

·生物信息学·

加权共表达网络分析在肝癌环状 RNA 研究中的应用

王俐勇¹,赵焕英¹,付文卓¹,杨予涛²

(首都医科大学中心实验室¹,基础医学院²,北京 100069)

摘要:目的 应用加权基因共表达网络分析(WGCNA)研究肝癌组织中环状 RNA,以期寻找合适的生物标志物作为肝癌临床诊断指标。方法 从 GEO 数据库下载肝癌数据,通过 WGCNA 方法深入挖掘分析得到与肝癌发展密切相关的环状 RNA,并结合环状 RNA-miRNA-mRNA 互作数据库筛选靶向 miRNA 和基因,利用 Kaplan-Meier plotter 数据库筛选出来 circRNA 靶向的 microRNA 进行生存分析验证,通过 DAVID 软件对靶向基因进行功能富集分析。结果 共检测到 3032 个 circRNA 分子;应用 WGCNA 方法共得到 7 个模块,其中 Turquoise 模块与肝癌组织呈正相关($P<0.05$);Turquoise 模块中 circRNA 表达量在肝癌组织中均表达上调,且经过在线数据库比对筛选得到 2 个环状 RNA 分子(hsa_circ_0064324 和 hsa_circ_0040827),Pathway 富集分析显示,这些靶向基因富集在鞘磷脂信号通路、癌症蛋白聚糖、血小板活化通路、胰腺癌通路、神经胶质细胞瘤、FoxO 信号通路、癌症中胆碱代谢通路等多个和癌症相关的通路上。肝癌细胞系中 hsa_circ_0064324 和 hsa_circ_0040827 均表达上调,且 Kaplan-Meier 生存分析发现,hsa-miR-146b-3p、hsa-miR-331-3p、hsa-miR-874-3p 和 hsa-miR-942-3p 高表达与肝癌总体生存期呈负相关($P<0.05$)。结论 WGCNA 方法筛选得到的环状 RNA 可能作为肝癌临床诊断生物标志物,有助于揭示肝癌发病机制,为研发药物和治疗诊断提供新依据。

关键词:环状 RNA;WGCNA;肝癌

中图分类号:R735.7

文献标识码:A

DOI:10.3969/j.issn.1006-1959.2023.04.001

文章编号:1006-1959(2023)04-0001-06

Application of Weighted Gene Co-expression Network Analysis in the Study of Circular RNA in Liver Cancer

WANG Li-yong¹,ZHAO Huan-ying¹,FU Wen-zhuo¹,YANG Yu-tao²

(Core of Facility¹,College of Basic Medicine²,Capital Medical University,Beijing 100069,China)

Abstract:Objective To study the circular RNA in liver cancer tissues by weighted gene co-expression network analysis (WGCNA), in order to find suitable biomarkers as clinical diagnostic indicators of liver cancer. Methods The data of liver cancer were downloaded from GEO database, and the circular RNA closely related to the development of liver cancer was obtained by WGCNA method. Combined with the circular RNA-miRNA-mRNA interaction database, the targeted miRNA and gene were screened. The circRNA-targeted microRNA was screened by Kaplan-Meier plotter database for survival analysis and verification, and the functional enrichment analysis of targeted genes was performed by DAVID software. Results A total of 3032 circRNA molecules were detected; a total of 7 modules were obtained by WGCNA method, among which Turquoise module was positively correlated with liver cancer tissues ($P<0.05$). The expression of circRNA in the Turquoise module was up-regulated in liver cancer tissues, and two circular RNA molecules (hsa_circ_0064324, hsa_circ_0040827) were screened by online database comparison. Pathway enrichment analysis showed that these targeted genes were enriched in sphingomyelin signaling pathway, proteoglycans in cancer, platelet activation pathway, pancreatic cancer pathway, glioma pathway, FoxO signaling pathway, choline metabolism in cancer pathway and other cancer-related pathways. The expression of hsa_circ_0064324 and hsa_circ_0040827 was up-regulated in hepatocellular carcinoma cell lines, and Kaplan-Meier survival analysis showed that the high expression of hsa-miR-146b-3p, hsa-miR-331-3p, hsa-miR-874-3p and hsa-miR-942-3p was negatively correlated with the overall survival of liver cancer ($P<0.05$). Conclusion The circular RNA screened by WGCNA method may be used as a biomarker for clinical diagnosis of liver cancer, which is helpful to reveal the pathogenesis of liver cancer and provide a new basis for the development of drugs and treatment diagnosis.

Key words: circRNA;WGCNA;Liver cancer

肝癌(liver cancer)是一种常见恶性肿瘤,占有癌症死亡 11%,排名第 2 位^[1]。肝癌发展过程中伴

随肝细胞增殖和脂肪变性、氧化应激、线粒体损伤和活性氧诱导,但其发生机制尚未完全明确。已有研究表明^[2],肝癌发展与癌细胞中核酸蛋白水平变化相关,找到合适分子生物标志物有助于研究肝癌发病机制及进行个性化治疗。1976 年首次在 RNA 病毒中发现环状 RNA(circRNAs),近年来随着 RNA 测序技术和生物信息学发展,发现 circRNAs 作为蛋白质复合物中的支架结构,从亚细胞定位隔离蛋白,调节

基金项目:国家自然科学基金项目(编号:82071527)

作者简介:王俐勇(1977.4-),男,四川成都人,博士,副主任技师,主要从事功能基因组学的研究

通讯作者:杨予涛(1972.7-),男,河南三门峡人,博士,副教授,主要从事抑郁症与功能基因组学研究

亲本基因表达,调节选择性剪接和 RNA-蛋白质相互作用,形成 microRNA 海绵结构。circRNAs 与线性 RNA 不同,其具有圆形共价键结构,对核酸外切酶耐受性更高。研究表明^[3,4],circRNAs 可能参与肿瘤发生发展。WGCNA 共表达网络分析可用于 microRNA 和非编码长链 RNA 研究,通过构建共表达模块可挖掘筛选关键基因^[5]。本研究探讨加权共表达网络分析在肝癌环状 RNA 研究中的应用价值,现报道如下。

1 材料与方法

1.1 数据下载和收集 通过 GEO 网站(<https://www.ncbi.nlm.nih.gov/geo/>) 下载肝癌数据集(<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97332>),将得到的 circRNA 表达原始值均一化。从 NCBI 数据库下载 circRNA 芯片数据 GSE97332(<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97332>),包括 14 例肝癌数据,其中 7 例肿瘤组织,7 例正常组织。

1.2 数据集差异基因筛选分析 利用 DESeq2 包(<http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html>) 按照差异倍数大于 2 倍, P 值小于 0.05 标准对肝癌 circRNA 进行差异表达分析。

1.3 WGCNA 共表达分析 该数据集共检测到 3032 个 circRNA 分子,均一化后应用 WGCNA 方法构建 circRNA 共表达网络。应用 R 软件包 WGCNA 构建

共表达网络,通过以下公式构建加权邻接矩阵 $S_{ij}=|cor(X_i,X_j)|^{a_{ij}}=S_{ij}^{\beta}$,其中 i 和 j 代表不同的 circRNA, X_i 和 X_j 代表 circRNA 的表达量, S_{ij} 表示皮尔逊相关系数, a_{ij} 表示 2 个 circRNA 之间的相关系数。本研究中选择软阈值 $\beta=10$,无尺度模型拟合系数 $R^2=0.91$ 。

1.4 确定关键 circRNA 将邻接矩阵转化成拓扑矩阵 (TOM),TOM 矩阵是定量网络节点之间加权关系相似性的方法,然后通过层次聚类确定模块,每个模块至少包含 30 个 circRNA,最后确定可能在肝癌起关键作用的 circRNA。

1.5 基因富集分析和聚类 应用 circRNA 在线数据库 (<https://circinteractome.nia.nih.gov/>) 筛选出 circRNA 可能结合的 miRNA,通过 miRNA 数据库 (<http://mirwalk.umm.uni-heidelberg.de/>) 得到 miRNA 对应的靶基因,利用 DAVID (<https://david.ncifcrf.gov/summary.jsp>) 在线分析工具对靶基因进行 KEGG 通路富集分析。

1.6 qPCR 验证 DMEM 培养液和 10%胎牛血清体外培养人肝癌细胞系 (HepG2) 和正常肝细胞系 (L02),收集细胞悬液使用 TRIzol 从这些细胞中分离总 RNA,通过 RNase R 处理提高 circRNA 的浓度,逆转录后使用实时定量 qPCR 分析定量 circRNA。以 β -actin 为内参照,PCR 的反应条件:95 °C 2 min,95 °C 15 s、58 °C 60 s,共 40 个循环,qPCR 引物序列见表 1。

表 1 circRNA 引物序列

环状 RNA	前引物(5'-3')	后引物(5'-3')
β -Actin	AGCGAGCATCCCCAAAGTT	GGGCACGAAGGCTCATCATT
hsa_circ_0064324	TTAACGTCAGTGCTTCTAAC	TGTCGCGAGGCAACCCTT
hsa_circ_0040827	TGTACGCAGGACACGAGCGAT	CTATGGAAGTAGCCAGTAGT

1.7 生存分析验证 通过 Kaplan-Meier plotter 数据库 (<http://kmplot.com/analysis/index.php>) 对筛选出来 circRNA 靶向的 microRNA 进行生存分析验证。

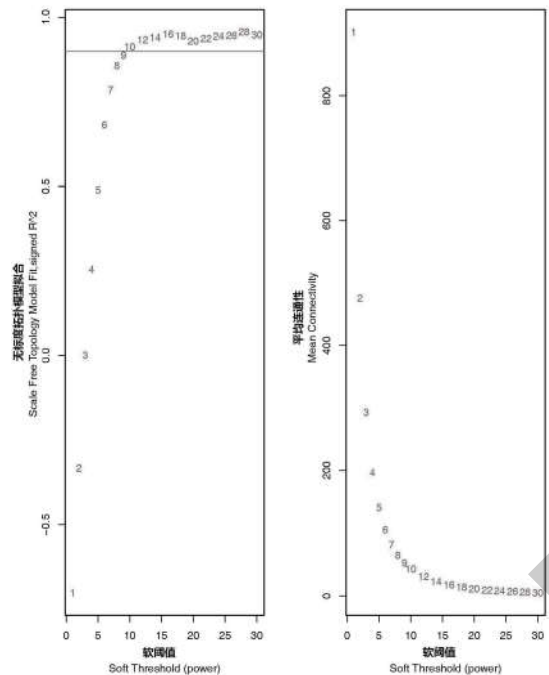
2 结果

2.1 数据均一化和选择软阈值 共检测到 3032 个 circRNA 分子,均一化后应用 WGCNA 方法构建 circRNA 共表达网络,当软阈值为 10 时,基因之间连通性符合无尺度网络分布,无尺度模型拟合系数 $R^2=0.91$,见图 1。

2.2 共表达模块构建与肝癌相关模块的关系 应用 WGCNA 方法共得到 7 个模块,其中 Turquoise 模块

与肝癌组织呈正相关($P<0.05$),见表 2、图 2。

2.3 circRNA 与 miRNA 关联基因筛选 Turquoise 模块中 circRNA 表达量在肝癌组织中均表达上调,见表 3。经过在线数据库比对筛选出 hsa_circ_0064324 和 hsa_circ_0040827 进一步分析,共筛选出 hsa-miR-146b-3p、hsa-miR-331-3p、hsa-miR-874-3p 和 hsa-miR-942-3p 四个 miRNA 分子,然后通过 miRWalk 数据库 miRDB、TargetScan、miRTarBase 软件筛选出这些 miRNA 对应的靶基因,构建 circRNA-miRNA-mRNA 网络,见图 3。



注:左图为不同软阈值 β 对应的无尺度拟合指数 R^2 ,右图为不同软阈值 β 对应的网络平均连通性

图 1 WGCNA 软阈值确定

表 2 共表达模块构建与肝癌相关模块的关系

模块	肝癌组织		正常组织	
	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
Turquoise 模块	0.913	9.22E-06	-0.913	9.22E-06
Brown 模块	-0.724	4.12E-03	0.724	4.12E-03
Yellow 模块	0.122	0.755	-0.122	0.755
Red 模块	-0.361	0.221	0.361	0.221
Blue 模块	-0.749	2.45E-03	0.749	2.45E-03
Green 模块	0.072	0.834	-0.072	0.834
Grey 模块	0.231	0.415	-0.231	0.415

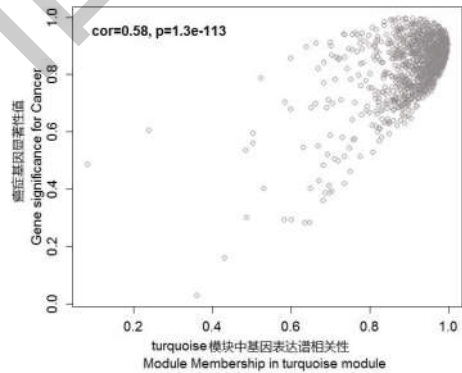


图 2 Turquoise 模块与肝癌组织的相关性

表 3 Turquoise 模块中 circRNA 量表达

环状 RNA	校正后 <i>P</i> 值	<i>P</i> 值	表达量比值
hsa_circ_0060063	3.91E-07	3.61E-09	2.352
hsa_circ_0064324	2.22E-07	1.69E-09	1.772
hsa_circ_0069086	8.72E-08	5.61E-10	1.993
hsa_circ_0031132	3.37E-07	2.96E-09	1.738
hsa_circ_0033408	8.30E-08	5.04E-10	1.362
hsa_circ_0040827	9.05E-07	1.10E-08	1.311
hsa_circ_0072386	6.71E-08	3.34E-10	1.162
hsa_circ_0007246	5.84E-07	6.33E-09	0.884
hsa_circ_0092369	1.64E-06	2.32E-08	0.751

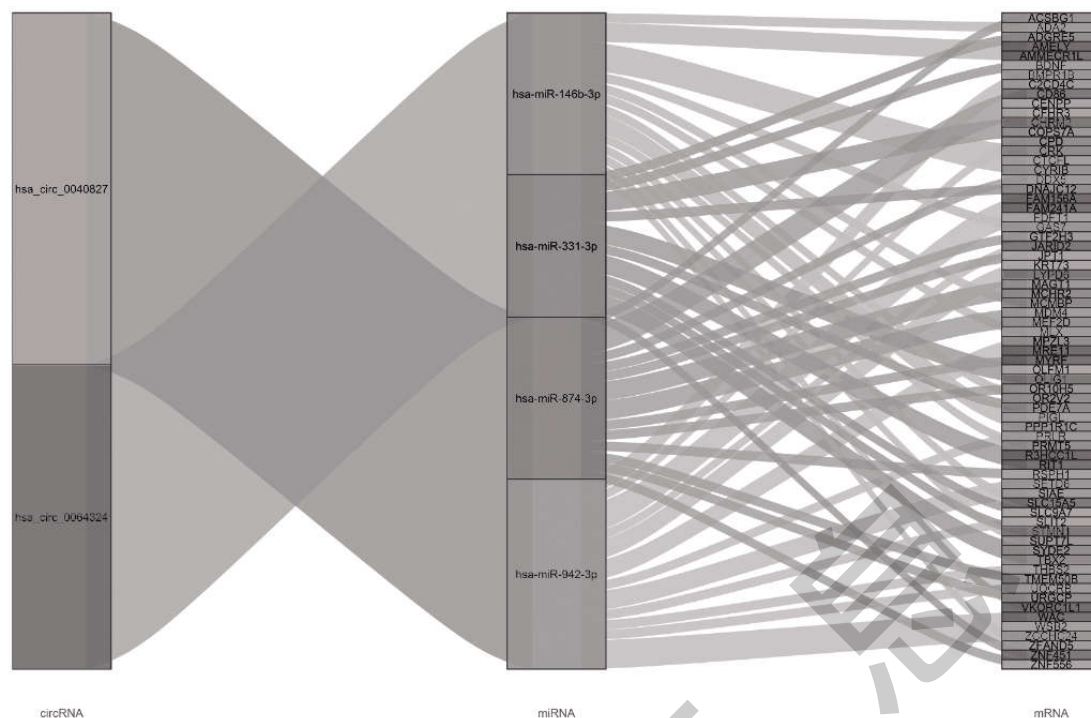


图 3 肝癌中 circRNA-miRNA-mRNA 相互作用网络图

2.4 功能富集分析验证 提取 miRWalk 数据库 miRDB,TargetScan,miRTarBase 软件筛选得到 miRNA 靶向基因,利用 David 在线分析软件对这些基因进行 Pathway 富集分析,结果显示这些靶向基因富集在鞘磷脂信号通路、癌症蛋白聚糖、血小板活化通路、胰腺癌通路、神经胶质细胞瘤、FoxO 信号通路、癌症中胆碱代谢通路等多个和癌症相关的通路上,见图 4。

2.5 功能验证 体外培养人肝癌细胞系(HepG2)和正

常肝细胞系(L02),并对筛选到的 hsa_circ_0064324 和 hsa_circ_0040827 定量检测,可见在肝癌细胞系中 hsa_circ_0064324 和 hsa_circ_0040827 均表达上调,见图 5。Kaplan-Meier 生存分析发现,hsa-miR-146b-3p、hsa-miR-331-3p、hsa-miR-874-3p 和 hsa-miR-942-3p 高表达与肝癌总体生存期呈负相关($P<0.05$),见图 6。

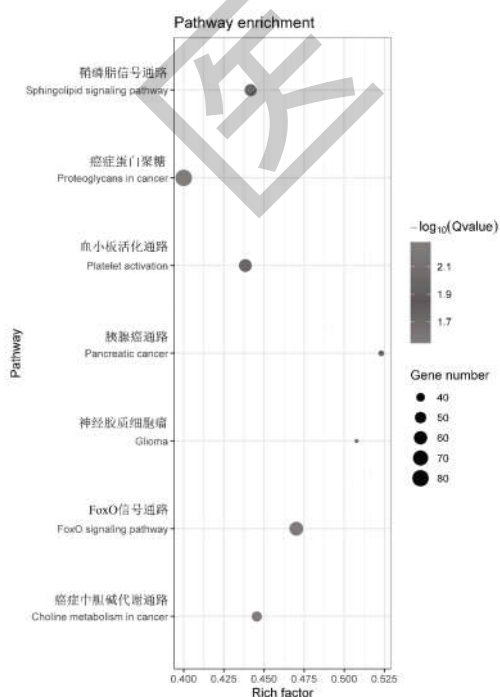


图 4 KEGG 富集通路分析

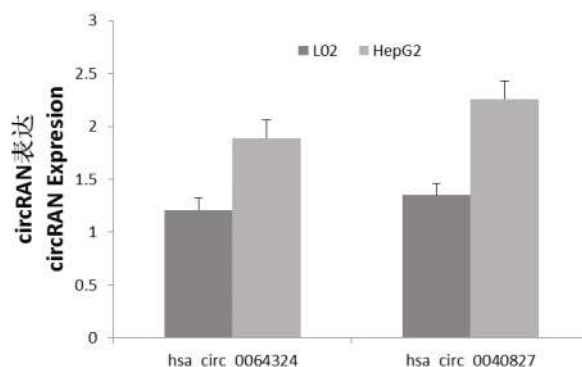


图 5 关键 circRNA 在不同细胞系中表达

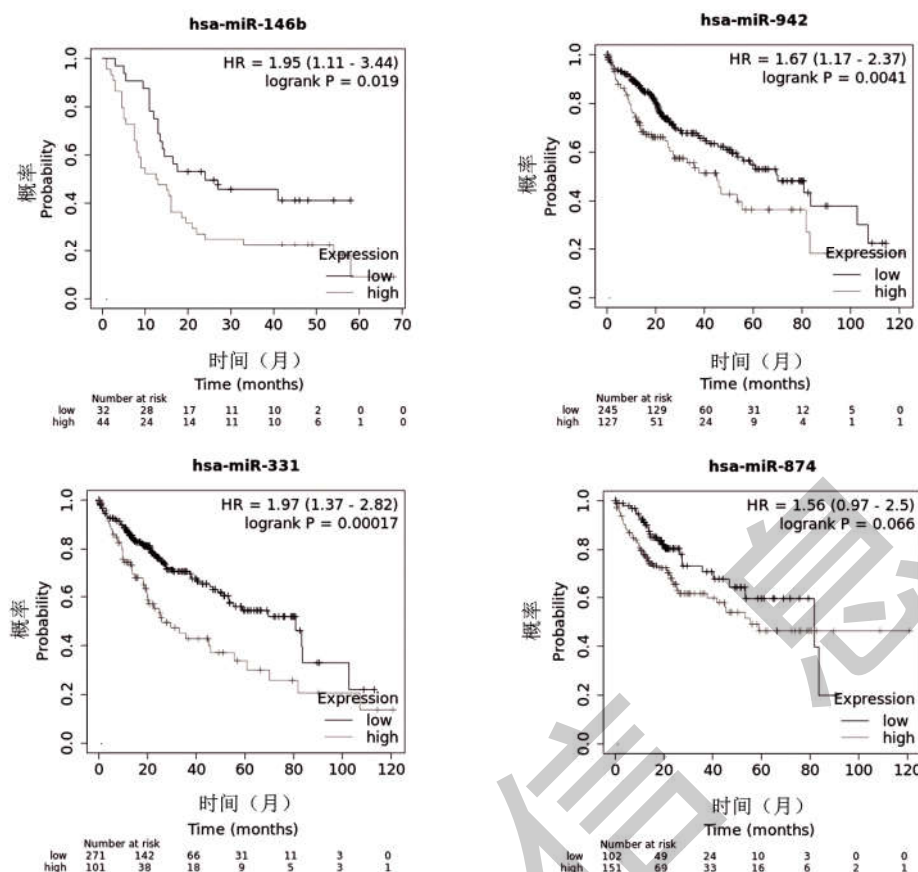


图 6 miRNA 生存分析验证

3 讨论

circRNA 是一类封闭环状非编码 RNA,分为外显子环状 RNA、内含子环状 RNA,通过和 miRNA 结合在转录、转录后以及翻译和翻译后水平调控基因表达发挥重要作用。circRNA 具有表达相对稳定、表达调控有时空特异性、成环序列保守等优势,在肝癌发生机制研究中备受关注^[6]。目前已发现 circRNA-5692^[7,8]、circRNA-100338^[9,10]、hsa_circRNA_104348^[11,12]与肝癌发生发展相关。这些环状 RNA 通过内源竞争 RNA(ceRNA)机制结合 miRNA 分子导致靶基因沉默,影响肝癌细胞侵袭转移。由于肝癌发生发展机制复杂,需要更好的方法挖掘更多肝癌相关环状 RNA 以揭示肝癌调控机制。

circRNA 芯片具有通量高检测灵敏度高特点,可以对人体内大多数 circRNA 同时高效检测。本研究利用芯片数据中 circRNA 之间相互关联绝对值定义共表达网络,通过一个邻接矩阵定义相互关联权重,采用软阈值参数构造加权网络,使用拓扑重叠度量(TOM)作为邻近度量,结合 2 个邻接 circRNA 相互作用强度,聚类成网络模块^[13,14]。另在网络模块

中,应用主成分分析计算出主成分和临床性状相互关系,确定与肝癌相关模块,并依据 circRNA 与主成分模块相互关系确定出关键 circRNA^[15,16]。本研究从与肝癌正相关 Turquoise 模块中挑选出上调表达 circRNAs,经文献数据库比对筛选出 hsa_circ_0064324、hsa_circ_0040827 进一步分析,利用 qPCR 定量检测发现在肝癌细胞系 HepG2 中它们均表达上调,并在数据库中发现这些 circRNA 可能通过竞争结合 hsa-miR-146b、hsa-miR-942、hsa-miR-331、hsa-miR-874;经生物信息学分析发现,这些 microRNA 分子下游靶基因多富集在与癌症发生发展相关通路。有研究证明^[17],FoxO 信号通路与肝癌转移密切相关,该通路中 FoxO 转录因子调节多种生物过程,包括细胞增殖、死亡、衰老、血管生成、细胞迁移和转移所必需的基因。

预后评价是制定临床方案重要参考之一,对于延长患者生存期意义重大。Li C 等^[18]研究发现,TRAF6(TNF 受体相关因子)为肝癌中 miR-146b-5p 靶向位点,通过抑制 TRAF6 介导 AKT 通路磷酸化可减少肝癌细胞生长和转移。Zhang QF 等^[19]研究发

现,与正常组织相比,肝癌组织和细胞系中 miR-942-3p 表达水平升高,并与病理分期和肿瘤淋巴结转移分期相关,是肝癌患者生存率低的独立预后因素。PH 域和亮氨酸丰富重复蛋白磷酸酶(PLPP)是 miR-331-3p 靶点,miR-331-3p 通过抑制 PLPP 介导 AKT 通路去磷酸化,促进肝癌细胞增殖。在肝癌细胞系 PLC/PRF/5 中过表达的 miR-874-3p 会抑制癌细胞生长,促进癌细胞凋亡。本研究 Kaplan-Meier 生存分析发现,hsa-miR-146b-3p、hsa-miR-331-3p、hsa-miR-874-3p 和 hsa-miR-942-3p 高表达与肝癌总体生存期呈负相关,即表达量越低,患者肝癌预后生存期越长,提示 hsa-miR-146b-3p、hsa-miR-331-3p、hsa-miR-874-3p 和 hsa-miR-942-3p 高表达有可能是患者不良预后独立预测因子。

综上所述,利用 WGCNA 方法筛选出 hsa_circ_0064324 和 hsa_circ_0040827,经生物信息学分析认为其可能结合 hsa-miR-146b-3p、hsa-miR-331-3p、hsa-miR-874-3p 和 hsa-miR-942-3p,间接调控下游靶基因表达,影响肿瘤细胞增殖侵袭,这两个 circRNA 分子可能作为肝癌的生物标记物和治疗靶点,对肝癌诊断和治疗有积极意义,然而如何调控靶基因作用于癌症发生与转移仍有待进一步探索。

参考文献:

- [1]Yin L,Cai ZH,Zhu BA,et al.Identification of key pathways and genes in the dynamic progression of HCC based on WGCNA[J].Genes (Basel),2018,9(2):92.
- [2]Cho HJ,Cheong JY.Role of immune cells in patients with hepatitis B virus-related hepatocellular carcinoma [J].Int J Mol Sci,2021,22(15):8011.
- [3]Chaabane M,Andreeva K,Hwang JY,et al.seekCRIT: Detecting and characterizing differentially expressed circular RNAs using high-throughput sequencing data [J].PLoS Comput Biol,2020,16(10):e1008338.
- [4]Zhang Y,Qian LL,Liu YY,et al.CircRNA-ceRNA network revealing the potential regulatory roles of CircRNA in alzheimer's disease involved the cGMP-PKG signal pathway[J].Front Mol Neurosci,2021,14:665788.
- [5]Tian ZL,He WX,Tang JN,et al.Identification of important modules and biomarkers in breast cancer based on WGCNA[J].Onco Targets Ther,2020,13:6805-6817.
- [6]Vo JN,Cieslik M,Zhang YJ,et al.The landscape of circular RNA in cancer[J].Cell 2019,176(4):869-881.e13.
- [7]Liu ZG,Yu YQ,Huang ZB,et al.CircRNA-5692 inhibits the progression of hepatocellular carcinoma by sponging miR-328-5p to enhance DAB2IP expression [J].Cell Death Dis,2019,10(12):900.
- [8]Huang ZY,Xia HM,Liu SQ,et al.The mechanism and clinical significance of circular RNAs in hepatocellular carcinoma [J].Front Oncol,2021,11:714665.
- [9]Huang XY,Huang ZL,Huang J,et al.Exosomal circRNA-100338 promotes hepatocellular carcinoma metastasis via enhancing invasiveness and angiogenesis [J].J Exp Clin Cancer Res,2020,39(1):20.
- [10]Huang XY,Huang ZL,Zhang PB,et al.CircRNA-100338 is associated with mTOR signaling pathway and poor prognosis in hepatocellular carcinoma[J].Front Oncol,2019,9:392.
- [11]Huang GQ,Liang M,Liu HY,et al.CircRNA hsa_circRNA_104348 promotes hepatocellular carcinoma progression through modulating miR-187-3p/RTKN2 axis and activating Wnt/ β -catenin pathway[J].Cell Death Dis,2020,11(12):1065.
- [12]Deldar Abad Paskeh M,Mirzaei S,Ashrafizadeh M,et al.Wnt/ β -catenin signaling as a driver of hepatocellular carcinoma progression: an emphasis on molecular pathways [J].J Hepatocell Carcinoma,2021,8:1415-1444.
- [13]Liang WW,Sun FF,Zhao YM,et al.Identification of susceptibility modules and genes for cardiovascular disease in diabetic patients using WGCNA analysis [J].J Diabetes Res,2020,2020:4178639.
- [14]Hou J,Ye XF,Feng WX,et al.Distance correlation application to gene co-expression network analysis [J].BMC Bioinformatics,2022,23(1):81.
- [15]Yang HY,Wang YK,Zhang ZY,et al.Identification of KIF18B as a hub candidate gene in the metastasis of clear cell renal cell carcinoma by weighted gene Co-expression network analysis[J].Front Genet,2020,11:905.
- [16]Bian YY,Huang JT,Zeng ZL,et al.Construction of survival-related co-expression modules and identification of potential prognostic biomarkers of osteosarcoma using WGCNA [J].Ann Transl Med,2022,10(6):296.
- [17]Jiramongkol Y,Lam EW.FOXO transcription factor family in cancer and metastasis [J].Cancer Metastasis Rev,2020,39(3):681-709.
- [18]Li C,Miao RC,Liu SS,et al.Down-regulation of miR-146b-5p by long noncoding RNA MALAT1 in hepatocellular carcinoma promotes cancer growth and metastasis[J].Oncotarget,2017,8(17):28683-28695.
- [19]Zhang QF,Zhu BL,Qian JP,et al.MiR-942 promotes proliferation and metastasis of hepatocellular carcinoma cells by inhibiting RRM2B[J].Onco Targets Ther,2019,12:8367-8378.

收稿日期:2022-04-14;修回日期:2022-06-15

编辑/杜帆