

本文引文格式:苏萌,张钊银,姚金光.基于 GEO 数据库筛选舌鳞状细胞癌顺铂耐药的关键基因及生物信息学分析[J].右江民族医学院学报,2023,45(4):557-562.

【论著与临床报道】

## 基于 GEO 数据库筛选舌鳞状细胞癌顺铂 耐药的关键基因及生物信息学分析

苏萌,张钊银,姚金光

(右江民族医学院口腔医学院,广西 百色 533000)

**摘要:**目的 通过生物信息学方法筛选舌鳞状细胞癌顺铂耐药的关键基因。方法 从 GEO 数据库下载 GSE111585 和 GSE115119 数据集,利用 limma 包进行差异分析,与 PRGs 取交集后获取 DEGs。利用 R 语言对 DEGs 进行 GO 功能富集及 KEGG 通路富集分析。String 数据库构建 PPI 网络,在 Cytoscape 中进行可视化,基于 Degree 拓扑分析算法筛选 Hub 基因。根据 Wilcox.test 法分析 Hub 基因与顺铂耐药性的联系。结果 ①共筛选出 32 个 DEGs。②GO 富集分析显示 DEGs 主要参与刺激反应调节、信号受体结合等功能;KEGG 富集分析显示 DEGs 富集在癌症通路和人 T 细胞白血病病毒 I 型感染信号通路。③筛选出 FN1、SOX2 和 COL1A1 3 个 Hub 基因,其中 COL1A1 具有成为潜在生物靶点的潜力。结论 COL1A1 可能是舌鳞状细胞癌顺铂耐药的潜在作用靶点。

**关键词:**舌肿瘤;生物信息学;顺铂;耐药性;靶点

中图分类号:R739.86

文献标识码:A

文章编号:1001-5817(2023)04-0557-06

doi:10.3969/j.issn.1001-5817.2023.04.003

### Screening the key genes of cisplatin resistance in tongue squamous cell carcinoma by bioinformatics analysis based on GEO database

Su Meng, Zhang Zhaoyin, Yao Jinguang

(School of Stomatology, Youjiang Medical University for Nationalities, Baise 533000, Guangxi, China)

**Abstract:** **Objective** To screen the key genes of cisplatin resistance in tongue squamous cell carcinoma by bioinformatics methods. **Methods** The GSE111585 and GSE115119 datasets were downloaded from the GEO database, differential analysis was performed using the limma package, and the DEGs were acquired after adopting the intersection with PRGs. The GO functional enrichment and KEGG pathway enrichment analysis of DEGs were performed by R language. The PPI network was constructed using the String database, and was visualized in Cytoscape. Hub genes were selected based on the Degree topological analysis algorithm. The relationship between Hub genes and cisplatin resistance was analyzed with the Wilcox.test method. **Results** ①A total of 32 DEGs were screened. ②The GO enrichment analysis showed that the DEGs were main involved in functions such as stimulus response regulation, signal receptor binding. KEGG enrichment analysis showed that the DEGs were enriched in the cancer pathways and type I infection signaling pathway of human T-cell leukemia virus. ③ FN1, SOX2, and COL1A1 were identified as 3 hub genes, with a potential biological target for COL1A1. **Conclusion** COL1A1 may be a potential target for cisplatin resistance in tongue squamous cell

基金项目:国家自然科学基金项目(81660495);广西研究生教育创新计划项目(YCSW2022459)

第一作者简介:苏萌(1996-),男,硕士,研究方向:舌鳞状细胞癌的侵袭及转移分子机制,E-mail:1453789170@qq.com

通讯作者简介:姚金光(1965-),男,博士,教授,博士、硕士研究生导师,研究方向:舌鳞状细胞癌的侵袭及转移分子机制,E-mail:yao7760698@126.com

carcinoma.

**Key words:** tongue squamous cell carcinoma; bioinformatics; cisplatin; drug resistance; target

舌鳞状细胞癌(tongue squamous cell carcinoma, TSCC)是一种舌体鳞状上皮组织来源的口腔恶性肿瘤,也是常见的头颈部恶性肿瘤<sup>[1]</sup>。舌头的前 2/3 被认为是舌鳞状细胞癌发生的主要区域,而舌头的后 1/3 为舌根癌的范畴<sup>[2]</sup>。流行病学资料显示,舌鳞状细胞癌约占口腔癌的 30%~50%和全身恶性肿瘤的 0.8%~1.5%<sup>[3-4]</sup>。舌体组织的淋巴管、血管丰富且机械运动频繁,出现区域淋巴结转移的可能性很高,即使在早期阶段也是如此,导致临床上治疗困难。

顺铂的联合化疗是包括舌鳞状细胞癌在内的多种癌症的标准治疗方法,在舌鳞状细胞癌的治疗中扮演极其重要的角色。然而,顺铂通常受到机体天然和获得性耐药的限制,许多舌鳞状细胞癌患者对顺铂固有抗性或经多次用药后对顺铂产生抗性,甚至对未使用过的化疗药物也产生抗性,从而严重限制了化疗效果,术后 5 年生存率仅为 50%<sup>[5-7]</sup>。目前,临床上仍缺乏针对舌鳞状细胞癌顺铂耐药的关键靶点,导致无法从根本上逆转其耐药性。

随着芯片技术的发展,学者们从专注于少数经典耐药基因的研究转向全基因组测序来挖掘特异的耐药基因,以更准确地评估机体基因调控耐药变化的过程。本研究基于 GEO 数据库中舌鳞状细胞癌顺铂耐药细胞的芯片数据,利用生物信息学分析方法筛选舌鳞状细胞癌顺铂耐药过程中的关键基因,探讨其与顺铂耐药性的联系,为揭示顺铂耐药的分子机制及逆转顺铂耐药性提供理论依据。

## 1 材料与方法

1.1 数据来源 登录 Gene Expression Omnibus 数据库(<https://www.ncbi.nlm.nih.gov/geo/>, GEO),以“Tongue Squamous Cell Carcinoma”和“Ciplatin Resistance”作为关键词,下载数据集 GSE111585 和 GSE115119 的表达数据及 GPL 平台文件。GSE111585 数据集基于 GPL14715 平台测序,共包括 3 例正常 SCC9 细胞样本和 3 例 SCC9 顺铂耐药细胞样本。GSE115119 数据集基于 GPL16955 平台测序,共包括 2 例正常 CAL27 细胞样本和 2 例 CAL27 顺铂耐药细胞样本。SCC9 与 CAL27 均为人舌鳞状细胞癌细胞系。根据 GPL 平台文件对探针 ID 进行注释转换,以获取对应的基因名称。从铂耐药基因数据库(<http://ptcr-ddr.cptac-data-view.org>)中下载 909 个与铂耐药密切相关的基因(platinum resistant genes, PRGs),该数据库包含近 30 年来 800 篇出版物关于铂耐药的研究,并对涉及铂耐药的 909 个基因进行高度

整理和注释<sup>[8]</sup>。

1.2 差异基因的筛选 用 R 语言 limma 包分别对 GSE111585 和 GSE115119 数据集进行差异分析<sup>[9]</sup>,以  $P < 0.05$  且  $|\log_2FC| \geq 2$  作为标准筛选差异表达基因(differentially expressed genes, DEGs),并绘制火山图。将两组数据集的 DEGs 与 PRGs 取交集,以降低筛选的假阳性,并绘制韦恩图。

1.3 GO 和 KEGG 富集分析 运用 R 语言对 DEGs 进行功能注释,以明确其参与的生物学功能及信号通路,并进行可视化,设置  $\text{Count} > 5$ ,  $P < 0.05$  及  $\text{FDR} < 0.1$ 。其中基因本体论(gene ontology, GO)主要包括生物学过程(biological processes, BP)、细胞组分(cellular components, CC)及分子功能(molecular functions, MF)。京都基因和基因组百科全书(kyoto encyclopedia of genes and genomes, KEGG)主要包括信号通路。

1.4 PPI 网络的构建及网络分析 String 数据库(<https://stringdb.org/>)可用于预测 DEGs 编码蛋白的相互作用关系<sup>[10]</sup>。将 DEGs 上传至 String 数据库,以互作分数  $\text{Score} > 0.400$  (中等置信度)为阈值构建蛋白互作网络(protein-protein interaction, PPI),去除游离的靶点,简化 PPI 网络。下载 PPI 网络的 TSV 文件,将其导入 Cytoscape\_v3.8.0 软件进行可视化,其中节点(Node)代表节点基因,边(Edge)代表基因之间的相互作用情况。基于 CytoHubba 插件中 Degree 拓扑分析算法对 PPI 网络进行分析,筛选出评分前 3 的基因作为 Hub 基因。

1.5 Hub 基因与顺铂耐药性分析 从 Cellminer 数据库(<https://discover.nci.nih.gov/cellminer/home.do>)中下载药物敏感性数据<sup>[11]</sup>,根据 Hub 基因表达量中位数分为高、低表达组,用 wilcox.test 计算两组顺铂敏感性 IC50 的表达差异,绘制箱线图,  $P < 0.05$  时差异有统计学意义。

## 2 结果

2.1 差异表达基因的鉴定 根据  $|\log_2FC| \geq 2$ ,  $P < 0.05$  标准筛选, GSE111585 数据集共筛选出 458 个 DEGs,其中上调基因 235 个,下调基因 223 个(见图 1A); GSE115119 数据集共筛选出 677 个 DEGs,其中上调基因 377 个,下调基因 300 个(见图 1B)。韦恩图交叉分析筛选出 32 个与顺铂耐药密切相关的 DEGs (见图 1C)。

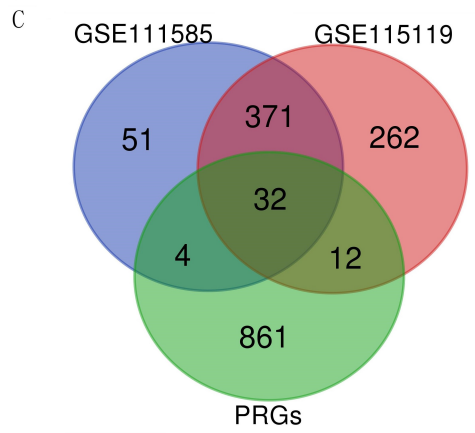
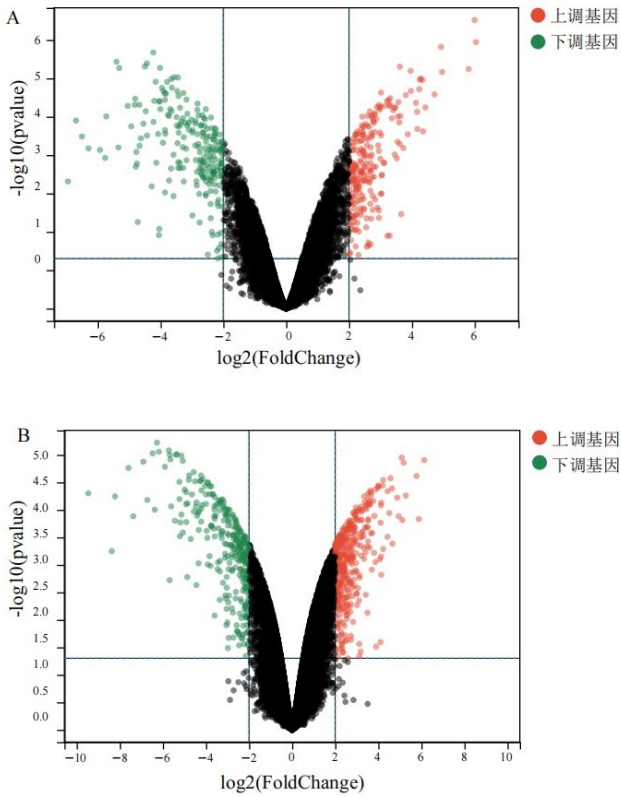


图 1 差异表达基因的火山图和韦恩图

2.2 GO 及 KEGG 富集分析 GO 功能富集分析显示,32 个 DEGs 主要富集在刺激反应调节、多细胞生物过程的调节及组织发育等生物学过程(见图 2A);细胞组成变化主要涉及细胞外区域、内膜系统及胞外区部分等(见图 2B);分子功能变化主要包括信号受体结合、酶结合、DNA 结合转录激活剂活性等(见图 2C)。KEGG 通路富集分析显示,DEGs 主要富集在癌症通路、人 T 细胞白血病病毒 I 型感染、人乳头瘤病毒感染及糖尿病并发症中的 AGE-RAGE 信号通路等(见图 2D)。

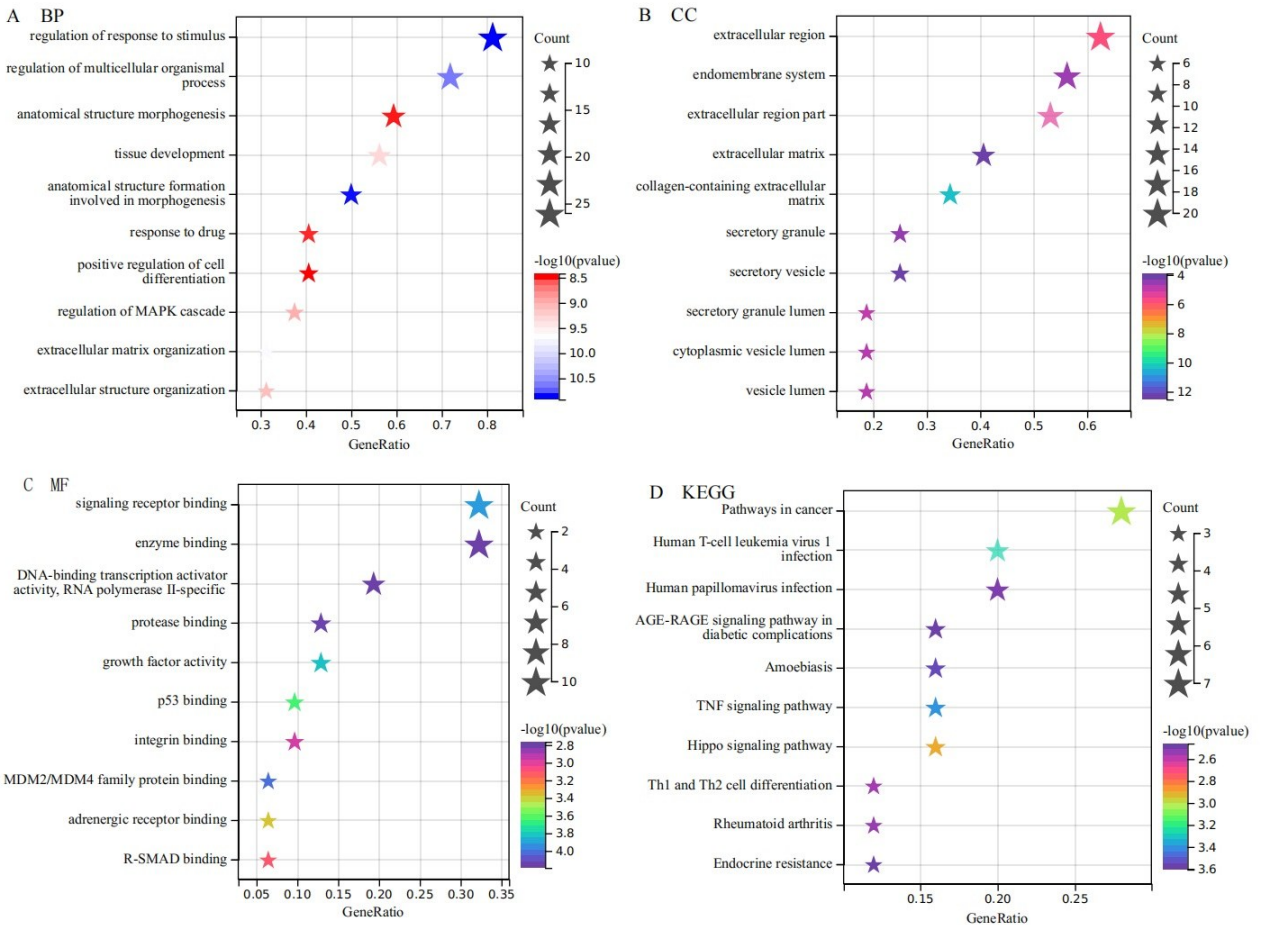


图 2 差异表达基因的 GO 及 KEGG 富集分析

2.3 PPI网络的构建及 Hub 基因的识别 将 32 个显著性 DEGs 上传至 STRING 数据库,去除游离节点,构建 32 个节点,60 条边数的 PPI 网络。使用 CytoHubba 插件 Degree 算法对 PPI 网络进行分析,评分排名前 3 的基因中纤维连接蛋白基因(fibronectin

1, FN1)、性别决定区 Y 框蛋白 2(sexdetermining region Y-Box2, SOX2)和 I 型胶原 α1 亚基因(collagen type I alpha 1, COL1A1)被鉴定为舌鳞状细胞癌顺铂耐药的关键基因即 Hub 基因,三者均为上调基因,见图 3。

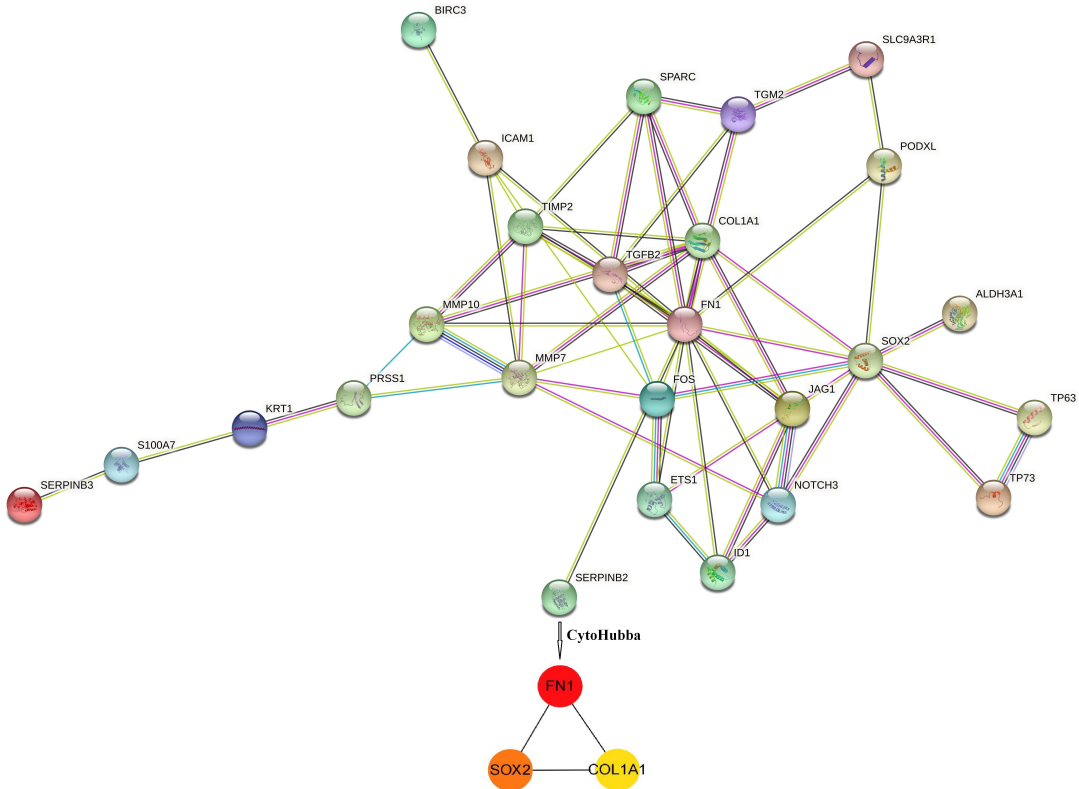


图 3 蛋白互作网络及 Hub 基因

2.4 Hub 基因与顺铂 IC50 的表达差异 根据  $P < 0.05$  为筛选阈值,发现其中 COL1A1 具有较大的研究意义。研究显示与 COL1A1 低表达组相比,高表达组的顺铂敏感性 IC50 较高,说明基因表达量越高,其对顺铂的抗药性越强。见图 4。

### 3 讨论

顺铂是舌鳞状细胞癌化疗的首选药物,但耐药性严重限制了其临床应用。尽管很多学者已经致力于顺铂抗性机制的研究,但是仍然缺乏逆转顺铂抗性和改善对基于顺铂的疗法响应的有效靶标<sup>[12]</sup>。因此,确定顺铂耐药变化过程中起主导作用的关键基因,运用新的分子靶向治疗方法来克服顺铂耐药性或针对性开发安全有效的药物靶标来逆转耐药性对舌鳞状细胞癌的治疗至关重要。本研究通过生物信息学方法筛选 GSE111585 和 GSE115119 数据集中 DEGs,但芯片数据中细胞样本量较少,为此引入了铂耐药基因数据库中的 PRGs,以排除筛选基因时出现的假阳性。经过多数据集及多基因的整合,共筛选出 32 个显著表达的 DEGs,其在正常人舌鳞癌细胞系和耐药细胞系中的表达存在显著差异,说明该基因与铂耐药性存密切的联系。GO 功能富集分析显示,DEGs 主要参与刺激反应调节及多细胞生物过程的调节等生物学过程,细胞组成变化主要在细胞外区域、内膜系统及胞外区等区域,与信号受体结合、酶结合及 DNA 结合转录激活剂活

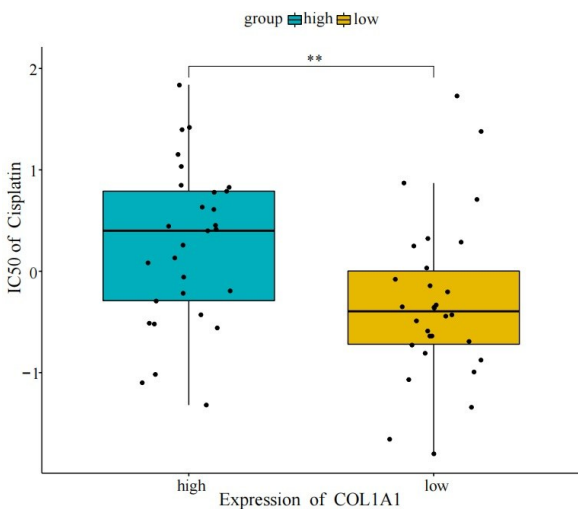


图 4 COL1A1 表达量与顺铂 IC50 的表达差异

性等分子功能相关。KEGG 通路富集分析显示,DEGs 主要富集在癌症通路、人 T 细胞白血病病毒 I 型感染等信号通路。GO 和 KEGG 富集分析提供了 DEGs 在顺铂耐药性中可能参与的功能及发挥作用的相关途径,将众多基因的表型与基因功能构建起联系,以明确研究的范围,为顺铂耐药性研究提供新的思路。

根据 CytoHubba 中 Degree 拓扑分析算法筛选出前 3 个候选 Hub 基因,即 FN1、SOX2 和 COL1A1,这些在舌鳞状细胞癌顺铂耐药细胞样本中均呈现出显著高表达,具有成为顺铂耐药关键靶点的潜力,存在极大的研究价值。FN1 是一种纤维连接蛋白,为 FN 蛋白家族的成员,主要位于胞外区的细胞外基质中。研究表明 FN1 与癌症的顺铂耐药过程密切相关,且 FN1 在顺铂耐药的卵巢癌中表达显著上调,但是其分子调控机制尚不明确<sup>[13]</sup>。目前,有学者在顺铂耐药的神经母细胞瘤中观察到上皮间质转化(epithelial-to-mesenchymal transitions, EMT)的表型,而 KURIMOTO R、CHENG M M 等<sup>[14-15]</sup>也发现 EMT 标志物在顺铂化疗耐药的 NSCLC 患者组织中表达显著上调,说明上皮间质转化途径可以介导顺铂耐药。FN1 是上皮间质转化的公认标志物,多项功能研究已将其与 EMT 的生物学行为联系起来<sup>[16]</sup>。因此,推测 FN1 的高表达可能改变舌鳞状细胞癌的 EMT 表型,介导恶性上皮细胞向间质细胞转化,从而促进耐药性改变。SOX2 位于 3 号染色体的长臂中,位于 3q26.3-27 区域,属于 SOX 基因家族,主要参与肿瘤细胞的增殖、分化、侵袭、转移、复发等过程<sup>[17-18]</sup>。作为肿瘤干细胞重要的标志物,SOX2 在介导耐药方面可能与肿瘤干细胞相关的耐药特征存在密切联系,被认为有助于耐药性的早期阶段。研究表明增强 SOX2 启动子的表达作用,可以促进 SOX2 表达并导致肿瘤生长和耐药性,提示其在改变耐药性方面存在重要作用<sup>[19]</sup>。COL1A1 是胶原蛋白家族的成员,主要参与细胞增殖、迁移及血管生成,而且 COL1A1 被认为与细胞外基质(extracellular matrix, ECM)的变化密切相关,特别是能增强 EMC 的重塑作用<sup>[20-21]</sup>。EMC 是分布在细胞表面多糖和蛋白,可以保护生物细胞免受外部环境的影响,其屏障作用有助于增加肿瘤细胞对抗微生物剂的耐药性<sup>[22]</sup>。推测 COL1A1 基因表达的改变可以导致异常的基质活化,引起强烈的细胞外基质重塑,出现过度的 ECM 积累,从而增加顺铂抗性。

本课题根据 Hub 基因表达量的中位数分为高、低表达组,用 wilcox.test 计算两组顺铂敏感性 IC50 的表达差异,期望从 3 个 Hub 基因中筛选出更具有预测价值的目标基因,发现 COL1A1 基因表达量与顺铂敏感性 IC50 存在显著正相关,而且其表达量越高,顺铂

的耐药性也就越强。有报道显示,冯晓杰等<sup>[22]</sup>通过干扰 COL1A1 基因表达成功逆转 A2780-cis/IGROV1-cis 细胞的耐药性,充分证明通过抑制 COL1A1 表达能增强肿瘤化疗的敏感性,可作为个性化逆转耐药的重要靶标,从而提高化疗效果。本研究与上述文献报道基本一致,表明该研究具有较大的前瞻性及其可信度。本研究利用基因表达谱技术初步证明 FN1、SOX2 和 COL1A1 具有作为舌鳞状细胞癌顺铂耐药潜在生物标志物的潜力,其中 COL1A1 最具研究价值,针对其开发新治疗药物或挖掘现有药物的有效成分,有望成为逆转顺铂耐药的关键。

综上所述,本研究通过生物信息学方法筛选出 32 个 DEGs,其中 FN1、SOX2 和 COL1A1 作为 Hub 基因,且顺铂耐药密切相关。此外,本课题组首次预测出 Hub 基因与顺铂敏感性 IC50 的联系,其中 COL1A1 基因具有极大的研究价值,可以作为逆转舌鳞状细胞癌顺铂耐药的潜在作用靶点。本研究为舌鳞状细胞癌顺铂耐药分子机制及开发敏感药物靶标提供强有力的理论依据,但仍需更多实验证据进一步证实。

#### 参考文献:

- [1] LI Y, ZHAO Z Y, LIU X X, et al. Nomograms to estimate long-term overall survival and tongue cancer-specific survival of patients with tongue squamous cell carcinoma[J]. Cancer Med, 2017, 6(5):1002-1013.
- [2] XU X L, LIU H, ZHANG Y, et al. SPP1 and FN1 are significant gene biomarkers of tongue squamous cell carcinoma[J]. Oncol Lett, 2021, 22(4):713.
- [3] HUANG Z X, ZHANG Q Y, WANG Y, et al. Inhibition of caspase-3-mediated GSDME-derived pyroptosis aids in noncancerous tissue protection of squamous cell carcinoma patients during cisplatin-based chemotherapy[J]. Am J Cancer Res, 2020, 10(12):4287-4307.
- [4] ZHU Y, ZHOU C M, HE Q X. Radiation therapy's efficacy on tongue cancer: a population-based survival analysis[J]. Onco Targets Ther, 2018, 11:7271-7276.
- [5] NIE C J, QIN X H, LI X Y, et al. CACNA2D3 enhances the chemosensitivity of esophageal squamous cell carcinoma to cisplatin via inducing Ca<sup>2+</sup>-mediated apoptosis and suppressing PI3K/Akt pathways[J]. Front Oncol, 2019, 9:185.
- [6] JIN L T, CHUN J, PAN C Y, et al. MAST1 drives cisplatin resistance in human cancers by rewiring cRaf-independent MEK activation[J]. Cancer Cell, 2018, 34(2):315-330.
- [7] 苏萌, 凌小芳, 张钊银, 等. 舌鳞状细胞癌关键基因的筛选及其潜在治疗药物的预测[J/OL]. 医学研究杂志, 1-9 [2023-03-22]. <http://kns.cnki.net/kcms/detail/11>.

5453. R. 20220905. 1301. 002. html.
- [8] HUANG D Q, SAVAGE S R, CALINAWAN A P, et al. A highly annotated database of genes associated with platinum resistance in cancer [J]. *Oncogene*, 2021, 40 (46): 6395-6405.
- [9] 李莎莎, 王琦, 钟斌, 等. 基于生物信息学筛选阿尔茨海默病核心基因及其功能分析 [J]. *右江民族医学院学报*, 2021, 43(6): 746-750, 756.
- [10] GE S X, JUNG D M, YAO R N, Shiny GO; a graphical gene-set enrichment tool for animals and plants [J]. *Bioinformatics*, 2020, 36(8): 2628-2629.
- [11] LUNA A, ELLOUMI F, VARMA S, et al. Cellminer cross-database (cellminer CDB) version 1. 2: exploration of patient-derived cancer cell line pharmacogenomics [J]. *Nucleic Acids Res*, 2021, 49(D1): D1083-D1093.
- [12] HOU S C, JIN W G, XIAO W M, et al. Integrin alpha5 promotes migration and cisplatin resistance in esophageal squamous cell carcinoma cells [J]. *Am J Cancer Res*, 2019, 9(12): 2774-2788.
- [13] WANG J T, DENG L, HUANG J P, et al. High expression of fibronectin 1 suppresses apoptosis through the NF-kappaB pathway and is associated with migration in nasopharyngeal carcinoma [J]. *Am J Transl Res*, 2017, 9 (10): 4502-4511.
- [14] KURIMOTO R, IWASAWA S, EBATA T, et al. Drug resistance originating from a TGF-beta/FGF-2-driven epithelial-to-mesenchymal transition and its reversion in human lung adenocarcinoma cell lines harboring an EGFR mutation [J]. *Int J Oncol*, 2016, 48(5): 1825-1836.
- [15] CHENG M M, JIANG Y N, YANG H, et al. FLNA promotes chemoresistance of colorectal cancer through inducing epithelial-mesenchymal transition and smad2 signaling pathway [J]. *Am J Cancer Res*, 2020, 10 (2): 403-423.
- [16] BERG H F, JU Z L, MYRVOLD M, et al. Development of prediction models for lymph node metastasis in endometrioid endometrial carcinoma [J]. *Br J Cancer*, 2020, 122(7): 1014-1022.
- [17] SWISTOWSKA M, GIL-KULIK P, KRZYZANOWSKI A, et al. Potential effect of SOX2 on the cell cycle of wharton's jelly stem cells (WJSCs) [J]. *Oxid Med Cell Longev*, 2019: 5084689.
- [18] LIU P, TANG H L, SONG C L, et al. SOX2 promotes cell proliferation and metastasis in triple negative breast cancer [J]. *Front Pharmacol*, 2018, 9: 942.
- [19] HU P S, LI T, LIN J F, et al. VDR-SOX2 signaling promotes colorectal cancer stemness and malignancy in an acidic microenvironment [J]. *Signal Transduct Target Ther*, 2020, 5(1): 183.
- [20] GUO Y, LU G C, MAO H H, et al. miR-133b suppresses invasion and migration of gastric cancer cells via the COL1A1/TGF-beta axis [J]. *Onco Targets Ther*, 2020, 13: 7985-7995.
- [21] WANG Q Y, SHI L S, SHI K, et al. CircCSPP1 Functions as a ceRNA to Promote Colorectal Carcinoma Cell EMT and Liver Metastasis by Upregulating COL1A1 [J]. *Front Oncol*, 2020, 10: 850.
- [22] ROSSONI R D, DE BARROS P P, MENDONCA I D C, et al. The postbiotic activity of *Lactobacillus paracasei* 28. 4 against *Candida auris* [J]. *Front Cell Infect Microbiol*, 2020, 10: 397.
- [22] 冯晓杰, 李雷. 干扰 COL11A1 基因表达对人卵巢癌细胞侵袭及耐药性影响 [J]. *中华肿瘤防治杂志*, 2019, 26 (13): 908-916.

收稿日期: 2022-10-07; 修回日期: 2022-10-24