

基于免疫相关基因构建神经母细胞瘤预后模型及潜在干预药物预测

马云飞¹, 史 学¹, 黄 静³, 尤圣杰¹, 刘 畅¹, 苏 雁², 赵 文², 黄 程², 秦 静², 胡 艳^{1*}, 倪 鑫^{4*}

1. 国家儿童医学中心 首都医科大学附属北京儿童医院 中医科, 北京 100045
2. 国家儿童医学中心 首都医科大学附属北京儿童医院 儿童肿瘤中心 肿瘤内科, 北京 100045
3. 中国中医科学院西苑医院 儿科, 北京 100091
4. 国家儿童医学中心 首都医科大学附属北京儿童医院 国家儿童肿瘤监测中心, 北京 100045

摘 要: **目的** 基于免疫相关基因构建神经母细胞瘤预后模型, 探讨其在神经母细胞瘤中的作用, 并预测挖掘潜在的干预中药。**方法** 通过 limma 包对 TARGET-NBL 数据集中高危患者组和非高危组表达谱数据进行差异分析。差异基因与 ImmPort 数据库中的免疫相关基因取交集, 对差异免疫相关基因进行富集分析。应用单因素/多因素 Cox 回归分析、Lasso 回归分析进一步筛选预后相关免疫差异基因, 据此建立预后模型。采用 CIBERSORT 和单样本基因集富集分析 (single-sample gene-set enrichment analysis, ssGSEA) 评估高风险评分组和低风险评分组样本中免疫细胞的丰度。通过 CTD 和 COREMINE 数据库预测与预后模型中基因高度相关的化合物和中药。**结果** 在 147 个差异免疫相关基因显著富集在 Janus 激酶 (Janus kinase, JAK) / 信号转导和转录激活因子 (signal transducer and activator of transcription, STAT)、Ras、B 细胞受体、环磷酸腺苷 (cyclic adenosine monophosphate, cAMP)、磷脂酰肌醇-3-羟激酶 (phosphatidylinositol-3-hydroxykinase, PI3K) / 蛋白激酶 B (protein kinase B, Akt)、Rap1、丝裂原活化蛋白激酶 (mitogen-activated protein kinase, MAPK)、转化生长因子- β (transforming growth factor- β , TGF- β) 等多个信号通路。单因素/多因素 Cox 回归分析显示促甲状腺激素 β 亚基 (thyroid stimulating hormone subunit beta, TSHB), 肽脱甲酰基酶、线粒体 (peptide deformylase, mitochondrial, PDF), 血管生成素样蛋白 3 (angiopoietin like 3, ANGPTL3), 前列腺素 E 受体 1 (prostaglandin E receptor 1, PTGER1) 是高危神经母细胞瘤预后的独立影响因素。基于 Lasso 分析构建的免疫预后模型中, 高风险评分组较低风险评分组预后差 (HR=4.49, 95% CI: 2.40~8.38, $P<0.001$), Calibration 校正曲线和决策曲线也证明了该模型具有较好的预后效能和一定的临床实用性。CIBERSORT 和 ssGSEA 免疫分析均显示 CD8⁺T 细胞在预后模型低风险评分组中浸润水平较高风险评分组显著升高 ($P<0.05$), 且与风险评分呈显著负相关 ($P<0.05$)。此外, 发现双酚 A (bisphenol A) 可同时调控 ANGPTL3、PDF、PTGER1、甲状旁腺激素受体 2 (parathyroid hormone receptor 2, PTH2R)、TSHB、神经生长因子诱导蛋白 (nerve growth factor inducible, VGF) 的表达, 进而抑制神经母细胞瘤的进展。基于免疫预后模型基因靶点预测得到菟丝子、茶树根等 130 味核心干预中药, 药味以苦、甘、辛为主, 药性以寒、温为主, 归经以肝、肺、脾、肾为主, 功效以清热、补虚为主。**结论** 在高危型神经母细胞瘤中, 免疫相关基因 ANGPTL3、PDF、PTGER1、PTH2R、TSHB、VGF 构建了具有较好预后效能和一定临床实用性的预后模型。通过整合多个数据库预测了作用于预后模型基因的药物, 为高危型神经母细胞瘤的治疗提供新的方向。

关键词: 神经母细胞瘤; 免疫相关基因; 预后模型; 双酚 A; 菟丝子; 茶树根; 生物信息学

中图分类号: R285 文献标志码: A 文章编号: 0253-2670(2024)13-4478-12

DOI: 10.7501/j.issn.0253-2670.2024.13.019

Construction of neuroblastoma prognostic model and prediction of potential intervention drugs based on immune-related genes

MA Yunfei¹, SHI Xue¹, HUANG Jing³, YOU Shengjie¹, LIU Chang¹, SU Yan², ZHAO Wen², HUANG Cheng², QIN Jing², HU Yan¹, NI Xin⁴

1. Department of Traditional Chinese Medicine, Beijing Children's Hospital, Capital Medical University, National Center for

收稿日期: 2024-02-19

基金项目: 北京“十四五”中医药重点专科项目 (领超类和示范类) (BJZKLC005)

作者简介: 马云飞 (1990—), 女, 在站博士后, 研究方向为中西医肿瘤学。E-mail: mayunfei2015@sina.com

*通信作者: 胡 艳 (1964—), 女, 博士生导师, 研究方向为中西医儿童风湿免疫学。E-mail: huyan1302@sina.com

倪 鑫 (1963—), 男, 博士生导师, 研究方向为儿童肿瘤。E-mail: nixin@bch.com.cn

Children's Health, Beijing 100045, China

2. Department of Medical Oncology, Pediatric Oncology Center, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China
3. Pediatric Department of Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing 100091, China
4. Department of National Center for Pediatric Cancer Surveillance, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China

Abstract: Objective To construct a neuroblastoma prognostic model based on immune-related genes, explore its role in neuroblastoma, and predict potential intervention Chinese medicine. **Methods** Limma package was used to analyze the difference between the expression profile data of high-risk patients and non-high-risk patients in the TARGET-NBL dataset. Differential genes were overlapped with immune-related genes from ImmPort database, and the resulting differential immune-related genes were enriched and analyzed. Univariate/multifactor Cox regression analysis and Lasso regression analysis were used to further screen prognostic-related immune differential genes, and the prognostic model was established. CIBERSORT and single-sample gene-set enrichment analysis (ssGSEA) were used to assess the abundance of immune cells in high risk score and low risk score samples. CTD and COREMINE databases were used to predict the compounds and Chinese medicines that highly interact with genes in the prognostic model. **Results** In high-risk patients, 147 differentially immune-related genes were significantly enriched in Janus kinase/signal transducer and activator of transcription (JAK/STAT), Ras, B cell receptor, cyclic adenosine monophosphate (cAMP), phosphatidylinositol-3-hydroxykinase/protein kinase B (PI3K/Akt), Rap1, mitogen-activated protein kinase (MAPK), transforming growth factor- β (TGF- β) and other signaling pathways. Univariate/multivariate Cox regression analysis showed that thyroid stimulating hormone subunit beta (TSHB), peptide deformylase, mitochondrial (PDF), angiopoietin like 3 (ANGPTL3), prostaglandin E receptor 1 (PTGER1) were independent prognostic factors of high-risk neuroblastoma. In the immune prognosis model constructed based on Lasso analysis, the high risk group had worse prognosis than the low risk group (HR = 4.49, 95% CI: 2.40—8.38, $P < 0.001$). Calibration curve and decision curve also prove that the model had good prognostic efficacy and certain clinical practicability. Both CIBERSORT and ssGSEA results showed that CD8⁺ T cells infiltration level was significantly increased in the low-risk group of prognostic model ($P < 0.05$) and significantly negatively correlated with risk score ($P < 0.05$). In addition, it was found that bisphenol A can simultaneously regulate the expression of *ANGPTL3*, *PDF*, *PTGER1*, parathyroid hormone receptor 2 (*PTH2R*), *TSHB* and nerve growth factor inducible nerve growth factor inducible (*VGF*), thus inhibiting the progression of neuroblastoma. Based on the gene target of immune prognosis model, 130 core traditional Chinese medicines such as Tusizi (*Cuscutae Semen*) and Chashugen (root of *Camellia sinensis*) were predicted, characterized by bitter, sweet and pungent tastes. The medicinal properties primarily focused on cold and warm, while the liver, lung, spleen and kidney meridians were the main channels affected, the function were mainly clearing heat and tonifying deficiency. **Conclusion** In high-risk neuroblastoma, immune-related genes *ANGPTL3*, *PDF*, *PTGER1*, *PTH2R*, *TSHB* and *VGF* construct a prognostic model with good prognostic efficacy and certain clinical practicability. By integrating multiple databases to predict drugs that act on prognostic model genes, it provides a new direction for the treatment of high-risk neuroblastoma. **Key words:** neuroblastoma; immune-related gene; prognostic model; bisphenol A; *Cuscutae Semen*; root of *Camellia sinensis*; bioinformatics

神经母细胞瘤 (neuroblastoma, NB) 是儿童最常见的颅外实体瘤, 起源于肾上腺髓质, 腹膜后、纵隔及脊柱旁交感神经节, 由于 NB 恶性程度高, 临床症状复杂, 近 50% 的 NB 属于预后不良的高危型, 对儿童的生命健康构成严重的威胁。最新研究数据表明, 在中国 NB 的发病率约为 5.9/100 万, 占儿童恶性肿瘤患者总数的 8%~10%, 但因 NB 死亡的人数占比却高达 15%^[1]。虽然高危 NB 经过放疗、化疗和手术、自体造血干细胞移植、顺式维甲酸诱导治疗、免疫治疗等多模式综合治疗后显著延长了

患者生存期, 但是高危型 NB 的 5 年生存率仍然低于 50%^[2]。近年来, 基于肿瘤微环境的治疗方法引起了人们的广泛关注, 而肿瘤微环境协助肿瘤细胞逃避免疫监视在肿瘤进展中发挥着重要作用。肿瘤浸润的免疫细胞和免疫相关基因 (immune-related genes, IRGs) 是肿瘤微环境的重要组成部分。研究表明, 肿瘤相关的巨噬细胞和癌症相关的成纤维细胞与 NB 的不良后果显著相关^[3]。然而, IRGs 在 NB 中的预后价值尚未阐明。目前, 抗双唾液酸神经节苷脂 (disialoganglioside, GD2) 单克隆抗体是唯一

用于高危 NB 患者的免疫治疗药物, 其可提高 2 年总生存期和无事件生存期, 但其具有显著的不良反应及长期疗效尚不明确的特点^[4]。而中医药作为多学科协作诊疗的一部分在肿瘤的治疗中发挥着重要的作用, 尤其是在调节免疫功能、增强抗肿瘤疗效上有着独特优势。因此, 本研究旨在挖掘高危型 NB 的预后 IRGs, 构建并探讨基于 IRGs 预后模型作为预测高危型 NB 人群的可行性, 以帮助临床医生优化个体治疗方案; 另外, 通过整合多个数据库预测作用于预后模型基因的药物, 为高危型 NB 的治疗提供新的方向。

1 材料与方法

1.1 数据来源

从 UCSU 数据库 (<https://xenabrowser.net/datapages/>) 中下载 TARGET-NBL 项目中 112 例 NB 患者的基因表达谱数据 (HTSeq-FPKM-UQ) 及相应的临床信息 (phenotype/survival data), 基于美国儿童肿瘤协作组 (Children's Oncology Group, COG) 危险度分层系统, 包括国际 NB 分期 (international neuroblastoma staging system, INSS)、年龄、MYCN 原癌基因扩增、国际神经母细胞瘤病理学分类和 DNA 倍型, 将 NB 分为低危、中危和高危 3 个组, 当前研究中高危患者 88 例, 非高危患者 24 例。从 ImmPort (<https://www.immport.org/home>) 数据库下载 IRGs, 最终整理得到了 1 793 个 IRGs。

1.2 差异基因 (differentially expressed genes, DEGs) 分析

根据 NB 患者样本分组情况, 将样本分为高危患者组和非高危组。通过 limma 包^[5]对表达谱数据进行差异分析。设定 $|\log_2FC| > 1$ 且调整后 P 值 (P_{adj}) < 0.05 为 DEGs 的阈值。其中, $\log_2FC > 1$ 且 $P_{adj} < 0.05$ 的基因为上调的 DEGs, $\log_2FC < -1$ 且 $P_{adj} < 0.05$ 的基因为下调的 DEGs。通过 ggplot2 包^[6]和 VennDiagram 包^[7]对 DEGs 和 IRGs 的交集进行可视化, 以及通过 ggplot2^[6]和 pheatmap (<https://CRAN.R-project.org/package=pheatmap>) 分别绘制火山图及热图对 DEGs 进行展示。

1.3 富集分析

基因本体论 (gene ontology, GO) 是一种主要的生物信息学工具, 用于注释基因并分析这些基因的生物学作用, 包括生物过程 (biological process, BP)、细胞组成 (cellular component, CC)、分子功能 (molecular function, MF) 3 个部分^[8]。京都基因

与基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 是一个数据库资源, 用于从通过高通量实验技术生成的大规模分子数据集中了解高级功能和生物系统^[9]。使用 R 包 clusterProfiler^[10]对差异免疫相关基因 (differential immune-related genes, DIRGs) 进行 GO 注释和 KEGG 途径富集分析, 其中 $P_{adj} < 0.05$ 被认为具有统计学差异。

1.4 预后模型构建

为了评估基因表达对高危患者组的预后能力, 以 TARGET-NBL 表达谱数据集为基础, 应用单/多因素 Cox 回归分析、最小绝对收缩和选择算子 (logistic least absolute shrinkage and selection operator, Lasso) 回归分析进一步筛选预后相关基因, 并据此建立预后模型。首先, 使用单变量 Cox 比例回归分析计算每个 DIRGs 的表达与总生存 (overall survival, OS) 之间的关联, 并保留 $P < 0.1$ 的 DIRGs。随后, 使用 Lasso 算法消除多重共线性并筛选单因素 Cox 回归分析中有意义的变量, 并根据给定的风险评分将患者分为高风险组和低风险组, 应用 survival 包^[11]进行 Kaplan-Meier 分析和 Cox 回归检验。基于 Lasso 回归分析结果, 通过 rms (<https://CRAN.R-project.org/package=rms>) 包构建预后 DIRGs 的列线图用于预测 NB 的 OS, 并通过 Calibration 校准曲线和决策曲线分析 (decision curve analysis, DCA) 来评估列线图的准确性及临床效用。

1.5 免疫分析

CIBERSORT^[12]是基于线性支持向量回归的原理对转录组表达矩阵进行反卷积, 利用 CIBERSORT 算法结合 22 种免疫细胞亚型的基因表达特征集, 从而在混合的细胞中估算出免疫细胞的组成和丰度, 得出免疫细胞浸润矩阵。单样本基因集富集分析 (single-sample gene-set enrichment analysis, ssGSEA)^[13]可量化每个免疫细胞浸润的相对丰度。通过标记各浸润免疫细胞类型利用 ssGSEA 分析计算的富集分数来表示各样本中各免疫细胞浸润的相对丰度。最后, 通过 pheatmap 绘制热图展示免疫细胞浸润在每个样本中的分布。采用 Spearman 分析数据免疫细胞浸润与 Lasso 风险评分之间相关性, 并通过 ggplot2 可视化结果。

1.6 潜在药物预测

比较毒理学数据库 (comparative toxicogenomics database, CTD)^[14]是将化学品、基因、表型、疾病和暴露的毒理学信息联系起来的公开数据库, 通

过该平台预测与预后 DIRGs 作用的单体化合物, 通过 igraph 包^[15]对相互作用 (interactions) >3 的互作网络进行可视化。然后在 COREMINE (<https://coremine.com/medical/>) 数据库匹配对应化合物的中药, 参考《中国药典》2020 年版、《中华本草》和《中药学》对筛选出的中药进行四气、五味、归经和功效统计分析, 并通过雷达图和柱状图进行可视化。

1.7 分子对接

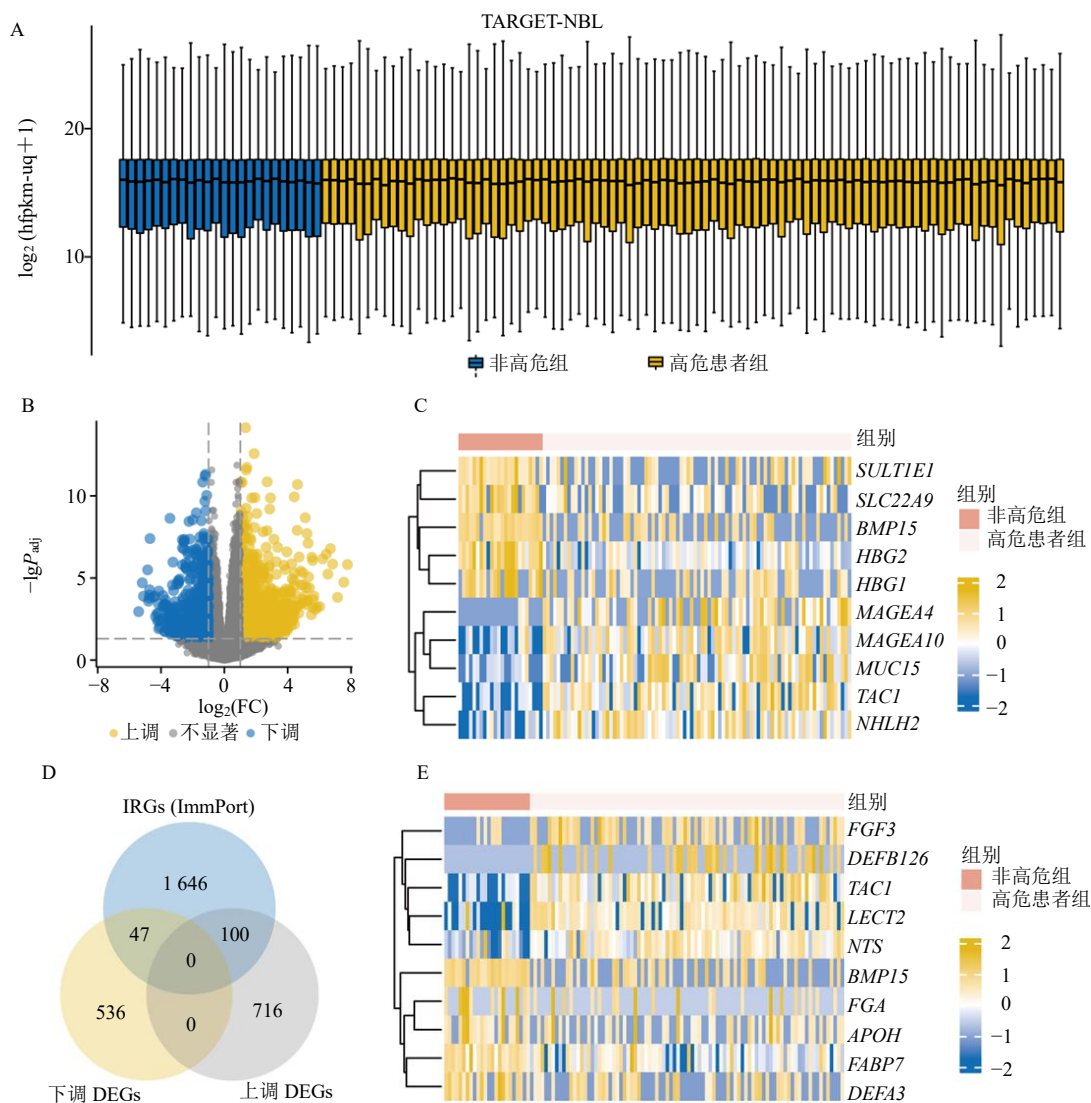
为了评估候选药物与其靶点之间的相互作用模式, 采用 AutodockVina 1.2.2 对蛋白-配体进行对

接^[16]。从 PubChem 化合物数据库 (<https://pubchem.ncbi.nlm.nih.gov/>) 中获得药物的分子结构^[17]。从 PDB (<http://www.rcsb.org/>) 下载蛋白的 3D 结构。首先将所有蛋白质和配体文件都转换为 PDBQT 格式, 去除了所有水分子, 并添加极性氢原子。由 Autodock Vina 1.2.2 (<http://autodock.scripps.edu/>) 对分子对接进行可视化。

2 结果

2.1 DEGs 和 DIRGs 分析

对 TARGET-NBL 数据集 (图 1-A) 中非高危组和高危患者组基因进行差异分析, 结果如图 1-B 所



A-TARGET-NBL 数据集样本分布; B-TARGET-NBL 数据集 DEGs 分析火山图; C-TARGET-NBL 数据集 top 5 差异上调、下调基因表达热图; D-DEGs 和 IRGs 韦恩图; E-TARGET-NBL 数据集 top 5 差异免疫相关的上调和下调基因表达热图。

A-sample distribution of TARGET-NBL dataset; B-volcano map of TARGET-NBL dataset DEGs analysis; C-heat map of top 5 differentially up-regulated and down-regulated gene expression in TARGET-NBL dataset; D-Venn diagram of DEGs and IRGs; E-heat map of top 5 differentially up-regulated and down-regulated immune-related genes expression in TARGET-NBL dataset.

图 1 TARGET-NBL 数据集差异分析

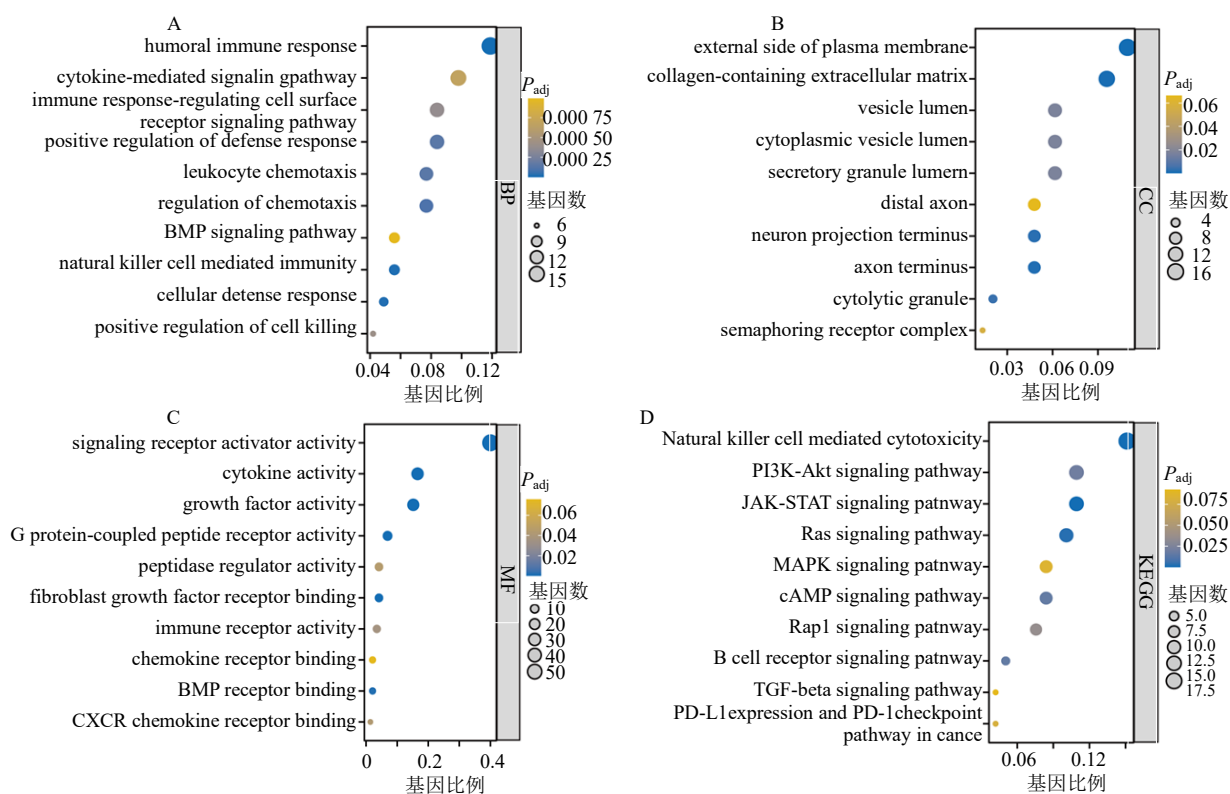
Fig. 1 Difference analysis of TARGET-NBL dataset

示, 满足 $|\log_2FC| > 1$ 且 $P_{adj} < 0.05$ 阈值的 DEGs 有 1 399 个, 高危患者组中上调基因 816 个, 下调基因 583 个, 排名前 5 的 DEGs 表达热图如图 1-C 所示。将 IRGs 与 DEGs 取交集 (图 1-D), 得到 147 个 DIRGs, 上调 DIRGs 100 个, 下调 DIRGs 47 个, 排名前 5 的 DIRGs 表达热图如图 1-E 所示。

2.2 GO 注释及 KEGG 信号通路富集分析

基于非高危组和高危患者组 147 个 DIRGs 进行 GO 注释及 KEGG 信号通路富集分析 (图 2, P_{adj} 均 < 0.05)。GO 注释结果表明, DIRGs 主要与体液免疫反应 (humoral immune response)、自然杀伤细胞介导的免疫 (natural killer cell mediated immunity)、细胞防御反应 (cellular defense response)、趋化性调控 (regulation of chemotaxis)、免疫反应调节细胞表面受体信号通路 (immune response-regulating cell surface receptor signaling pathway) 等生物学过程有关 (图 2-A); 细胞组分主要富集在质膜的外侧 (external side of plasma membrane)、含胶原蛋白的细胞外基质 (collagen-

containing extracellular matrix)、轴突末端 (axon terminus)、细胞溶解的颗粒 (cytolytic granule)、细胞质囊腔 (cytoplasmic vesicle lumen) 等 (图 2-B); 分子功能主要富集在信号受体激活剂活性 (signaling receptor activator activity)、生长因子活性 (growth factor activity)、细胞因子活性 (cytokine activity)、成纤维细胞生长因子受体结合 (fibroblast growth factor receptor binding)、免疫受体活性 (immune receptor activity)、趋化因子受体结合 (chemokine receptor binding) 等 (图 2-C)。而 KEGG 信号通路富集分析结果显示, DIRGs 主要富集于自然杀伤细胞介导的细胞毒性 (natural killer cell mediated cytotoxicity)、JAK-STAT 信号通路 (JAK-STAT signaling pathway)、Ras 信号通路 (Ras signaling pathway)、B 细胞受体信号通路 (B cell receptor signaling pathway)、cAMP 信号通路 (cAMP signaling pathway)、PI3K-Akt 信号通路 (PI3K-Akt signaling pathway)、Rap1 信号通路 (Rap1 signaling pathway)、肿瘤中 PD-L1 表达和 PD-1 检查点通路



A~C-差异免疫相关基因的 GO 注释 (BP-生物过程, CC-细胞成分, MF-分子功能); D-差异免疫相关基因的 KEGG 富集分析。

A—C—GO annotation of differential immune-related genes (BP-biological process, CC-cellular component, MF-molecular function); D—KEGG enrichment analysis of differential immune-related genes.

图 2 差异免疫相关基因的 GO 注释及 KEGG 富集分析

Fig. 2 GO annotation and KEGG enrichment analysis of differential immune-related genes

(PD-L1 expression and PD-1 checkpoint pathway in cancer)、MAPK 信号通路 (MAPK signaling pathway)、TGF- β 信号通路 (TGF-beta signaling pathway) 等 (图 2-D)。

2.3 预后模型构建

在 TARGET-NBL 数据集中, 对 147 个 DIRGs 进行了单/多因素 Cox 回归分析 (图 3-A、B), Cox 回归分析显示促甲状腺激素 β 亚基 (thyroid stimulating hormone subunit beta, TSHB), 肽脱甲酰基酶、线粒体 (peptide deformylase, mitochondrial, PDF), 血管生成素样蛋白 3 (angiopoietin like 3, ANGPTL3), 前列腺素 E 受体 1 (prostaglandin E receptor 1, PTGER1) 是高危神经母细胞瘤预后的独立影响因素。将单因素 Cox 回归分析中 $P < 0.1$ 的 DIRGs 包括 C-X-C 模体趋化因子配体 9 (C-X-C motif chemokine ligand 9, CXCL9)、TSHB、胰岛素样肽 5 (insulin-like peptide 5, INSL5)、甲状旁腺激素受体 2 (parathyroid hormone receptor 2, PTH2R)、PDF、ANGPTL3、PTGER1、CXCL11、杀伤细胞凝集素样受体亚家族 C4 (killer cell lectin like receptor C4, KLRC4)、速激肽前体 1 (tachykinin precursor 1, TAC1)、神经生长因子诱导蛋白 (nerve growth factor inducible, VGF) 进一步进行 Lasso 分析, 最终共鉴定出 6 个核心 DIRGs 用于 NB 风险评分模型的构建 (图 3-C、D), 包括 TSHB、PTH2R、PDF、ANGPTL3、PTGER1 和 VGF (表 1)。根据 Lasso 风险评分将患者分为高风险组和低风险组进行生存预后分析, 结果 (图 3-E) 显示高风险组较低风险组预后差, 其中风险比 (hazard ratio, HR) 为 4.49, 95% 置信区间 (confidence interval, CI) 为 2.40~8.38, $P < 0.001$ 。由 TSHB、PTH2R、PDF、ANGPTL3、PTGER1 和 VGF 构成的风险因子图 (图 3-F) 也显示 Lasso 风险评分越高预后越差。此外, 基于 Lasso 分析筛选的预后变量绘制了列线图 (图 3-G), Calibration 分析显示实际概率和模型预测的概率拟合较好, 表明该模型具有较好的预后效能 (图 3-H), 决策曲线 (图 3-I、J) 结果表明列线图预后模型具有一定的临床实用性。

2.4 免疫分析

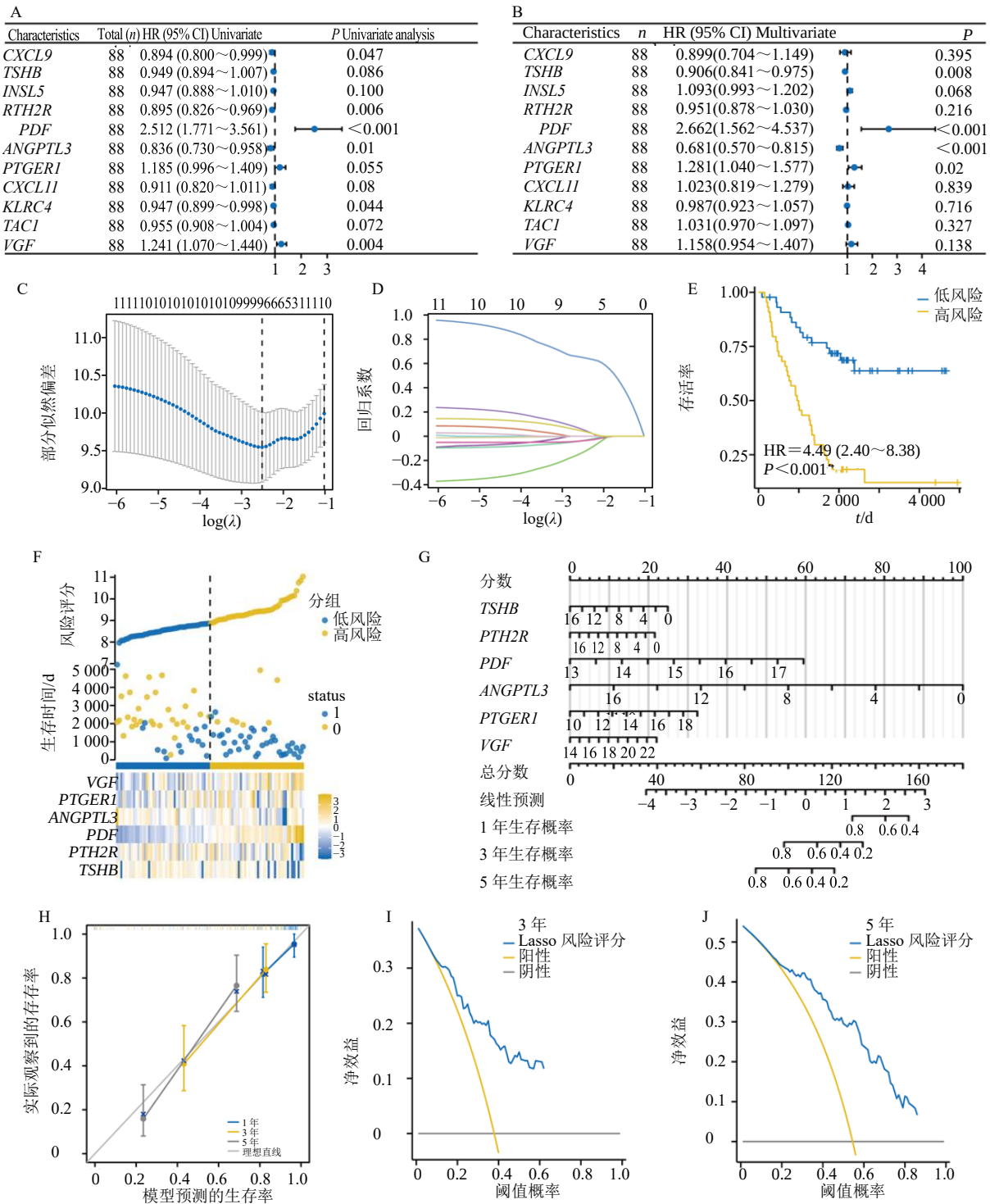
采用 ssGSEA 评估了高风险组和低风险组样本中免疫细胞的丰度 (图 4-A), 其中, CD8⁺T 细胞在低风险组的浸润水平较高风险组显著升高 ($P < 0.05$, 图 4-B), 样本中风险评分与 CD8⁺T 细胞浸润

水平呈显著负相关 ($P < 0.05$, Correlation = -0.221, 图 4-C)。此外, 采用 CIBERSORT 免疫分析也评估了高风险组和低风险组样本中免疫细胞的丰度 (图 4-D), 结果也显示 CD8⁺T 细胞在低风险组的浸润水平较高风险组显著升高 ($P < 0.05$, 图 4-E), 样本中风险评分与 CD8⁺T 细胞浸润水平呈显著负相关 ($P < 0.05$, Correlation = -0.215, 图 4-F)。

2.5 潜在药物预测

基于 CTD 数据库预测与预后核心 DIRGs 相关的化合物所构建的网络如图 5-A 所示, 其中, 与 ANGPTL3 相关的化合物有双酚 A (bisphenol A)、环孢霉素 (cyclosporine) 和黄曲霉毒素 B1 (aflatoxin B1), 与 PDF 相关的化学药物有双酚 A、氟尿嘧啶 (fluorouracil) 和奥沙利铂 (oxaliplatin); 与 PTGER1 相关的化合物有地诺前列酮 (dinoprostone)、苯并芘 [benzo (a) pyrene]、雌二醇 (estradiol); 与 PTH2R 相关的化合物有双酚 A、苯并芘、雌二醇; 与 TSHB 相关的化学药物有维罗非尼 (vemurafenib)、双酚 A、甲状腺素 (thyroxine); 与 VGF 相关的化合物有丙戊酸 (valproic acid)、苯并芘、双酚 A。随后将这些化学成分通过 COREMINE 数据库检索含有这些化学成分的中草药, 结果显示 130 味中药作用于预后核心 DIRGs (表 2)。为了进一步探究这些药物的特点, 根据《中国药典》2020 年版、《中华本草》和《中药学》对这些中药进一步统计分析, 结果发现中药药味以苦 (70 次)、甘 (53 次)、辛 (51 次) 为主 (图 5C), 药性以寒 (45 次)、温 (44 次) 为主 (图 5B), 归经以肝经 (66 次)、肺经 (56 次)、脾 (49 次)、肾 (46 次) 为主 (图 5-D), 功效以清热药 (33 次)、补虚药 (30 次) 为主 (图 5-E)。

在药物预测结果中, 发现双酚 A 与多个靶点相关, 为了评估双酚 A 对其靶点的亲和力, 本研究随后进行了分子对接分析。用 Autodock Vina v.1.2.2 获得候选药物双酚 A 与 ANGPTL3 (PDB id: 6EUA)、PDF (PDB id: 3G5K)、PTH2R (PDB id: 7F16)、TSHB (PDB id: 7UTZ)、VGF (PDB id: 7D16) 蛋白质的结合姿势和相互作用 (图 6), 候选药物双酚 A 通过可见的氢键和强静电相互作用与 ANGPTL3、PDF、PTH2R、TSHB、VGF 蛋白质靶点相结合。ANGPTL3、PDF、PTH2R、TSHB、VGF 与双酚 A 的结合能依次为 -6.781、-7.134、-6.627、-7.390、-77.913 kcal/mol (1 kcal = 4.2 kJ), 表明其结合均具有较好的稳定性 (表 3)。



A-基于 147 个 DIRGs 进行了单因素 Cox 回归分析; B-基于单因素 Cox 回归分析中 $P<0.1$ 的 DIRGs 进行多因素 COX 回归分析; C-基于单因素 COX 回归分析中 $P<0.1$ 的 DIRGs 进行多因素 Lasso 系数筛选; D-Lasso 变量轨迹图; E-基于 Lasso 风险评分将患者分为高风险组和低风险组的生存分析; F-基于 Lasso 筛选的变量绘制风险因子图; G-基于 Lasso 筛选的变量绘制列线图; H-基于预后模型绘制 Calibration 校准曲线; I-基于预后模型绘制 3 年决策曲线; J-基于预后模型绘制 5 年决策曲线。

A-univariate Cox regression analysis was performed based on 147 DIRGs; B-univariate Cox regression analysis of DIRGs ($P<0.1$) for multivariate COX regression analysis; C-univariate Cox regression analysis of DIRGs ($P<0.1$) for Lasso; D-Lasso variable trajectory diagram; E-survival analysis of patients divided into high risk group and low risk group based on Lasso risk score; F-draw risk factor map based on Lasso-screened variables; G-draw nomogram based on Lasso-filtered variables; H-calibration curve based on nomogram; I-3-year decision curve based on prognostic model; J-5-year decision curve based on prognostic model.

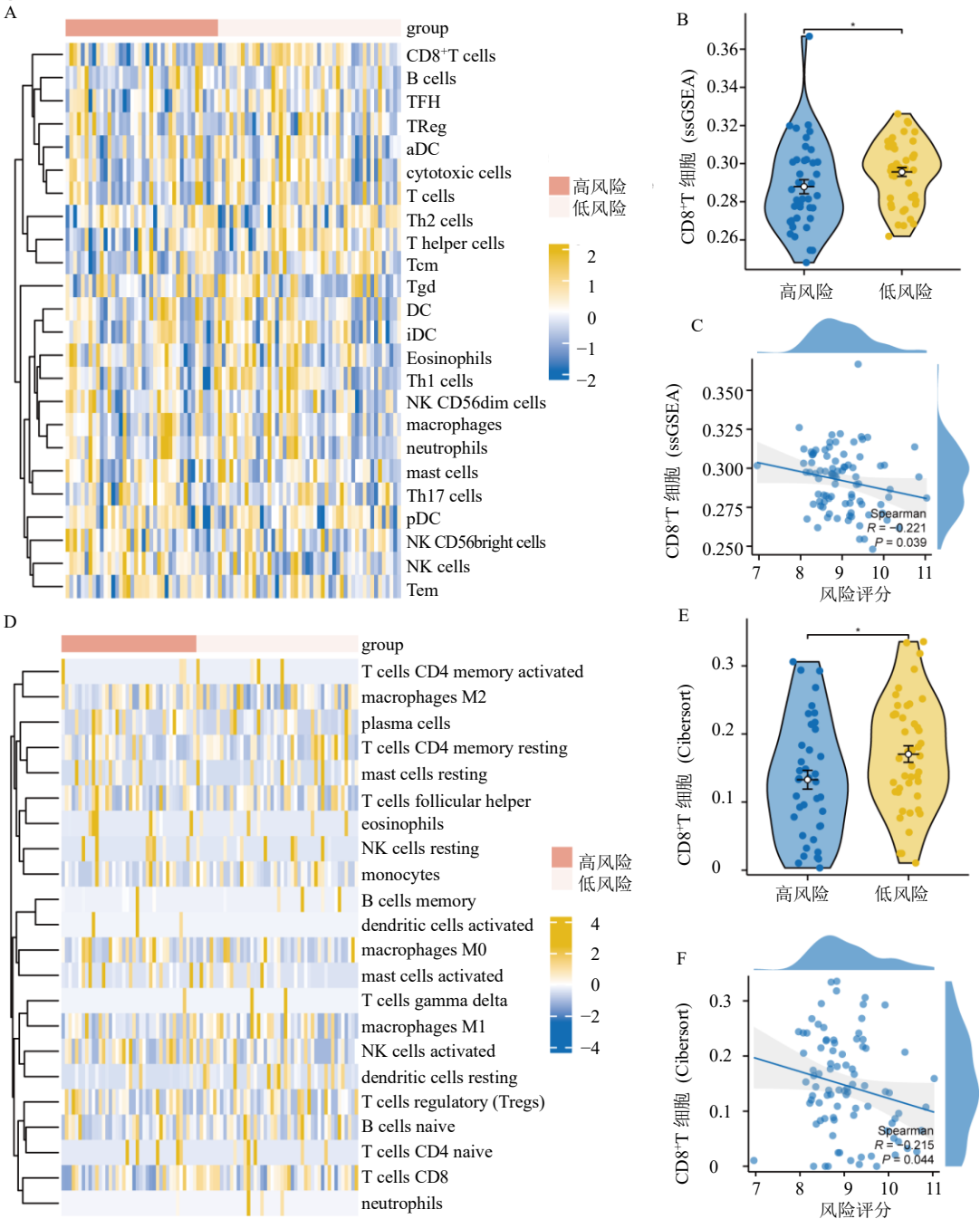
图 3 基于 Lasso 分析构建神经母细胞瘤的预后模型

Fig. 3 Construction of neuroblastoma prognostic model based on Lasso analysis

表 1 神经母细胞瘤预后模型中的差异免疫相关基因

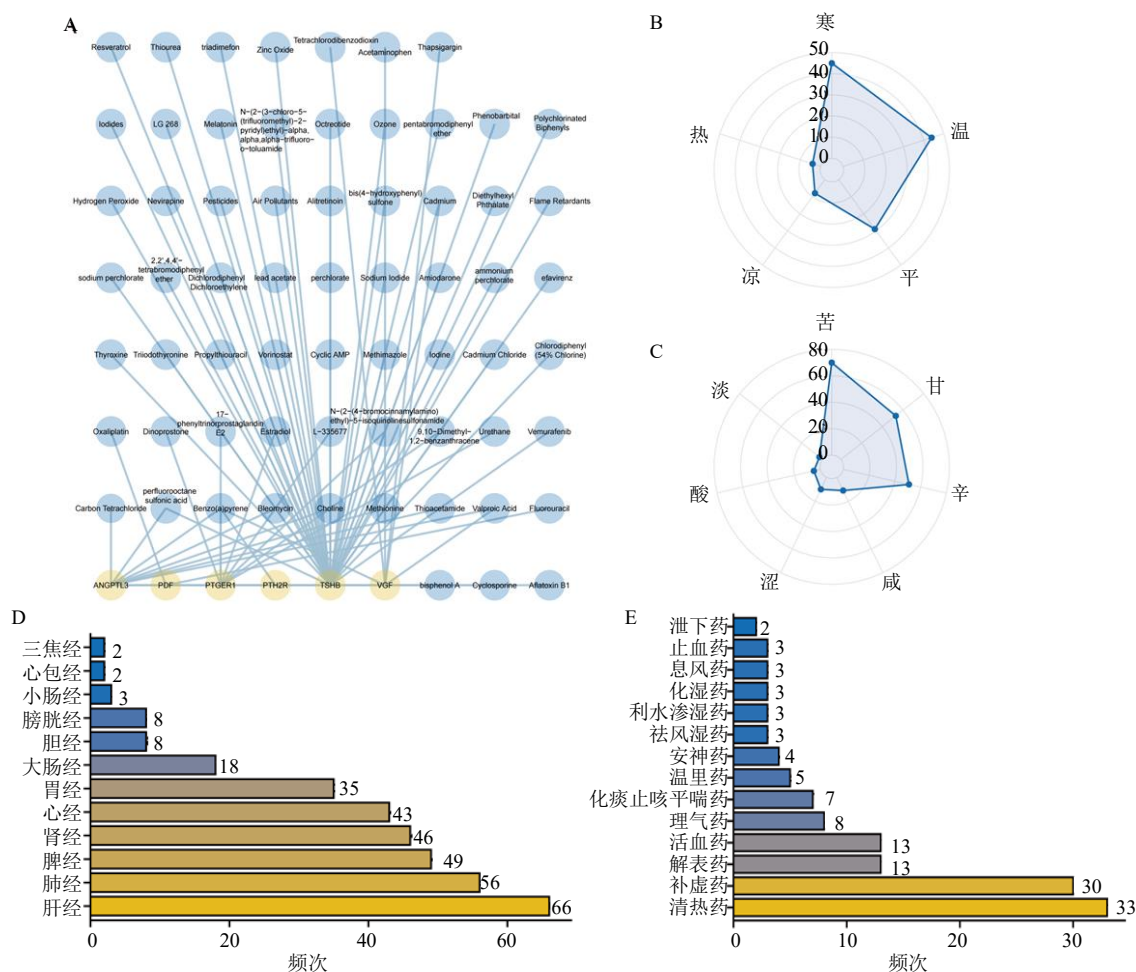
Table 1 Differential immune-associated genes in neuroblastoma prognostic models

序号	基因	基因名称	序号	基因	基因名称
1	ANGPTL3	血管生成素样蛋白 3 (angiopoietin like 3)	4	PTH2R	甲状旁腺激素受体 2 (parathyroid hormone receptor 2)
2	PDF	肽脱甲酰酶, 线粒体 (peptide deformylase, mitochondrial)	5	TSHB	促甲状腺激素 β 亚基 (thyroid stimulating hormone subunit beta)
3	PTGER1	前列腺素 E 受体 1 (prostaglandin E receptor 1)	6	VGF	神经生长因子诱导型 (nerve growth factor inducible)



A-采用 ssGSEA 评估高风险组和低风险组样本中免疫细胞的丰度; B-CD8⁺T 细胞在低风险和高风险组中的浸润水平; C-风险评分与 CD8⁺T 细胞浸润水平的相关性; D-采用 CIBERSORT 分析高风险组和低风险组样本中免疫细胞的丰度; E-CD8⁺T 细胞在低风险组和高风险组中的浸润水平; F-风险评分与 CD8⁺T 细胞浸润水平的相关性; * $P < 0.05$ 。
A-abundance of immune cells in samples of high risk score group and low risk score group by ssGSEA; B-infiltration level of CD8⁺T cells in low risk score group and the high risk score group; C-correlation between risk score and infiltration level of CD8⁺T cells; D-abundance of immune cells in samples of high risk score group and low risk score group by CIBERSORT; E-infiltration level of CD8⁺T cells in low risk scoregroup and high risk score group; F-correlation between risk score and infiltration level of CD8⁺T cells; * $P < 0.05$.

图 4 ssGSEA 和 CIBERSORT 免疫分析
Fig. 4 Immune analysis of ssGSEA and CIBERSORT



A-预后核心 DIRGs 及相关化合物的网络图; B-中药四气统计分析雷达图; C-中药五味统计分析雷达图; D-中药归经统计分析柱状图; E-中药功能统计分析柱状图。

A-network diagram of prognostic hub DIRGs and chemical compounds; B-radar diagram for statistical analysis of four natures of traditional Chinese medicine; C-radar diagram for statistical analysis of five flavours of traditional Chinese medicine; D-column diagram for statistical analysis of channel tropism of traditional Chinese medicine; E-column diagram for statistical analysis of functions of traditional Chinese medicine.

图 5 潜在药物预测分析

Fig. 5 Predictive analysis of potential drugs

3 讨论

目前,虽然研究人员已经发现了可能有助于临床医生治疗 NB 的预后标志物,但很少有患者从这些标志物中获益,特别是那些处于亚组的患者^[18-19]。IRGs 以及肿瘤浸润的免疫细胞在高危型 NB 的预后价值尚未阐明。本研究发现,在高危患者中 147 个差异 IRGs 显著富集在 JAK/STAT、Ras、B 细胞受体、cAMP、PI3K/Akt、Rap1、MAPK、TGF- β 等多个信号通路。基于 Lasso 分析,最终共鉴定出 *TSHB*、*PTH2R*、*PDF*、*ANGPTL3*、*PTGER1* 和 *VGF* 6 个核心 DIRGs 用于神经母细胞瘤免疫预后模型的构建。*TSHB* 是一种编码介导甲状腺激素的特异性高亲和力核受体的基因,研究表明, *THRB* 可能与肿瘤的发

生、进展有关^[20-21]。*PTH2R* 在卵巢癌中高表达,与树突细胞、中性粒细胞、NK 细胞呈负相关,敲低 *PTH2R* 可以抑制卵巢癌细胞的增殖、侵袭和迁移^[22]。*PDF* 是存在于线粒体中的一种含有亚铁离子的金属蛋白酶,其广泛过表达于肺癌、乳腺癌和结肠癌等癌细胞中,且通过放线菌素抑制或敲除 *PDF*,可以显著降低这些癌细胞的增殖^[23]。目前已有多种抗菌活性的 *PDF* 抑制剂,如放线菌素、LBM415、GSK1322322 以及 BB-3497 等。*ANGPTL3* 是脂质和葡萄糖代谢以及胰岛素敏感性的重要调节因子,但在口腔癌、肝癌和卵巢癌中均观察到 *ANGPTL3* 过表达,并且可以诱导口腔癌肿瘤微环境中肿瘤相关成纤维细胞的分化^[24]。目前针对 *ANGPTL3* 靶点

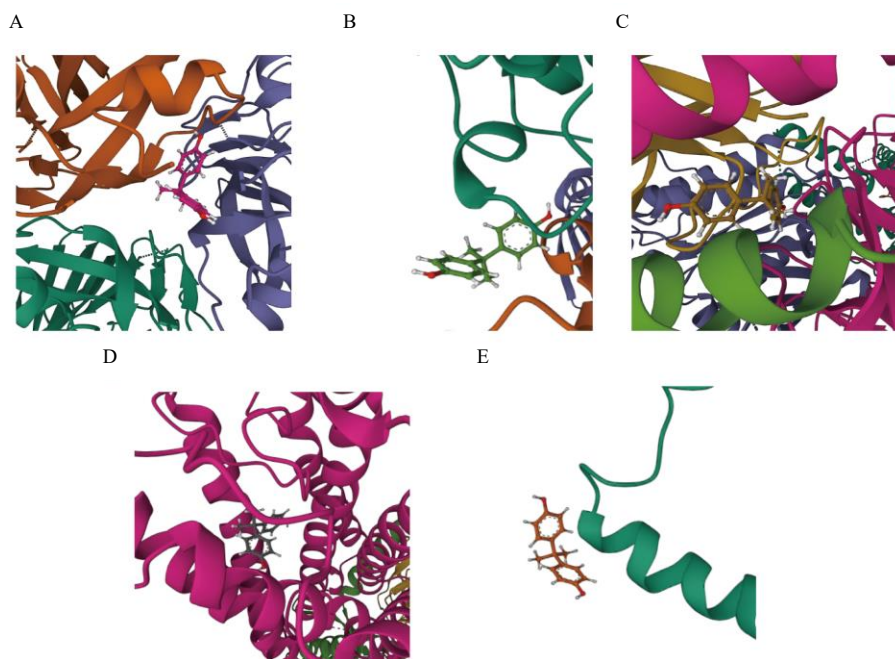
表 2 作用于神经母细胞瘤预后模型中的差异免疫相关基因的化学成分及中药

Table 2 Chemical components and traditional Chinese medicines acting on differential immune-related genes of neuroblastoma prognostic models

基因	化学成分（相互作用数）	代表中药（ $P<0.05$ ）
ANGPTL3	双酚 A（7）	菟丝子、石决明、芦根、马齿苋、瓦楞子、紫梢花、干漆、茶树根
ANGPTL3	环孢霉素（5）	雷公藤、冬虫夏草、蜂蜜
ANGPTL3	黄曲霉毒素 B1（4）	肉豆蔻、肉桂、甜杏仁、白花蛇舌草、茶树根、干姜、丁香、姜黄、莲子、红参、玉米须、半枝莲、紫芝、栀子花、人参、薏苡仁、山药、萆薢、大腹皮、干漆、栀子、土鳖虫、覆盆子、橘核、青皮、番泻叶、陈皮、太子参、甘草
PDF	双酚 A（3）	菟丝子、石决明、芦根、马齿苋、瓦楞子、干漆、紫梢花
PDF	氟脲嘧啶（3）	阿胶、重楼、地榆、半枝莲、白花蛇舌草、败酱草、喜树果、土贝母、云芝、薏苡仁、苍术、锁阳
PDF	奥沙利铂（3）	重楼、川乌、茯苓、肉桂、鸦胆子、桂枝、栀子、白术
PTGER1	地诺前列酮（27）	红花、旋覆花、生姜、干姜、射干、五味子、黄芩、海藻、鹅不食草、赤芍、五味子、山茱萸、郁金、威灵仙、姜黄、远志、苏木、苍术、香薷、乳香、甘草、苍术、昆布、当归、桂枝、肉桂、干漆、白术、金银花、野菊花、黄连、人参、款冬花、骨碎补、山药、淫羊藿、白芷、大青叶、钩藤、野菊花、升麻、重楼、茯神、白花蛇舌草、苍术、黄柏、肿节风、白豆蔻、菝葜、远志、丹参、鱼腥草、细辛、侧柏叶、柴胡、柴胡、旋覆花、茯苓、夏枯草、黄柏、板蓝根、栀子、香薷、石见穿、麻黄根、麻黄、羌活、木香、苍耳子、平贝母、茶树根、十大功劳叶、佩兰、石斛、黄芪、蛇床子、白果、银杏叶、穿山龙、鸦胆子、夏天无、红参
PTGER1	苯并芘（5）	茶树根、瓦楞子、石决明、红参、石上柏、蛇莓、人参、大腹皮、鹅不食草、苦楝子、郁金、姜黄、菝葜、番泻叶、肉桂、仙鹤草、树舌灵芝、桂枝、大黄、麻黄根
PTGER1	雌二醇（4）	水牛角、菟丝子、麦芽、葛根、淫羊藿、升麻、当归、川牛膝、女贞子、补骨脂、地黄、仙茅、玉米须、茯苓、菟丝子、茯神、首乌藤、蜂蜜、桂枝
PTH2R	双酚 A（5）	菟丝子、石决明、芦根、马齿苋、瓦楞子、干漆、紫梢花
PTH2R	苯并芘（4）	茶树根、瓦楞子、石决明、红参、石上柏、蛇莓、人参、大腹皮、鹅不食草、苦楝子、郁金、姜黄、菝葜、番泻叶、肉桂、仙鹤草、树舌灵芝、桂枝、大黄、麻黄根
PTH2R	雌二醇（2）	水牛角、菟丝子、麦芽、葛根、淫羊藿、升麻、当归、川牛膝、女贞子、补骨脂、地黄、仙茅、玉米须、茯苓、菟丝子、茯神、首乌藤、蜂蜜、桂枝
TSHB	维罗非尼（22）	木鳖子
TSHB	双酚 A（19）	菟丝子、石决明、芦根、马齿苋、瓦楞子、干漆、紫梢花
TSHB	甲状腺素（19）	水牛角、大黄、玄参、蛤蚧、蜂蜜
VGF	丙戊酸（6）	紫草、缬草、钩藤、天麻、积雪草
VGF	苯并芘（5）	茶树根、瓦楞子、石决明、红参、石上柏、蛇莓、人参、大腹皮、鹅不食草、苦楝子、郁金、姜黄、菝葜、番泻叶、肉桂、仙鹤草、树舌灵芝、桂枝、大黄、麻黄根
VGF	双酚 A（4）	菟丝子、石决明、芦根、马齿苋、瓦楞子、干漆、紫梢花

用于高脂血症治疗的靶向药物 Evinacumab 已获批上市。EP2（由 PTGER1 编码）和 EP4（由 PTGER4 编码）是肺中性粒细胞中表达水平较高的 PGE2 受体，在间充质细胞与中性粒细胞共培养体系中，同时抑制 EP2 和 EP4 可明显降低间充质细胞诱导肺中性粒细胞的免疫抑制能力，并减轻乳腺癌的肺转移^[25]。VGF 在小细胞肺癌和大细胞神经内分泌瘤中显著高表达，靶向 VGF 可以减缓癌细胞的生长和

肿瘤的形成^[26]。本研究显示，ANGPTL3、THRB、PTH2R 高表达的 NB 较低表达的 NB 预后好，而 PDF、PTGER1、VEG 高表达的 NB 较低表达的 NB 预后差，其所构建的预后模型中高风险组较低风险组预后差（HR=4.49, 95% CI: 2.40-8.38, $P<0.001$ ）。Calibration 校正曲线和决策曲线也证明了该模型具有较好的预后效能和一定的临床实用性。ssGSEA 和 CIBERSORT 免疫分析分别评估了高风险组和低风险



A-双酚 A 与 ANGPTL3 的结合方式; B-双酚 A 与 PDF 的结合方式; C-双酚 A 与 PTH2R 的结合方式; D-双酚 A 与 TSHB 的结合方式; E-双酚 A 与 VGF 的结合方式。

A-binding mode of bisphenol A with ANGPTL3; B-binding mode of bisphenol A with PDF; C-binding mode of bisphenol A with PTH2R; D-binding mode of bisphenol A with TSHB; E-binding mode of bisphenol A with VGF.

图 6 筛选药物双酚 A 与其靶标的分子对接

Fig. 6 Molecular docking of screening drug bisphenol A and its targets

表 3 候选药物双酚 A 与靶标的结合能

Table 3 Binding energy for targets with bisphenol A

靶标	结合能/(kcal·mol ⁻¹)
ANGPTL3	-6.781
PDF	-7.134
PTH2R	-6.627
TSHB	-7.390
VGF	-77.913

险组样本中免疫细胞的丰度,其中,CD8⁺T 细胞在低风险组的浸润水平较高风险组显著升高 ($P < 0.05$),样本中风险评分与 CD8⁺T 细胞浸润水平呈显著负相关 ($P < 0.05$)。CD8⁺T 细胞也被称为细胞毒性 T 淋巴细胞,是人体免疫系统的重要组成部分,尤其是 CD8⁺T 细胞能够识别和杀伤某些异常的自身细胞,如癌细胞,并且进一步分泌细胞因子激活其他细胞来进一步杀死癌细胞。这也证实了低风险组患者中 CD8⁺T 细胞的浸润可能参与了抗肿瘤过程。

NB 多属中医“癥瘕”范畴,患儿自幼年起发病,外因影响甚微,内因为主要致病因素,因其先天禀赋素弱,脏器虚损,内生邪气,故表现此病。亦如《幼幼集成》中描述:“儿之初生有病,亦惟胎弱、胎毒

两者而已矣”。本研究基于免疫预后模型基因靶点预测得到 130 味核心中药,其功效以清热解毒、扶正补虚为主,这也体现了 NB 的主要病机是以“虚和毒”为主要致病因素。其中白花蛇舌草、半枝莲、鸦胆子、菝葜、喜树果、肿节风、木鳖子、重楼等均为临床常用的清热类抗肿瘤中药。由于 NB 患儿脏腑娇嫩,形气未充,气血本已虚弱,化疗药物加之又损伤脾胃,所以临床应用清热苦寒药物抗肿瘤的同时常常会配合甘温之品白术、山药等顾护脾胃,同时也多用黄芪、太子参、云芝、树舌灵芝等扶正补虚协同抗肿瘤。这样通过整合生物学信息预测作用于靶基因的药物,为高危型 NB 的治疗提供了新的方向。中医作为多学科协作诊疗的一部分,其注重整体和局部的结合,通过扶正与祛邪相结合在肿瘤的治疗上有着自身的特色和优势,现将中医宏观整体观念、辨证论治与现代医学、生物信息学的微观分子相结合,多学科交叉融合,从微观角度阐释中医药科学的原理,也为大数据和人工智能时代中医药的传承、创新和发展提供了新支撑^[27]。

4 结论与展望

本研究在高危型 NB 中,以 ANGPTL3、PDF、

PTGER1、PTH2R、TSHB、VGF 构建了具有较好的预后效能和一定的临床实用性的预后模型。通过整合多个数据库预测作用于预后模型基因的药物,为高危型 NB 的治疗提供了新的方向。后续可基于此理论进一步通过功能实验探讨和验证预后模型中 6 个核心 DIRGs 与 NB 发生、转移、预后关系的分子机制以及中药抗肿瘤的药效机制。同时,为了使这种预后预测模型更广泛地适用于 NB,其临床效能和临床实用性也可通过前瞻性、多中心、大规模的临床试验进行进一步佐证。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Ni X, Li Z, Li X P, *et al.* Socioeconomic inequalities in cancer incidence and access to health services among children and adolescents in China: A cross-sectional study [J]. *Lancet*, 2022, 400(10357): 1020-1032.
- [2] Park J R, Kreissman S G, London W B, *et al.* Effect of tandem autologous stem cell transplant vs single transplant on event-free survival in patients with high-risk neuroblastoma: A randomized clinical trial [J]. *JAMA*, 2019, 322(8): 746-755.
- [3] Hashimoto O, Yoshida M, Koma Y, *et al.* Collaboration of cancer-associated fibroblasts and tumour-associated macrophages for neuroblastoma development [J]. *J Pathol*, 2016, 240(2): 211-223.
- [4] Szanto C L, Cornel A M, Vijver S V, *et al.* Monitoring immune responses in neuroblastoma patients during therapy [J]. *Cancers*, 2020, 12(2): 519.
- [5] Ritchie M E, Phipson B, Wu D, *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies [J]. *Nucleic Acids Res*, 2015, 43(7): e47.
- [6] Wickham H. *ggplot2: Elegant Graphics for Data Analysis* [M]. New York: Springer-Verlag New York, 2009: 147-168.
- [7] Chen H B, Boutros P C. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R [J]. *BMC Bioinformatics*, 2011, 12: 35.
- [8] Gene Ontology Consortium. Gene ontology consortium: Going forward [J]. *Nucleic Acids Res*, 2015, 43: D1049-D1056.
- [9] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes [J]. *Nucleic Acids Res*, 2000, 28(1): 27-30.
- [10] Yu G C, Wang L G, Han Y Y, *et al.* clusterProfiler: An R package for comparing biological themes among gene clusters [J]. *OMICS*, 2012, 16(5): 284-287.
- [11] Therneau T M, Grambsch P M. *Modeling Survival Data: Extending the Cox Model* [M]. New York: Springer New York, 2000: 37-77.
- [12] Chen B B, Khodadoust M S, Liu C L, *et al.* Profiling tumor infiltrating immune cells with CIBERSORT [J]. *Methods Mol Biol*, 2018, 1711: 243-259.
- [13] Xiao B, Liu L Y, Li A Y, *et al.* Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma [J]. *Front Oncol*, 2020, 10: 607622.
- [14] Davis A P, Grondin C J, Johnson R J, *et al.* The comparative toxicogenomics database: Update 2019 [J]. *Nucleic Acids Res*, 2019, 47(D1): D948-D954.
- [15] Csárdi G, Nepusz T. The igraph software package for complex network research [J]. *Inter J Compl Syst*, 2006, 1695(5): 1-6.
- [16] Trott O, Olson A J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading [J]. *J Comput Chem*, 2010, 31(2): 455-61.
- [17] Wang Y L, Bryant S H, Cheng T J, *et al.* PubChem BioAssay: 2017 update [J]. *Nucleic Acids Res*, 2017, 45(D1): D955-D963.
- [18] Gao L, Lin P, Chen P, *et al.* A novel risk signature that combines 10 long noncoding RNAs to predict neuroblastoma prognosis [J]. *J Cell Physiol*, 2020, 235(4): 3823-3834.
- [19] Suo C, Deng W J, Vu T N, *et al.* Accumulation of potential driver genes with genomic alterations predicts survival of high-risk neuroblastoma patients [J]. *Biol Direct*, 2018, 13(1): 14.
- [20] González-Sancho J M, García V, Bonilla F, *et al.* Thyroid hormone receptors/THR genes in human cancer [J]. *Cancer Lett*, 2003, 192(2): 121-132.
- [21] Wu Z H, Xi Z H, Xiao Y X, *et al.* TSH-TSHR axis promotes tumor immune evasion [J]. *J Immunother Cancer*, 2022, 10(1): e004049.
- [22] Wang X W, Lu T, Qu Y J, *et al.* PTH2R is related to cell proliferation and migration in ovarian cancer: A multi-omics analysis of bioinformatics and experiments [J]. *Cancer Cell Int*, 2022, 22(1): 148.
- [23] Escobar-Alvarez S, Gardner J, Sheth A, *et al.* Inhibition of human peptide deformylase disrupts mitochondrial function [J]. *Mol Cell Biol*, 2010, 30(21): 5099-5109.
- [24] Kim J Y, Moon S, Kim D. Oral squamous cell carcinoma-derived ANGPTL3 induces cancer associated fibroblastic phenotypes in surrounding fibroblasts [J]. *Asian Pac J Cancer Prev*, 2022, 23(12): 4315-4322.
- [25] Gong Z, Li Q, Shi J Y, *et al.* Immunosuppressive reprogramming of neutrophils by lung mesenchymal cells promotes breast cancer metastasis [J]. *Sci Immunol*, 2023, 8(80): eadd5204.
- [26] Yang L H, Lee R K, Kuo M H, *et al.* Neuronal survival factor VGF promotes chemoresistance and predicts poor prognosis in lung cancers with neuroendocrine feature [J]. *Int J Cancer*, 2022, 151(9): 1611-1625.
- [27] 杨会军, 郭响, 李伟青, 等. 靶向铜死亡相关基因治疗类风湿关节炎生物信息学分析及干预中药的预测 [J]. *中草药*, 2023, 54(13): 4253-4265.

[责任编辑 潘明佳]