

## • 数据挖掘与循证医学 •

## 基于生物信息学的乳腺癌细胞焦亡相关基因多组学分析及相关中药筛选预测

刘毅, 谢雁鸣\*, 黎元元\*, 崔鑫, 席俊羽

中国中医科学院中医临床基础医学研究所, 北京 100700

**摘要:** **目的** 基于生物信息学探究细胞焦亡相关基因在乳腺癌中的作用及临床意义, 并筛选可对细胞焦亡起调控作用的中药。**方法** 从TCGA、GEO数据库中获取乳腺癌相关数据集; 使用R、Perl语言对细胞焦亡相关基因在乳腺癌细胞中的表达、变异进行评估; 对数据集分别进行细胞焦亡基因分型及预后差异基因分型, 并对各分型进行多组学分析; 构建预后模型、进行风险评分, 按照风险评分分组进行亚组多组学分析并检验对生存的预测能力; 以细胞焦亡基因为靶点, 利用TCMSP数据库及Cytoscape软件构建“细胞焦亡靶点-成分-中药”网络并进行拓扑学分析, 进而得出核心中药及其相关属性。**结果** 大多数细胞焦亡基因在乳腺癌中表达异常并具有拷贝数变异, 细胞肿瘤抗原 (cellular tumor antigen p53, TP53) 等11个基因发生了体细胞突变; 细胞焦亡分型中A亚型预后好、细胞焦亡基因及免疫相关通路高表达、免疫细胞含量高; B型与之相反, 且A、B两亚型差异明显, 其分型差异基因有1223个。将不同的焦亡基因重新聚类, 得到3个基因型, 乳腺癌细胞焦亡预后差异基因分型中I组预后最好, II组最差, I组主要为基因及细胞焦亡基因高表达, II组主要呈现低表达。预后模型风险评分将患者分为高、低风险组, 低风险组细胞焦亡基因高表达、预后好, 高风险组相反, 通过列线图可对患者不同临床特征进行预后评估; 免疫细胞与风险评分相关性大多呈负相关; 肿瘤微环境中基质细胞、免疫细胞及总评分均为高风险组更低, 肿瘤突变负荷则高风险组更高; 干细胞与风险得分呈正相关。“细胞焦亡靶点-成分-中药”网络得出木蝴蝶、红花等17味中药为核心药物, 其性味主要为苦寒, 次以辛温, 辅以甘平之品, 主要调节肝、肺、脾、胃等脏腑。**结论** 细胞焦亡基因在乳腺癌的免疫中发挥重要作用, 与乳腺癌患者预后具有明显相关性, 调控细胞焦亡基因的中药主要有木蝴蝶、红花等17味药物, 可为乳腺癌的诊疗及中药干预研究提供思路与参考。

**关键词:** 细胞焦亡; 生物信息学; 中药; 乳腺癌; 多组学分析; 木蝴蝶; 红花**中图分类号:** R285 **文献标志码:** A **文章编号:** 0253-2670(2022)18-5768-18**DOI:** 10.7501/j.issn.0253-2670.2022.18.020**Multi-omics analysis of pyroptosis-related genes in breast cancer and screening prediction of related traditional Chinese medicines based on bioinformatics**

LIU Yi, XIE Yan-ming, LI Yuan-yuan, CUI Xin, XI Jun-yu

Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China

**Abstract: Objective** To explore the role and clinical significance of pyroptosis related genes in breast cancer and screen traditional Chinese medicines (TCMs) that can regulate pyroptosis based on bioinformatics. **Methods** Breast cancer related data sets were obtained from TCGA and GEO databases; The expression and variation of pyroptosis related genes in breast cancer was evaluate by R and Perl language; Pyroptosis genotyping and prognostic differential genotyping were performed on the data sets, and multi-omics analysis was performed for each genotype; The prognostic model were constructed, risk scores were performed, and sub-group multi-omics analysis was carried out according to risk scores to test the predictive power of survival. Finally, the pyroptosis related genes were used as the targets, and the

**收稿日期:** 2022-02-20**基金项目:** 国家重点研发计划“中医药现代化研究”项目(2018YFC1707400); 中国中医科学院科技创新工程中医临床基础学科创新团队项目(CI2021B003); 2021年岐黄学者支持项目**作者简介:** 刘毅(1993—), 男, 山西长治人, 博士研究生在读, 研究方向为基于中成药真实世界证据与价值评估研究。

E-mail: 709473245@qq.com

**\*通信作者:** 谢雁鸣(1959—), 女, 教授, 研究员, 博士生导师, 中国中医科学院首席研究员, 研究方向为中药上市后再评价方法学研究、中西医结合临床、老年病学。E-mail: ktzu2018@163.com

黎元元, 博士, 研究员, 研究方向为中药上市后再评价。Tel: (010)64014411-3351 E-mail: chibjyy@163.com

TCMSP database and Cytoscape software were used to construct “target-component-Chinese herbal medicine” network and topological analysis to obtain core TCMS and their related attributes. **Results** Most genes related to pyroptosis were abnormally expressed in breast cancer and had copy number variation; Eleven genes such as cellular tumor antigen p53 (*TP53*) had somatic mutations. The A subtype of pyroptosis had good prognosis, high expression of pyroptosis genes and immune-related pathways, and high content of immune cells. Type B was the opposite, and type A and B had obvious differences, the number of differential genes was 1223. The differential genes of pyroptosis were reclustered to generate three genotypes. Among them, group I had the best prognosis, group II had the worst prognosis, and pyroptosis related genes were mainly highly expressed in group I, and lowly expressed in group II. The patients were divided into high and low risk groups according to risk scores of prognostic model. The low risk group showed high expression of pyroptosis gene and good prognosis, while the high risk group showed the opposite. The prognostic evaluation of different clinical characteristics of patients could be carried out through the nomogram; The correlation analysis of immune cells and risk scores showed that most of them were negatively correlated; In the tumor microenvironment, the stromal cell score, immune cell score and total score were lower in all high risk group, and tumor mutation burden was higher in the high risk group. Stem cells were also positively correlated with the risk score. The network of “target of pyroptosis-component-TCMs” obtained 17 TCMS, such as *Muhudie* (*Oroxylis Semen*) and *Honghua* (*Carthami Flos*). Their nature and flavor were mainly bitter and cold, followed by pungent and warm, supplemented by sweet and light medicines, which mainly regulate the liver, lungs, spleen and stomach, etc. **Conclusion** The pyroptosis genes play an important role in the immunity of breast cancer, and have obvious correlation with the prognosis of breast cancer patients. The main TCMS that regulate pyroptosis include 17 medicines such as *Oroxylis Semen* and *Carthami Flos*. The conclusion can provide ideas and references for further research.

**Key words:** pyroptosis; bioinformatics; traditional Chinese medicine; breast cancer; multi-omics analysis; *Oroxylis Semen*; *Carthami Flos*

细胞焦亡 (pyroptosis) 是一种由炎症小体引发的裂解性程序性细胞死亡形式<sup>[1]</sup>, 该过程由促炎性半胱氨酸蛋白酶 (Caspase) -1、4、5、11 的底物 gasdermin D 介导并伴随炎症和免疫反应<sup>[2]</sup>。研究发现细胞焦亡可以影响肿瘤的增殖、侵袭和转移<sup>[3]</sup>。因此, 有报道认为诱导肿瘤细胞焦亡是一种潜在的癌症治疗策略<sup>[4]</sup>。

乳腺癌 (breast cancer) 是目前最常见的恶性肿瘤之一, 是女性癌症死亡的最常见原因<sup>[5]</sup>, 其终生患病概率为 12.3%<sup>[6]</sup>, 对女性群体有极大的危害性。受益于诊断及影像技术发展及早期筛查, 乳腺癌的死亡率已经大幅降低<sup>[7]</sup>, 但由于存在一定复发及转移风险, 乳腺癌患者的预后仍然不佳<sup>[8-9]</sup>。目前乳腺癌的治疗主要包括手术、内分泌治疗、放疗、化疗、靶向治疗等<sup>[10-12]</sup>, 但这些治疗往往存在较严重的不良反应<sup>[13-15]</sup>。因此, 鉴于细胞焦亡对癌症存在调控作用, 以此研究方向为指导或可开发更加安全、有效的新药物。

中药对于乳腺癌治疗的有效性被广泛报道<sup>[16]</sup>, 而挖掘与细胞焦亡相关的中药则更有利于指导相关研究与临床应用。本研究通过生物信息学方法从多组学角度全面分析细胞焦亡相关基因的表达及其与乳腺癌生存、预后、免疫、肿瘤微环境等方面的相关性, 并在此基础上筛选出可能对细胞焦亡有调控作用的中药。

## 1 资料与方法

### 1.1 数据收集及处理

在 TCGA 数据库 (<https://portal.gdc.cancer.gov/>)

中分别获取乳腺癌相关转录组、临床、突变及拷贝数数据。对其数据进行解压、ID 转换并将其格式由 FPKM (fragments per kilobase per million) 转换为 TPM (transcripts per million)。在 GEO 数据库 (<https://www.ncbi.nlm.nih.gov/geo/>) 中检索 “breast cancer survival” 并限定数据类型 (expression profiling by array) 及物种 (homo sapiens) 进行筛选, 获取基因表达及临床数据, 然后对其数据进行基因符号注释。

### 1.2 乳腺癌细胞焦亡基因突变、拷贝数变异、差异、生存及预后分析

通过检索细胞焦亡相关文献获得细胞焦亡相关基因, 将样品突变数据中的细胞焦亡相关基因利用 R 语言 “maftools” 包及 Perl 语言进行突变类型、频率、数目、碱基改变及负荷的统计分析, 并将结果以瀑布图表示; 然后检索 UCSC Xena (<https://xena.ucsc.edu/>) 数据库并下载乳腺癌拷贝数相关数据, 使用 Perl 语言提取样品中细胞焦亡相关基因的拷贝数, 再使用 R 语言将样品细胞焦亡相关基因拷贝数结果绘制成拷贝数频率图及圈图; 将 TCGA 及 GEO 数据库中的样品表达量数据使用 R 语言中的 “limma” 和 “sva” 包进行合并, 并提取细胞焦亡基因的表达量, 再将 2 个数据库所得临床数据中的生存时间、生存状态数据进行合并, 使用 R 语言的 “survival” 和 “surminer” 包进行生存和预后分析并可视化。

### 1.3 乳腺癌细胞焦亡分型及分型基因生存分析、基因集变异分析 (gene set variation analysis, GSVA)、免疫细胞差异分析 (single sample gene set enrichment analysis, ssGSEA) 和主成分分析 (principal component analysis, PCA) 和差异性分析

根据细胞焦亡的表达量使用 R 语言的“ConsensuClusterPlus”对乳腺癌细胞焦亡基因进行聚类分型, 其中聚类算法为 K-means, 距离类型为欧氏距离; 然后使用 R 语言的“RcolorBrewer”“GSVA”“PRGcluster”等将分型后的两组细胞焦亡基因进行生存分析、细胞焦亡基因表达分析、GSVA、ssGSEA、PCA 并可视化, 最后将细胞焦亡分型的两组样品基因表达量进行差异分析, 并获得分型之间的差异基因。

### 1.4 乳腺癌细胞焦亡分型差异基因基因本体 (gene ontology, GO) 富集分析和京都基因与基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 通路富集分析

将“1.3”项所得的细胞焦亡分型之间的差异基因使用 R 语言的“clusterProfiler”和“enrichplot”包分别进行 GO 的生物过程 (biological process, BP)、分子功能 (molecular function, MF) 及细胞成分 (cellular component, CC) 富集分析和 KEGG 通路富集分析, 按照  $P < 0.005$  进行筛选, 并绘制 GO 及 KEGG 分析的气泡图。

### 1.5 乳腺癌细胞焦亡分型预后相关差异基因再分型及其生存与表达分析

根据“1.3”项所得分型差异基因的生存分析结果探寻预后相关的差异基因, 再通过 R 语言聚类分析将预后相关差异基因进行聚类分型, 选取准确性最高的聚类结果作为分型结果。根据分型结果比较不同分型之间的生存情况, 并绘制生存曲线与差异基因热图。然后将“1.3”项所得基因分型的结果与“1.2”项所得细胞焦亡基因表达量数据使用 R 语言进行差异分析并绘制箱线图。

### 1.6 预后模型的构建、分析及其可视化

①将包含单因素显著基因表达数据和临床生存数据的集合进行随机分组, 分为训练 (train) 组及验证 (test) 组; 然后构建基因表达量及生存时间、状态的套索算法 (least absolute shrinkage and selection operator, Lasso) 回归模型并进行交叉验证, 选择其中误差最小的点对应的基因作为 Lasso 回归模型显著基因, 结合显著基因与其表达量构建 COX 回归

模型。②利用训练组得出模型计算公式, 即基因表达量乘以系数再求和, 进一步根据公式计算训练组风险得分, 根据得分中位数区分高、低风险组; 验证组同理, 即通过公式计算风险得分, 并根据训练组的中位数区分验证组的高、低风险组; 进行生存差异分析并得出差异的  $P$  值、受试者工作特征曲线 (receiver operating characteristic curve, ROC)、风险曲线及列线图。③将细胞焦亡分型、基因分型结果通过 R 语言分别进行预后分析, 比较不同分型之间的差异性并可视化; 再分析高、低风险组之间细胞焦亡相关基因的差异性, 并绘制箱线图。

### 1.7 免疫细胞浸润与免疫细胞含量分析

将“1.2”项中 TCGA 和 GEO 数据库所得数据的并集利用 R 语言 CIBERSORT 命令获得总量为 1 的免疫细胞相对含量; 将免疫细胞相对含量的数据结果根据  $P < 0.05$  进行过滤, 并与“1.6”项风险得分的结果取交集, 最后对所有免疫细胞进行循环计算, 得到风险得分、基因与免疫细胞的相关性结果。

### 1.8 肿瘤微环境分析、肿瘤突变负荷分析及干细胞分析

①将“1.2”项中 TCGA 和 GEO 数据库所得数据的并集使用 R 语言“estimate”包对肿瘤微环境进行基质细胞、免疫细胞、肿瘤纯度分析以及综合评分; 利用评分结果与“1.6”项所得高、低风险组进行肿瘤微环境差异分析, 并采用小提琴图表示。②将所有样品的风险数据与染色体突变位置数据利用 Perl 语言分析得出其按照高、低风险组区分的肿瘤突变数据, 并使用 R 语言绘制瀑布图; 利用 R 语言的“reshape2”包结合“1.7”项中样品风险数据及“1.5”项中分型数据获取肿瘤突变负荷数据, 并以箱线图表示。③将泛癌干细胞打分数据结合“1.7”项中样品风险数据, 利用 R 语言分析获得相关性结果并绘制散点图。

### 1.9 统计分析

使用 R v.4.1.2 软件进行统计分析。2 组独立样品数据比较采用  $t$  检验。3 组及以上的数据, 则采用单因素方差分析 (ANOVA) 及 Kruskal-Wallis 秩和检验, 以  $P < 0.05$  表示存在统计学意义。

### 1.10 “细胞焦亡相关靶点-有效成分-中草药”网络构建与药物筛选

将细胞焦亡相关靶基因使用 TCMSp 数据库 (<https://old.tcmsp-e.com/tcmsp.php>) 查找靶点相关的有效成分并根据口服生物利用度 (oral

bioavailability, OB)  $\geq 30\%$ 、类药性 (drug likeness, DL)  $\geq 0.18$  进行筛选, 然后通过有效成分筛选相关中药, 使用 CytoscapeV 3.8.0 构建“靶点-成分-中药”网络并通过拓扑学分析获取核心中药。

## 2 结果

### 2.1 数据收集与处理结果

通过检索 TCGA 及 GEO 数据库, 获得 TCGA-乳腺癌中包含正常样品 113 个, 肿瘤样品 1109 个; 获得的 GEO 数据 (GSE146558) 中包含肿瘤样品 109 个。

### 2.2 乳腺癌细胞焦亡基因突变分析、拷贝数变异分析、差异分析、生存分析及预后分析结果

经过检索文献获得细胞焦亡基因 52 个, 具体如表 1 所示。通过分析乳腺癌中细胞焦亡相关基因的拷贝数变异和基因突变的发生率, 可知在 986 个突变数据样品中有 394 个样品发生了变异, 突变比例占 39.96%, 其中 *TP53* 突变率最高, 达 34%, 其他为 *CASP8* (2%)、*NLRC4* (1%)、*NLRP3* (1%)、*NLRP1* (1%)、*NLRP2* (1%)、*PLCG1* (1%)、*NLRP7* (1%)、*SCAF11* (1%)、*GSDMC* (1%)、*NOD1* (1%), 如图 1-A 所示; 图 1-B 则显示大多数细胞焦亡相关基因存在拷贝数的变化, 而拷贝增加较缺失频率略大; 由图 1-C 可进一步看出拷贝数变化的细胞焦亡相关基因所在染色体的位置; 由图 1-D 可以看出 71% 的细胞焦亡基因在正常样品和肿瘤样品间表达存在显著性差异, 其中 *BAK1*、*BAX*、*CSCP3*、*CHMP2A*、*CHMP2B*、*CHMP4B*、*CHMP4C*、*CHMP6*、*CYCS*、*CSDMD*、*IL18*、*IRF1*、*IRF2*、*AIM2*、*CASP6*、*CASP8*、*CSDMC*、*NLRP6*、*NLRP7*、*NOD2*、*PYCARD*、*GZMA* 为肿瘤组表达上调, *CASP1*、*CASP4*、*CHMP3*、*ELANE*、*GSDMD*、*IL1B*、*TP63*、*GPX4*、*GSDMB*、*IL6*、*NLRP1*、*NLRP3*、*NOD1*、*PJKK*、*SCAF11* 为肿瘤组表达下调。将 TCGA 与 GEO 数据合并后进行生存分析, 结果见表 2, 其中风险率 (hazard ratio, HR)  $> 1$  的基因为高风险基因,  $P < 0.05$  则说明该基因为预后相关, 根据图 2-A 可知, 生存曲线中的预后相关基因有 *CASP1*、*CASP4*、*CASP8*、*CASP9*、*CHMP2A*、*CHMP2B*、*CHMP4B*、*CHMP4C*、*CHMP6*、*CHMP7*、*CYCS*、*GPX4*、*GSDMC*、*GSDMD*、*GZMA*、*GZMB*、*IL1A*、*IL18*、*IRF1*、*IRF2*、*NLRC4*、*NLRP1*、*NLRP3*、*NLRP6*、*NLRP7*、*PYCRD1*、*SCAF11*、*TIRAP*、*TP53*、*TP63*, 其大多为高表达则预后较好; 然后进一步构建预后网络, 如图 2-B 所示, 其中紫色节点

表示高风险基因, 绿色节点表示低风险基因、节点大小表示  $P$  值, 红色连线表示正向共表达, 蓝色连线表示负向共表达, 由连线密度可知其中存在密切的联系。

### 2.3 乳腺癌细胞焦亡分型及分型基因生存分析、GSVA、ssGSEA、PCA 和差异分析结果

通过对细胞焦亡相关基因表达量数据进行聚类分析, 可将样品分为 A 和 B 2 个亚型 (图 3-A); 对 2 个亚型进行生存分析, 如图 3-B 所示, 可知其存在统计学差异 ( $P < 0.05$ ), 且 A 亚型预后较好; 对 2 个亚型基因表达进行分析, 如图 3-C 所示, 可得 A 亚型多数呈现高表达, B 亚型则相反; 通过 GSVA 分析, 则能看出 A 亚型主要对免疫相关通路呈现高表达, B 亚型主要表现为低表达 (图 3-D); 从 ssGSEA 分析结果 (图 3-E) 可知, 22 个在两亚型间存在统计学差异的免疫细胞, 其中活化 B 细胞、活化的 CD4 细胞、活化的 CD8 细胞、活化的树突状细胞、CD56bright 自然杀伤细胞、嗜酸性粒细胞、 $\gamma\delta$ T 细胞、未成熟 B 细胞、未成熟的树突状细胞、骨髓源性抑制细胞、巨噬细胞、肥大细胞、单核细胞、自然杀伤 T 细胞、浆细胞样树突状细胞、调节性 T 细胞、滤泡辅助 T 细胞、1 型 T 辅助细胞、17 型 T 辅助细胞、2 型 T 辅助细胞均在 A 亚型中含量高; 唯有 CD56dim 自然杀伤细胞在 B 亚型中含量高; PCA 分析 (图 3-F) 显示, 根据细胞焦亡相关基因的表达量可将 A 与 B 亚型明显区分; 最后通过设置  $|\log_2FC| > 0.585$ 、 $P < 0.05$ , 筛选出细胞焦亡分型的显著差异基因有 1223 个。

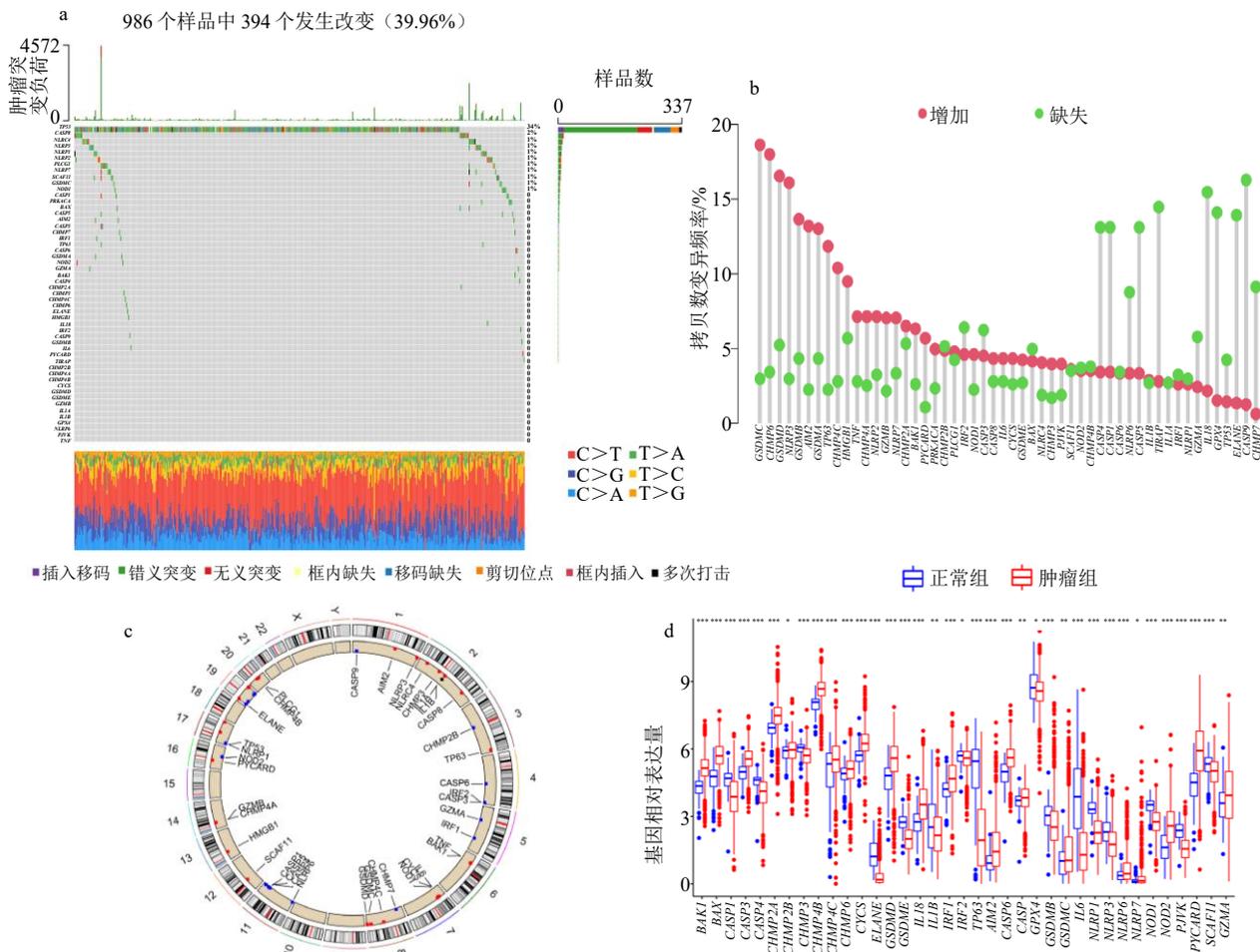
### 2.4 乳腺癌细胞焦亡分型差异基因 GO 富集分析和 KEGG 通路富集分析

通过对 1223 个细胞焦亡分型的显著差异基因进行 GO 分析, 结果表明, BP 主要涉及 T 细胞活化、细胞间黏附的调节、白细胞-细胞黏附、细胞因子产生的正调控、单核细胞分化、细胞活化的正调控、T 细胞黏附的调节等免疫过程; MF 主要涉及受体配体活性、信号受体激活剂活性、细胞因子受体结合、细胞因子活性、免疫受体等; CC 主要涉及质膜筏的外侧、内吞囊泡、膜筏、膜微区等 (图 4-A)。KEGG 通路富集分析显示, 主要包括细胞因子-细胞因子受体相互作用、趋化因子信号通路、细胞黏附分子、爱泼斯坦-巴尔病毒感染、人类 T 细胞白血病病毒 1 感染、病毒蛋白与细胞因子和细胞因子受体的相互作用、造血细胞谱系、结核、吞噬体、破骨

表 1 细胞焦亡相关基因

Table 1 Information of pyroptosis-related genes

序号	基因符号	基因名称
1	<i>BAK1</i>	B 淋巴细胞瘤-2 (B-cell lymphoma-2, Bcl-2) 同源拮抗剂/杀伤剂 (Bcl-2 homologous antagonist/killer)
2	<i>BAX</i>	Bcl-2-相关 X 蛋白 (Bcl-2 associated X protein)
3	<i>CASP1</i>	半胱氨酸天冬氨酸酶-1 (Caspase-1)
4	<i>CASP3</i>	半胱氨酸天冬氨酸酶-3 (Caspase-3)
5	<i>CASP4</i>	半胱氨酸天冬氨酸酶-4 (Caspase-4)
6	<i>CASP5</i>	半胱氨酸天冬氨酸酶-5 (Caspase-5)
7	<i>CHMP2A</i>	带电多泡体蛋白 2a (charged multivesicular body protein 2a)
8	<i>CHMP2B</i>	带电多泡体蛋白 2b (charged multivesicular body protein 2b)
9	<i>CHMP3</i>	带电多泡体蛋白 3 (charged multivesicular body protein 3)
10	<i>CHMP4A</i>	带电多泡体蛋白 4a (charged multivesicular body protein 4a)
11	<i>CHMP4B</i>	带电多泡体蛋白 4b (charged multivesicular body protein 4b)
12	<i>CHMP4C</i>	带电多泡体蛋白 4c (charged multivesicular body protein 4c)
13	<i>CHMP6</i>	带电多泡体蛋白 6 (charged multivesicular body protein 6)
14	<i>CHMP7</i>	带电多泡体蛋白 7 (charged multivesicular body protein 7)
15	<i>CYCS</i>	细胞色素 C (cytochrome C)
16	<i>ELANE</i>	中性白细胞弹性蛋白酶 (neutrophil elastase)
17	<i>GSDMD</i>	焦孔素 D (gasdermin D)
18	<i>GSDME</i>	焦孔素 E (gasdermin E)
19	<i>GZMB</i>	粒酶 B (granzyme B)
20	<i>HMGB1</i>	高迁移率基团蛋白 B1 (high mobility group protein B1)
21	<i>IL18</i>	白细胞介素-18 (interleukin-18)
22	<i>IL1A</i>	白细胞介素-1 $\alpha$ (interleukin-1alpha)
23	<i>IL1B</i>	白细胞介素-1 $\beta$ (interleukin-1beta)
24	<i>IRF1</i>	干扰素调节因子 1 (interferon regulatory factor 1)
25	<i>IRF2</i>	干扰素调节因子 2 (interferon regulatory factor 2)
26	<i>TP53</i>	细胞肿瘤抗原 p53 (cellular tumor antigen p53)
27	<i>TP63</i>	肿瘤蛋白 63 (tumor protein 63)
28	<i>AIM2</i>	干扰素诱导蛋白 AIM2 (interferon-inducible protein AIM2)
29	<i>CASP6</i>	半胱氨酸天冬氨酸酶-6 (Caspase-6)
30	<i>CASP8</i>	半胱氨酸天冬氨酸酶-8 (Caspase-8)
31	<i>CASP9</i>	半胱氨酸天冬氨酸酶-9 (Caspase-9)
32	<i>GPX4</i>	磷脂过氧化氢谷胱甘肽过氧化物酶 (phospholipid hydroperoxide glutathione peroxidase)
33	<i>GSDMA</i>	焦孔素 A (gasdermin A)
34	<i>GSDMB</i>	焦孔素 B (gasdermin B)
35	<i>GSDMC</i>	焦孔素 C (gasdermin C)
36	<i>IL6</i>	白细胞介素 6 (interleukin-6)
37	<i>NLRC4</i>	含 NLR 家族 CARD 域蛋白 4 (NLR family CARD domain-containing protein 4)
38	<i>NLRP1</i>	核苷酸结合寡聚化结构域样受体蛋白 1 (NACHT, LRR and PYD domains-containing protein 1)
39	<i>NLRP2</i>	核苷酸结合寡聚化结构域样受体蛋白 2 (NACHT, LRR and PYD domains-containing protein 2)
40	<i>NLRP3</i>	核苷酸结合寡聚化结构域样受体蛋白 3 (NACHT, LRR and PYD domains-containing protein 3)
41	<i>NLRP6</i>	核苷酸结合寡聚化结构域样受体蛋白 6 (NACHT, LRR and PYD domains-containing protein 6)
42	<i>NLRP7</i>	核苷酸结合寡聚化结构域样受体蛋白 7 (NACHT, LRR and PYD domains-containing protein 7)
43	<i>NOD1</i>	核苷酸结合寡聚化结构域蛋白 1 (nucleotide-binding oligomerization domain-containing protein 1)
44	<i>NOD2</i>	核苷酸结合寡聚化结构域蛋白 2 (nucleotide-binding oligomerization domain-containing protein 2)
45	<i>PJVK</i>	pejvakin
46	<i>PLCG1</i>	1-磷脂酰肌醇 4,5-二磷酸磷酸二酯酶 $\gamma$ -1 (1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1)
47	<i>PRKACA</i>	cAMP 依赖性蛋白激酶催化亚基 $\alpha$ (cAMP-dependent protein kinase catalytic subunit alpha)
48	<i>PYCARD</i>	含有 CARD 的凋亡相关斑点样蛋白 (apoptosis-associated speck-like protein containing a CARD)
49	<i>SCAF11</i>	富含丝氨酸/精氨酸的剪接因子 2 相互作用蛋白 11 (serine/arginine-rich splicing factor 2-interacting protein 11)
50	<i>TIRAP</i>	受体结构域的衔接蛋白 (Toll/interleukin-1 receptor domain-containing adapter protein Toll/interleukin-1)
51	<i>TNF</i>	肿瘤坏死因子 (tumor necrosis factor)
52	<i>GZMA</i>	颗粒酶 A (granzyme A)



a-细胞焦亡基因突变频率瀑布图 (图中 A 为腺嘌呤, T 为胸腺嘧啶, C 为胞嘧啶, G 为鸟嘌呤) b-拷贝数变异频率统计图 c-细胞焦亡基因拷贝数圈图 d-细胞焦亡基因差异分析箱线图 \* $P < 0.05$  \*\* $P < 0.01$  \*\*\* $P < 0.001$ , 下同

a-waterfall chart of mutation frequency of pyroptosis genes (A is adenine, T is thymine, C is cytosine, G is guanine) b-statistics chart of copy number variation frequency c-circle chart of pyroptosis genes' copy numbers d-box plot of differential analysis of pyroptosis genes \* $P < 0.05$  \*\* $P < 0.01$  \*\*\* $P < 0.001$ , same as below

图 1 乳腺癌中细胞焦亡基因突变分析、拷贝数变异分析

Fig. 1 Diagram of mutation analysis and copy number variation analysis of pyroptosis genes in breast cancer

细胞分化、类风湿关节炎、核因子- $\kappa$ B (nuclear factor- $\kappa$ B, NF- $\kappa$ B) 信号通路、Th17 细胞分化、TNF 信号通路等炎症、免疫相关通路 (图 4-B)。

### 2.5 乳腺癌细胞焦亡分型预后相关差异基因再分型及其生存分析与表达分析结果

通过对 1223 个细胞焦亡分型预后显著差异基因进行再分型, 可得到 3 组分型基因, 分别为 I、II、III, 如图 5-A 所示; 对这 3 组分型基因进行生存分析, 3 组分型差异存在显著性差异 ( $P < 0.05$ ), 且其中 I 组预后最好, II 组预后最差 (图 5-B); 对 3 组分型基因进行表达分析, 可见 I 组主要表现为高表达, II 组主要呈现低表达 (图 5-C), 结合生存分析及基因表达分析可知, 相关基因呈现高表达则预

后较好, 反之则预后较差, 与预期一致; 接下来对 3 组细胞焦亡基因分型进行差异分析, 可知 3 组之间超过 80% 细胞焦亡基因存在显著性差异, 其中超过 60% 细胞焦亡基因为 I 组高表达, 超过 70% 细胞焦亡基因为 II 组低表达 (图 5-D)。

### 2.6 预后模型构建、预后分析及其可视化结果

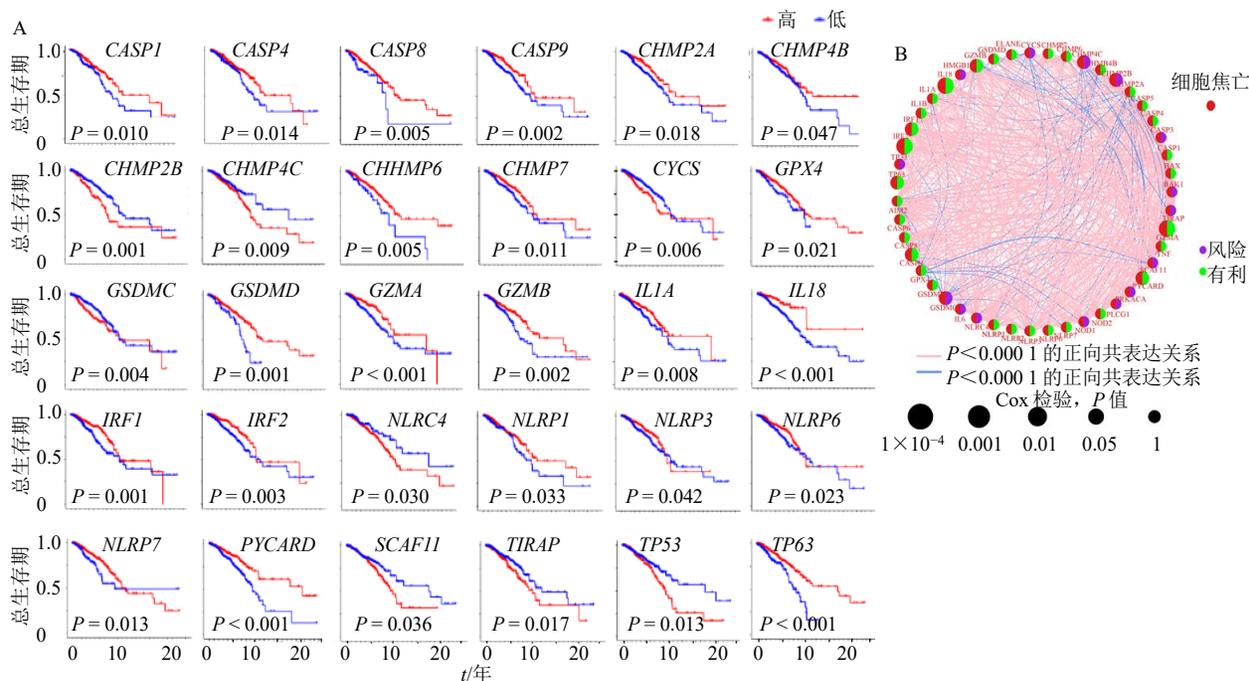
#### 2.6.1 预后模型 Lasso 回归结果及交叉验证

结果如图 6-A、B 所示, COX 回归模型得出 2 组及总体的风险值; 预后模型构建过程如图 6-C 所示, 即首先根据细胞焦亡相关基因的表达量将样品分为 2 个亚型 (A、B 亚型), 然后通过差异分析找出该 2 亚型中存在的差异基因, 通过分析差异基因的表达量对基因进行分型 (分型 I、分型 II 和分型 III);

表 2 细胞焦亡基因生存分析

Table 2 Survival analysis of pyroptosis genes

基因	HR	P 值	基因	HR	P 值
BAK1	1.086 233 416	0.087 484 657	AIM2	0.931 151 062	0.066 065 479
BAX	0.970 691 354	0.105 234 434	CASP6	0.910 201 293	0.084 018 105
CASP1	0.896 254 879	0.010 444 668	CASP8	0.818 232 047	0.005 053 259
CASP3	1.067 483 840	0.158 124 658	CASP9	0.697 598 354	0.002 007 942
CASP4	0.826 676 995	0.014 357 829	GPX4	0.852 816 890	0.020 811 101
CASP5	0.928 306 572	0.109 390 972	GSDMB	0.967 202 344	0.064 985 900
CHMP2A	0.848 139 875	0.018 369 884	GSDMC	1.113 768 964	0.004 412 956
CHMP2B	1.345 011 503	0.001 323 049	IL6	1.004 507 198	0.280 644 973
CHMP4B	0.874 446 548	0.046 762 979	NLRC4	1.172 839 955	0.030 111 467
CHMP4C	1.190 235 273	0.008 719 777	NLRP1	0.942 709 208	0.032 939 873
CHMP6	0.813 665 039	0.005 475 246	NLRP2	0.956 027 022	0.068 839 075
CHMP7	0.943 839 458	0.010 553 598	NLRP3	0.944 446 008	0.042 430 619
CYCS	1.262 775 294	0.005 881 638	NLRP6	0.893 862 704	0.022 686 052
ELANE	0.877 124 510	0.084 815 313	NLRP7	0.928 231 695	0.012 910 692
GSDMD	0.917 067 575	0.001 474 721	NOD1	1.013 837 236	0.087 182 922
GZMB	0.880 111 255	0.001 741 422	NOD2	0.986 949 281	0.054 333 534
HMGB1	1.001 715 022	0.244 834 971	PLCG1	1.035 148 318	0.109 209 786
IL18	0.802 492 407	0.000 842 706	PRKACA	1.082 406 990	0.115 309 162
IL1A	0.943 171 588	0.008 360 707	PYCARD	0.872 856 354	0.000 215 098
IL1B	0.976 908 522	0.050 530 822	SCAF11	1.111 688 261	0.036 312 269
IRF1	0.813 292 016	0.001 267 743	TIRAP	1.223 149 413	0.016 585 961
IRF2	0.636 756 119	0.002 620 915	TNF	0.972 806 341	0.129 396 207
TP53	1.125 176 566	0.012 799 973	GZMA	0.867 120 107	0.000 788 690
TP63	0.863 959 655	0.000 120 009			

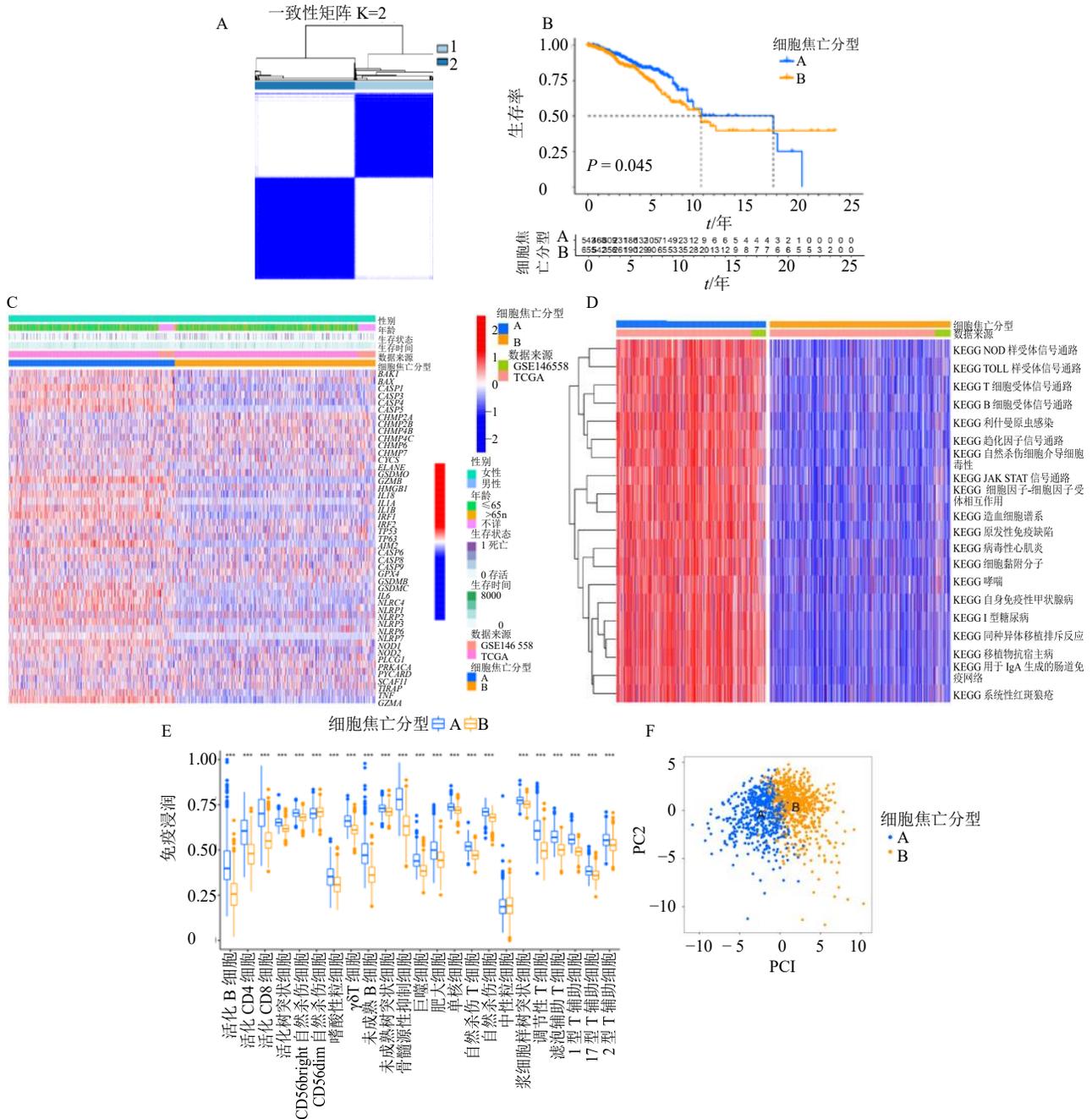


A-预后存在统计学差异的细胞焦亡基因生存曲线 B-细胞焦亡基因预后网络

A-survival curve of pyroptosis genes with statistical difference in prognosis B-prognosis network diagram of pyroptosis-related genes

图 2 乳腺癌细胞焦亡基因生存及预后分析

Fig. 2 Results of survival and prognosis analysis of pyroptosis-related genes in breast cancer



A-BRCA 细胞焦亡聚类分型结果 (1-聚类 1, 2-聚类 2) B-细胞焦亡分型生存曲线 C-细胞焦亡分型热图 D-细胞焦亡分型 GSEA 分析热图 E-细胞焦亡分型免疫细胞差异分析箱线图 F-细胞焦亡分型 PCA 散点图

A-pyroptosis cluster typing results (1-cluster 1, 2-cluster 2) B-survival curve of clusters of pyroptosis genes in BRCA C-heat map of clusters of pyroptosis genes in BRCA D-GSEA heat map of clusters of pyroptosis genes E-box plot of ssGSEA of clusters of pyroptosis genes F-scatter plot of PCA of clusters of pyroptosis genes

图 3 BRCA 细胞焦亡分型及相关分析结果

Fig. 3 Results of pyroptosis typing in breast cancer and related analysis results

对差异基因进行生存分析以找出预后相关的差异基因并构建预后相关模型；再根据预后模型将患者分为高、低风险组，最终根据预后模型预测患者的生存期。

2.6.2 细胞焦亡分型风险得分的差异分析 结果表

明 A、B 两亚型间差异有统计学意义，其中分型 B 的风险评分更高，如图 6-D 所示；基因分型风险得分的差异分析结果表明分型 I、II 及分型 II、III 之间差异有统计学意义，分型 I、III 之间则无统计学意义，同时可以看出分型 II 的风险评分最高，如图 6-E 所示。

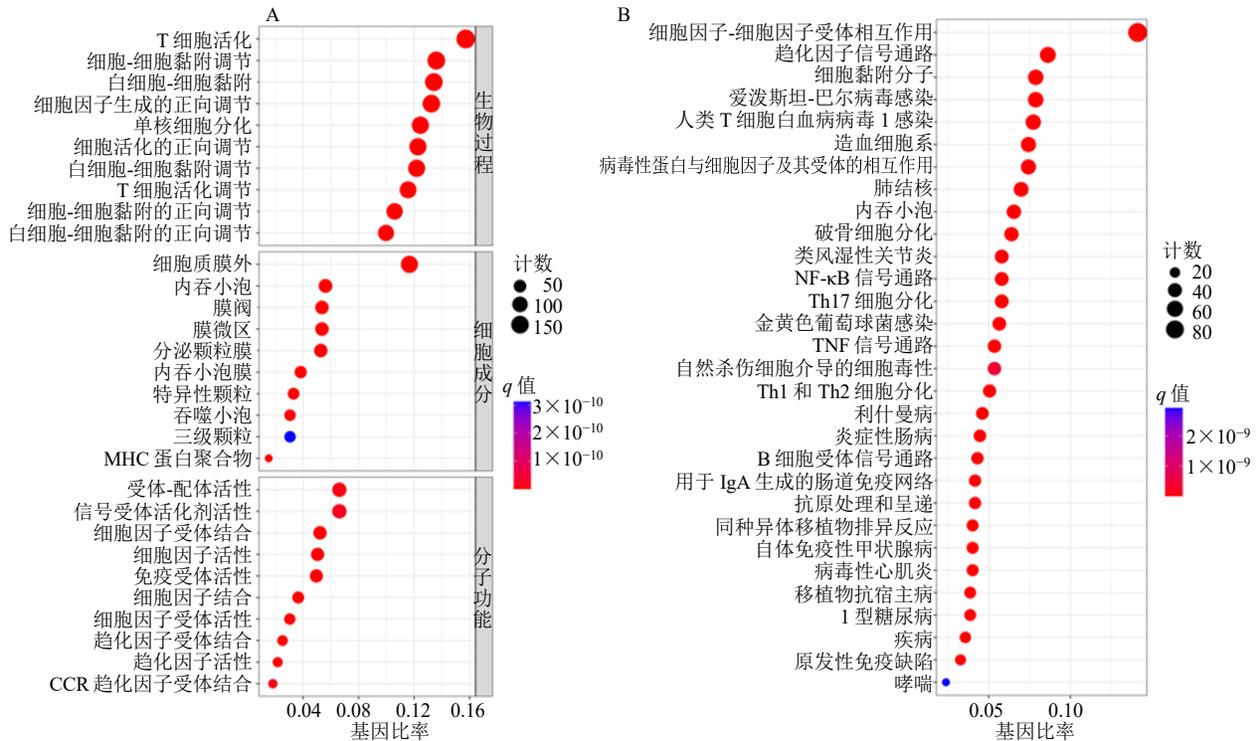


图 4 细胞焦亡分型差异基因 GO 及 KEGG 富集分析气泡图

Fig. 4 Bubble charts of GO and KEGG enrichment analysis of pyroptosis typing differential genes

**2.6.3** 高、低风险组细胞焦亡基因表达量的差异分析结果如图 6-F 所示，高、低风险组中细胞焦亡基因绝大多数 (41/52) 存在显著差异，且这些基因中除 *GSDMC*、*CYCS*、*BAK1* 3 个基因在高风险组为高表达外，其余均表现为低表达；此外由生存曲线 (图 6-G) 可知，不论训练组、验证组还是总体组 (训练组及验证组的总和) 其高、低风险组的生存时间都存在显著差异，说明构建的预后模型可区分高、低风险组的患者；由 ROC 曲线 (图 6-H) 可知，训练组线下面积最大而验证组线下面积最小，因此通过预后模型预测生存期训练组准确性最大而验证组最小，而不论训练组、验证组还是总体组线下面积都大于 0.65，因此预后模型预测生存期准确性均较高；列线图可提供预后模型对患者不同临床特征生存期的预测，由图 6-I 可知女性、年龄 62 岁、高风险评分的患者预后差，其 1 年、3 年、5 年生存率分别为 96.7%、83.4%、70.9%，图 6-J 则显示了列线图生存率预测的准确性；最后根据患者高、低风险排序绘制风险曲线 (图 6-K)，不论训练组、验证组还是总体组 3 者都是随风险值增高而死亡人数增多，与预测相符，而风险热图则提示所列基因皆为低风险基因。

**2.7 免疫细胞浸润与免疫细胞含量分析结果**

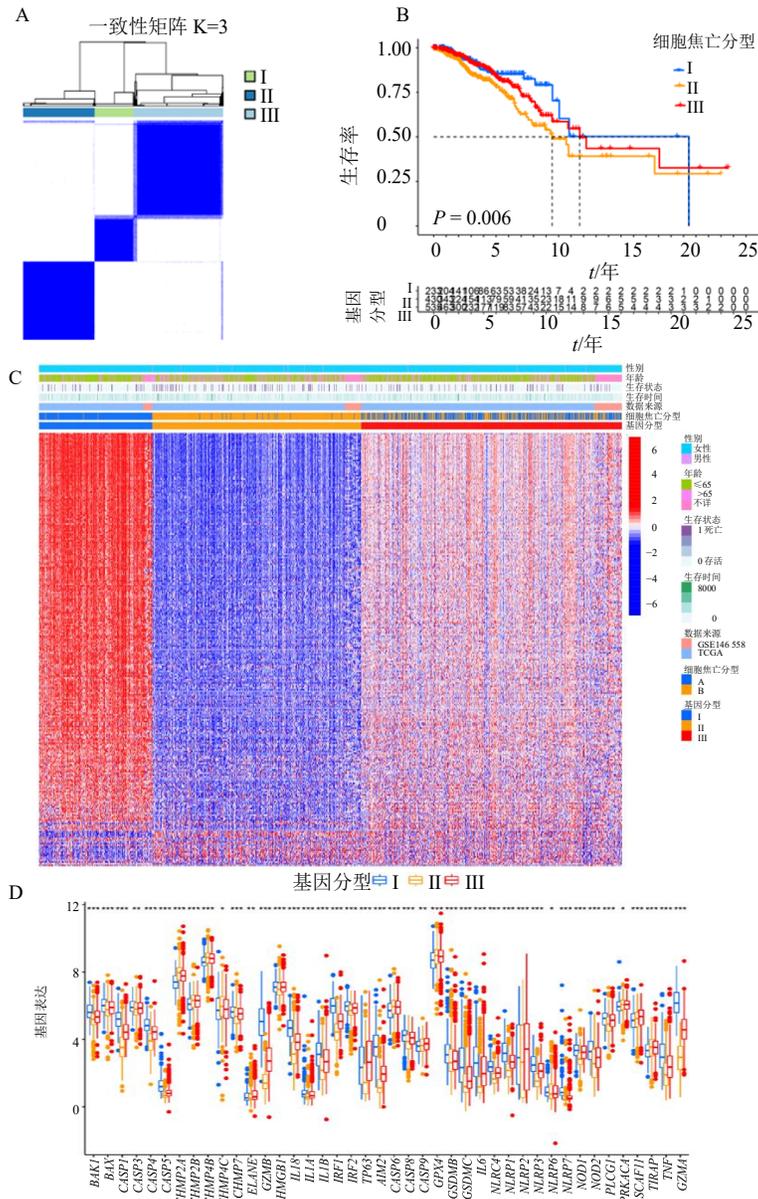
通过 CIBERSORT 计算求得每种免疫细胞在每

个样品中的相对含量作为免疫细胞浸润结果；利用该结果进一步得出风险得分与免疫细胞相关性，其中存在显著负相关性的有幼稚 B 细胞、静息树突状细胞、巨噬细胞 M1、单核细胞、 $\gamma\delta$ T 细胞、CD8 T 细胞、静息肥大细胞、静息 CD4 记忆 T 细胞、浆细胞；存在显著正相关性的有活化树突状细胞、M0 巨噬细胞、M2 巨噬细胞、活肥大细胞、静息自然杀伤细胞，见图 7-A。参与模型构建的基因与免疫细胞相关性如图 7-B 所示，其中大部分基因与免疫细胞存在相关性，且呈负相关性较正相关性者多。

**2.8 肿瘤微环境分析、肿瘤突变负荷分析及干细胞分析结果**

**2.8.1 肿瘤微环境分析** 首先获得肿瘤微环境 TCGA、GEO 两数据库临床样品的评分结果，包含基质细胞评分、免疫细胞评分、总评分以及肿瘤纯度 4 个方面；然后将肿瘤微环境评分结果按照高、低风险分组进行差异分析，由图 8-A 可知基质细胞评分、免疫细胞评分及总评分在高、低风险组间均具有统计学差异，且肿瘤微环境评分均为高风险组更低；使用瀑布图展示高、低风险组的突变数据，如图 8-B、C 所示，2 组的突变率较接近 (84.12%、85.89%)，而高风险组突变率略高于低风险组。

**2.8.2 肿瘤突变负荷相关性分析** 结果显示高、低



A-细胞焦亡分型预后差异基因再聚类分型结果 B-预后差异基因分型生存曲线 C-预后差异基因分型热图 D-预后差异基因分型的细胞焦亡基因差异分析箱线图

A-results of re-clustering prognostic differential genes of pyroptosis clusters B-survival curves of prognostic differential genes' clusters C-heat map of clusters of prognostic differential genes D-box plot of pyroptosis genes' differential analysis from prognostic differential genes

图5 BRCA 细胞焦亡分型相关预后差异基因分型及其生存分析与表达分析结果

Fig. 5 Results of clusters of prognostic differential genes from pyroptosis genes' clusters in BRCA, survival analysis and expression analysis

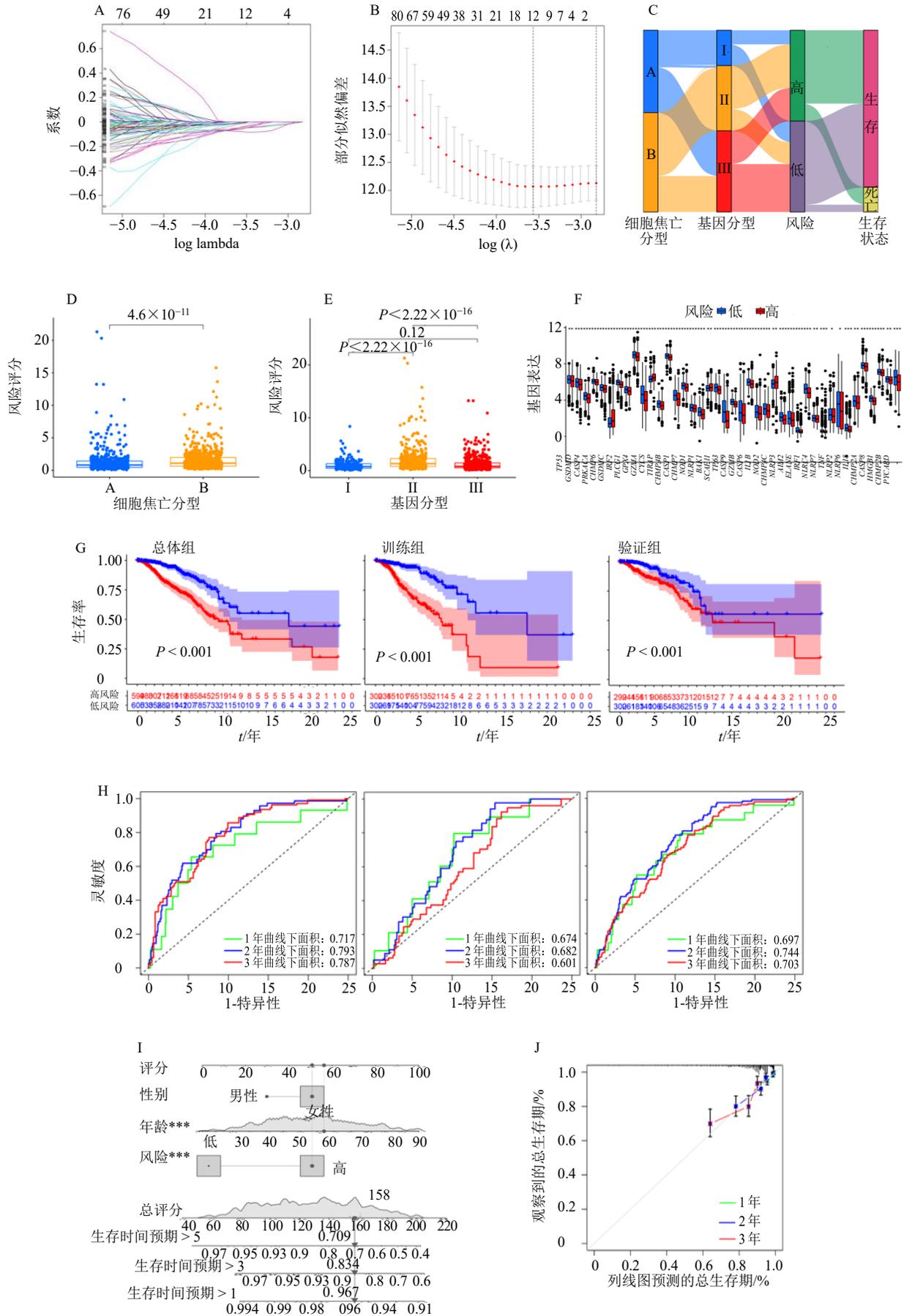
风险组间肿瘤突变负荷存在统计学差异，其中高风险组患者的肿瘤突变负荷较高（图 8-D）；对 3 组基因分型样品的风险评分及肿瘤突变负荷进行相关性分析，由图 8-E 可知二者相关性有统计学意义，且呈正相关。

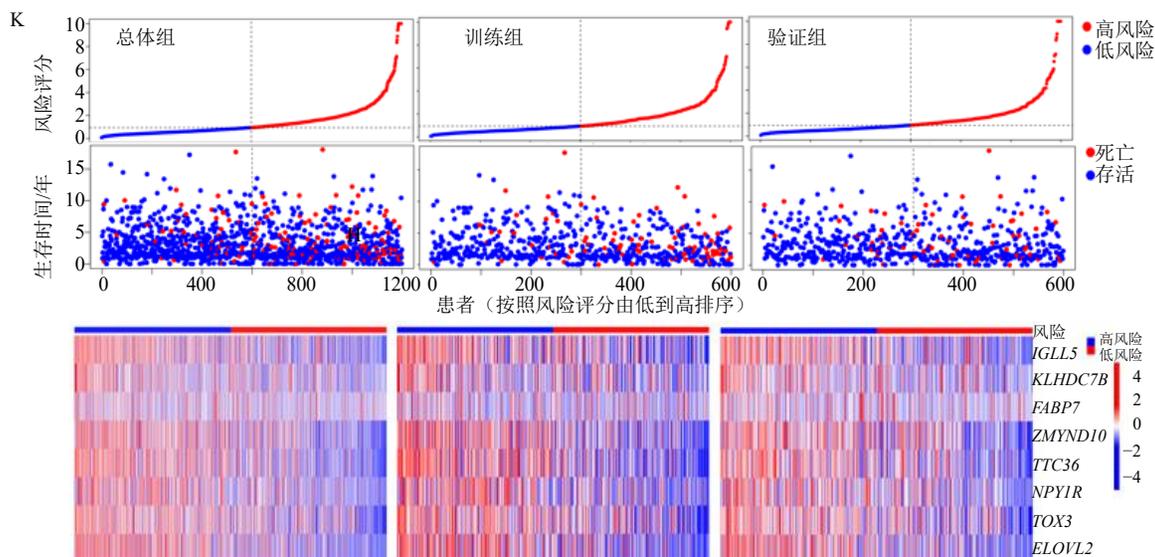
2.8.3 干细胞相关性分析 结果（图 8-F）显示干细胞指数与风险评分之间存在相关性（ $P < 0.05$ ），且

二者呈正相关。

### 2.9 “细胞焦亡靶点-成分-中药”网络构建与中药筛选结果

通过检索 TCMSP 数据库，找到 18 个细胞焦亡靶点数据，将与该 18 个细胞焦亡靶点相关联的 34 种有效成分（表 3）及 380 味中药进行网络构建（图 9-A）并进行分析，分别通过节点度（degree）值筛选





A-Lasso 回归模型图 B-交叉验证图 C-预后模型构建过程桑基图 D-细胞焦亡分型风险得分差异分析图 E-预后差异基因分型风险得分差异分析图 F-高、低风险组细胞焦亡基因表达差异分析箱线图 G-训练组、验证组、总体组的高低风险生存分析曲线 H-训练组、验证组、总体组 ROC 曲线 I-生存期预测列线图 J-列线图校准曲线 K-训练组、验证组、总体组风险曲线  
 A-diagram of model of lasso regression B-cross-validation results C-alluvial diagram of prognostic model construction process D-diagram of differential analysis of risk score in different clusters of pyroptosis genes E-diagram of differential analysis of risk score in different clusters of prognostic differential genes F-box plot of differential analysis of pyroptosis genes' expression in group of high or low risk G-survival curves of high or low risk in train, test and total groups H-ROC curves in train, test and total groups I-nomogram of life time prediction J-nomogram calibration curve K-risk curves of train, test and total groups

图 6 预后模型构建、预后分析及其可视化结果

Fig. 6 Results of construction of prognostic model, prognostic analysis and their visualization

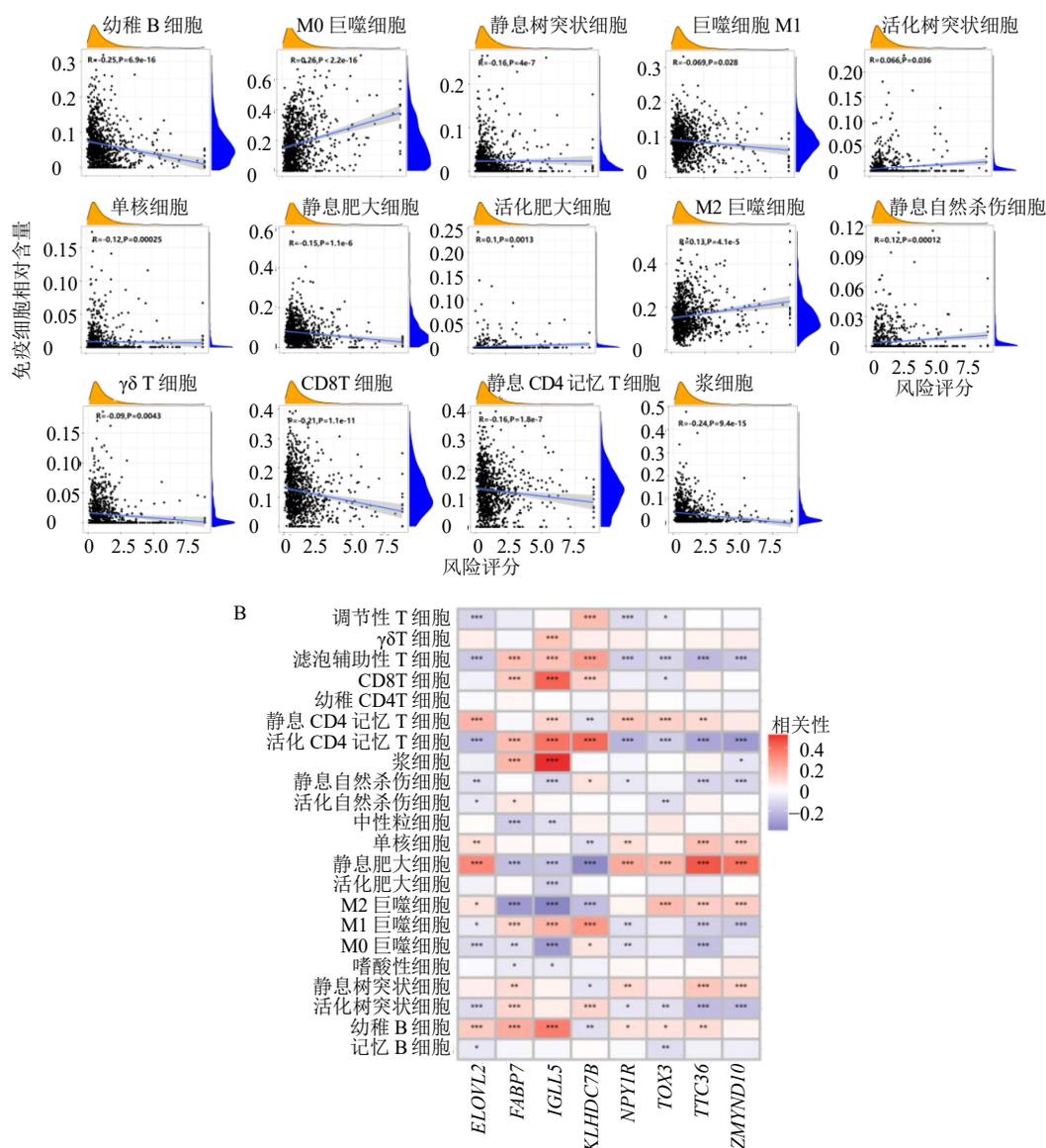
核心成分和核心中药，核心成分（度值>2.5 倍中位数）主要包括槲皮素（quercetin）、β-谷甾醇（β-sitosterol）、木犀草素（luteolin）、山柰酚（kaempferol）、β-胡萝卜素（β-carotene）、金合欢素（acacetin）、黄芩素（baicalein）、汉黄芩素（wogonin）、芦荟大黄素（aloe-emodin）、表没食子儿茶精没食子酸酯 [(-)-epigallocatechin-3-gallate]、川陈皮素（nobiletin）、人参皂苷 Rh<sub>2</sub>（ginsenoside Rh<sub>2</sub>）、柚皮素（naringenin）；核心中药（度值>2.5 倍中位数）主要包括木蝴蝶、红花、余甘子、白果、半枝莲、牛膝、菊花、香薷、马齿苋、连翘、连钱草、败酱草、枇杷叶、关黄柏、皂角刺、马鞭草、金银花 17 味。对 380 味中药进行性味归经统计分析，其中四气以温、寒、微寒、平为主；五味以苦、辛、甘为主；归经主要归肝、肺经，见图 9-B、C、D。

### 3 讨论

细胞焦亡作为一种受调节的促炎形式的细胞死亡，其在形态学、机制学和病理生理学上不同于包括细胞凋亡和坏死在内的其他形式的细胞死亡，其特点是质膜快速破裂，随后释放细胞内容物和促炎

介质，包括 IL-1β、IL-18 和警报素高迁移率族蛋白 1（high mobility group protein B1, HMGB-1）等<sup>[17]</sup>。细胞焦亡已被确认与心血管疾病、传染病、卵巢癌、肺癌等疾病有关，但是其与乳腺癌的关系却仍缺少相关研究证实<sup>[18-20]</sup>，且相关的生物信息学研究也缺乏涉及中药方面的分析，因此本研究对 3 者的关系进行了一定程度探究。

本研究首先根据乳腺癌细胞焦亡基因突变分析可知 TP53 突变频率远高于其他基因，而研究证实 TP53 的突变可能是除乳腺癌 1 型易感蛋白（breast cancer type 1 susceptibility protein, BRCA1）、乳腺癌 2 型易感蛋白（Breast cancer type 2 susceptibility protein, BRCA2）外的另一个导致乳腺癌的关键因素，其与三阴性乳腺癌有关<sup>[21]</sup>；肿瘤组和正常组的对比则可发现细胞焦亡相关基因的表达大多数在二者中存在显著差异，而肿瘤组表达上调者略多于表达下调者，研究认为<sup>[4]</sup>细胞焦亡与癌症的关系复杂，不同组织和遗传背景的细胞焦亡对癌症的影响不同而呈现促癌和抑癌 2 种作用。乳腺癌细胞焦亡基因生存分析显示，生存曲线差异显著且高风险的基



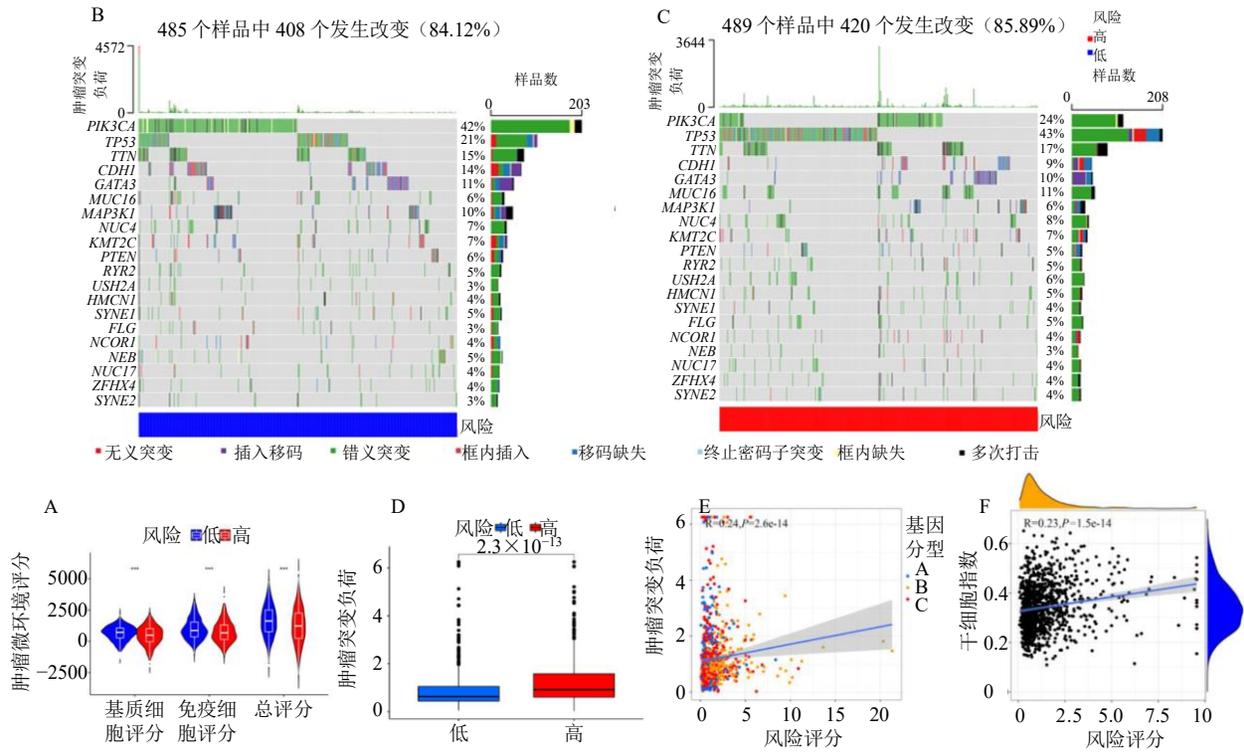
A-风险得分与免疫细胞相关性图 B-模型构建基因与免疫细胞相关性热图  
 A-diagram of correlation of risk score and immune cells B-heat map of correlation of genes of model construction and immune cells

图 7 免疫细胞浸润与免疫细胞含量分析结果

Fig. 7 Results of immune cell infiltration and immune cells' content analysis

因中 *GSDMC* 高表达与乳腺癌患者较差的存活率相关<sup>[22]</sup>，并可介导癌细胞中的非典型细胞焦亡途径，导致肿瘤坏死；*CASP3* 在肿瘤组中低表达，有动物实验证明其表达上调可促进乳腺癌细胞凋亡<sup>[23]</sup>；*PLCG1* 作为磷脂酶 C 的主要亚型，是细胞信号传导的重要介质，其在各种癌症中经常富集和突变，并参与肿瘤发生的过程，包括增殖、迁移和侵袭<sup>[24]</sup>；*CYCS* 有文献报道称其可作为乳腺癌细胞凋亡的生物标志物在化疗后表达增加<sup>[25]</sup>；分析研究认为 *BAK1* 与乳腺癌的骨转移有关<sup>[26]</sup>；*PICAKA* 所参与的通路则有促成雌激素诱导的乳腺癌细胞增殖和内分泌抵抗的作用<sup>[27]</sup>；*IL-6* 在包括乳腺癌在内的多种肿瘤组织中过度表达，且在乳腺癌生长、转移及治疗抵抗方面起作用<sup>[28]</sup>。总体来说细胞焦亡基因在乳腺癌中突变频率不高，但其突变大部分与预后相关。

通过将乳腺癌细胞焦亡基因进行聚类及分型，发现 A 亚型患者生存结局优于 B 亚型，且 A 亚型中细胞焦亡基因多数呈现高表达；GSVA、ssGSEA 分析显示，炎症相关通路、免疫细胞均在 A 亚型中富集，A 亚型特征为细胞焦亡基因高表达、免疫激活和免疫细胞浸润、有生存优势，B 亚型与之相反。对细胞焦亡基因分型进行预后差异分析，得出差异



A-高、低风险组肿瘤微环境差异分析小提琴图 B-低风险组基因突变频率瀑布图 C-高风险组基因突变频率瀑布图 D-高、低风险组肿瘤突变负荷差异分析箱线图 E-风险评分与肿瘤突变负荷相关性散点图 F-风险评分与干细胞相关性分析散点图

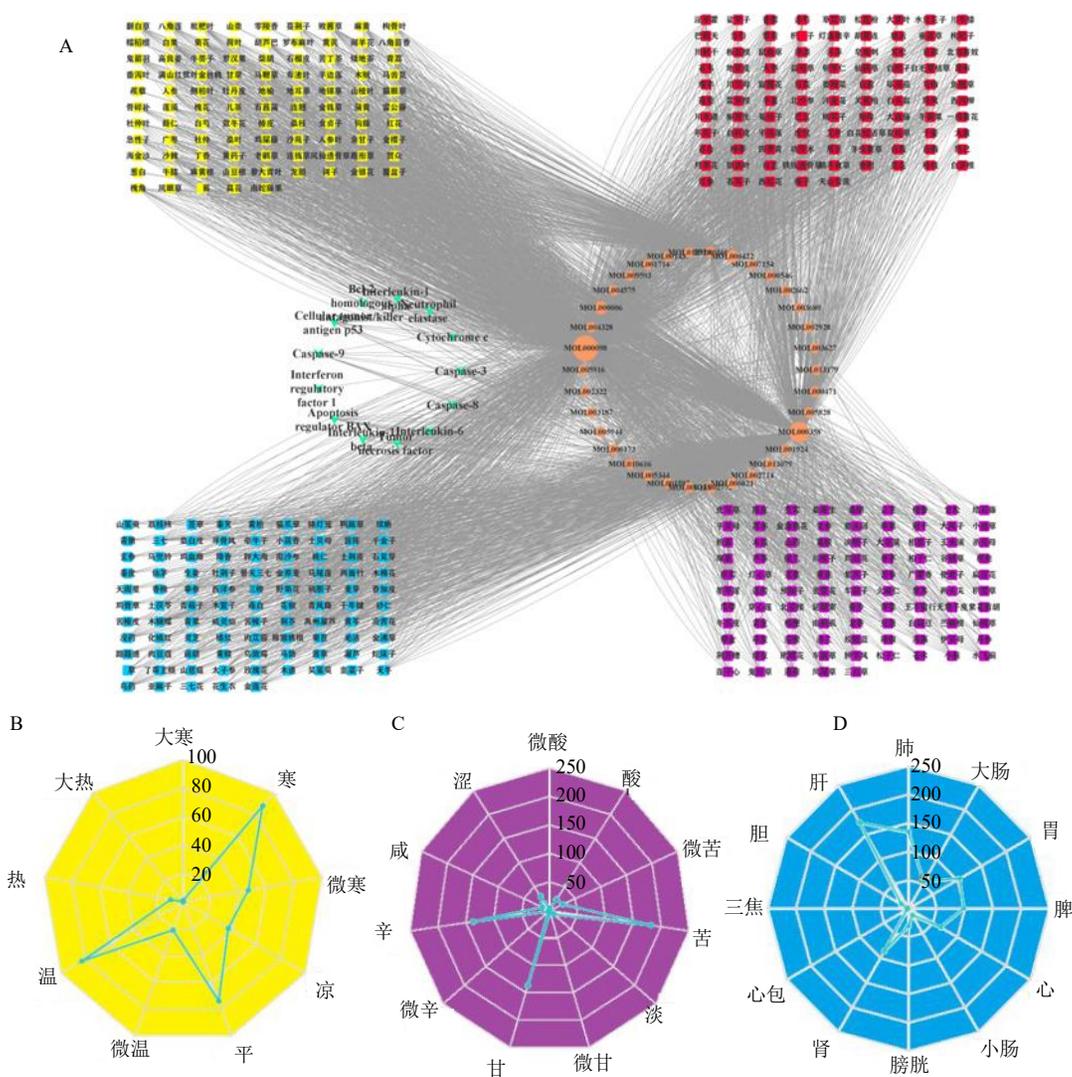
A-violin plot of tumor microenvironment differential analysis in group of high and low risk B-waterfall chart of gene mutation frequency in low-risk group C-waterfall chart of gene mutation frequency in high-risk group D-box plot of tumor mutation burden differential analysis in groups of high and low risk E-scatter plot of correlation between risk score and tumor mutation burden F-scatter plot of risk score and stem cell correlation analysis

图 8 肿瘤微环境分析、肿瘤突变负荷分析及干细胞分析结果

Fig. 8 Results of tumor microenvironment analysis, tumor mutation burden analysis, and stem cell analysis

表 3 18 个细胞焦亡靶点相关成分  
Table 3 Related ingredients of 18 pyroptosis genes

MOLID	化学成分	OB/%	DL	MOLID	化学成分	OB/%	DL
MOL006821	(-)-epigallocatechin-3-gallate	55.09	0.77	MOL003187	triptolide	51.29	0.68
MOL009135	ellipticine	30.82	0.28	MOL005944	matrine	63.77	0.25
MOL000422	kaempferol	41.88	0.24	MOL007154	tanshinone iia	49.89	0.40
MOL010616	eckol	87.06	0.63	MOL002773	β-carotene	37.18	0.58
MOL000358	β-sitosterol	36.91	0.75	MOL002928	oroxylin A	41.37	0.23
MOL005828	nobiletin	61.67	0.52	MOL001592	piperine	42.52	0.23
MOL013179	fisetin	52.60	0.24	MOL001714	podophyllotoxin	59.94	0.86
MOL000471	aloe-emodin	83.38	0.24	MOL005916	irisolidone	37.78	0.30
MOL001689	acacetin	34.97	0.24	MOL000546	diosgenin	80.88	0.81
MOL002714	baicalein	33.52	0.21	MOL012920	sinomenine	30.98	0.46
MOL009593	verticinone	60.07	0.67	MOL001924	paeoniflorin	53.87	0.79
MOL000173	wogonin	30.68	0.23	MOL003627	sophocarpine	64.26	0.25
MOL005344	ginsenoside Rh <sub>2</sub>	36.32	0.56	MOL003680	sophoridine	60.07	0.25
MOL000098	quercetin	46.43	0.28	MOL013079	dl-praeruptorin A	46.46	0.53
MOL004328	naringenin	59.29	0.21	MOL004575	astilbin	36.46	0.74
MOL000006	luteolin	36.16	0.25	MOL002662	rutaecarpine	40.30	0.60
MOL001439	arachidonic acid	45.57	0.20	MOL002322	isovitexin	31.29	0.72



A-“细胞焦亡靶点-成分-中药”网络图 B-细胞焦亡相关中药四气分布 C-细胞焦亡相关中药五味分布 D-细胞焦亡相关中药归经分布  
 A-“pyroptosis gene-related ingredient-traditional Chinese medicine” network B-distribution of four-qi in traditional Chinese medicines related to pyroptosis genes C-distribution of five-flavour in traditional Chinese medicines related to pyroptosis genes D-distribution of channels in traditional Chinese medicines related to pyroptosis genes

图9 “细胞焦亡靶点-成分-中药”网络构建与中药筛选结果

Fig. 9 “Pyroptosis gene-related ingredient-traditional Chinese medicine” network construction and results of screening of traditional Chinese medicine

基因 1223 个，并将之再次聚类获得 3 组基因分型，同样得到 I 组细胞焦亡基因表达最高、生存预期最好，II 组细胞焦亡基因表达最低、生存预期最差，2 次分型结果一致。通过构建预后模型并进行风险评估，得出高风险组预后差，低风险组细胞焦亡基因高表达，且有较高的准确性。不论免疫浸润还是肿瘤微环境分析都得出同一个结论，即高风险组预后较差、细胞焦亡基因低表达、生存期较短、免疫细胞浸润少；低风险组则与之相反，这也反映了细胞焦亡与乳腺癌及其肿瘤微环境、免疫反应之间的

关系。研究认为，肿瘤微环境与乳腺癌的产生与进展密切相关<sup>[29-30]</sup>，而其则表现为大量炎性细胞浸润，如 B 细胞浸润与晚期乳腺癌相关、CD4<sup>+</sup>、CD8<sup>+</sup> T 细胞的浸润与乳腺癌预后相关<sup>[30-31]</sup>。

从上述分析可知，细胞焦亡与乳腺癌有显著的相关性，因此挖掘可调控细胞焦亡靶基因的中药则可为乳腺癌治疗药物的研发提供参考。乳腺癌属中医“乳岩”范畴，其发病主要为情志失调、脾胃损伤、冲任失调，痰浊、瘀血乘结乳房而生<sup>[32]</sup>。从数据挖掘结果来看，所得核心中药木蝴蝶、红花、余

甘子、白果、半枝莲、牛膝、菊花、香薷、马齿苋、连翘、连钱草、败酱草、枇杷叶、关黄柏、皂角刺、马鞭草、金银花<sup>[33-49]</sup>等均具有抗肿瘤、调节免疫及抗炎作用，其所含成分中山柰酚可通过抑制 Caspase-1 的表达和活性、增加 Toll 样受体 4 和 NLRP3 的蛋白表达以及抑制 GSDMD 的分解来减少细胞焦亡<sup>[50]</sup>；动物实验证明槲皮素可阻断氧化应激诱导的软骨细胞焦亡来抑制骨关节炎<sup>[51]</sup>；木犀草素则被报道有助于减轻脓毒症诱导的肺损伤小鼠模型肺组织中 Caspase-11 依赖的细胞焦亡<sup>[52]</sup>；黄芩素可上调微小 RNA-192-5p (miR-192-5p) /硫氧还蛋白互作蛋白 (thioredoxin-interacting protein, TXNIP) 轴而抑制 NLRP3/Caspase-1 通路，从而调控细胞焦亡和炎症<sup>[53]</sup>；汉黄芩素可通过抑制 GSDMD 介导的细胞焦亡，减轻顺铂诱导的小鼠心脏毒性<sup>[54]</sup>；表没食子儿茶精没食子酸酯可抑制 Caspase-1 激活和 IL-1 $\beta$  分泌<sup>[55]</sup>；柚皮素可通过调节微小 RNA-200b (miR-200b) /锌指蛋白 802 (zinc finger protein 802, JAZF1) 轴抑制乳腺癌的发生<sup>[56]</sup>。根据生物信息学分析表明，乳腺癌中细胞焦亡过程涉及免疫浸润与炎症因子聚集，而炎症因子聚集多造成局部炎症发热，该过程与中医的“火毒”存在一定的相似性<sup>[57]</sup>，因此其治疗用药以苦、寒为主；而内火多因“郁”而成，毒则为“聚而不散”所致，故用药应当辛散温通；中医历来强调顾护正气，正气存内，邪不可干，故甘平补益之品当为辅。从归经来看细胞焦亡相关药物主要归肝、肺、脾、胃经，因龙虎回环、脾升胃降皆为周身气机之枢，故细胞焦亡、乳腺癌此类“郁”所致的病理过程或疾患应以此为归经。

#### 4 结论

本研究采用生物信息学方法对 TCGA、GEO 数据库的乳腺癌数据进行细胞焦亡相关多组学分析，从而得出细胞焦亡与乳腺癌有显著的相关性，其涉及肿瘤微环境免疫细胞浸润相关机制。而对细胞焦亡起调控作用的中药主要为木蝴蝶、红花、余甘子、白果、半枝莲、牛膝、菊花、香薷、马齿苋、连翘、连钱草、败酱草、枇杷叶、关黄柏、皂角刺、马鞭草、金银花，其主要为苦寒，次为辛温，辅以甘平之品，而主要调节肝、肺、脾、胃等脏腑，发挥清热解毒、调和肝脾之功。本研究结果可为中药的临床应用与进一步研究提供指导与参考。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

- [1] Shi J J, Gao W Q, Shao F. Pyroptosis: gasdermin-mediated programmed necrotic cell death [J]. *Trends Biochem Sci*, 2017, 42(4): 245-254.
- [2] Kovacs S B, Miao E A. Gasdermins: Effectors of pyroptosis [J]. *Trends Cell Biol*, 2017, 27(9): 673-684.
- [3] Fang Y, Tian S W, Pan Y T, et al. Pyroptosis: A new frontier in cancer [J]. *Biomed Pharmacother*, 2020, 121: 109595.
- [4] Xia X J, Wang X, Cheng Z, et al. The role of pyroptosis in cancer: Pro-cancer or pro- "host"? [J]. *Cell Death Dis*, 2019, 10(9): 650.
- [5] Nagini S. Breast cancer: Current molecular therapeutic targets and new players [J]. *Anticancer Agents Med Chem*, 2017, 17(2): 152-163.
- [6] Rojas K, Stuckey A. Breast cancer epidemiology and risk factors [J]. *Clin Obstet Gynecol*, 2016, 59(4): 651-672.
- [7] Li T, Mello-Thoms C, Brennan P C. Descriptive epidemiology of breast cancer in China: Incidence, mortality, survival and prevalence [J]. *Breast Cancer Res Treat*, 2016, 159(3): 395-406.
- [8] Fan L, Strasser-Weippl K, Li J J, et al. Breast cancer in China [J]. *Lancet Oncol*, 2014, 15(7): e279-e289.
- [9] Chetlen A, Mack J, Chan T. Breast cancer screening controversies: Who, when, why, and how? [J]. *Clin Imaging*, 2016, 40(2): 279-282.
- [10] Pan H C, Gray R, Braybrooke J, et al. 20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years [J]. *N Engl J Med*, 2017, 377(19): 1836-1846.
- [11] Peart O. Metastatic breast cancer [J]. *Radiol Technol*, 2017, 88(5): 519M-539M.
- [12] McDonald E S, Clark A S, Tchou J, et al. Clinical diagnosis and management of breast cancer [J]. *J Nucl Med*, 2016, 57(Suppl 1): 9S-16S.
- [13] Condorelli R, Vaz-Luis I. Managing side effects in adjuvant endocrine therapy for breast cancer [J]. *Expert Rev Anticancer Ther*, 2018, 18(11): 1101-1112.
- [14] Taylor C W, Kirby A M. Cardiac side-effects from breast cancer radiotherapy [J]. *Clin Oncol (R Coll Radiol)*, 2015, 27(11): 621-629.
- [15] Lacouture M, Sibaud V. Toxic side effects of targeted therapies and immunotherapies affecting the skin, oral mucosa, hair, and nails [J]. *Am J Clin Dermatol*, 2018, 19(Suppl 1): 31-39.
- [16] 李云祥, 梁引库, 高飞雄, 等. 中药治疗乳腺癌疾病研究进展 [J]. *中国实验方剂学杂志*, 2019, 25(3): 211-219.
- [17] Xu Y J, Zheng L, Hu Y W, et al. Pyroptosis and its relationship to atherosclerosis [J]. *Clin Chim Acta*, 2018, 476: 28-37.

- [18] Zeng Z L, Li G H, Wu S Y, *et al.* Role of pyroptosis in cardiovascular disease [J]. *Cell Prolif*, 2019, 52(2): e12563.
- [19] Man S M, Karki R, Kanneganti T D. Molecular mechanisms and functions of pyroptosis, inflammatory caspases and inflammasomes in infectious diseases [J]. *Immunol Rev*, 2017, 277(1): 61-75.
- [20] Zheng S S, Xie X Y, Guo X K, *et al.* Identification of a pyroptosis-related gene signature for predicting overall survival and response to immunotherapy in hepatocellular carcinoma [J]. *Front Genet*, 2021, 12: 789296.
- [21] Sheikh A, Hussain S A, Ghorri Q, *et al.* The spectrum of genetic mutations in breast cancer [J]. *Asian Pac J Cancer Prev*, 2015, 16(6): 2177-2185.
- [22] Hou J W, Zhao R C, Xia W Y, *et al.* PD-L1-mediated gasdermin C expression switches apoptosis to pyroptosis in cancer cells and facilitates tumour necrosis [J]. *Nat Cell Biol*, 2020, 22(10): 1264-1275.
- [23] Ke H L, Wang X P, Zhou Z, *et al.* Effect of weimaining on apoptosis and Caspase-3 expression in a breast cancer mouse model [J]. *J Ethnopharmacol*, 2021, 264: 113363.
- [24] Jang H J, Suh P G, Lee Y J, *et al.* PLC $\gamma$ 1: Potential arbitrator of cancer progression [J]. *Adv Biol Regul*, 2018, 67: 179-189.
- [25] Kadam C Y, Abhang S A. Apoptosis markers in breast cancer therapy [J]. *Adv Clin Chem*, 2016, 74: 143-193.
- [26] Zohny S F, El-Shinawi M. Significance of survivin and Bcl-2 homologous antagonist/killer mRNA in detection of bone metastasis in patients with breast cancer [J]. *Med Oncol*, 2011, 28(Suppl 1): S108-S114.
- [27] Yang F, Xie H Y, Yang L F, *et al.* Stabilization of MORC2 by estrogen and antiestrogens through GPER1-PRKACA-CMA pathway contributes to estrogen-induced proliferation and endocrine resistance of breast cancer cells [J]. *Autophagy*, 2020, 16(6): 1061-1076.
- [28] Masjedi A, Hashemi V, Hojjat-Farsangi M, *et al.* The significant role of interleukin-6 and its signaling pathway in the immunopathogenesis and treatment of breast cancer [J]. *Biomed Pharmacother*, 2018, 108: 1415-1424.
- [29] Soysal S D, Tzankov A, Muenst S E. Role of the tumor microenvironment in breast cancer [J]. *Pathobiology*, 2015, 82(3/4): 142-152.
- [30] Mittal S, Brown N J, Holen I. The breast tumor microenvironment: Role in cancer development, progression and response to therapy [J]. *Expert Rev Mol Diagn*, 2018, 18(3): 227-243.
- [31] Wu J B, Zhu Y Y, Luo M M, *et al.* Comprehensive analysis of pyroptosis-related genes and tumor microenvironment infiltration characterization in breast cancer [J]. *Front Immunol*, 2021, 12: 748221.
- [32] Kresovich J K, O'Brien K M, Xu Z L, *et al.* Prediagnostic immune cell profiles and breast cancer [J]. *JAMA Netw Open*, 2020, 3(1): e1919536.
- [33] 谭新华, 陆德铭. 中医外科学 [M]. 北京: 人民卫生出版社, 1999: 396.
- [34] Cao Y, Cao W J, Qiu Y M, *et al.* Oroxylin A suppresses ACTN1 expression to inactivate cancer-associated fibroblasts and restrain breast cancer metastasis [J]. *Pharmacol Res*, 2020, 159: 104981.
- [35] Wu J Y C, Yu Z L, Fong W F, *et al.* Chemotherapeutic activities of *Carthami Flos* and its reversal effect on multidrug resistance in cancer cells [J]. *Afr J Tradit Complement Altern Med*, 2013, 10(4): 36-40.
- [36] Zhao H J, Liu T, Mao X, *et al.* Fructus *Phyllanthi* tannin fraction induces apoptosis and inhibits migration and invasion of human lung squamous carcinoma cells *in vitro* via MAPK/MMP pathways [J]. *Acta Pharmacol Sin*, 2015, 36(6): 758-768.
- [37] 夏梦雨, 张雪, 王云, 等. 白果的炮制方法、化学成分、药理活性及临床应用的研究进展 [J]. 中国药房, 2020, 31(1): 123-128.
- [38] Powell C B, Fung P, Jackson J, *et al.* Aqueous extract of *Herba Scutellaria Barbatae*, a Chinese herb used for ovarian cancer, induces apoptosis of ovarian cancer cell lines [J]. *Gynecol Oncol*, 2003, 91(2): 332-340.
- [39] 罗懿钊, 欧阳文, 唐代凤, 等. 牛膝中皂苷和甾酮类物质基础及药理活性研究进展 [J]. 中国现代中药, 2020, 22(12): 2122-2136.
- [40] Li X F, Xu L, Jiang H, *et al.* Flash extraction of the active constituents and its antitumor activity from *Rehmanniae Radix*, *Achyranthes Bidentatae Radix*, *Dioscoreae Rhizoma* and *Chrysanthemi Flos* [J]. *Pak J Pharm Sci*, 2021, 34(5): 1777-1782.
- [41] 姚奕, 许浚, 黄广欣, 等. 香薷的研究进展及其质量标志物预测分析 [J]. 中草药, 2020, 51(10): 2661-2670.
- [42] Zhou Y X, Xin H L, Rahman K, *et al.* *Portulaca oleracea* L.: A review of phytochemistry and pharmacological effects [J]. *Biomed Res Int*, 2015, 2015: 925631.
- [43] Wang Z Y, Xia Q, Liu X, *et al.* Phytochemistry, pharmacology, quality control and future research of *Forsythia suspensa* (Thunb.) Vahl: A review [J]. *J Ethnopharmacol*, 2018, 210: 318-339.
- [44] 吴西, 周雷罡, 邓维, 等. 连钱草化学成分及药理作用研究进展 [J]. 药品评价, 2021, 18(1): 4-7.
- [45] Gong L N, Zou W, Zheng K Y, *et al.* The *Herba Patriniae* (Caprifoliaceae): A review on traditional uses,

- phytochemistry, pharmacology and quality control [J]. *J Ethnopharmacol*, 2021, 265: 113264.
- [46] Zar P P K, Yano S, Sakao K, *et al.* *In vitro* anticancer activity of loquat tea by inducing apoptosis in human leukemia cells [J]. *Biosci Biotechnol Biochem*, 2014, 78(10): 1731-1737.
- [47] 李先宽, 冯杉, 郑艳超, 等. 黄柏与关黄柏的化学成分及生物活性研究进展 [J]. *药物评价研究*, 2019, 42(5): 1033-1037.
- [48] 李文基, 马骏, 赵文秀. 皂角刺抗肿瘤药理作用及化学成分研究进展 [J]. *甘肃中医药大学学报*, 2020, 37(6): 85-88.
- [49] 刘雅琳, 苗晋鑫, 田硕, 等. 马鞭草化学成分及药理作用研究进展 [J]. *河南中医*, 2021, 41(2): 294-299.
- [50] He C L, Yang J, Jiang X L, *et al.* Kaempferol alleviates LPS-ATP mediated inflammatory injury in splenic lymphocytes via regulation of the pyroptosis pathway in mice [J]. *Immunopharmacol Immunotoxicol*, 2019, 41(5): 538-548.
- [51] Wang Q, Ying L J, Wei B, *et al.* Effects of quercetin on apoptosis and extracellular matrix degradation of chondrocytes induced by oxidative stress-mediated pyroptosis [J]. *J Biochem Mol Toxicol*, 2022, 36(2): e22951.
- [52] Zhang Z T, Zhang D Y, Xie K, *et al.* Luteolin activates Tregs to promote IL-10 expression and alleviating caspase-11-dependent pyroptosis in *Sepsis*-induced lung injury [J]. *Int Immunopharmacol*, 2021, 99: 107914.
- [53] Wang X Y, Cai H, Chen Z Y, *et al.* Baicalein alleviates pyroptosis and inflammation in hyperlipidemic pancreatitis by inhibiting NLRP3/Caspase-1 pathway through the miR-192-5p/TXNIP axis [J]. *Int Immunopharmacol*, 2021, 101(Pt B): 108315.
- [54] Xu J J, Zhang B, Chu Z L, *et al.* Wogonin alleviates cisplatin-induced cardiotoxicity in mice via inhibiting gasdermin D-mediated pyroptosis [J]. *J Cardiovasc Pharmacol*, 2021, 78(4): 597-603.
- [55] Zhang C, Li X, Hu X, *et al.* Epigallocatechin-3-gallate prevents inflammation and diabetes-Induced glucose tolerance through inhibition of NLRP3 inflammasome activation [J]. *Int Immunopharmacol*, 2021, 93: 107412.
- [56] Wang J G, Jian W J, Li Y, *et al.* Nobiletin promotes the pyroptosis of breast cancer via regulation of miR-200b/JAZF1 axis [J]. *Kaohsiung J Med Sci*, 2021, 37(7): 572-582.
- [57] 梁子成, 陈泓秀, 文璐, 等. 基于网络药理学及数据挖掘探讨中药调节细胞焦亡用药规律 [J]. *中国中医药信息杂志*, 2021, 28(11): 27-33.

[责任编辑 潘明佳]