

· 论 著 ·

基于生物信息学方法对胰腺癌血浆差异 miRNA 和基因表达的分析研究

陈雪萍¹, 刘 航², 张阳丽¹, 彭 健¹, 吴 江¹, 张玉洪¹, 府伟灵^{3△}

1. 重庆医科大学附属第一医院临床分子医学检测中心, 重庆 400042; 2. 重庆医科大学生命科学院, 重庆 400042; 3. 陆军军医大学第一附属医院检验科, 重庆 400038

摘要:目的 采用生物信息学方法进一步筛选和验证胰腺癌血浆标志物, 为明确其分子机制提供依据。

方法 通过关键词“胰腺癌、血浆、miRNA”进行检索, 搜集文献, 并下载原始数据, 进行生物信息学分析, 寻找具有特异性的血浆 miRNA 标志物, 预测其靶基因, 进行通路富集和互作分析, 通过文献数据和数据库进行关键基因的验证, 分析其差异表达谱及其与胰腺癌预后的关系。**结果** 通过对 GSE59856 和 GSE124158 进行差异 miRNA 的筛选, 最后筛选得出 miR-122-5p。通过对 miR-122-5p 预测得到了 517 个靶基因, 并对靶基因作 GO/KEGG 富集分析。进一步对靶基因做蛋白互作图, 并通过相关文献的查询与筛查, 筛选出 UBA52、ADAM10、CBL、VAMP3 这 4 个关键基因。根据 TCGA 表达谱及文献数据论证, 与健康人相比, 胰腺癌患者高表达 UBA52、ADAM10、CBL、VAMP3。VAMP3、ADAM10 基因低表达的胰腺癌患者总体生存率更高 ($P < 0.05$)。**结论** 采用生物信息学方法对胰腺癌患者血浆差异 miRNA 进行分析, 筛选出的 miR-122-5p 具有作为胰腺癌诊断标志物的潜力。miR-122-5p 作用的关键靶基因 VAMP3、ADAM10 与胰腺癌的预后密切相关。

关键词: 胰腺癌; miR-122-5p; 靶基因

DOI: 10.3969/j.issn.1673-4130.2021.21.020

中图法分类号: R735.9

文章编号: 1673-4130(2021)21-2644-06

文献标志码: A

Analysis of differential miRNA and gene expression in plasma of pancreatic cancer based on bioinformatics method

CHEN Xueping¹, LIU Hang², ZHANG Yangli¹, PENG Jian¹,
WU Jiang¹, ZHANG Yuhong¹, FU Weiling^{3△}

1. Center for Clinical Molecular Medical Detection, First Affiliated Hospital of Chongqing Medical University, Chongqing 400042, China; 2. Institute of Life Sciences, Chongqing Medical University, Chongqing 400042, China; 3. Department of Clinical Laboratory, First Affiliated Hospital, Army Medical University, Chongqing 400038, China

Abstract: Objective The bioinformatics method were used to further screen and verify the plasma markers of pancreatic cancer, which provide a basis for clarifying its molecular mechanism. **Methods** The retrieval was performed by the Keyword "pancreatic cancer" "serum" "miRNA" for collecting the literatures and downloading the original data. The bioinformatics analysis was conducted to find specific serum miRNAs markers, and predict their target genes. The pathway enrichment and transactional analysis were performed. The key genes were verified by the literature data and databases. The differential expression profile and its relationship with the prognosis of pancreatic cancer were analyzed. **Results** The differential miRNA screening of GSE59856 and GSE124158 were performed, miR-122-5p was finally screened out. Then, 517 target genes were obtained by predicting miR-122-5p and the target gene conducted the GO/KEGG enrichment analysis. The targeted genes further conducted the protein interaction mapping. By the related literature inquiring and screening, the four key genes were screened out, including UBA52, ADAM10, CBL and VAMP3. According to TCGA expression profile and literature data argument, compared with the healthy people, UBA52, ADAM10, CBL and VAMP3 were highly expressed in the patients with pancreatic cancer. The pancreatic cancer patients with low expression of VAMP3 and ADAM10 genes had higher overall survival rate. **Conclusion** Adopting

作者简介: 陈雪萍, 女, 技师, 主要从事临床检验诊断新技术研究。△ 通信作者, E-mail: weiling_fu@126.com。

本文引用格式: 陈雪萍, 刘航, 张阳丽, 等. 基于生物信息学方法对胰腺癌血浆差异 miRNA 和基因表达的分析研究[J]. 国际检验医学杂志, 2021, 42(21): 2644-2649.

the bioinformatics method to conduct the plasma differential miRNA analysis, the screened miR-122-5p has the potential as the biomarker for the diagnosis of pancreatic cancer. The key target genes VAMP3 and ADAM10 of miR-122-5p are closely related to the prognosis of pancreatic cancer.

Key words: pancreatic cancer; miR-122-5p; target gene

胰腺癌被称为癌中之王,放眼全球,胰腺癌都是最致命的癌症之一^[1-2]。然而大多数胰腺癌患者在早期并没有任何特殊的临床症状,因此大多数胰腺癌患者都错过了最佳的治疗时期。在过去的 40 年中并没有发现可以显著提升胰腺癌患者生存率的方法。胰腺癌患者的 5 年生存率为 5%~15%,总体生存率约为 6%,且仅有 20%患者可以进行手术切除^[3]。因而早期诊断在胰腺癌防治过程中扮演着重要的角色。

目前胰腺癌的主要筛查诊断方法有超声检查、CT 以及核磁共振,各自有优缺点。超声检查很难区分出恶性组织和非恶性组织;CT 和核磁共振也是诊断胰腺癌的辅助手段,目前也无法满足胰腺癌早期筛查的临床需求。除此之外,许多的肿瘤抗原也被用作胰腺癌诊断的相关指标。其中最有效和临床应用最广的生物标志物是 CA19-9。然而 CA19-9 在机体具有其他炎症的状态下水平也会升高,特异性较低,这也限制了其作为胰腺癌术后检测指标的应用发展^[4-5]。目前还没有有效且准确的生物标志物用于胰腺癌早期筛查和诊断。

miRNA 是一类大小约为 22 bp 的非编码单链 RNA 分子,它们在细胞增殖、蛋白质代谢和肿瘤形成等生理、病理过程中扮演着重要的功能。有研究报道 miRNA 参与实体瘤的发生、发展。在胰腺癌中,miRNA 的上调或下调均与胰腺癌的进展过程密切相关^[6]。

既往有研究表明,miR-21 及其下游靶标在胰腺癌的细胞增殖、凋亡和细胞周期中起关键作用,miR-21 可能是诊断胰腺癌的有潜力的生物标志物^[7]。研究表明 miR-25 对早期胰腺癌的诊断灵敏度和特异度分别达到 76%和 93%^[8]。miR-25 有助于鉴别慢性胰腺炎与胰腺癌,这也大大增加 miRNA 在胰腺癌中的应用潜力,目前关于 miR-25 的临床诊断效能评估也还在进行中。这些报告表明,miRNA 有作为胰腺癌早期筛查指标的应用前景。然而,迄今为止,这些基于 miRNA 的研究均未进入临床实践,还需要更多的研究验证。因此,筛选胰腺癌的血浆肿瘤标志物,对进一步分析其作用的靶基因与胰腺癌发生、发展的关系具有重要的临床意义。

1 资料与方法

1.1 一般资料 以“胰腺癌、miRNA、血浆”为关键词检索文献,然后在 NCBI 数据库 GEO 中进行检索,按

包含胰腺癌标本、其他肿瘤标本、良性胰腺疾病、胆道疾病标本及相应正常胰腺标本的基因芯片为标准,经过筛选,最终研究团队使用了编号为 GSE124158 和 GSE59856 的基因芯片数据。

1.2 方法

1.2.1 显著差异 miRNA 的筛选 首先,采用 GEO 数据库所提供的基于 R 语言的在线分析软件 GEO2R,将编号为 GSE59856 的基因芯片数据分为胰腺癌与健康对照组、胰腺癌与良性疾病对照组、胰腺癌与其他癌症对照组 3 个对照组进行初步筛选。然后,筛选标准 $|\log_2 FC| \geq 1$,经过初步的分析筛选,分别得出 3 个对照组胰腺癌组织差异表达的 miRNA。接着,通过同样的处理步骤将编号为 GSE124158 中的数据也分为同样的 3 个对照组进行处理,设置相同的参数进行筛选,分别得到相应的胰腺癌组织差异表达的 miRNA。进一步使用 Venn diagram webtool 分析,分别得到每个对照组中初步分析所得到的差异表达的 miRNA 的重叠部分,取每个重叠部分的交集,最后得到的重叠部分为显著差异表达的 miRNA。

1.2.2 靶基因预测 将上一步分析所得到的显著差异 miRNA 运用在线分析软件 DIANA tools 进行靶基因的预测^[9]。

1.2.3 关键基因的 GO 功能及 KEGG 通路富集分析 采用 DAVID 和 KOBAS3.0 对预测获得的靶基因开展 GO/KEGG 通路富集分析,以进一步分析靶基因对胰腺癌患者的作用部位及途径。

1.2.4 靶基因的蛋白质相互作用网络的构建 通过 STRING 数据库对上述预测得到的靶基因进行关键蛋白之间的相互作用关系分析,得到 PPI 相互作用图,然后通过查询相关的文献,最后筛选出关键基因^[10]。

1.2.5 关键基因的 TCGA 验证分析 将上述筛选分析得出的关键基因导入在线分析网站 GEPIA、cBioPortal 进行 TCGA 表达谱的分析,再通过查找相关的文献验证所筛选的关键基因与胰腺癌患者预后的相关性,进一步分析其在胰腺癌发生、进展及预后中的作用。

2 结果

2.1 差异表达 miRNA 筛选 本次研究筛选了两个基因芯片(GSE59856、GSE124158)(表 1)进行差异分析。在 GSE59856 中,得到胰腺癌与健康对照组内差

异表达 miRNA 21 个,胰腺癌与其他癌症对照组内差异表达 miRNA 91 个,胰腺癌与良性疾病对照组内差异表达 miRNA 60 个。在 GSE124158 中,胰腺癌与健康对照组中差异表达 miRNA 1 877 个,胰腺癌与其他癌症对照组中差异表达 miRNA 43 个,胰腺癌与良性疾病对照组中差异表达 miRNA 47 个。再用 GSE124158 的数据分别对 GSE59856 每个对照组的数据进行验证,作 VEEN 图,每个对照组分别得到

13、9、6 个差异表达 miRNA(图 1),最后取 3 个对照组数据的交集,最终得到 miR-122-5p 为显著差异表达的 miRNA。

表 1 GEO 数据库胰腺癌基因芯片样本(n)

基因芯片编号	胰腺癌 样本量	健康对照 样本量	其他癌症 样本量	良性疾病 样本量	总量
GSE59856	100	150	300	21	571
GSE124158	87	148	230	21	486

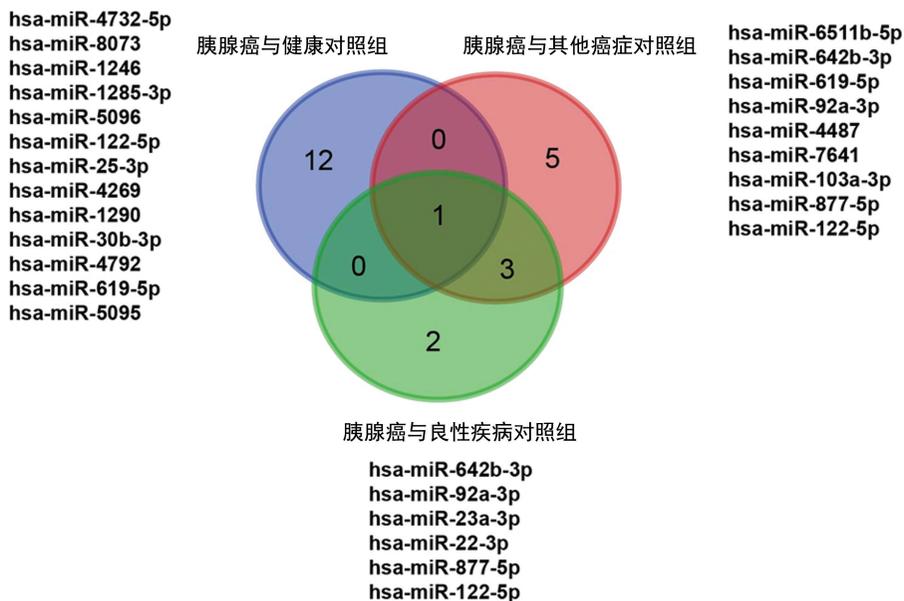


图 1 VEEN 分析

2.2 差异表达 miRNA 的靶基因预测 为了阐明差异表达 miRNA 的生物学功能及其具体的作用机制及通路,将上一步分析所得到的差异表达 miRNA 运用在线数据库 DIANA tools 和 ENCORI 预测 miR-122-5p 的靶基因,最终得到 517 个预测靶基因。

2.3 差异表达 miRNA 的靶基因的 GO 富集和 KEGG 途径分析 使用 DAVID 在线数据库对 517 个靶基因的功能和通路进行富集分析。GO 功能富集分析结果表明,靶基因在生物过程中主要参与胶原蛋白分解代谢、GTP 酶活性的调节、细胞间黏附、RNA 聚合酶 II 启动子转录的正调控、转录的正调控。细胞成分主要是聚集在细胞薄膜、细胞内区、细胞间黏附连接、细胞顶端质膜、细胞质膜、细胞表面。而分子功能

主要与 RNA 聚合酶 II 核心启动子近端区序列特异性 DNA 结合、泛素连接酶 E3 结合域、金属肽酶活性、锌离子结合、金属内肽酶活性、GTP 酶活性有关。见表 2。

KEGG 的结果(图 2)显示,靶基因富集于抗原处理和呈递、囊泡转运中的相互作用、黏蛋白型 O-聚糖生物合成、甲状腺激素合成、自然杀伤细胞介导的细胞毒性、移植物抗宿主病、多巴胺能突触、胰岛素分泌、胆碱能突触、糖基磷脂酰肌醇(GPI)-锚定物、心肌细胞肾上腺素能信号传导、幽门螺杆菌的上皮细胞信号传导、1 型糖尿病的介导、AMPK 信号通路、MAPK 信号通路、黏着连接、心肌收缩、HIF-1 信号通路、内吞作用等。

表 2 GO 功能富集分析

分类	GO 术语	名称	富集基因数	P
生物过程(BP)	GO:0007420	胶原蛋白分解代谢	13	0.002 450
	GO:0043087	GTP 酶活性的调节	7	0.004 953
	GO:0098609	细胞间黏附	3	0.005 637
	GO:0045944	RNA 聚合酶	37	0.009 351
	GO:0045893	RNA 聚合酶 II 启动子转录的正调控	22	0.015 259

续表 2 GO 功能富集分析

分类	GO 术语	名称	富集基因数	P
细胞成分(CC)	GO:0016020	转录的正调控	68	0.017 52
	GO:0005622	细胞薄膜	44	0.023 362
	GO:0005913	细胞内区	14	0.044 417
	GO:0016324	细胞间黏附连接	13	0.044 777
	GO:0005887	细胞顶端质膜	44	0.053 759
	GO:0009986	细胞质膜的成分	20	0.057 935
分子功能(MF)	GO:0000978	RNA 聚合酶 II 核心启动子近端区序列特异性 DNA 结合	20	0.001 091
	GO:0032399	泛素连接酶 E3 结合域	3	0.001 720
	GO:0008237	金属肽酶活性	8	0.003 426
	GO:0008270	锌离子结合	44	0.003 853
	GO:0004222	金属内肽酶活性	9	0.006 132
	GO:0005096	GTP 酶活性	15	0.008 262

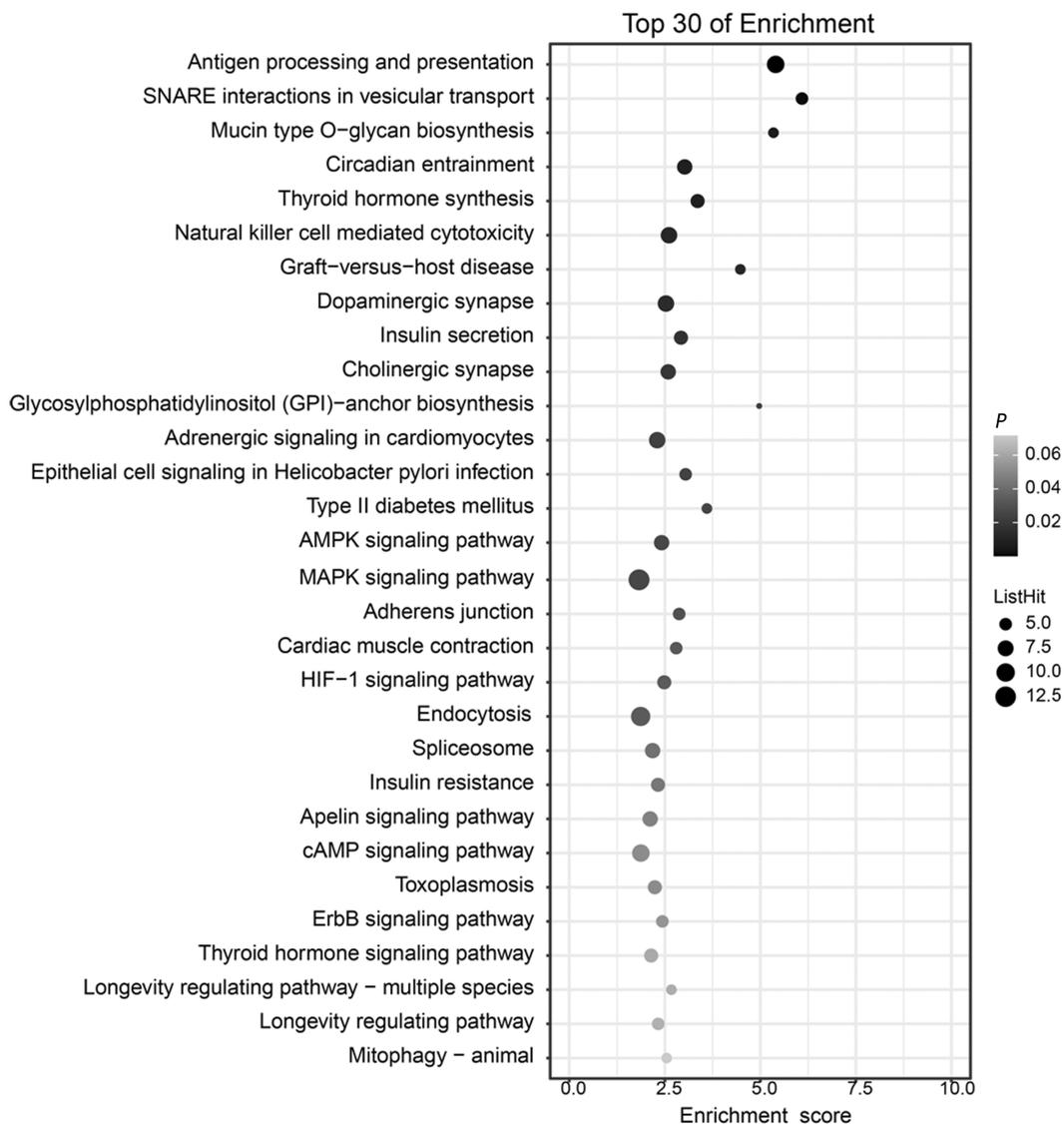


图 2 KEGG 通路富集分析

2.4 从 PPI 网络中筛选关键基因 通过将所预测的靶基因利用在线数据库 STRING 进行数据挖掘及绘

制,得到了蛋白质之间的互相作用关系图。进一步通过相关文献的查询与筛查,筛选得出 UBA52、AD-

AM10、CBL、VAMP3 这 4 个关键基因。然后将得到的结果导入到 Cytoscape 中,作出 4 个关键基因的互作图(图 3)。

2.5 TCGA 验证分析 将筛选所得的 4 个关键基因使用在线分析网站 GEPIA、cBioPortal 分析发现,与健康对照相比,胰腺癌患者体内 UBA52、CBL、VAMP3、ADAM10 的表达均远远高于健康人。然后在 GEPIA、cBioPortal 网站中分析得出胰腺癌患者的生存曲线(图 4),发现与高表达 VAMP3 和 ADAM10 的胰腺癌患者相比,这 2 个基因低表达的患者总生存

率更高($P < 0.05$)。UBA52 和 CBL 的调节对胰腺癌患者的生存没有影响($P > 0.05$)。

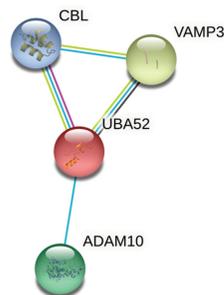
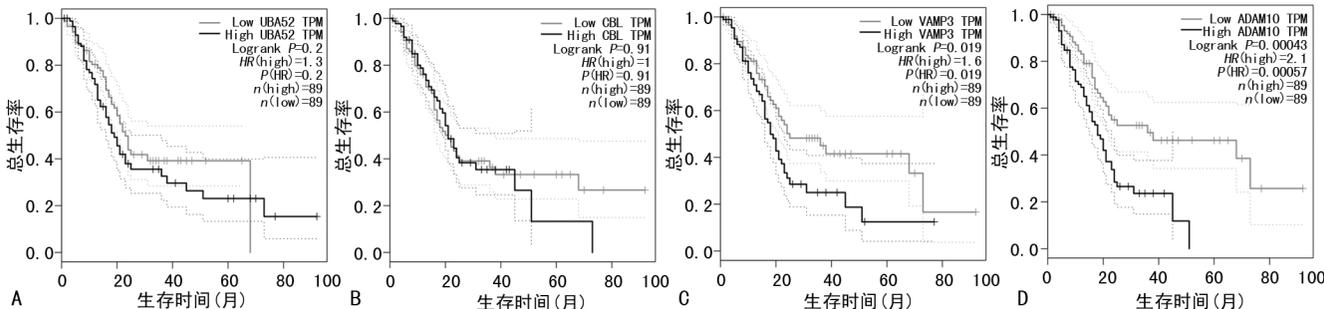


图 3 关键基因互作图



注:A 和 B 分别为 UBA52 和 CBL 基因,两基因与胰腺癌患者的总生存率无相关性;C 和 D 分别为 VAMP3 和 ADAM10 基因,两基因与胰腺癌患者总生存率有关。

图 4 靶基因的表达与胰腺癌患者预后的相关性分析

3 讨 论

胰腺癌的临床表现隐蔽,大多数胰腺癌患者在确诊时已经处于晚期,因为癌细胞已经扩散并无法通过手术切除治疗,晚期胰腺癌患者 5 年生存率低于 5%,因此胰腺癌是恶性肿瘤中病死率最高的肿瘤之一。然而胰腺癌的早期诊断仍面临巨大的挑战。在本次研究中,笔者团队通过 NCBI 的在线分析软件 GEO2R 及 VEEN 对 GSE59856 的数据按照胰腺癌与健康对照组、胰腺癌与其他癌症对照组、胰腺癌与良性疾病对照组进行在线分析,然后再用 GSE124158 的数据进行验证,最后得出 miR-122-5p 这一明显差异表达的 miRNA,miR-122-5p 在胰腺癌患者体内表达下降。此前有研究将健康人群、慢性胰腺炎患者与胰腺癌患者相比较,发现 miR-122-5p 在胰腺癌患者体内表达下降。这与既往的研究结果相符^[11]。另外 miR-122-5p 可作为评估胰腺癌预后的标志物,即 miR-122-5p 血浆水平升高与胰腺癌患者预后较差之间具有一定的关联性^[12]。另有研究报道 miR-122-5p 通过直接靶向细胞周期蛋白 1 抑制了胰腺癌细胞的生长、侵袭和转移^[13]。因此,miR-122-5p 可能是胰腺导管腺癌的治疗靶标。

为继续探索 miR-122-5p 在调节人胰腺癌进展过程中的作用机制,使用 DIANA tools 对 miR-122-5p 进行靶基因预测,得到 517 个靶基因。对关键基因的功能富集分析显示,它们主要参与胶原蛋白分解代

谢、GTP 酶活性的调节、转录的调控等过程。以往有研究和证据表明,本文筛选出的靶基因的 GO、KEGG 富集通路在胰腺癌的进展过程中可能具有重要的功能。再运用结合 DAVID、STRING 等生物信息数据库以及查询文献进行筛选,最后得到 UBA52、ADAM10、CBL、VAMP3 这 4 个关键基因。这些关键基因在癌组织中的表达水平与正常胰腺组织相比显著上调或下降。

对筛选到的这 4 个关键基因进行 TCGA 验证分析,胰腺癌患者体内 UBA52、ADAM10、CBL、VAMP3 的表达均远远高于健康人。再验证 4 个基因的预后价值,与高表达 VAMP3、ADAM10 的胰腺癌患者相比,这 2 个基因低表达的患者总体生存率更高($P < 0.05$)。而 UBA52、CBL 上调对胰腺癌患者生存无明显影响($P > 0.05$)。有研究表明,VAMP3 在胰腺癌细胞的外泌体分泌中发挥积极作用,其中 VAMP3 对于 PVT1 介导的外泌体分泌至关重要^[14]。而胰腺癌组织中 ADAM10 表达比正常组织中更高。ADAM10 沉默会显著降低癌细胞的迁徙与转移能力,但对正常细胞却没有影响^[15]。然而目前这些关键基因在胰腺癌患者中的具体作用机制还不清楚,还需进一步研究。

综上所述,本研究是基于生物信息学方法对胰腺癌患者血浆差异 miRNA 进行分析,由此筛选出 miR-122-5p,并发现其具有作为胰腺癌诊断标志物的潜

力。通过数据库验证,miR-122-5p 作用的关键靶基因 VAMP3、ADAM10 在胰腺癌组织中差异表达,可能与胰腺癌的侵袭与转移相关。但上述 miRNA 及其靶基因与胰腺癌的相关性还需要进一步的大量临床样本和随访数据验证。

参考文献

- [1] ZHAO X, REN Y, CUI N, et al. Identification of key microRNAs and their targets in exosomes of pancreatic cancer using bioinformatics analysis[J]. *Medicine (Baltimore)*, 2018, 97(39): e12632.
- [2] MIZRAHI J D, SURANA R, VALLE J W, et al. Pancreatic cancer[J]. *Lancet*, 2020, 395(10242): 2008-2020.
- [3] GAIDA M M, HAAG N, GUNTHER F, et al. Expression of a disintegrin and metalloprotease 10 in pancreatic carcinoma[J]. *Int J Mol Med*, 2010, 26(2): 281-288.
- [4] KOJMA M, SUDO H, KAWAUCHI J, et al. MicroRNA markers for the diagnosis of pancreatic and biliary-tract cancers[J]. *PLoS One*, 2015, 10(2): e0118220.
- [5] BUSCAIL L, BOURNET B, CORDELIER P. Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer[J]. *Nat Rev Gastroenterol Hepatol*, 2020, 17(3): 153-168.
- [6] ASANO N, MATSUZAKI J, ICHIKAWA M, et al. A serum microRNA classifier for the diagnosis of sarcomas of various histological subtypes[J]. *Nat Commun*, 2019, 10(1): 1299-1302.
- [7] YOKIO A, MATSUZAKI J, YAMAMOTO Y, et al. Integrated extracellular microRNA profiling for ovarian cancer screening[J]. *Nat Commun*, 2018, 9(1): 4319-4322.
- [8] FRANCINE B, LIU X Y, GREGORY J T, et al. MiRNA

analysis in pancreatic cancer: the dartmouth experience[J]. *Clin Chem Lab Med*, 2017, 55(5): 755-762.

- [9] SINGHI A D, KOAY E J, CHARI S T, et al. Early detection of pancreatic cancer: opportunities and challenges[J]. *Gastroenterology*, 2019, 156(7): 2024-2040.
- [10] WANG J, CHEN J, CHANG P, et al. MicroRNAs in plasma of pancreatic ductal adenocarcinoma patients as novel blood-based biomarkers of disease[J]. *Cancer Prev Res*, 2009, 2(9): 807-813.
- [11] SUN C, WANG P, DONG W, et al. LncRNA PVT1 promotes exosome secretion through YKT6, RAB7, and VAMP3 in pancreatic cancer[J]. *Aging*, 2020, 12(11): 10427-10440.
- [12] SINGHI A D, KOAY E J, CHARI S T, et al. Early detection of pancreatic cancer: opportunities and challenges[J]. *Gastroenterology*, 2019, 156(7): 2024-2040.
- [10] YU Y, TONG Y, ZHONG A, et al. Identification of Serum microRNA-25 as a novel biomarker for pancreatic cancer[J]. *Medicine (Baltimore)*, 2020, 99(52): e23863.
- [13] MAZZA T, GIOFFREDA D, FONTANA A, et al. Clinical significance of circulating miR-1273g-3p and miR-122-5p in pancreatic cancer[J]. *Front Oncol*, 2020, 10: 44-48.
- [14] DAI C, ZHANG Y, XU Z, et al. MicroRNA-122-5p inhibits cell proliferation, migration and invasion by targeting CCNG1 in pancreatic ductal adenocarcinoma[J]. *Cancer Cell Int*, 2020, 20: 98-103.
- [15] ZHOU X, LU Z, WANG T, et al. Plasma miRNAs in diagnosis and prognosis of pancreatic cancer: a miRNA expression analysis[J]. *Gene*, 2018, 673: 181-193.

(收稿日期: 2021-02-11 修回日期: 2021-08-09)

(上接第 2643 页)

- adenocarcinoma[J]. *Oncotarget*, 2017, 8(63): 106685-106692.
- [12] 林静, 王善欢, 李美琼, 等. 纤支镜联合细胞学检测在肺肿瘤诊断中的应用[J]. *中国内镜杂志*, 2019, 25(7): 7-11.
- [13] 王学敏, 施琳, 李时荣, 等. 纤维支气管镜刷检液基细胞学检查在肺癌诊断中的价值分析[J]. *重庆医学*, 2018, 47(8): 1040-1043.
- [14] 赵琳琳, 孙悦, 王聪, 等. 纤维支气管镜刷检液基薄片细胞学联合免疫细胞化学在肺癌病理分型诊断中的临床价值[J]. *中华肿瘤杂志*, 2019, 41(5): 326-330.
- [15] LOU S K, KO H M, KINOSHITA T, et al. Implementation of PD-L1 22C3 IHC pharmDx™ in cell block preparations of lung cancer: concordance with surgical resec-

tions and technical validation of cytolyt prefixation[J]. *Acta Cytol*, 2020, 64(6): 577-587.

- [16] 张功学, 丁凯, 齐峰, 等. 胸腔积液细胞块免疫组织化学检测对疑似肺癌患者的诊断价值[J]. *新乡医学院学报*, 2018, 35(2): 129-132.
- [17] 陈一峰, 陈婷, 陈玮珊, 等. 胸水细胞蜡块在肺癌诊断和 EGFR 基因检测中的应用研究[J]. *中国现代医学杂志*, 2018, 28(29): 43-46.
- [18] 赵业, 张继新, 梁丽, 等. 细胞块技术在胸腔积液病理诊断评估中的应用[J]. *中国组织工程研究*, 2018, 22(12): 1934-1938.

(收稿日期: 2020-12-21 修回日期: 2021-08-16)