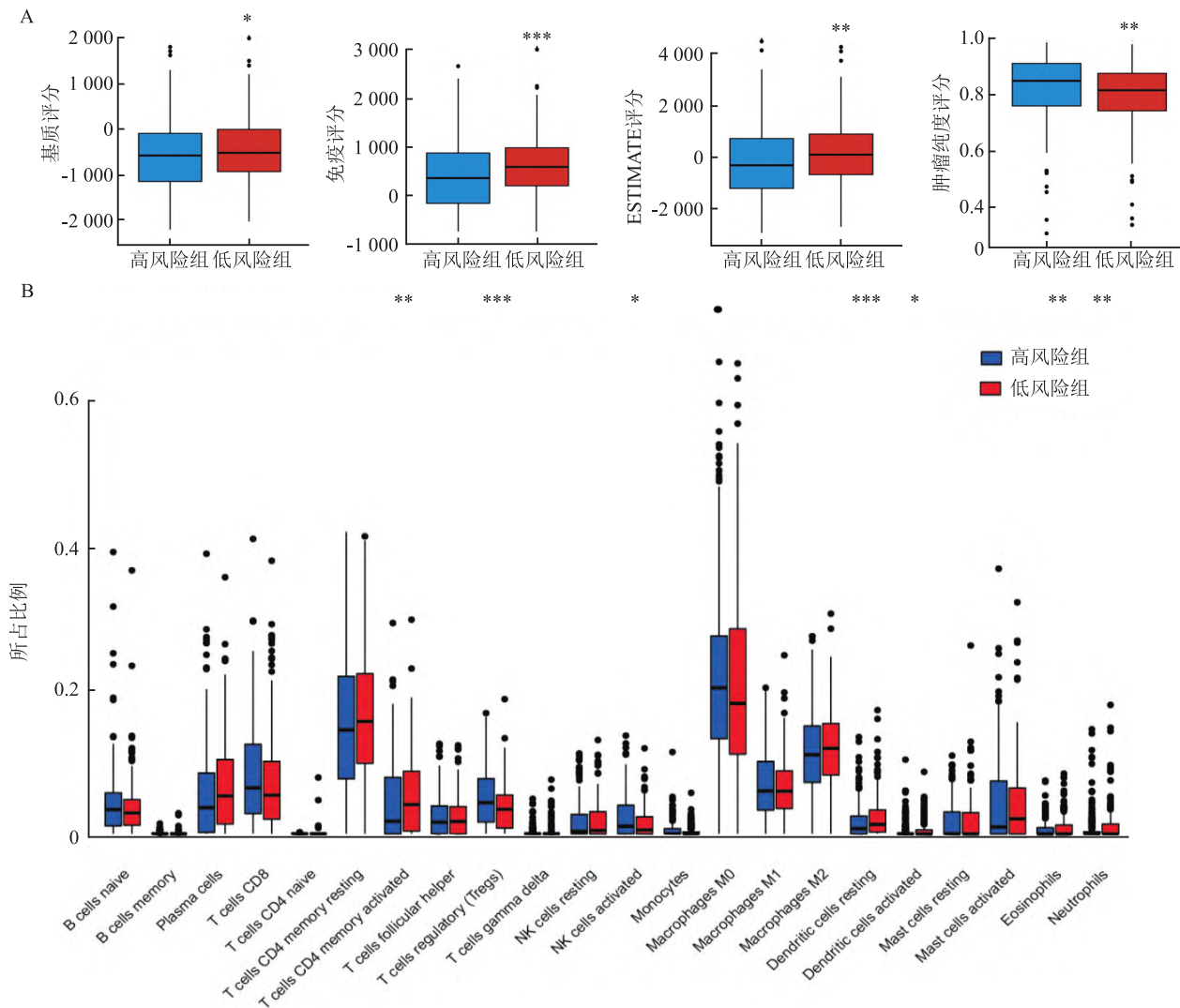


图5 两组患者的免疫特征比较

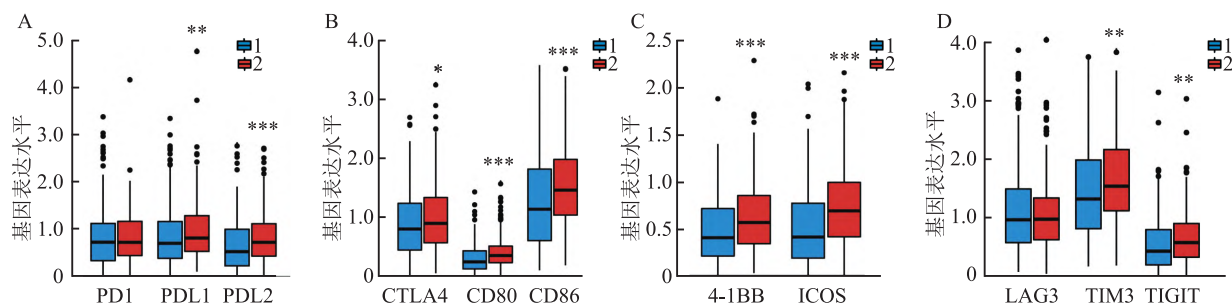
A: ESTIMATE; B: CIBERSORT; 与高风险组比较: \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ 

图6 低危组和高危组患者的一些免疫检查点基因表达水平

A-D: PD1 相关免疫检查点、CTLA4 相关免疫检查点、活化 T 细胞激动剂免疫检查点、其他免疫检查点基因表达水平; 1: 高风险组; 2: 低风险组; 与高风险组比较: \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ 

建了一个新的预后模型。并且相较于之前的研究,该研究着重揭示了基于该免疫基因的预后模型的患者免疫细胞的表达水平以及对药物的敏感性。

通过 ESTIMATE 分析了高低风险组患者的免疫特征,发现低风险组患者的免疫、基质和 ESTIMATE 得分更高,这意味着 TME 的活力可能会使患者有更

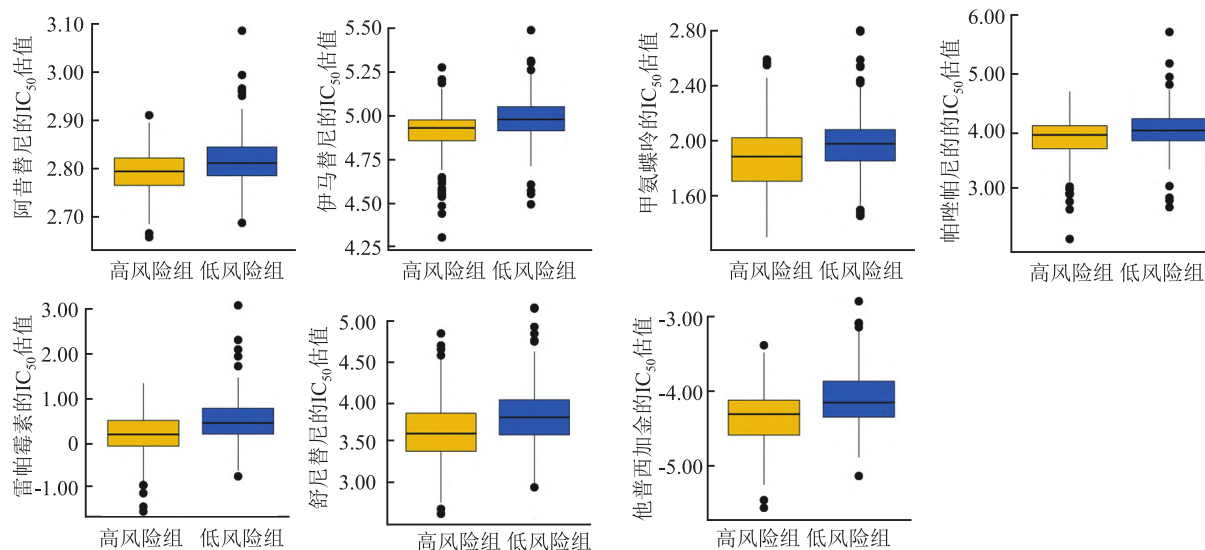


图7 两组风险组患者之间对化疗药物敏感性的比较

大的生存潜力。多项研究<sup>[10-11]</sup>表明肿瘤浸润 T 细胞与结肠癌患者的生存和预后密切相关。此外, CIBESORT 分析进一步证明了低风险组患者的 T 细胞浸润水平更高。这些结果显示, T 细胞的浸润在决定 CRC 患者的预后中扮演者重要作用。

除了肿瘤浸润 T 细胞对 CRC 的生存有着密切影响外, 单核细胞和中性粒细胞等髓系细胞也被证明会影响患者的生存<sup>[12-13]</sup>。在该研究中, 通过 CIBESORT 分析显示低风险组患者嗜酸性粒细胞、静止 DC 细胞、活化 DC 细胞和中性粒细胞的比例较高, 而调节性 T 细胞的比例较低。该研究表明 DC 细胞的活化有助于提高患者的生存, 并且肿瘤相关巨噬细胞和肿瘤相关中性粒细胞更倾向于限制低风险组患者的 CRC 进展。

此外, 该研究还分析了两种风险组患者的一些免疫检查点基因的表达水平。该研究表明 PDL1、PDL2 与 CTLA4 相关的 CD80 和 CD86、4-1BB、TIM3、TIGIT 和 ICOS 在低风险组中的表达量呈上升趋势。已有多项研究与该研究结果一致, 有研究<sup>[14-15]</sup>表明免疫检查点的表达增加往往会改善 CRC 患者的临床预后, CTLA4、PD-L1、ICOS 和 4-1BB 的高表达可通过促进 T 细胞增殖来改善预后。这些研究均认为高表达免疫检查点基因的 CRC 患者, 他们有可能从 ICI 疗法中获益, 这也是该研究结果显示低风险组患者长期生存率更高的原因。因此, 该研究假设活跃的免疫细胞浸润有助于提高免疫检查点的表达水平, 从而使 CRC 患者对免疫疗法产生更大的反应, 并获得良好的预后。

目前仍有许多患者对免疫疗法反应不佳。因此, 通过全球最大的药物敏感性信息数据库 GDSC 中确认了两风险组患者对化疗药物的药物敏感性。该结果提示了 7 种可能的药物(阿昔替尼、伊马替尼、甲氨蝶呤、帕唑帕尼、雷帕霉素、舒尼替尼和他普西加金), 这些药物对高风险组 CRC 患者可能更有效。然而, 该研究还存在一些不足之处。最主要的是这些免疫相关预后基因在 CRC 中的功能和机制还需要进一步的实验验证。此外, 还需要对大量临床样本进行研究确认。

### 参考文献

- [1] Sung H, Ferlay J, Siegel R L, 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): 209-49.
- [2] Siegel R L, Miller K D, Sauer A G, et al. Colorectal cancer statistics, 2020 [J]. CA Cancer J Clin, 2020, 70(3): 145-64.
- [3] Kennedy L B, Salama A K S. A review of cancer immunotherapy toxicity [J]. CA Cancer J Clin, 2020, 70(2): 86-104.
- [4] Zhang X, Wu T, Cai X, et al. Neoadjuvant immunotherapy for MSI-H/dMMR locally advanced colorectal cancer: New strategies and unveiled opportunities [J]. Front Immunol, 2022, 17(13): 795972.
- [5] Ganesh K, Stadler Z K, Cercek A, et al. Immunotherapy in colorectal cancer: Rationale, challenges and potential [J]. Nat Rev Gastroenterol Hepatol, 2019, 16(6): 361-75.
- [6] Gajewski T F, Schreiber H, Fu Y X. Innate and adaptive immune cells in the tumor microenvironment [J]. Nat Immunol, 2013, 14(10): 1014-22.
- [7] Ceelen W, Ramsay R G, Narasimhan V, et al. Targeting the

- tumor microenvironment in colorectal peritoneal metastases [J]. *Trends Cancer*, 2020, 6(3):236–46.
- [8] Zhu C, Xia Q, Gu B, et al. Esophageal cancer associated immune genes as biomarkers for predicting outcome in upper gastrointestinal tumors [J]. *Front Genet*, 2021, 19(12):707299.
- [9] Zhang H, Li X, Li Y, et al. An immune-related signature for predicting the prognosis of lower-grade gliomas [J]. *Front Immunol*, 2020, 8(11):603341.
- [10] Malka D, Lièvre A, André T, et al. Immune scores in colorectal cancer: Where are we? [J]. *Eur J Cancer*, 2020, 140:105–18.
- [11] Kamal Y, Dwan D, Hoehn H J, et al. Tumor immune infiltration estimated from gene expression profiles predicts colorectal cancer relapse [J]. *Oncoimmunology*, 2021, 10(1):1862529.
- [12] Ho W W, Gomes-Santos I L, Aoki S, et al. Dendritic cell paucity in mismatch repair-proficient colorectal cancer liver metastases limits immune checkpoint blockade efficacy [J]. *Proc Natl Acad Sci U S A*, 2021, 118(45):e2105323118.
- [13] Galdiero M R, Bonavita E, Barajon I, et al. Tumor associated macrophages and neutrophils in cancer [J]. *Immunobiology*, 2013, 218(11):1402–10.
- [14] Kitsou M, Ayiomamitis G D, Zaravinos A. High expression of immune checkpoints is associated with the TIL load, mutation rate and patient survival in colorectal cancer [J]. *Int J Oncol*, 2020, 57(1):237–48.
- [15] Zhang G, Xu Y, Zhou H. The infiltration of ICOS<sup>+</sup> cells in nasopharyngeal carcinoma is beneficial for improved prognosis [J]. *Pathol Oncol Res*, 2020, 26(1):365–70.

## The construction of a prognostic model for colorectal cancer based on immune gene correlation and drug screening

Zheng Wei<sup>1</sup>, Zhao Jiajia<sup>2</sup>, Cheng Xiang<sup>1</sup>, Tan Hongxin<sup>1</sup>, Huang Qi<sup>1</sup>

(<sup>1</sup>Dept of Oncology, The Second Affiliated Hospital of Anhui Medical University, Hefei 230601 ;

<sup>2</sup>Dept of Neurosurgery, The Second Affiliated Hospital of Anhui Medical University, Hefei 230601)

**Abstract Objective** To search for new biomarkers to predict prognosis in colorectal cancer (CRC) patients.

**Methods** A prognostic model was developed for colorectal cancer with immune-related genes from the cancer genome atlas (TCGA) database using one-way Cox regression analysis and least absolute shrinkage and selection operator (LASSO) regression analysis. Moreover, the immune infiltration characteristics of patients in high and low risk groups was compared by estimation of stromal and immune cells in malignant tumor tissues using expression data (ESTIMATE) and cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT). In addition, the expression levels of immune checkpoints were analyzed in patients from different risk groups. The sensitivity of patients in the two risk groups to chemotherapeutic agents was also compared based on genomics of drug sensitivity in cancer (GDSC). **Results** It was found that the prognostic model constructed based on immune genes could better predict the overall survival (OS) of CRC patients, and the results showed area under curve (AUC) values of 0.764 (95% CI:0.751–0.793), 0.773 (95% CI:0.761–0.779), and 0.760 (95% CI:0.742–0.774) for 1-, 3-, and 5-year OS, respectively. Patients in the low-risk group had higher expression levels of immune checkpoints and more abundant immune cells such as T cells ( $P < 0.001$ ), dendritic cells ( $P < 0.001$ ), macrophages ( $P < 0.001$ ), neutrophils ( $P < 0.001$ ). Patients in the high-risk group might be more sensitive to some chemotherapeutic agents such as axitinib, imatinib, methotrexate, pazopanib, rapamycin, sunitinib and tasigarnib. **Conclusion** A prognostic model based on 19 immune genes was effective in predicting the prognosis of CRC patients. The number and activity of immune cells in the immune microenvironment in different patients may be an important factor influencing their response to immunocheck inhibitors and chemotherapeutic agents.

**Key words** colorectal cancer; immune checkpoint inhibitors; tumour microenvironment; prognostic models; immune-related genes