

本文引文格式:王想,方波,袁昊,等.基于转录组测序探究 TGF- $\beta$  对成纤维细胞作用的机制研究[J].右江民族医学院学报,2023,45(2):224-229.

【论著与临床报道】

## 基于转录组测序探究 TGF- $\beta$ 对成纤维细胞作用的机制研究

王想<sup>1,2</sup>,方波<sup>1,3</sup>,袁昊<sup>1,2</sup>,米俊豪<sup>1,3</sup>,谢毓丽<sup>1,2</sup>,莫泳锋<sup>4</sup>,杨日荣<sup>1,2</sup>

(1. 广西医科大学基因组与个体化医学研究中心,广西 南宁 530021;

2. 广西医科大学基础医学院,广西 南宁 530021;

3. 广西医科大学再生医学与医用生物资源开发应用协同创新中心,广西 南宁 530021;

4. 广西医科大学第二附属医院,广西 南宁 530007)

**摘要:**目的 基于转录组 RNA 测序筛选转化生长因子  $\beta$ (transforming growth factor-beta, TGF- $\beta$ ) 刺激的成纤维细胞与正常成纤维细胞的差异 mRNA。方法 对正常人成纤维 MRC-5 细胞和经 20 ng/mL TGF- $\beta$  刺激 24 h 后的 MRC-5 细胞进行转录组 RNA 测序,筛选出两组之间的差异表达基因(DEGs),对 DEGs 进行生物信息学分析,包括基因功能(GO)、京都基因百科全书(KEGG)富集分析、蛋白质相互作用网络分析(PPI)以及相关基因对肿瘤生存预后的影响。**结果** 筛选出在 33 692 个基因中差异表达的 mRNAs 共 161 个,其中上调 39 个,下调 122 个;GO 分析表明,DEGs 主要富集的生物学过程包括核苷酸生物合成、代谢和神经元凋亡、坏死等过程;KEGG 通路富集发现这些信号主要涉及核苷酸生物合成过程,核苷酸代谢过程,神经元凋亡等过程,其中核苷酸生物合成在 6 条信号通路中均被富集到;利用 TCGA 数据库对膀胱癌(BLCA)进行生存分析发现 TGFBI1、HYOU1 和 CAT 基因同时高表达时,与患者较差预后显著相关。**结论** mRNA 在 TGF- $\beta$  刺激后的成纤维细胞和正常成纤维细胞中存在表达差异,筛选出在 BLCA 病人中表达的 TGFBI1、HYOU1 和 CAT,有望作为患者的预后标志物。

**关键词:**成纤维细胞;转化生长因子  $\beta$ ;转录组测序;膀胱肿瘤

中图分类号:R737.14

文献标识码:A

文章编号:1001-5817(2023)02-0224-06

doi:10.3969/j.issn.1001-5817.2023.02.010

### Study on the mechanism of TGF- $\beta$ action on fibroblasts based on transcriptome sequencing

Wang Xiang<sup>1,2</sup>, Fang Bo<sup>1,3</sup>, Yuan Hao<sup>1,2</sup>, Mi Junhao<sup>1,3</sup>,  
Xie Yuli<sup>1,2</sup>, Mo Yongfeng<sup>4</sup>, Yang Rirong<sup>1,2</sup>

(1. Center for Genomics and Personalized Medicine of Guangxi Medical University, Nanning 530021, Guangxi, China; 2. School of Basic Medicine, Guangxi Medical University, Nanning 530021, Guangxi, China; 3. Collaborative Innovation Center of Regenerative Medicine and Medical Bioresource Development and Application of Guangxi Medical University, Nanning 530021, Guangxi, China; 4. The Second Affiliated Hospital of Guangxi Medical University, Nanning 530007, Guangxi, China)

**Abstract:** **Objective** To screen the differential mRNA between transforming growth factor-beta (TGF- $\beta$ ) stimulated fibroblasts and normal fibroblasts based on transcriptome RNA sequencing. **Methods** Transcriptome RNA sequencing was performed on normal human fibroblasts MRC-5 cells and MRC-5 cells stimulated with 20 ng/mL TGF- $\beta$  for 24 hours, and differentially expressed genes (DEGs) were screened between the two groups. Bioinformatics analysis was performed on DEGs, including gene ontology (GO), kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis, protein-protein interaction network analysis (PPI) and

基金项目:国家自然科学基金项目(82260575)

第一作者简介:王想(1996-),女,在读硕士研究生,研究方向:肿瘤免疫,E-mail:wangxiang0312@163.com

通讯作者简介:杨日荣(1982-),男,教授,硕士研究生导师,研究方向:肿瘤免疫,E-mail:yangrirong@sr.gxmu.edu.cn

the effects of related genes on tumor survival and prognosis. **Results** A total of 161 differentially expressed mRNAs were screened out from 33 692 genes, including 39 up-regulated and 122 down-regulated genes; GO analysis showed that biological processes mainly enriched by DEGs included nucleotide biosynthesis, metabolism, neuronal apoptosis and necrosis; KEGG pathway enrichment found that these signals were mainly involved in nucleotide biosynthesis, nucleotide phosphate metabolism and neuronal apoptosis, among which nucleotide biosynthesis was enriched in six signal pathways; survival analysis of bladder cancer (BLCA) by TC-GA database found that simultaneous high expression of TGFBI1, HYOU1, and CAT genes were significantly associated with poor prognosis in patients with BLCA. **Conclusion** mRNAs have differential expression between TGF- $\beta$  stimulated fibroblasts and normal fibroblasts, and the screened TGFBI1, HYOU1 and CAT genes are expected to be used as prognostic markers in patients with BLCA.

**Key words:** fibroblasts; transforming growth factor  $\beta$  (TGF- $\beta$ ); transcriptome sequencing; bladder cancer

转化生长因子  $\beta$  (transforming growth factor-beta, TGF- $\beta$ ) 家族是调节细胞生长和分化的蛋白质家族。TGF- $\beta$  信号紊乱不仅是炎症性疾病的基础,还可以转化正常的成纤维细胞表型,促进肿瘤的发生<sup>[1]</sup>。在肿瘤微环境 (tumor microenvironment, TME) 中,大量的基质细胞、肿瘤细胞、免疫细胞以及这些细胞分泌的细胞因子和趋化因子嵌入其中<sup>[2]</sup>。有研究报道,在胰腺癌、乳腺癌、结直肠癌中,TGF- $\beta$  可以促进正常成纤维细胞向肿瘤相关成纤维细胞 (cancer association fibroblast, CAF) 的转化,与预后不良相关<sup>[3-5]</sup>。CAFs 作为 TME 中基质的主要成分,分泌多种生长因子、炎症因子、基质金属蛋白酶和细胞外基质蛋白,促结缔组织再生反应,导致不良预后<sup>[6-7]</sup>。由于 CAFs 主要来源于成纤维细胞、肌成纤维细胞和内皮细胞,并且肺中有丰富的成纤维细胞,癌细胞常倾向于扩散到肺<sup>[8]</sup>。因此,本课题组评价了 TGF- $\beta$  因子刺激下的人成纤维细胞 MRC-5 的形态变化、差异基因的表达以及参与的主要通过程。

本次研究采用细胞培养、转录组 RNA 测序技术和生物信息学分析方法,筛选出差异表达基因 (differentially expressed genes, DEGs),对 DEGs 的功能和富集的相关通路进行分析,同时构建 DEGs 的蛋白相互作用网络图谱,筛选出受 TGF- $\beta$  调控的关键基因在膀胱癌和正常组织中的差异表达以及生存预后相关性,有望作为膀胱癌患者的预后标志物。

## 1 材料与方法

1.1 材料 人成纤维细胞 MRC-5 购自上海盈湾生物科技有限公司;MEM 培养基和胎牛血清均购自美国 Gibco 公司;人 TGF- $\beta$ 1 重组蛋白购自美国 Pepro-Tech 公司;逆转录酶、建库试剂盒均购于瑞士罗氏生物公司。

### 1.2 方法

1.2.1 细胞刺激实验 常规细胞培养后,取对数生长

期细胞按照每孔  $5 \times 10^4$  个细胞种于 48 孔培养板中,分别设置 Mock 培养基组、20 ng/mL TGF- $\beta$  刺激组进行 24 h 药物反应,拍照记录,收取同批细胞沉淀,放入  $-80^\circ\text{C}$  冰箱冻存。实验重复 3 次后进行全转录组检测。

1.2.2 转录组 RNA 测序 细胞经裂解后使用包含细胞条形码片段 (barcode) 和唯一分子识别符 (UMI) 序列的 25 个核苷酸核苷引物 Oligo(dT) 进行逆转录。在第一链 cDNA 合成后,经过 PCR 扩增,得到互补链 DNA。将新合成的双链 DNA 退火到索引引物,致 3' 末端的 DNA 生物素化。然后,用超声波将合成的 DNA 片段切成 300 bp 的片段,并用链霉亲和素 C1 磁珠孵育连接 3' 端,随后将终产物进行 Qubit 浓度检测,并利用 Agilent 2100 进行进一步质量检测。所有测序数据在 Illumina HiSeq 4000 平台上使用 150 个碱基对的末端测序产生。

1.2.3 差异表达基因分析 利用 R 软件 (v4.0.2) 进行生物信息学分析。利用 limma (v3.44.3) 分析两组间差异基因表达情况,当差异倍数 (fold change) 大于 2,即  $|\log_2\text{FC}| > 1$  且校正  $P < 0.05$  时被认为是差异显著的基因。

1.2.4 功能富集分析与蛋白质相互作用网络 利用 cluster Profiler (v3.16.1) 软件进行生物学过程的注释 (gene ontology, GO) 和京都基因和基因组通路富集分析 (kyoto encyclopedia of genes and genomes, KEGG),并通过蛋白相互作用网络 (protein-protein interaction, PPI) 筛选出关键基因 (<https://string-db.org/>, v11.5, 2023-2-17)。将 PPI 导出图通过 cytoscape (v3.9.0) 软件对结果进行美化和分析,按节点的互作分数设置色阶度进行排序,设置为不同的表现颜色。

1.2.5 组织表达量差异和预后生存分析 在 GEPIA2 在线数据库中分析其中 5 个差异基因在膀胱癌

组织与正常组织间 RNA 表达量的差异,并对膀胱癌中的 3 个基因和基因集的表达进行总体生存(overall survival, OS)分析,  $P < 0.05$  视为具有统计学差异(<http://gepia2.cancer-pku.cn/>)。

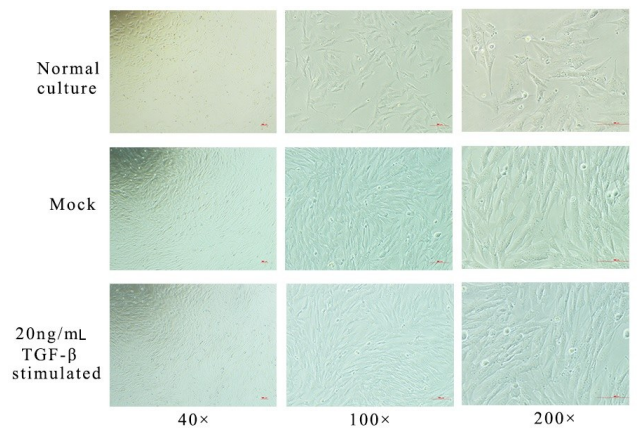
1.3 统计学方法 采用 R(v4.0.2) 软件进行统计分析和作图;采用非参数检验或  $t$  检验分析两组间差异,以矫正  $P$  或  $P < 0.05$  为差异具有统计学意义。

## 2 结果

2.1 正常成纤维细胞经 TGF- $\beta$  的刺激情况 光镜下观察体外培养的人成纤维细胞 MRC-5 为梭形,其形态尚可依细胞的功能变化和物理性状不同而发生改变。经过 24 h 培养后, TGF- $\beta$  处理组细胞较 Mock 组形态细长,生长速度快,但贴壁情况差异不明显(见图 1)。随后,我们收集 3 次重复实验的两组细胞沉淀进行转录组测序。

2.2 转录组文库的成功构建 Qubit 检测样本 cDNA 浓度为  $23.8 \text{ ng}/\mu\text{L}$ , Agilent 2100 检测样本文库,主峰为 354 bp,同时在  $900 \sim 2000 \text{ bp}$ ;  $1000 \sim 5000 \text{ bp}$  占比大于 15%; 300 bp 以下片段占比小于 40%。质量合格,进行上机检测(见图 2)。

2.3 DEGs 分析 通过对两组细胞的转录组测序数据分析,以差异倍数  $\geq 2$  且矫正  $P < 0.05$  绘制火山图和层次聚类图,共筛选出 161 个差异基因,包括 39 个上调基因和 122 个下调基因(见图 3)。



注:图中比例尺为  $100 \mu\text{m}$

图 1 光镜下 MRC-5 细胞刺激前形态及条件培养 24 h 后的细胞刺激形态

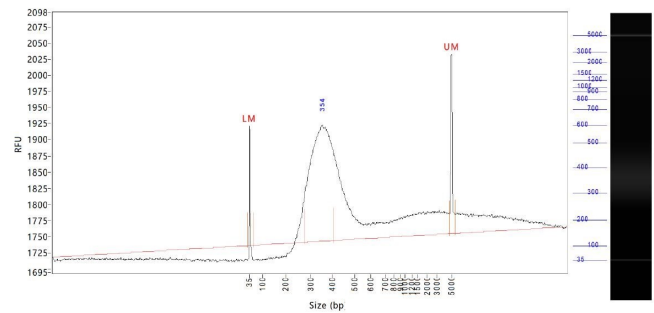
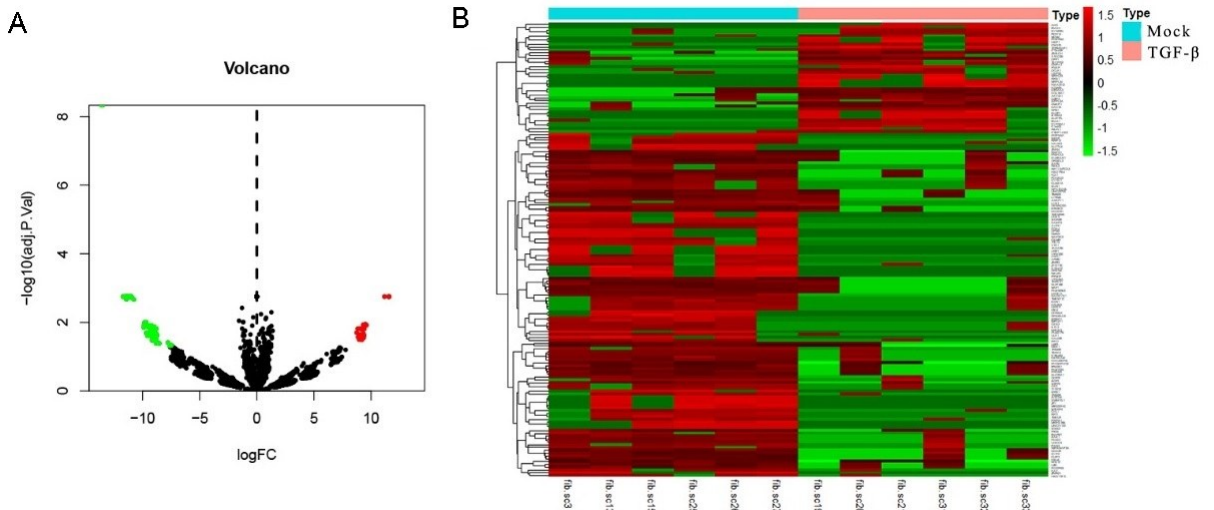


图 2 构建细胞文库质检图谱



注:A. 差异基因火山图(绿色表示下调,红色表示上调),图中基因均为矫正  $P$  有意义的基因;B. 差异表达基因层次聚类热图(红色表示上调,绿色表示下调),颜色越深,表示该基因表达量越高或越低。

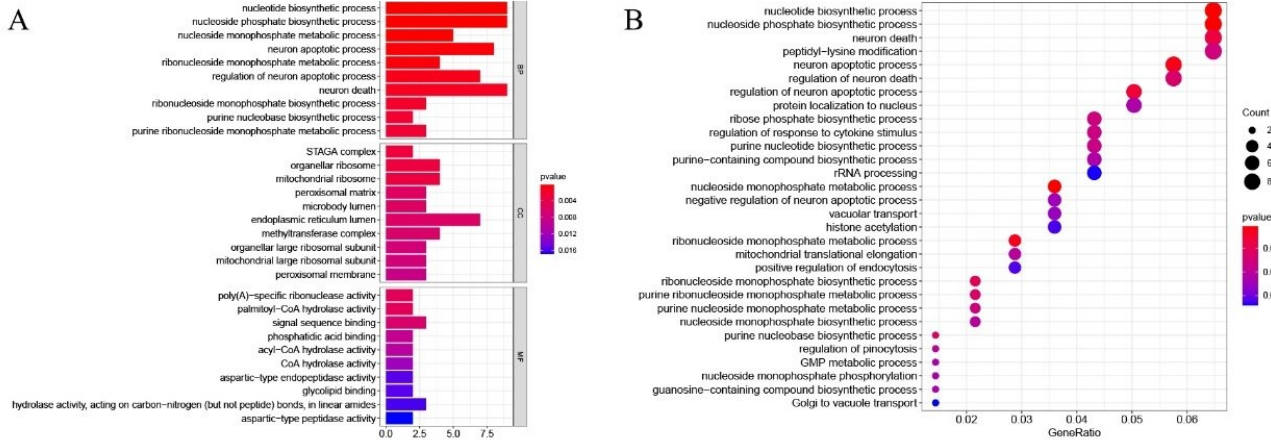
图 3 差异基因火山图和层次聚类图

2.4 GO 功能富集分析 通过 GO 富集分析结果发现,在 GO 生物学过程(biological process, BP)中 TGF- $\beta$  主要影响的是神经元凋亡、坏死过程,核苷酸生物合成过程,嘌呤核糖核苷酸代谢过程,核糖核苷酸生物合成过程和核糖核苷酸代谢过程。富集基

因数量相对较多的细胞成分(cellular component, CC)主要包括内质网、线粒体、核糖体、甲基转移酶复合体、过氧化体基质和酶体、微体管腔以及细胞器、线粒体大核糖体亚基。生物学功能(molecular function, MF)主要是结合信号序列(见图 4A)。

### 2.5 KEGG 通路富集分析 对 DEGs 进行 KEGG 通路富集,发现这些信号主要涉及核苷酸生物合成,核苷

磷酸代谢,神经元凋亡、坏死和肽基-赖氨酸修饰等过程(见图 4B)。



注:A. 差异基因 GO 生物学过程富集分析排名前 10 位的注释情况;B. 差异基因 KEGG 通路富集分析排名前 10 位的通路情况。GeneRatio 表示这些基因在该通路或者 GO 注释条目中富集程度;颜色深浅代表了 P 值大小,颜色越深代表差异越显著;点大小代表了富集在该通路中的基因数目。

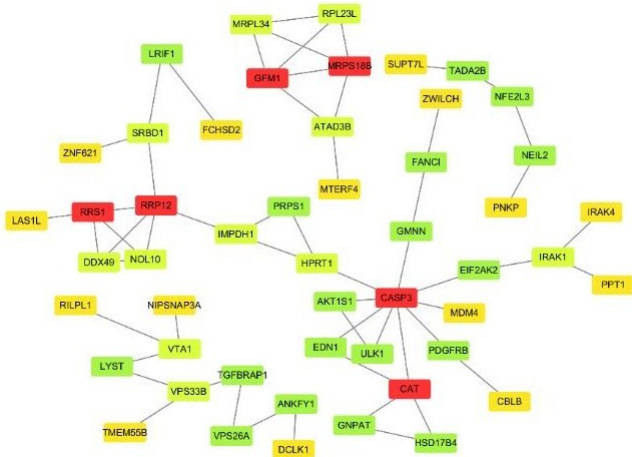
图 4 GO、KEGG 富集结果

### 2.6 蛋白质相互作用网络分析 通过在线数据库 STRING 中的 Multiple Protein 构建蛋白相互作用网络,以 P < 0.05 为筛选条件时,154 个 DEGs 编码的蛋白能够形成复杂的 PPI(见图 5)。而根据相互作用网络中蛋白与蛋白之间连接的节点数,筛选出受 TGF-β 调控的 6 个潜在关键基因 CAT、GFM1、MRPS18B、RRP12、RRS1、CASP3。

组织和正常组织对比,发现 TGFB11I、COL16A1 在 BLCA 组织表达量较低;HYOU1、ZWLCH 基因在 BLCA 组织相对表达上调(见图 6A; P < 0.01)。其中,TGFB11I 和 CAT 高表达时,对 BLCA 患者生存有统计学意义(见图 6B; P < 0.05),并且当 3 个基因同时高表达时,BLCA 患者的生存期较短(见图 6C; P < 0.05),潜在说明 TGF-β 刺激的成纤维细胞群与膀胱癌患者生存预后显著相关,可能是治疗的潜在靶点。

### 3 讨论

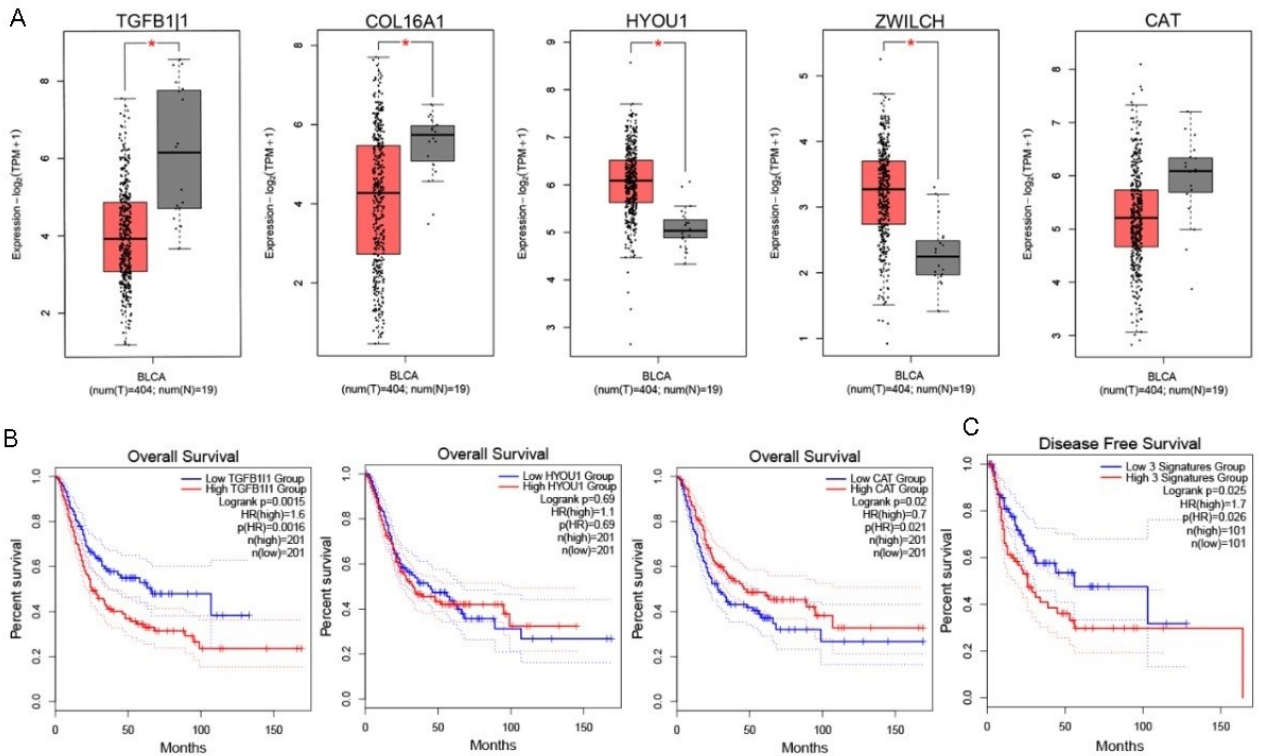
本研究通过对差异表达基因进行 GO 和 KEGG 的富集分析,发现 DEGs 主要富集在核苷酸合成、代谢的信号通路。由于核苷酸在人体内广泛分布,其主要生物学功能应当不容忽视,主要包括:①核苷酸是核酸的基本单位,可以储存能量,主要通过三磷酸核苷酸实现,特别是 ATP,是给予细胞能量的主要形式。另外,一些活化的中间产物,如 UDP-葡萄糖,亦含有核苷酸成分。②核苷酸可以参与生理和代谢调节:cAMP(或 cGMP)是多种细胞膜激素受体的调节作用的第二信使,许多代谢过程也要受到体内 ATP、ADP 或 AMP 水平的调节。③核苷酸可以组成辅酶。如腺苷酸可作为 NAD<sup>+</sup>、ANDP<sup>+</sup>、FMN、FAD 及 CoA 等的组成成分<sup>[11]</sup>。此外,核苷酸还可在转录水平上受一系列主要转录因子的调控,亦在酶水平上受到变构调节和反馈抑制的调节,同时也参与细胞周期中的代谢和调节机制<sup>[12]</sup>。近年来,药物对肿瘤微环境的影响还尚不清楚,非肿瘤细胞在肿瘤免疫微环境中的代谢脆弱性被尤为关注<sup>[13]</sup>。因此,针对癌症核苷酸的合成与代谢的有关成纤维细胞的研究仍值得探索。



注:交互分数越高,颜色标记越红。

图 5 基于差异基因的 PPI 图

### 2.7 TCGA 数据库中差异基因表达和 BLCA 患者预后的关系 根据文献查询,DEGs 中筛选出 TGF-β 调控的潜在关键基因 TGFB11I、HYOU1 与膀胱癌的恶性程度和预后显著相关<sup>[9-10]</sup>。将差异表达分析和 PPI 筛选的潜在关键基因放入 TCGA 数据库进行膀胱癌



注:A. TCGA 数据中 TGFBI1、COL16A1、HYOU1、ZWLCH、CAT 5 个基因在膀胱癌与正常组织中的表达水平差异(红色表示膀胱癌,灰色表示正常);B. GEPIA2 分析 TGFBI1、HYOU1、AT 单基因高低表达对膀胱癌患者生存率的影响;C. TGFBI1、HYOU1、CAT 基因组合高低表达对膀胱癌患者生存率与生存分析关系。

图 6 膀胱癌差异基因的临床意义

本研究通过对 PPI 和生存预后的分析,还发现 TGFBI1、HYOU1 和 CAT 等基因出现在多条富集的信号通路上,与多种生物学机制相关。TGFBI1(也称为 Hic-5)编码一种局部黏附支架蛋白,先前已被证明 TGFBI1 与膀胱癌血管发育、细胞黏附及细胞外基质的形成有关,并与细胞外基质受体途径相关的基因有很强的相关性<sup>[10]</sup>。除此之外,TGFBI1 的表达还与巨噬细胞和 NK 细胞呈正相关。巨噬细胞在上调 VI 型胶原的合成和组装,诱导细胞外基质沉积等方面发挥着重要作用,研究者推测 TGFBI1 还参与了浸润性免疫细胞和细胞外基质沉积之间的相互作用<sup>[14]</sup>。这可能与 TGFBI1 在 BLCA 中都较正常组织更低表达,影响患者预后显著相关联;HYOU1 是一种内质网伴侣蛋白,属于 HSP70 家族,在缺氧、缺糖、还原剂和衣霉素等应激条件下诱导表达<sup>[15]</sup>。HYOU1 抗细胞凋亡作用主要包括:①抑制内质网应激相关的细胞凋亡通路;②HYOU1 和 ER 应激感应器的结合可能使它们保持不活跃,从而抑制 UPR;③激活 Akt 磷酸化作为生成磷脂酰肌醇 3-激酶/Akt 信号转导的一部分<sup>[16]</sup>,膀胱癌中该基因的表达很可能也参与了以上的生物学过程;作为过氧化氢酶的 CAT 基因是人体内一种天然的、很重要的抗氧化酶,在活性氧的清除过程中作为超

氧化物歧化酶的下游抗氧化蛋白,通过催化超氧阴离子自由基的自身氧化还原反应,有效清除自由基,维持体内自由基和氧化还原状态的动态平衡,来避免损伤机体的组织细胞<sup>[17]</sup>。同时,心血管危险因素常常导致由 NADPH 氧化酶、黄嘌呤氧化酶、线粒体电子传输链和功能障碍的内皮型一氧化氮合酶产生的活性氧物种的增加,超过抗氧化防御系统(如过氧化氢酶、超氧化物歧化酶、谷胱甘肽过氧化物酶、血红素加氧酶、对氧磷酶)的能力时,会导致氧化应激,从而促进动脉粥样硬化的形成,目前 CAT 也已成为心血管疾病研究的候选基因<sup>[18]</sup>。除此之外,过氧化氢酶 C262T 的变异与不同癌症之间的联系也被广泛报道,CAT C262T 变异与总体癌症易感性有关,还可能与血液、骨髓、胃肠道、前列腺癌、皮肤癌和妇科癌症有关<sup>[19-20]</sup>。这些数据更加提示了差异基因 TGFBI1、HYOU1 和 CAT 的表达可以调控肿瘤的发生发展,影响癌症患者预后,可能是 BLCA 一个潜在的肿瘤治疗靶点或预后监测指标,但还需多中心的临床数据进一步验证。

#### 参考文献:

- [1] BATLLE E, MASSAGUE J. Transforming growth factor-beta signaling in immunity and cancer[J]. Immunity,

- 2019,50(4):924-940.
- [2] KALLURI R. The biology and function of fibroblasts in cancer[J]. Nat Rev Cancer,2016,16(9):582-598.
- [3] BIFFI G, ONI T E, SPIELMAN B, et al. IL1-induced JAK/STAT signaling is antagonized by TGF $\beta$  to shape CAF heterogeneity in pancreatic ductal adenocarcinoma [J]. Cancer Discov,2019,9(2):282-301.
- [4] YU Y, XIAO C H, TAN L D, et al. Cancer-associated fibroblasts induce epithelial-mesenchymal transition of breast cancer cells through paracrine TGF- $\beta$  signalling[J]. Br J Cancer,2014,110(3):724-732.
- [5] HENRY L R, LEE H O, LEE J S, et al. Clinical implications of fibroblast activation protein in patients with colon cancer[J]. Clin Cancer Res,2007,13(6):1736-1741.
- [6] COX T R. The matrix in cancer[J]. Nat Rev Cancer, 2021,21(4):217-238.
- [7] DEASY S K, EREZ N. A glitch in the matrix: organ-specific matrixesomes in metastatic niches[J]. Trends Cell Biol, 2022,32(2):110-123.
- [8] DING S M, LU J F, EDOO M I A, et al. MRC-5 cancer-associated fibroblasts influence production of cancer stem cell markers and inflammation-associated cell surface molecules, in liver cancer cell lines[J]. Int J Med Sci,2019,16(8):1157-1170.
- [9] RAO S, OYANG L, LIANG J, et al. Biological function of HYOU1 in tumors and other diseases[J]. Onco Targets Ther,2021,14:1727-1735.
- [10] DI Y, CHEN D S, YU W, et al. Bladder cancer stage-associated hub genes revealed by WGCNA co-expression network analysis[J]. Hereditas,2019,156:7.
- [11] VEDULA P, KASHINA A. The makings of the 'actin code': regulation of actin's biological function at the amino acid and nucleotide level[J]. J Cell Sci,2018,131(9): jcs215509.
- [12] MOFFATT B A, ASHIHARA H. Purine and pyrimidine nucleotide synthesis and metabolism [J]. Arabidopsis Book,2002,1:e0018.
- [13] STINE Z E, SCHUG Z T, SALVINO J M, et al. Targeting cancer metabolism in the era of precision oncology [J]. Nat Rev Drug Discov,2022,21(2):141-162.
- [14] HE Y, PAN Z F, SHI Q, et al. Bioinformatics profiling and experimental validation of 4 differentially-expressed LIM genes in the course of colorectal-adenoma-carcinoma [J]. Med Sci Monit,2022,28:e937081.
- [15] PAGARE P P, WANG H Q, WANG X Y, et al. Understanding the role of glucose regulated protein 170 (GRP170) as a nucleotide exchange factor through molecular simulations [J]. J Mol Graph Model, 2018, 85: 160-170.
- [16] KUSACZUK M, CECHOWSKA-PASKO M. Molecular chaperone ORP150 in ER stress-related diseases [J]. Curr Pharm Des,2013,19(15):2807-2818.
- [17] ZHANG Y, ZHENG L J, YUN L, et al. Catalase (CAT) Gene Family in Wheat (*Triticum aestivum* L.): evolution, expression pattern and function analysis [J]. Int J Mol Sci,2022,23(1):542.
- [18] LI H G, HORKE S, FORSTERMANN U. Oxidative stress in vascular disease and its pharmacological prevention [J]. Trends Pharmacol Sci, 2013, 34(6): 313-319.
- [19] JONES G D, BRANDT W S, SHEN R L, et al. A genomic-pathologic annotated risk model to predict recurrence in early-stage lung adenocarcinoma [J]. JAMA Surg, 2021,156(2):e205601.
- [20] SHEN Y C, LI D D, TIAN P W, et al. The catalase C-262T gene polymorphism and cancer risk: a systematic review and meta-analysis [J]. Medicine (Baltimore), 2015,94(13):e679.

收稿日期:2023-02-01;修回日期:2023-03-08