

基于转录组学、单细胞虚拟敲除和分子模拟提示CHST11在三阴性乳腺癌中的潜在调控作用

于 梦, 杨 慧, 许卓华, 江 舟, 陆 颖*

柳州市工人医院, 广西医科大学第四附属医院肿瘤科, 广西 柳州

收稿日期: 2026年7月13日; 录用日期: 2026年8月6日; 发布日期: 2026年8月17日

摘 要

目的: 本研究旨在系统解析三阴性乳腺癌(TNBC)的分子特征, 并探讨CHST11在TNBC发生发展中的潜在作用。方法: 基于GEO转录组数据, 采用差异表达分析和WGCNA筛选TNBC相关候选基因, 并通过LASSO回归和随机森林算法筛选核心候选基因。随后进行GO、KEGG、GSEA和Metascape功能富集分析。利用HPA数据库评估核心基因的蛋白表达, 并结合CIBERSORT分析其与免疫浸润的关系。通过单细胞RNA测序评估核心基因在TNBC肿瘤微环境细胞亚群中的表达分布, 并采用scTenifoldKmk进行核心基因虚拟敲除及扰动基因富集分析。最后, 通过分子对接和分子动力学模拟评估槲皮素与核心基因编码蛋白的结合稳定性。结果: 共筛选获得678个TNBC相关候选基因, 并筛选CHST11为核心候选基因。功能富集结果显示, 候选基因主要参与细胞增殖、迁移、上皮-间质转化、细胞外基质重塑及炎症反应, 并显著富集于PI3K/AKT、IL-6/JAK-STAT3、TGF- β 和NF- κ B等通路。HPA验证显示CHST11在乳腺癌组织中表达上调。免疫浸润分析提示其与多类免疫细胞亚群密切相关。单细胞分析进一步显示, CHST11主要表达于巨噬细胞、单核细胞、成纤维细胞、B细胞、CD8⁺ T细胞、树突状细胞和内皮细胞。CHST11虚拟敲除后共有27个基因发生显著扰动, 扰动基因主要富集于炎症信号、免疫调节、细胞黏附迁移、细胞外基质重塑及血管通透性调控等过程。分子对接和动力学模拟显示, 槲皮素与CHST11蛋白具有稳定结合能力。结论: CHST11可能通过参与多信号通路调控促进TNBC发生发展, 而槲皮素可能作为靶向CHST11的潜在治疗分子。

关键词

三阴性乳腺癌, CHST11, 单细胞测序, 虚拟敲除, 免疫浸润, 槲皮素, 分子动力学模拟

*通讯作者。

Transcriptomic, Single-Cell Virtual Knockout, and Molecular Simulation Analyses Suggest a Potential Regulatory Role of CHST11 in Triple-Negative Breast Cancer

Meng Yu, Hui Yang, Zhuohua Xu, Zhou Jiang, Ying Lu*

Department of Oncology, The Fourth Affiliated Hospital of Guangxi Medical University, Liuzhou Workers' Hospital, Liuzhou Guangxi

Received: July 13, 2026; accepted: August 6, 2026; published: August 17, 2026

Abstract

Objective: This study aimed to systematically characterize the molecular features of triple-negative breast cancer (TNBC) and explore the potential role of CHST11 in TNBC development and progression. **Methods:** Based on GEO transcriptomic data, TNBC-related candidate genes were screened using differential expression analysis and weighted gene co-expression network analysis, and core genes were identified using LASSO regression and random forest algorithms. GO, KEGG, GSEA, and Metascape functional enrichment analyses were then performed. Protein expression of the core gene was assessed using the HPA database, and its association with immune infiltration was evaluated using CIBERSORT. Single-cell RNA sequencing was further used to evaluate the expression distribution of the core gene across different cell subsets in the TNBC tumor microenvironment. scTenifoldKnn was applied for virtual knockout of the core gene and enrichment analysis of perturbed genes. Finally, molecular docking and molecular dynamics simulation were performed to assess the binding stability between quercetin and the protein encoded by the core gene. **Results:** A total of 678 TNBC-related candidate genes were identified, and CHST11 was screened as a core candidate gene. Functional enrichment analysis showed that the candidate genes were mainly involved in cell proliferation, migration, epithelial-mesenchymal transition, extracellular matrix remodeling, and inflammatory responses, and were significantly enriched in the PI3K/AKT, IL-6/JAK-STAT3, TGF- β , and NF- κ B pathways. HPA validation showed that CHST11 was upregulated in breast cancer tissues, and immune infiltration analysis indicated that CHST11 was closely associated with multiple immune cell subsets. Single-cell analysis further revealed that CHST11 was mainly expressed in macrophages, monocytes, fibroblasts, B cells, CD8⁺ T cells, dendritic cells, and endothelial cells. After CHST11 virtual knockout, 27 genes were significantly perturbed, and these perturbed genes were mainly enriched in inflammatory signaling, immune regulation, cell adhesion and migration, extracellular matrix remodeling, and vascular permeability regulation. Molecular docking and molecular dynamics simulation suggested that quercetin may stably bind to CHST11. **Conclusion:** CHST11 may promote TNBC development and progression by participating in multiple signaling pathways, while quercetin may serve as a potential therapeutic molecule targeting CHST11.

Keywords

Triple-Negative Breast Cancer, CHST11, Single-Cell Sequencing, Virtual Knockout, Immune Infiltration, Quercetin, Molecular Dynamics Simulation



1. 前言

三阴性乳腺癌(Triple-Negative Breast Cancer, TNBC)是乳腺癌的重要亚型, 约占全部乳腺癌的15%~20%, 其特点是缺乏雌激素受体、孕激素受体和 HER2 表达[1]。与其他分子亚型相比, TNBC 的发病人群更趋年轻化, 往往在诊断时已处于较晚分期, 且肿瘤呈现高度异质性和侵袭性[2]。临床研究表明, TNBC 对化疗较为敏感, 但易出现早期复发和远处转移, 尤其是脑转移和肺转移的发生率显著高于非 TNBC 患者[3][4]。尽管近年来, 免疫治疗和 PARP 抑制剂等新兴疗法在部分患者中显示出一定疗效, 但总体获益仍有限, 且耐药问题普遍存在[5]。因此, 系统解析 TNBC 的分子特征并寻找潜在治疗靶点, 对于改善该类患者的预后具有重要临床意义。然而, TNBC 的高度异质性提示其发生发展并非由单一基因或单一通路驱动, 而是涉及肿瘤细胞内在异常、免疫微环境重塑及基质细胞互作等多层次机制。传统单一维度分析难以全面提示这些复杂调控网络, 因此有必要整合多组学数据、机器学习和计算模拟方法, 从不同层面筛选并验证 TNBC 的关键分子。

随着转录组学、高通量测序及公共数据库资源的发展, 基于 GEO 等数据库的多层次生物信息学分析已成为解析 TNBC 分子机制的重要手段。差异表达分析和加权基因共表达网络分析(Weighted Gene Co-expression Network Analysis, WGCNA)可筛选与疾病表型相关的候选基因。基因集富集分析(Gene Set Enrichment Analysis, GSEA)、基因本体论(Gene Ontology, GO)、京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)及 Metascape 分析可从通路层面提示肿瘤增殖、侵袭迁移和微环境重塑等关键过程。机器学习方法可进一步从高维转录组数据中识别核心基因。单细胞 RNA 测序可明确核心基因在不同细胞群中的表达分布, 虚拟敲除分析则可预测目标基因扰动后的下游转录变化。分子对接和分子动力学模拟用于评估候选药物与靶蛋白的结合模式及稳定性, 为潜在治疗分子筛选提供结构生物学依据。

在此背景下, 本研究基于 GEO 数据库的 TNBC 表达数据, 整合差异表达分析、WGCNA、GO、KEGG、GSEA 和 Metascape 功能富集, 系统解析 TNBC 的关键分子通路。通过最小绝对收缩与选择算子(Least Absolute Shrinkage and Selection Operator, LASSO)回归与随机森林算法筛选核心基因, 并利用人类蛋白图谱(Human Protein Atlas, HPA)数据库对核心基因的编码蛋白进行验证。同时, 结合免疫浸润分析探讨其与肿瘤微环境的关系。在此基础上, 纳入单细胞 RNA 测序数据, 从单细胞水平评估核心基因在 TNBC 肿瘤微环境不同细胞亚群中的表达分布, 并通过 scTenifoldKnk 虚拟敲除分析评估核心基因扰动后显著变化的调控基因及其功能富集特征。最后, 本研究采用分子对接与分子动力学模拟, 验证潜在中药单体对核心基因蛋白的结合稳定性以探究其潜在治疗作用。通过这一研究框架, 本研究旨在从转录组、免疫浸润、单细胞表达分布、虚拟敲除调控网络及分子模拟等多个层面系统评估 CHST11 在 TNBC 发生发展中的潜在作用, 并预测可能靶向 CHST11 的候选治疗分子, 为 TNBC 分子机制研究和潜在干预策略提供理论依据。

2. 方法

2.1. 数据集筛选与标准化处理

本研究于 2025 年 1 月在 GEO 中检索并筛选适用于 TNBC 研究的转录组数据集。纳入标准为: ① 样本需明确标注为 TNBC 与正常乳腺组织; ② 每组样本量不少于 10 例; ③ 样本采集时间均在放化疗前;

④ 数据集需包含完整的基因表达矩阵及临床信息。原始表达矩阵下载后,首先采用 K-nearest neighbor 方法插补缺失值,随后进行 $\log_2$ 转换,并使用 R 语言 limma 包的 normalizeBetweenArrays 函数进行分位数标准化,以减少系统偏差并提高样本间的可比性。

2.2. 差异表达基因获取

在完成标准化后,首先进行主成分分析以评估 TNBC 与正常组织在整体转录模式上的分布差异和组间分离情况。随后基于分组信息构建表达矩阵,并利用 limma 包拟合线性模型。差异检验结果经经验贝叶斯方法校正,以提高统计效能。差异表达基因(Differentially Expressed Genes, DEGs)的筛选标准设定为 $|\log_2FC| > 1$ 且假发现率(False Discovery Rate, FDR) < 0.05 。

2.3. 加权基因共表达网络的构建与模块识别

在差异分析的基础上,进一步应用 WGCNA 探索基因间的协同表达关系并识别关键模块。首先选取变异度最高的前 25%基因构建网络,并通过样本聚类剔除异常值。随后确定软阈值以保证网络符合无尺度特性,计算拓扑重叠矩阵并进行层次聚类。模块划分采用动态剪切树方法,并对高度相似的模块进行合并。最后,将模块特征向量与临床表型进行相关性分析。

2.4. 候选基因的交集分析与筛选

为获得与 TNBC 高度相关的候选基因,将差异分析得到的 DEGs 与 WGCNA 显著相关模块基因取交集。交集结果通过 Venn 图展示。该部分基因既具备差异表达特征,又体现网络模块属性,因此被定义为 TNBC 相关差异基因。

2.5. GO 与 KEGG 功能富集分析及 Metascape 注释

候选基因进一步进行 GO 与 KEGG 富集分析,以提示其潜在生物学功能和关键信号通路。GO 分析涵盖生物过程、细胞组分和分子功能三个维度;KEGG 分析用于探索信号通路富集特征。两者均以 FDR < 0.05 作为显著性阈值,并对与 TNBC 密切相关的条目进行可视化展示。为增强功能注释的广度,本研究同时将候选基因输入 Metascape 数据库,物种选项均设为 H.sapiens,并导出显著性最高的前 20 个条目作为补充参考。

2.6. 基因集富集分析

所有基因按照 $\log_2FC$ 值由高到低排序,参考基因集为 MSigDB Hallmark (h.all.v2024.1.Hs.symbols.gmt)。结果通过 Benjamini-Hochberg 方法进行 FDR 校正,并以 FDR < 0.05 为显著性阈值。最终选取在 TNBC 样本中显著富集的前五条通路进行可视化展示。

2.7. 免疫浸润分析

为探索 TNBC 与正常乳腺组织免疫微环境差异,本研究采用 CIBERSORT 算法对标准化表达矩阵进行免疫浸润推断。分析以 LM22 基因矩阵为参考,置换次数设为 1000,仅保留 $P < 0.05$ 的样本以确保推断结果的可靠性。免疫细胞整体分布通过堆叠条形图展示,并采用 Wilcoxon 秩和检验比较两组间差异。显著差异的免疫细胞亚群进一步通过箱线图进行可视化。

2.8. 核心基因筛选与验证

在候选基因的基础上,结合两种机器学习方法筛选 TNBC 的核心基因。首先,利用 LASSO 回归构

建二分类模型,并通过十折交叉验证确定最优惩罚参数 λ ,筛选出回归系数不为零的基因。其次,采用随机森林模型训练基因集,并根据平均基尼指数减少值评估基因重要性,筛选出得分大于1的基因。两种方法的结果取交集,最终获得核心基因集。为评估核心基因在蛋白水平的差异表达,本研究检索 HPA 数据库,选择乳腺癌和正常乳腺组织免疫组化切片,对比核心基因编码的蛋白在两种组织中的表达模式。

2.9. 核心基因与免疫浸润的关联及单细胞表达特征分析

为探讨核心基因在 TNBC 免疫微环境中的潜在作用,本研究首先分析核心基因在 TNBC 样本中的表达水平与免疫细胞浸润比例之间的相关性。从标准化表达矩阵中提取核心基因表达数据,并与 CIBERSORT 免疫浸润结果进行样本匹配。随后,仅保留 TNBC 样本,采用 Spearman 相关检验计算核心基因表达水平与不同免疫细胞亚群比例之间的相关系数及 P 值,并通过 R 语言 linkET 包中的 qcorrplot 函数进行可视化,以展示核心基因与免疫细胞浸润之间的潜在关联。

为进一步从单细胞水平评估核心基因在 TNBC 肿瘤微环境中的表达分布,本研究纳入 GEO 数据库中的单细胞 RNA 测序数据集 GSE263995,并选取样本 GSM8391113 进行分析。利用 Seurat 包读取 10X Genomics 表达矩阵并构建 Seurat 对象。经细胞质控、LogNormalize 标准化、高变基因筛选、PCA 降维、聚类及 UMAP 可视化后,采用 SingleR 包进行细胞类型自动注释。主要参考数据库包括 Blueprint/ENCODE、Human Primary Cell Atlas、Monaco Immune Data、Novershtern Hematopoietic Data 和 Database Immune Cell Expression Data。随后从标准化表达矩阵中提取核心基因在不同细胞亚群中的表达水平,并通过气泡图展示其细胞类型特异性分布。

2.10. 核心基因虚拟敲除分析及功能富集

为进一步评估核心基因在 TNBC 单细胞转录调控网络中的潜在作用,本研究基于 GSE263995 数据集中 GSM8391113 样本开展单基因虚拟敲除分析。首先读取已注释的 Seurat 对象,并以该样本全部细胞的 RNA counts 矩阵作为输入数据。随后对表达矩阵进行标准化处理,筛选前 5000 个高变基因,并将核心基因纳入分析基因集,去除全零表达基因后构建输入矩阵。虚拟敲除分析采用 scTenifoldKnk 算法完成,主要参数设置为 qc_mtThreshold = 0.1、qc_minLSize = 1000、nc_nNet = 10, nc_nCells 根据实际细胞数量自动设定,最大值为 500。敲除后根据 diffRegulation 结果筛选显著扰动基因,并以校正后 P 值 $p_{adj} < 0.05$ 作为筛选标准。随后对显著扰动基因进行 GO、KEGG 和 Reactome 富集分析,以解析核心基因相关调控网络涉及的生物学过程、细胞组分、分子功能及关键信号通路。

2.11. 分子对接分析

为筛选可能靶向核心基因编码蛋白的候选药物,本研究进一步开展分子对接分析。首先通过文献检索获得在 TNBC 中具有潜在治疗价值的中药单体,并从 PubChem 数据库下载其化学结构文件。随后使用 ChemBio3D 对中药单体结构进行能量最小化处理,以获得更稳定的分子构象。核心基因编码蛋白的三维结构文件从 AlphaFold 数据库获取,并在 AutoDockTools 中进行受体与配体的预处理。对接过程中通过设置对接盒子确定活性口袋范围,最终利用 AutoDockVina 完成分子对接。结合能最低的构象被视为最优结合模式,并通过 PyMOL 与 LigPlot 进行可视化。

2.12. 分子动力学模拟

为进一步评估分子对接复合物的结合稳定性,本研究采用 GROMACS2023.2 软件进行分子动力学模拟。体系基于 CHARMM36 力场参数化,水模型选择 TIP3P。同时添加 Na^+ 与 Cl^- 离子以中和体系电荷。体系经过最速下降法与共轭梯度法两阶段能量最小化后,依次进行 100 ps 的 NVT 平衡(V-rescale 温控,

300 K)与 NPT 平衡(Parrinello-Rahman 压控, 1 bar)。随后进行 125 ns 正式模拟。模拟结束后, 通过计算均方根偏差(Root Mean Square Deviation, RMSD)、均方根波动(Root Mean Square Fluctuation, RMSF)、回转半径(Radius of Gyration, Rg)、溶剂可及表面积(Solvent Accessible Surface Area, SASA)及氢键数等指标, 以评估复合物的结构稳定性。

3. 结果

3.1. 数据集筛选结果

根据预设纳入标准, 本研究最终筛选获得 GSE53752 数据集, 共包含 25 例正常乳腺组织样本与 51 例 TNBC 组织样本。

3.2. 差异表达基因的鉴定

主成分分析结果显示 TNBC 与正常乳腺组织在整体转录模式上存在明显分离, 提示二者在基因表达谱方面具有明显差异(图 1(A))。进一步利用 limma 包进行差异分析, 共鉴定出 1332 个 DEGs, 其中 507 个在 TNBC 中显著上调, 825 个显著下调(图 1(B))。

3.3. 共表达网络模块与 TNBC 的相关性

WGCNA 分析以 softPower = 7 为阈值共划分出 5 个基因模块。其中 MEblue 模块与 TNBC 的相关性最强($r = -0.93$, $P < 0.05$), 包含 1376 个基因, 提示该模块可能在 TNBC 的发生发展中发挥核心作用(图 1(C))。

3.4. 差异表达基因与关键模块基因的交集

将 DEGs 与 MEblue 模块基因进行交集, 共获得 678 个候选基因(图 1(D))。这些基因同时具备差异表达特征和共表达网络属性, 因此被定义为与 TNBC 高度相关的关键候选基因集。

3.5. 候选基因的功能富集结果

GO 富集分析显示候选基因主要涉及细胞增殖、迁移、细胞外基质(Extracellular Matrix, ECM)重塑、缺氧应答和炎症反应等生物过程, 细胞组分集中于 ECM 与基底膜, 分子功能富集于 ECM 结构成分、整合素结合及胶原结合(图 1(E))。KEGG 分析进一步提示这些基因显著富集于 PI3K/AKT、IL-17、ECM-receptor interaction、focal adhesion 及 proteoglycans in cancer 等关键通路, 同时涉及药物代谢和能量代谢(图 1(F))。Metascape 结果补充强调 ECM 降解、缺氧反应及炎症相关过程在 TNBC 发生发展中的核心作用(图 1(G))。

3.6. 基因集富集分析提示 TNBC 中显著激活的信号通路

为评估整体转录水平的功能差异, 进行了 GSEA 分析。结果显示 TNBC 样本中显著富集多条肿瘤相关通路, 包括上皮-间质转化(Epithelial-Mesenchymal Transition, EMT)、IL6/JAK-STAT3 信号通路、炎症反应、TNF α /NF- κ B 信号通路以及 KRAS 上调信号通路(图 2(A))。这些结果提示 TNBC 的进展伴随 EMT 激活、炎症与免疫信号增强以及多条促增殖信号的共同驱动。

3.7. TNBC 与正常组织的免疫浸润差异

CIBERSORT 分析结果表明 TNBC 与正常组织在免疫细胞浸润构成方面存在显著差异(图 2(B))。差异检验结果进一步提示, TNBC 样本中浆细胞比例升高($P = 0.047$, 图 2(C)), 而嗜酸性粒细胞($P = 0.00093$,

图 2(D))、单核细胞($P = 0.0024$, 图 2(E))及静息记忆 $CD4^+$ T 细胞($P = 0.00039$, 图 2(F))比例均显著下降。这些发现提示 TNBC 的免疫微环境存在重塑, 其特征为多类免疫效应和调节细胞减少, 可能促进免疫逃逸和肿瘤进展。

3.8. 机器学习筛选核心候选基因集

为进一步缩小候选范围, 本研究结合 LASSO 回归与随机森林模型进行特征筛选。LASSO 模型在十折交叉验证下筛选出 19 个候选基因(图 3(A), 图 3(B))。随机森林模型筛选出 13 个重要性评分大于 1 的

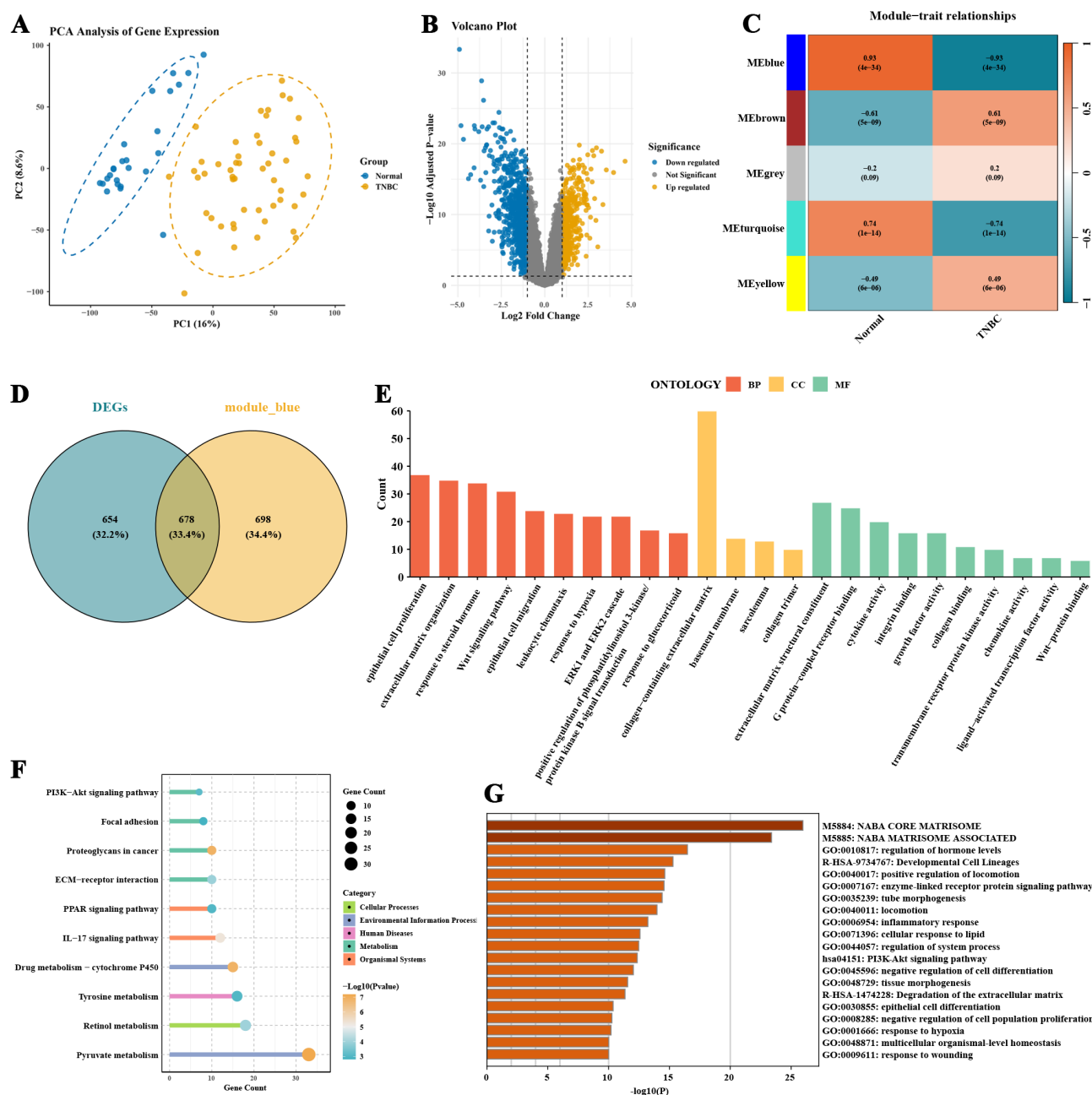


Figure 1. Identification of candidate genes in TNBC
图 1. TNBC 候选基因的鉴定

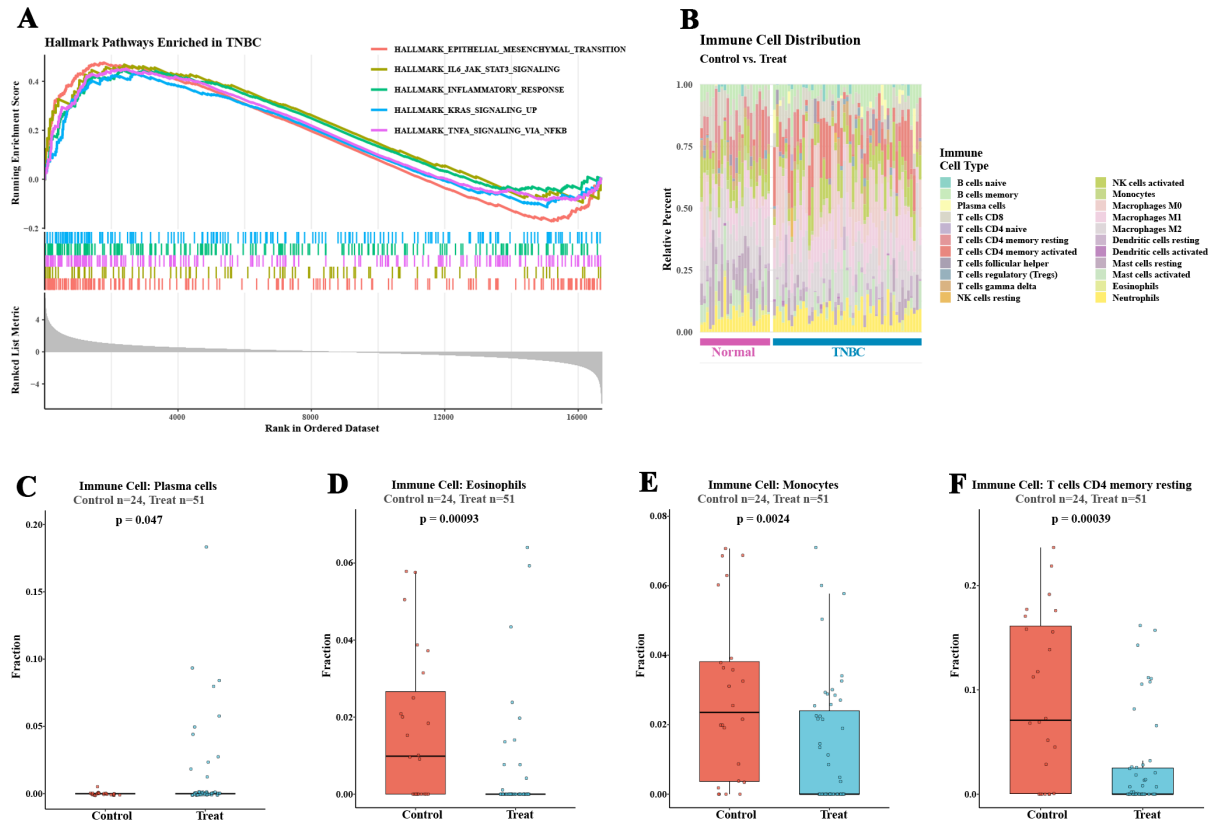


Figure 2. Functional enrichment and immune infiltration characteristics of TNBC

图 2. TNBC 功能富集与免疫浸润特征

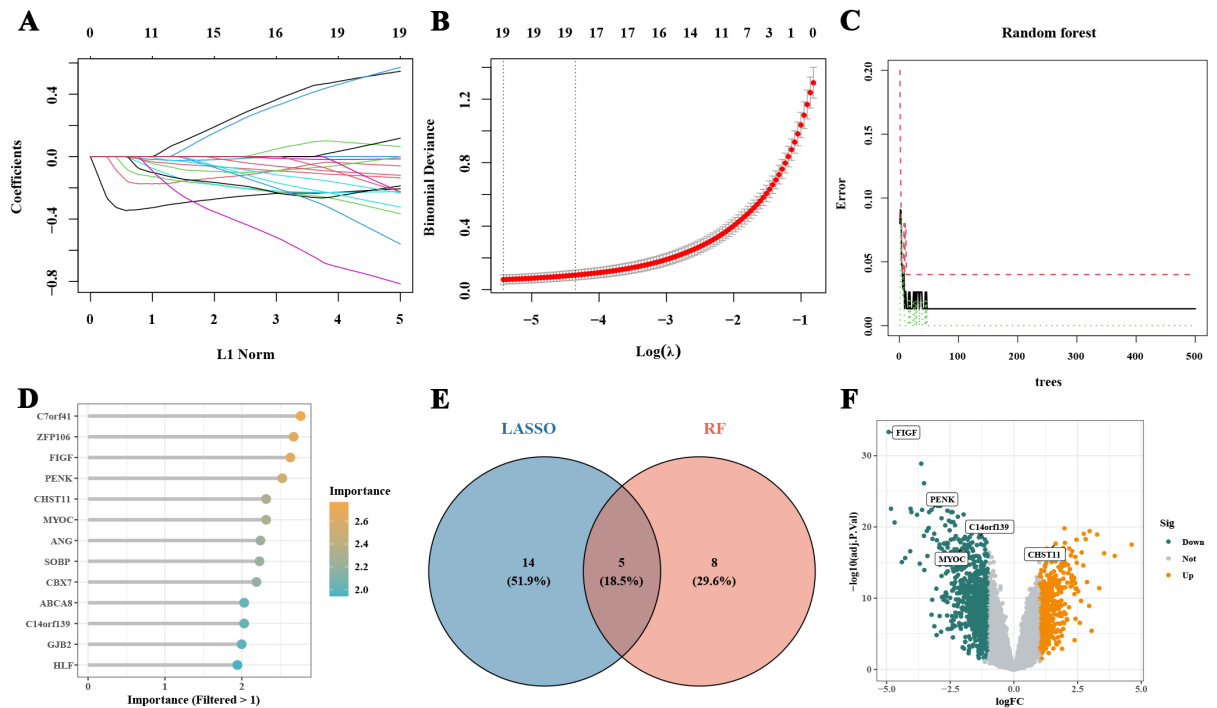


Figure 3. Machine learning-based identification of TNBC core genes

图 3. 机器学习筛选 TNBC 核心基因

基因(图 3(C), 图 3(D))。两种方法的交集结果最终确定 FIGF、PENK、CHST11、MYOC 和 C14orf139 为核心基因, 且 5 个核心基因在 TNBC 与正常组织之间均表现为显著差异表达(图 3(E), 图 3(F))。

3.9. CHST11 的表达验证、免疫相关性及单细胞表达特征

HPA 数据库免疫组化结果显示, CHST11 在乳腺癌组织中呈中等强度的细胞质及膜染色, 而在正常乳腺腺体细胞中仅表现为低水平弱染色, 提示 CHST11 在乳腺癌组织中存在蛋白水平上调趋势(图 4(A))。GSEA 分析表明, CHST11 高表达样本显著富集于 EMT、IL6/JAK-STAT3 信号通路、PI3K/AKT/mTOR 信号通路、TGF- β 信号通路以及 TNF α /NF- κ B 信号通路等关键途径, 提示 CHST11 可能与 TNBC 细胞侵袭、炎症信号增强、免疫抑制及耐药性形成密切相关(图 4(B))。

免疫浸润相关性分析显示, CHST11 表达水平与多类免疫细胞亚群存在显著相关性, 尤其与 CD8⁺ T 细胞、 $\gamma\delta$ T 细胞及嗜酸性粒细胞的浸润水平关系密切(图 4(C))。单细胞测序数据分析结果显示, CHST11 主要表达于巨噬细胞、单核细胞、成纤维细胞、B 细胞、CD8⁺ T 细胞、树突状细胞和内皮细胞等细胞群中(图 4(D)~(F))。这些细胞群分别涉及炎症反应、抗原呈递、细胞外基质重塑、免疫调节和血管内皮功能, 提示 CHST11 可能参与 TNBC 肿瘤微环境中免疫细胞、基质细胞及内皮细胞之间的相互作用。

3.10. CHST11 虚拟敲除提示其潜在调控网络及功能特征

为进一步评估 CHST11 在 TNBC 单细胞转录调控网络中的作用, 本研究基于 GSE263995 数据集中 GSM8391113 样本进行 CHST11 虚拟敲除分析。scTenifoldKnn 结果显示, CHST11 虚拟敲除后共有 27 个基因发生显著扰动, 其中包括 TYROBP、AIF1、LAPTM5 和 CD74 等与髓系细胞活化、炎症反应、溶酶体功能及抗原加工呈递密切相关的基因(图 5(A)-(C))。进一步对 27 个显著扰动基因进行 GO、KEGG 和 Reactome 富集分析。结果显示, 这些基因主要富集于细胞因子和趋化因子产生、TNF/NF- κ B、MAPK、TGF- β 、STAT 相关信号、细胞外基质结合与重塑、细胞黏附分子相互作用、白细胞跨内皮迁移、吞噬体和溶酶体功能、PD-1 共抑制、TLR 级联反应以及 VEGFR2 介导的血管通透性等过程(图 5(D)-(F))。

3.11. 槲皮素与 CHST11 的分子对接结果

文献检索提示槲皮素在 TNBC 中具有一定的抗肿瘤活性, 并可通过抑制 PI3K/AKT/mTOR 通路发挥作用。为评估其潜在作用靶点, 本研究对槲皮素与 CHST11 蛋白进行分子对接。结果显示二者结合能为 -8.5 kcal/mol, 提示具备较强亲和力。结合模式分析发现槲皮素可与 Pro302、Tyr304 和 Thr308 形成稳定氢键, 并伴随疏水作用共同稳定复合物(图 6(A))。这些作用位点处于蛋白的功能相关区域, 可能影响其结构及信号转导活性, 为槲皮素介导的抗 TNBC 作用提供了结构学依据。

3.12. 分子动力学模拟提示复合物的稳定结合

分子动力学模拟进一步支持槲皮素与 CHST11 蛋白复合物的稳定性。RMSD 分析显示复合物在约 50 ns 后趋于稳定, 主链与配体整体波动均维持在合理范围内(图 6(B))。RMSF 结果表明大多数残基波动较低(图 6(C))。氢键分析显示模拟过程中二者存在稳定而持久的氢键相互作用(图 6(D))。Rg 在模拟期间保持稳定, 提示蛋白整体结构未发生显著变化(图 6(E))。SASA 值随时间略有下降, 反映复合物趋于更紧密的构象(图 6(F))。自由能地形图显示体系在低能态区域聚集, 进一步支持复合物在能量学上可能处于相对稳定状态(图 6(G))。

4. 讨论

本研究通过整合生物信息学、机器学习、单细胞虚拟敲除与分子动力学模拟初步提示 TNBC 存在

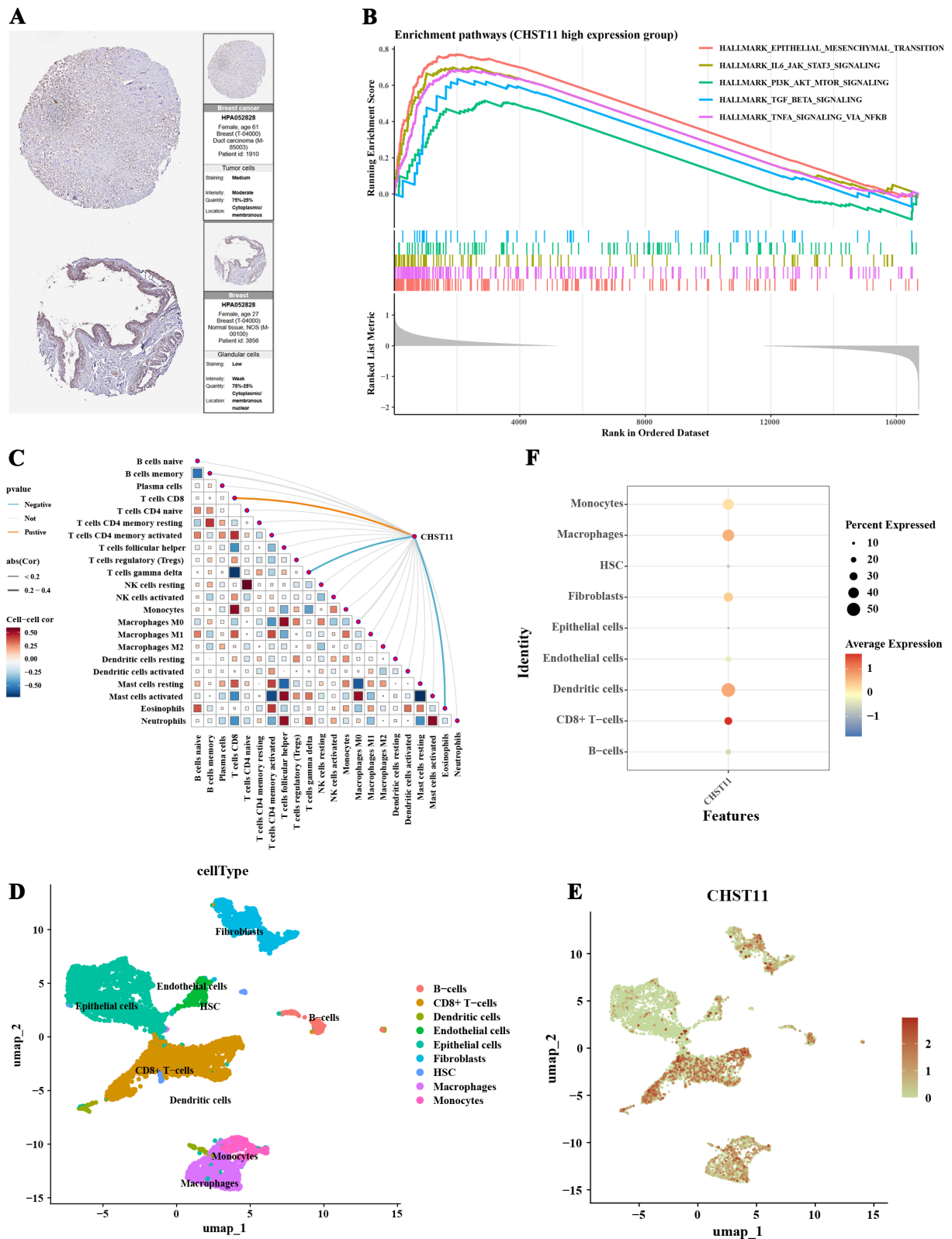


Figure 4. Expression validation, functional enrichment, immune correlation, and single-cell expression characteristics of CHST11

图 4. CHST11 的表达验证、功能富集、免疫相关性及单细胞表达特征

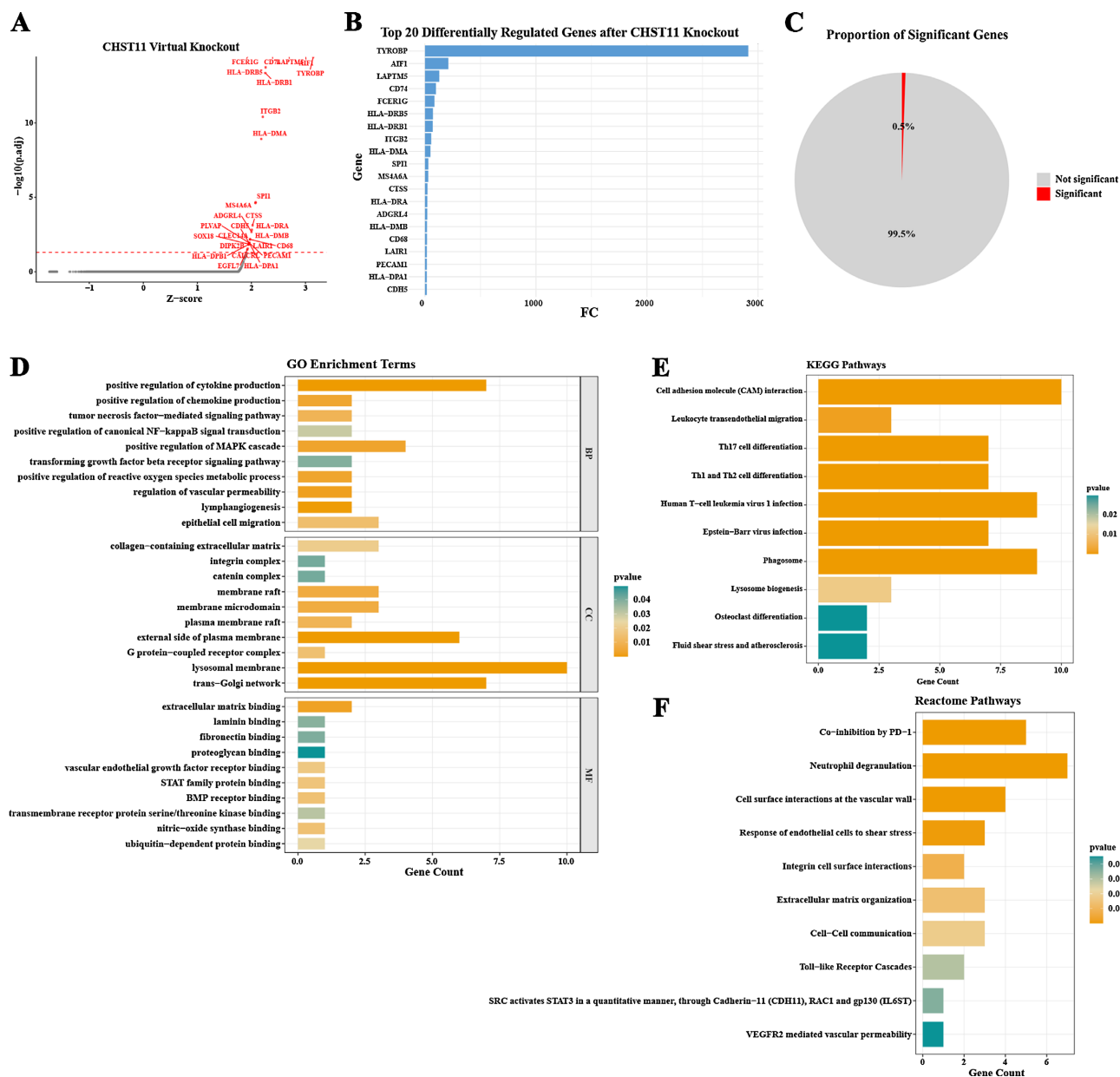
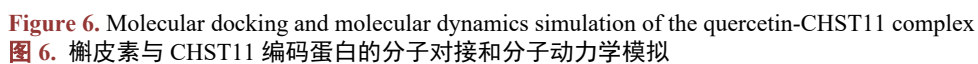


Figure 5. Virtual knockout of CHST11 and functional enrichment analysis of significantly perturbed genes

图 5. CHST11 虚拟敲除及显著扰动基因的功能富集分析



多通路异常及 CHST11 的潜在作用, 并为槲皮素作为潜在治疗候选分子提供了结构层面的预测依据。

在信号通路层面, 本研究结果提示 TNBC 并非由单一机制驱动, 而是与多条肿瘤相关通路的协同异常有关。从机制上推测, CHST11 作为糖胺聚糖/硫酸软骨素修饰相关基因, 可能通过改变细胞表面和细胞外基质中蛋白聚糖的硫酸化模式, 影响生长因子、炎症因子及黏附受体的可及性和受体-配体相互作用。PI3K/AKT/mTOR 通路在乳腺癌中普遍高度活化, 是维持细胞生长、代谢和耐药性的关键环节, 其失调与预后不良密切相关[6] [7]。IL-6/JAK-STAT3 通路则与炎症驱动和免疫逃逸相关, 其持续激活不仅促进肿瘤细胞增殖, 还削弱抗肿瘤免疫反应[8]。TGF- β 信号早期具有抑癌作用, 但在晚期则通过诱导 EMT 增强细胞迁移和转移能力[9] [10]。同时, TNF α /NF- κ B 通路维持了促炎微环境和抗凋亡状态, 为 TNBC 细胞提供持续的生存优势[11]。这些结果支持 TNBC 高侵袭性和耐药性与多通路异常有关, 也为后续探索多通路联合干预提供了参考。

在核心基因层面, CHST11 的高表达具有提示意义。CHST11 参与硫酸软骨素的合成和修饰, 从而影响细胞外基质的稳定性以及细胞表面糖胺聚糖的构象[12]-[14]。蛋白聚糖可作为肿瘤细胞表面及细胞微环境中的信号调节分子, 其与配体、受体及基质蛋白的相互作用可影响肿瘤生长、血管生成和微环境重塑。已有研究提示, CHST11 在多种高侵袭性肿瘤中呈上调趋势, 其表达升高与肿瘤细胞迁移和转移潜能增强有关[12] [15]。其潜在作用可通过多条途径体现。首先, CHST11 介导的糖胺聚糖修饰可促进 ROR1 受体活化, 并进一步影响 JNK 信号通路及细胞侵袭能力[16]。并且基底样乳腺癌模型研究还提示, 硫酸软骨素通过 N-cadherin/ β -catenin 相关信号影响细胞外基质黏附和侵袭表型。其次, CHST11 调节的硫酸软骨素与血小板和内皮细胞的 P-选择素结合, 从而为循环肿瘤细胞黏附及免疫监视逃逸提供条件[17] [18]。进一步的单细胞表达和虚拟敲除分析从细胞异质性和调控网络层面支持了 CHST11 在 TNBC 微环境中的潜在作用。单细胞结果提示, CHST11 不仅在免疫细胞中表达, 也分布于成纤维细胞和内皮细胞等基质及血管相关细胞群, 提示其可能参与肿瘤细胞与免疫、基质和血管内皮成分之间的相互作用。虚拟敲除后的扰动方向与髓系细胞活化、抗原加工呈递、白细胞跨内皮迁移和血管通透性调节等过程相吻合, 提示 CHST11 可能通过免疫-血管-基质互作影响 TNBC 微环境, 但该推测仍需进一步体内外实验验证。结合本研究结果推测, CHST11 可能通过影响炎症信号维持、免疫逃逸、细胞黏附迁移、ECM 重塑和血管通透性变化, 参与 TNBC 肿瘤微环境重塑和侵袭进展。

在药物干预探索方面, 本研究重点评估了天然化合物槲皮素。作为一种具有多重药理活性的黄酮类物质, 槲皮素在多项研究中被报道具有抗肿瘤潜能[19]。其机制包括抑制 PI3K/AKT/mTOR 通路以限制乳腺癌细胞的增殖并诱导凋亡[20] [21], 以及降低 IL-6/JAK-STAT3 通路活性以缓解炎症环境并增强抗肿瘤免疫[22]。此外, 槲皮素能够抑制 Wnt/ β -catenin 信号通路并下调 BCL2 蛋白表达, 从而削弱肿瘤细胞增殖并促进凋亡[23]。在本研究中, 分子对接结果提示槲皮素与 CHST11 的结合能为 -8.5 kcal/mol, 关键残基的氢键与疏水相互作用共同维持了复合物的稳定。分子动力学模拟进一步支持这一结合在长时间尺度下处于低能态并保持稳定构象。这些结果提示, 槲皮素可能通过结合 CHST11 编码的蛋白影响相关信号网络, 从而在 TNBC 中产生潜在抑制效应。但分子模拟结果仅提供结构层面的预测依据, 仍需细胞内靶点结合、酶活性变化和药效学实验进一步验证。

综上, 本研究初步提示 TNBC 存在多通路异常富集模式, 且 CHST11 与肿瘤进展存在潜在关联, 并提出槲皮素可能作为潜在治疗分子。研究框架为理解 TNBC 分子机制和开发新型靶向策略提供了参考。然而, 本研究仍存在局限性。首先, 分析主要依赖于公开数据库和计算模拟, 缺乏细胞及动物实验的验证。其次, TNBC 具有高度异质性, CHST11 的作用模式是否在不同亚型中保持一致仍需进一步探讨。最后, 槲皮素在体内的低生物利用度可能限制其临床应用。未来研究应结合体外与体内实验进一步检验 CHST11 的潜在功能及槲皮素的抑制作用, 并探索改进药物递送方式以提高其生物利用度。总体而言,

CHST11 可能成为 TNBC 候选预后标志物和潜在治疗靶点, 而槲皮素则为天然产物在 TNBC 治疗中的应用提供了潜在方向。

5. 结论

本研究通过整合生物信息学分析筛选鉴定了 TNBC 潜在关键基因 CHST11, 提示其可能参与肿瘤进展及免疫微环境调控, 并为 TNBC 分子机制研究和潜在治疗策略探索提供理论依据。

参考文献

- [1] Almansour, N.M. (2022) Triple-Negative Breast Cancer: A Brief Review about Epidemiology, Risk Factors, Signaling Pathways, Treatment and Role of Artificial Intelligence. *Frontiers in Molecular Biosciences*, **9**, Article ID: 836417. <https://doi.org/10.3389/fmolb.2022.836417>
- [2] Jie, H., Ma, W. and Huang, C. (2025) Diagnosis, Prognosis, and Treatment of Triple-Negative Breast Cancer: A Review. *Breast Cancer: Targets and Therapy*, **17**, 265-274. <https://doi.org/10.2147/bctt.s516542>
- [3] Chen, Z., Liu, Y., Lyu, M., Chan, C.H., Sun, M., Yang, X., *et al.* (2025) Classifications of Triple-Negative Breast Cancer: Insights and Current Therapeutic Approaches. *Cell & Bioscience*, **15**, Article No. 13. <https://doi.org/10.1186/s13578-025-01359-0>
- [4] Marra, A. and Curigliano, G. (2021) Adjuvant and Neoadjuvant Treatment of Triple-Negative Breast Cancer with Chemotherapy. *The Cancer Journal*, **27**, 41-49. <https://doi.org/10.1097/ppo.0000000000000498>
- [5] Corti, C., Koca, B., Rahman, T., Mittendorf, E. and Tolane, S. (2025) Recent Advances in Immune Checkpoint Inhibitors for Triple-Negative Breast Cancer. *ImmunoTargets and Therapy*, **14**, 339-357. <https://doi.org/10.2147/itt.s495751>
- [6] Garg, P., Ramisetty, S., Nair, M., Kulkarni, P., Horne, D., Salgia, R., *et al.* (2025) Strategic Advancements in Targeting the PI3K/AKT/mTOR Pathway for Breast Cancer Therapy. *Biochemical Pharmacology*, **236**, Article ID: 116850. <https://doi.org/10.1016/j.bcp.2025.116850>
- [7] Zhu, K., Wu, Y., He, P., Fan, Y., Zhong, X., Zheng, H., *et al.* (2022) PI3K/AKT/mTOR-Targeted Therapy for Breast Cancer. *Cells*, **11**, Article No. 2508. <https://doi.org/10.3390/cells11162508>
- [8] Manore, S.G., Doheny, D.L., Wong, G.L. and Lo, H.W. (2022) IL-6/JAK/STAT3 Signaling in Breast Cancer Metastasis: Biology and Treatment. *Frontiers in Oncology*, **12**, Article ID: 866014. <https://doi.org/10.3389/fonc.2022.866014>
- [9] Hao, Y., Baker, D. and ten Dijke, P. (2019) TGF- β -Mediated Epithelial-Mesenchymal Transition and Cancer Metastasis. *International Journal of Molecular Sciences*, **20**, Article No. 2767. <https://doi.org/10.3390/ijms20112767>
- [10] Sheikh, K.A., Amjad, M., Irfan, M., Anjum, S., Majeed, T., Riaz, M., *et al.* (2025) Exploring TGF- β Signaling in Cancer Progression: Prospects and Therapeutic Strategies. *OncoTargets and Therapy*, **18**, 233-262. <https://doi.org/10.2147/ott.s493643>
- [11] Shinde, A., Chandak, N., Singh, J., Roy, M., Mane, M., Tang, X., *et al.* (2024) TNF- α induced NF- κ B Mediated LYRM7 Expression Modulates the Tumor Growth and Metastatic Ability in Breast Cancer. *Free Radical Biology and Medicine*, **211**, 158-170. <https://doi.org/10.1016/j.freeradbiomed.2023.12.018>
- [12] Chang, W.M., Li, L.J., Chiu, I.A., Lai, T., Chang, Y., Tsai, H., *et al.* (2022) The Aberrant Cancer Metabolic Gene Carbohydrate Sulfotransferase 11 Promotes Non-Small Cell Lung Cancer Cell Metastasis via Dysregulation of Ceruloplasmin and Intracellular Iron Balance. *Translational Oncology*, **25**, Article ID: 101508. <https://doi.org/10.1016/j.tranon.2022.101508>
- [13] Li, C.H., Chan, M.H., Chang, Y.C. and Hsiao, M. (2022) The CHST11 Gene Is Linked to Lung Cancer and Pulmonary Fibrosis. *The Journal of Gene Medicine*, **24**, e3451. <https://doi.org/10.1002/jgm.3451>
- [14] Lin, Y.C., Chu, Y.H., Liao, W.C., *et al.* (2023) CHST11-Modified Chondroitin 4-Sulfate as a Potential Therapeutic Target for Glioblastoma. *American Journal of Cancer Research*, **13**, 2998-3012.
- [15] Behrens, A., Jousheghany, F., Yao-Borengasser, A., Siegel, E.R., Kieber-Emmons, T. and Monzavi-Karbassi, B. (2020) Carbohydrate (Chondroitin 4) Sulfotransferase-11-Mediated Induction of Epithelial-Mesenchymal Transition and Generation of Cancer Stem Cells. *Pharmacology*, **105**, 246-259. <https://doi.org/10.1159/000506710>
- [16] Nadanaka, S., Tamura, J. and Kitagawa, H. (2022) Chondroitin Sulfates Control Invasiveness of the Basal-Like Breast Cancer Cell Line MDA-MB-231 through ROR1. *Frontiers in Oncology*, **12**, Article ID: 914838. <https://doi.org/10.3389/fonc.2022.914838>
- [17] Cooney, C.A., Jousheghany, F., Yao-Borengasser, A., Phanavanh, B., Gomes, T., Kieber-Emmons, A.M., *et al.* (2011)

- Chondroitin Sulfates Play a Major Role in Breast Cancer Metastasis: A Role for CSPG4 and CHST11 Gene Expression in Forming Surface P-Selectin Ligands in Aggressive Breast Cancer Cells. *Breast Cancer Research*, **13**, R58. <https://doi.org/10.1186/bcr2895>
- [18] Xiong, D.D., Li, J.D., He, R.Q., Li, M., Pan, Y., He, X., *et al.* (2023) Highly Expressed Carbohydrate Sulfotransferase 11 Correlates with Unfavorable Prognosis and Immune Evasion of Hepatocellular Carcinoma. *Cancer Medicine*, **12**, 4938-4950. <https://doi.org/10.1002/cam4.5186>
- [19] Tang, S.M., Deng, X.T., Zhou, J., Li, Q., Ge, X. and Miao, L. (2020) Pharmacological Basis and New Insights of Quercetin Action in Respect to Its Anti-Cancer Effects. *Biomedicine & Pharmacotherapy*, **121**, Article ID: 109604. <https://doi.org/10.1016/j.biopha.2019.109604>
- [20] Jiang, J., Yang, Y., Wang, F., Mao, W., Wang, Z. and Liu, Z. (2024) Quercetin Inhibits Breast Cancer Cell Proliferation and Survival by Targeting Akt/mTOR/PTEN Signaling Pathway. *Chemical Biology & Drug Design*, **103**, e14557. <https://doi.org/10.1111/cbdd.14557>
- [21] Li, X., Zhou, N., Wang, J., Liu, Z., Wang, X., Zhang, Q., *et al.* (2018) Quercetin Suppresses Breast Cancer Stem Cells (CD44(+)/CD24(-)) by Inhibiting the PI3K/Akt/mTOR-Signaling Pathway. *Life Sciences*, **196**, 56-62. <https://doi.org/10.1016/j.lfs.2018.01.014>
- [22] Liao, Y., Xie, X., Zhang, C., Zhong, H., Shan, L., Yu, P., *et al.* (2024) Quercetin Exerts Anti-Tumor Immune Mechanism by Regulating IL-6/JAK2/STAT3 Signaling Pathway to Deplete Treg Cells. *Toxicon*, **243**, Article ID: 107747. <https://doi.org/10.1016/j.toxicon.2024.107747>
- [23] Srivastava, N.S. and Srivastava, R.A.K. (2019) Curcumin and Quercetin Synergistically Inhibit Cancer Cell Proliferation in Multiple Cancer Cells and Modulate Wnt/ β -Catenin Signaling and Apoptotic Pathways in A375 Cells. *Phytomedicine*, **52**, 117-128. <https://doi.org/10.1016/j.phymed.2018.09.224>