

肝细胞癌相关差异基因的生物信息学及预后分析

朱亚玲¹, 赵洪波², 蒋保三¹, 王 刚¹, 张越美¹, 刁 勇^{1*}

(1. 华侨大学 医学院, 中国福建 泉州 362021; 2. 昆明医科大学 分子临床医学研究院暨云南省干细胞和再生医学重点实验室, 中国云南 昆明 650500)

摘 要: 肝细胞癌是全球癌症相关死亡的主要原因, 目前对肝细胞癌的发病机制研究尚不完善, 探索肝细胞癌发生、发展相关的分子标志物及其预后具有重要意义。从 GEO 数据库获得肝细胞癌组织和非癌组织的基因表达阵列数据 GSE84402, 利用 GEO2R 筛选差异表达基因; 采用 DAVID 数据库对差异基因进行 GO 富集分析和 KEGG 通路分析; 通过 STRING 数据库和 Cytoscape 软件构建差异表达基因对应的蛋白质相互作用网络, 并从网络中筛选出核心基因(hub genes); 结合 KM plotter 数据库的临床信息对 hub genes 进行预后分析。结果显示: 共得到 1 307 个差异表达基因, 其中上调基因 741 个, 下调基因 566 个, 这些差异表达基因主要涉及细胞分裂、细胞周期、DNA 复制及物质代谢等生物学过程及生物通路。通过 GO、KEGG 及蛋白质相互作用网络筛选出 *BUB1*、*BUB1B*、*CCNA2*、*CCNB1*、*CCNB2*、*CDC20*、*CDK1*、*MAD2L1*、*PLK1* 等 9 个 hub genes, 进一步分析发现 hub genes 均与细胞周期的调控相关, 表明细胞周期的调控失常在肝细胞癌的发生、发展过程中具有重要作用。生存分析显示 9 个 hub genes 在肝细胞癌患者中均为表达上调的基因, 且与患者预后不良相关, 这为寻找肝细胞癌患者预后相关生物标志物的研究提供了线索。

关键词: 肝细胞癌(HCC); 差异表达基因(DEGs); 生物信息学; 预后

中图分类号: Q811.4, R735.7

文献标识码: A

文章编号: 1007-7847(2019)01-0020-08

Bioinformatics and Prognostic Analysis of Differentially Expressed Genes in Hepatocellular Carcinoma

ZHU Ya-ling¹, ZHAO Hong-bo², JIANG Bao-san¹, WANG Gang¹,
ZHANG Yue-mei¹, DIAO Yong^{1*}

(1. School of Medicine, Huaqiao University, Quanzhou 362021, Fujian, China; 2. Yunnan Key Laboratory of Stem Cell and Regenerative Medicine, Institute of Molecular and Clinical Medicine, Kunming Medical University, Kunming 650500, Yunnan, China)

Abstract: Hepatocellular carcinoma (HCC) is the leading cause of cancer-related death worldwide, but the research on the pathogenesis of HCC is still imperfect. It is important to explore its molecular markers and prognosis. Bioinformatics was used to screen HCC related genes. Gene expression profile data GSE84402, which includes data coming from HCC and non-cancer tissues, were obtained from Gene Expression Omnibus (GEO) database. GEO2R was used to screen differentially expressed genes (DEGs). GO function and KEGG pathway analyses of DEGs were performed using DAVID database, and the protein-protein interaction (PPI) network was constructed using STRING and Cytoscape software, and the hub genes were screened. The prognostic significance of hub genes was analyzed by KM plotter database. A total of 1 307 DEGs were screened, including 741 up-regulated and 566 down-regulated genes. Biological processes and pathway enrichment analyses showed that the DEGs were mainly enriched in cell division, cell cycle, DNA replication and material metabolism. From GO, KEGG and PPI network, *BUB1*, *BUB1B*, *CCNA2*, *CCNB1*, *CCNB2*,

收稿日期: 2018-11-05; 修回日期: 2018-12-10

基金项目: 福建省第四批创业创新人才引进计划资助项目(Z16X0079); 泉州市桐江学者特聘教授奖励计划资助项目(Z17X0232); 华侨大学科研基金资助项目(17BS501)

作者简介: 朱亚玲(1987-), 女, 河南开封人, 硕士, 主要从事肿瘤流行病学研究; 朱亚玲和赵洪波对本文的贡献相同, 为本文共同第一作者; * 通讯作者: 刁勇(1967-), 男, 山东运城人, 博士, 华侨大学教授, 博士研究生导师, 主要从事基因药物研究, Tel: 0595-22692516, E-mail: diaoyong@hqu.edu.cn.

CDC20, *CDK1*, *MAD2L1* and *PLK1* were screened as hub genes. Further analysis showed that the 9 hub genes were all related to the regulation of cell cycle, indicating that the abnormal regulation of cell cycle plays an important role in the development of HCC. Survival analysis showed that the 9 hub genes were found to be up-regulated in HCC patients and correlated with poor prognosis. The study would help us find out the biomarkers related to the prognosis of HCC patients.

Key words: hepatocellular carcinoma (HCC); differentially expressed genes (DEGs); bioinformatics; prognosis (*Life Science Research*, 2019, 23(1): 020~027)

肝细胞癌(hepatocellular carcinoma, HCC)是常见的恶性肿瘤之一,占原发性肝癌的 70%~90%,是全球范围内致死率排名第二的恶性肿瘤^[1]。遗传因素^[2]和环境因素(酒精、黄曲霉毒素^[3]、慢性乙型肝炎病毒^[4]等)被认为是导致肝细胞癌发生和发展的主要原因^[5-6]。虽然诊断技术、外科治疗、化疗和分子靶向治疗技术有所改善,但由于肝细胞癌患者被诊断时多为中晚期,因此患者的 5 年生存率仍不尽如人意^[7-8]。

肝细胞癌的发生是一个多基因参与、多因子协调的复杂的生物学过程。研究发现,血管生成功能失调、慢性炎症、内分泌激素、脂肪因子和细胞代谢的改变均可能参与肿瘤的形成过程^[9]。尽管已经发现了一些肝细胞癌相关的重要驱动基因,但其发生、发展和转移的机制尚不清楚,临床上仍然缺乏与患者预后相关的生物标志物^[10-11]。因此,深入研究肝细胞癌发病及进展过程潜在的生物学机制,对于开发新的诊断标志物和治疗策略具有重要的意义。

近年来,高通量基因芯片和测序技术作为基因表达分析的工具,已被广泛用于识别肿瘤发生过程中遗传信息的改变,为研究肝细胞癌基因表达情况及发现关键基因的变化规律提供了基础^[12]。本研究利用基因表达数据库(Gene Expression Omnibus, GEO)中肝细胞癌组织和非癌组织的基因表达阵列数据筛选肝细胞癌相关差异表达基因(differentially expressed genes, DEGs),并对差异表达基因进行生物信息学分析,同时结合 KM plotter (Kaplan Meier plotter)数据库进行预后分析,以期揭示肝细胞癌发生、发展的分子机制和探索肝细胞癌诊断、治疗和预后相关的潜在候选生物标志物提供理论依据。

1 材料与方法

1.1 数据获取

数据 GSE84402^[13]从 NCBI 基因表达数据库

GEO (<https://www.ncbi.nlm.nih.gov/geo>)中下载得到。该数据采用的芯片平台是 GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array),包括 14 例肝细胞癌组织和 14 例非癌组织。

1.2 HCC 差异表达基因筛选

采用 GEO 数据库中基于 R 语言的网络分析工具 GEO2R^[14] (<https://www.ncbi.nlm.nih.gov/geo/info/geo2r>)分析 GSE84402 中肝细胞癌组织和非癌组织的差异表达基因。该工具利用 Bioconductor 项目中的 R 语言软件包 GEOquery 和 Limma,根据 Benjamini 和 Hochberg 提出的假阳性率控制法^[15],进行多重检验校正。差异表达基因同时满足以下条件: FDR (false discovery rate) <0.05 , $\log_2FC > 1$ 。

1.3 差异表达基因的生物信息学分析

生物学信息注释及可视化数据库(Database for Annotation, Visualization and Integrated Discovery, DAVID)是基因功能分析的在线生物信息学工具,可对差异表达基因进行 GO (Gene Ontology)富集分析及 KEGG (Kyoto Encyclopedia of Genes and Genomes)通路分析。将差异表达基因数据导入 DAVID 6.8 (<https://david.ncifcrf.gov/>)^[16],以 FDR <0.05 为显著性基因富集的临界值,对差异显著的基因进行功能注释,并分析其参与的生物学过程及通路。

1.4 蛋白质互作网络分析及 hub genes 的筛选

蛋白质相互作用数据库 STRING (<https://string-db.org/>)可利用已知或预测蛋白质之间的相互作用构建互作网络^[17]。将差异基因导入 STRING 10.5 分析工具,将最低互作分值(minimum required interaction score)设置成最高可信(highest confidence 0.9),删除与其他蛋白质无相互作用的节点后,分析差异基因所编码的蛋白质之间的相互作用。将互作网络数据导入 Cytoscape 3.6.1 软件^[18],计算节点的度(degree),列出 degree 排名前 20 位的中心节点蛋白质,并与 GO 和 KEGG 富集分析得到的显著性基因取交集,得到的基因即为核心基因(hub genes)。

1.5 Hub genes 与预后的相关性分析

KM plotter 数据库(<http://kmplot.com/>)从 GEO、EGA 及 TCGA 数据库下载基因表达数据、无复发和总生存信息,利用集成基因表达和临床数据的 PostgreSQL 服务器对数据进行处理,依据所提出的生物标志物的不同分位数表达式将患者样本分成两组^[9]。本研究根据 hub genes 的表达水平,以中位数值将 KM plotter 数据库中纳入的 364 例肝癌患者分成高表达组和低表达组,分析两组患者预后的相关性。通过 Kaplan Meier 生存图比较两组病人队列,计算危险比(hazard ratio, HR)及95%置信区间(confidence interval, CI),差异比较用log-rank *P* 检验, logrank *P*<0.05 表示差异具有统计学意义。

2 结果

2.1 差异表达基因筛选结果

通过对肝细胞癌组织和非癌组织的基因表达数据进行分析,共筛选出 1 307 个表达差异明显的基因,其中上调基因 741 个,下调基因 566 个。排名前 10 的差异表达基因见表 1。

2.2 差异表达基因的 GO 富集分析结果

采用 DAVID 6.8 对 741 个上调基因和 566 个下调基因进行 GO 生物学过程富集分析。上调基因主要富集到细胞分裂、姐妹染色单体黏着和分离、有丝分裂核分裂、DNA 复制、细胞周期等 25 个生物学过程;下调基因主要涉及氧化还原、急性期反应、环氧酶 P450 通路、异源代谢过程、药物代谢过程等 23 个生物学过程。上、下调基因前 5 位的生物学过程富集见表 2。

2.3 差异表达基因的 KEGG 通路分析结果

采用 DAVID 6.8 对 741 个上调基因和 566 个下调基因进行 KEGG 通路富集分析。上调基因中有 2 条富集通路呈显著性(FDR<0.05),下调基因中有 10 条富集通路呈显著性(FDR<0.05)。上调基因主要涉及细胞周期和 DNA 复制信号通路,下调基因主要涉及代谢途径、补体途径、视黄醇代谢、化学物致癌作用、胆汁分泌、脂肪酸降解等。上调及下调基因前 5 位的显著富集通路见表 3。

2.4 差异表达基因的蛋白质相互作用网络分析及 hub genes 的筛选

通过 STRING 10.5 及 Cytoscape 3.6.1 软件分

表 1 排名前 10 的差异表达基因

Table 1 Top 10 DEGs

Up-regulated gene	Log ₂ FC	FDR	Down-regulated gene	Log ₂ FC	FDR
<i>MCM10</i>	3.81	9.86E-04	<i>CLEC4M</i>	-4.50	1.28E-06
<i>NUF2</i>	3.72	6.07E-04	<i>MT1M</i>	-4.41	1.90E-04
<i>CDC25A</i>	3.63	8.59E-04	<i>CXCL14</i>	-4.38	1.66E-04
<i>HJURP</i>	3.61	1.53E-04	<i>SLC22A1</i>	-4.35	2.76E-04
<i>IGF2BP3</i>	3.61	1.64E-03	<i>NAT2</i>	-4.35	2.75E-04
<i>CDKN3</i>	3.57	2.94E-05	<i>FOSB</i>	-4.25	2.94E-05
<i>GPSM2</i>	3.55	1.44E-04	<i>TTC36</i>	-4.23	2.23E-05
<i>TTK</i>	3.51	2.84E-04	<i>CFP</i>	-4.21	4.32E-05
<i>GPC3</i>	3.47	5.27E-04	<i>GLYAT</i>	-4.10	8.83E-05
<i>BUB1</i>	3.33	1.22E-03	<i>CRHBP</i>	-4.03	3.02E-04

表 2 前 5 位差异表达基因的生物学过程富集

Table 2 Biological process enrichment results of DEGs (Top 5)

GO ID	GO term	Count	FDR
Up-regulated			
GO: 0051301	Cell division	69	7.271 73E-29
GO: 0007062	Sister chromatid cohesion	36	1.779 25E-22
GO: 0007067	Mitotic nuclear division	46	4.713 43E-17
GO: 0006260	DNA replication	37	6.389 54E-17
GO: 0000082	G1/S transition of mitotic cell cycle	29	8.620 48E-15
Down-regulated			
GO: 0055114	Oxidation-reduction process	57	2.271 52E-11
GO: 0006953	Acute-phase response	15	4.540 67E-09
GO: 0019373	Epoxygenase P450 pathway	11	2.975 75E-08
GO: 0006805	Xenobiotic metabolic process	18	1.672 1E-07
GO: 0017144	Drug metabolic process	11	4.510 22E-06

表 3 前 5 位差异表达基因的 KEGG 通路富集
Table 3 KEGG pathway analysis of DEGs (Top 5)

Term	Count	Genes	FDR
Up-regulated			
hsa04110: Cell cycle	28	<i>YWHAZ, DBF4, TTK, CHEK1, PTTG1, CCNE2, MCM7, CDKN2A, BUB1, ORC6, CCNA2, ORC1, CDK1, CDC6, CDC20, MCM2, CDK4, CDC25C, MCM3, MCM4</i>	7.927E-12
hsa03030: DNA replication	15	<i>LIG1, POLA2, MCM2, RNASEH2A, MCM3, MCM4, MCM5, MCM6, PRIM1, RFC3, RFC4, MCM7, POLE2, PRIM2, FEN1</i>	5.912E-09
Down-regulated			
hsa01100: Metabolic pathways	100	<i>CYP3A4, PTGS2, CNDP1, ADH1C, ADH1B, ADH1A, AGXT, ASPA, GOT1, MAT1A, ST3GAL6, RGN, AADAT, ALDH6A1, CYP1A1, HAL, FBP1, CYP26A1, CYP1A2, TAT</i>	1.068E-09
hsa04610: Complement and coagulation cascades	20	<i>F11, MBL2, F12, C7, C9, MASP1, MASP2, C6, F8, F9, C1R, SERPING1, C1S, PLG, C8A, C8B, FGA, KLKB1, SERPINE1, F2</i>	2.082E-08
hsa00830: Retinol metabolism	18	<i>CYP3A4, CYP1A1, CYP2B6, CYP2C9, CYP2C8, ADH1C, ADH1B, CYP26A1, ADH1A, CYP1A2, RDH5, CYP4A11, ADH4, CYP2A6, HSD17B6, RDH16, UGT2B15, UGT2B28</i>	6.139E-07
hsa05204: Chemical carcinogenesis	17	<i>CYP3A4, PTGS2, CYP1A1, CYP2C9, CYP2C8, NAT2, ADH1C, ADH1B, ADH1A, CYP1A2, CYP3A43, SULT1A1, ADH4, HSD11B1, CYP2A6, UGT2B15, UGT2B28</i>	0.000 127 2
hsa04976: Bile secretion	15	<i>ADCY1, AQP9, ABCB11, SLC22A7, AQP4, NROB2, SLC10A1, ABCB4, SLC01B3, BAAT, KCNN2, CA2, SLC4A4, SLC27A5, SLC22A1</i>	0.000 679 1

注: 每个通路富集分类仅列出前 20 位的基因。

Note: If there were more than twenty genes enriched in this term, only the top twenty were listed in the table.

析, 得到包含 567 个差异表达基因及 3 095 条互作关系的蛋白质相互作用网络(图 1)。结果显示最大 degree 值为 107。Degree 值排名前 20 位的中心节点蛋白质见表 4。将编码前 20 位中心节点蛋白质的差异基因与 GO 和 KEGG 富集得到的显著性基因取交集, 得到 *BUB1* (*mitotic checkpoint serine/threonine kinase*, 丝裂原检查点丝氨酸/苏氨酸激酶)、*BUB1B* (*mitotic checkpoint serine/threonine kinase B*, 丝裂原检查点丝氨酸/苏氨酸激酶 B)、*CCNA2* (*cyclin A2*, 细胞周期蛋白 A2)、*CCNB1* (*cyclin B1*, 细胞周期蛋白 B1)、*CCNB2* (*cyclin B2*, 细胞周期蛋白 B2)、*CDC20* (*cell division cycle 20*, 细胞分裂周期蛋白 20)、*CDK1* (*cyclin-dependent kinase 1*, 周期蛋白依赖激酶 1)、*MAD2L1* (*mitotic arrest deficient 2 like 1*)、*PLK1* (*Polo-like kinase 1*, Polo 样激酶 1) 等 9 个 hub genes, 其相互作用网络包括 36 条互作关系(图 2A)。KEGG 分析发现 hub genes 主要富集在细胞周期等通路(图 2B)。

2.5 Hub genes 与预后的相关性

利用 KM plotter 数据库中纳入的 364 例肝癌患者的基因表达及临床数据, 对 hub genes 进行总生存期的分析, 相关数据如下: *BUB1* (HR=1.85; 95%CI=1.3~2.64; logrank $P=0.000\ 52$), *BUB1B* (HR=1.82; 95%CI=1.28~2.59; logrank $P=0.000\ 71$), *CC-*

表 4 蛋白质互作网络中排名前 20 的中心节点蛋白质
Table 4 The central node proteins in the PPI network (Top 20)

Gene	Degree	Gene	Degree
<i>CDK1</i>	107	<i>BUB1B</i>	54
<i>CCNB1</i>	82	<i>KIF2C</i>	54
<i>CCNB2</i>	73	<i>CENPA</i>	53
<i>PLK1</i>	70	<i>CDCA8</i>	52
<i>CDC20</i>	69	<i>H2AFX</i>	52
<i>AURKB</i>	67	<i>NDC80</i>	51
<i>BUB1</i>	67	<i>HIST1H2BN</i>	51
<i>MAD2L1</i>	67	<i>NUP133</i>	51
<i>CCNA2</i>	61	<i>CKAP5</i>	48
<i>CENPE</i>	55	<i>BIRC5</i>	47

NA2 (HR=1.69; 95%CI=1.19~2.4; logrank $P=0.002\ 9$), *CCNB1* (HR=1.92; 95%CI=1.34~2.74; log-rank $P=0.000\ 26$), *CCNB2* (HR=1.62; 95%CI=1.14~2.29; logrank $P=0.006\ 7$), *CDC20* (HR=2.3; 95%CI=1.6~3.3; logrank $P=3.4E-06$), *CDK1* (HR=1.69; 95%CI=1.19~2.4; logrank $P=0.002\ 9$), *MAD2L1* (HR=1.88; 95%CI=1.33~2.68; logrank $P=0.000\ 33$), *PLK1* (HR=1.89; 95%CI=1.33~2.71; logrank $P=0.000\ 36$), 以上结果显示 9 个 hub genes 在肝癌患者中的表达均上调, 各基因 logrank $P<0.05$, 说明 hub genes 高表达组与低表达组肝癌患者生存期差异具有统计学意义, 提示 hub genes 的高表达与肝癌患者预后

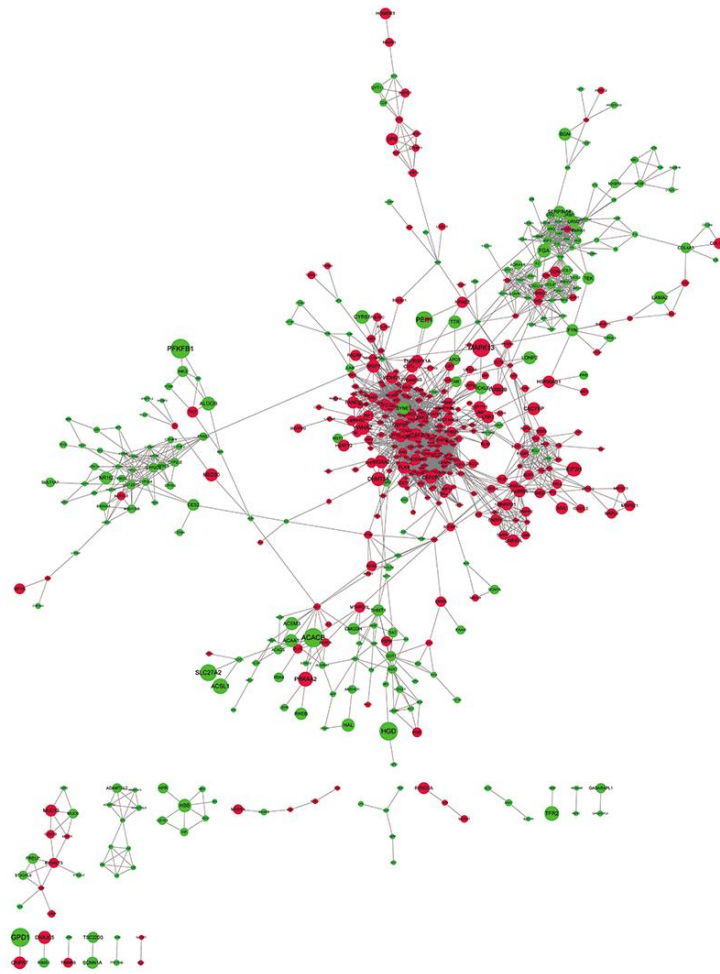


图 1 差异表达基因的蛋白质相互作用网络
红色节点为表达上调的基因, 绿色节点为表达下调的基因。

Fig.1 PPI network of DEGs

Red nodes denote up-regulated genes, while green nodes denote down-regulated genes.

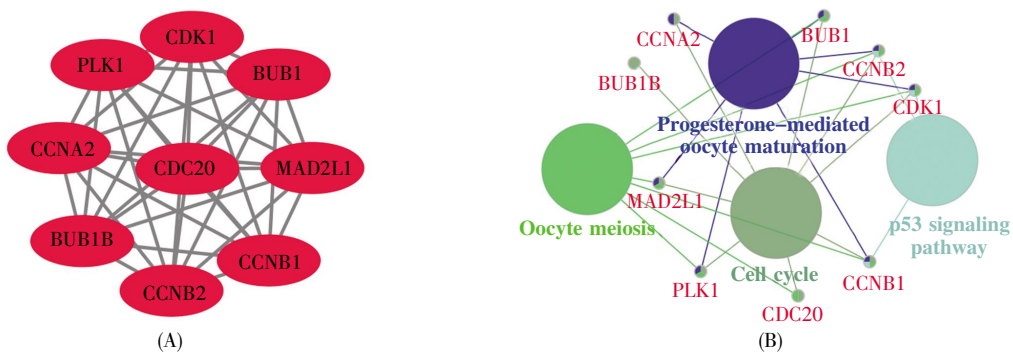


图 2 Hub genes 分析
(A) Hub genes 的蛋白质相互作用网络; (B) Hub genes 的 KEGG 通路分析。

Fig.2 Result analysis of hub genes

(A) PPI network of hub genes; (B) KEGG pathway analysis of hub genes.

不良相关(图 3)。

3 讨论

肝细胞癌是世界范围内常见恶性肿瘤之一,

每年全世界约有 778 000 例新发病例和 745 000 例死亡病例, 其中仅中国就占一半^[1]。虽然肝细胞癌的诊断和治疗技术取得了显著进展, 但因其具有浸润、转移和复发的特性, 最终导致肝细胞癌患

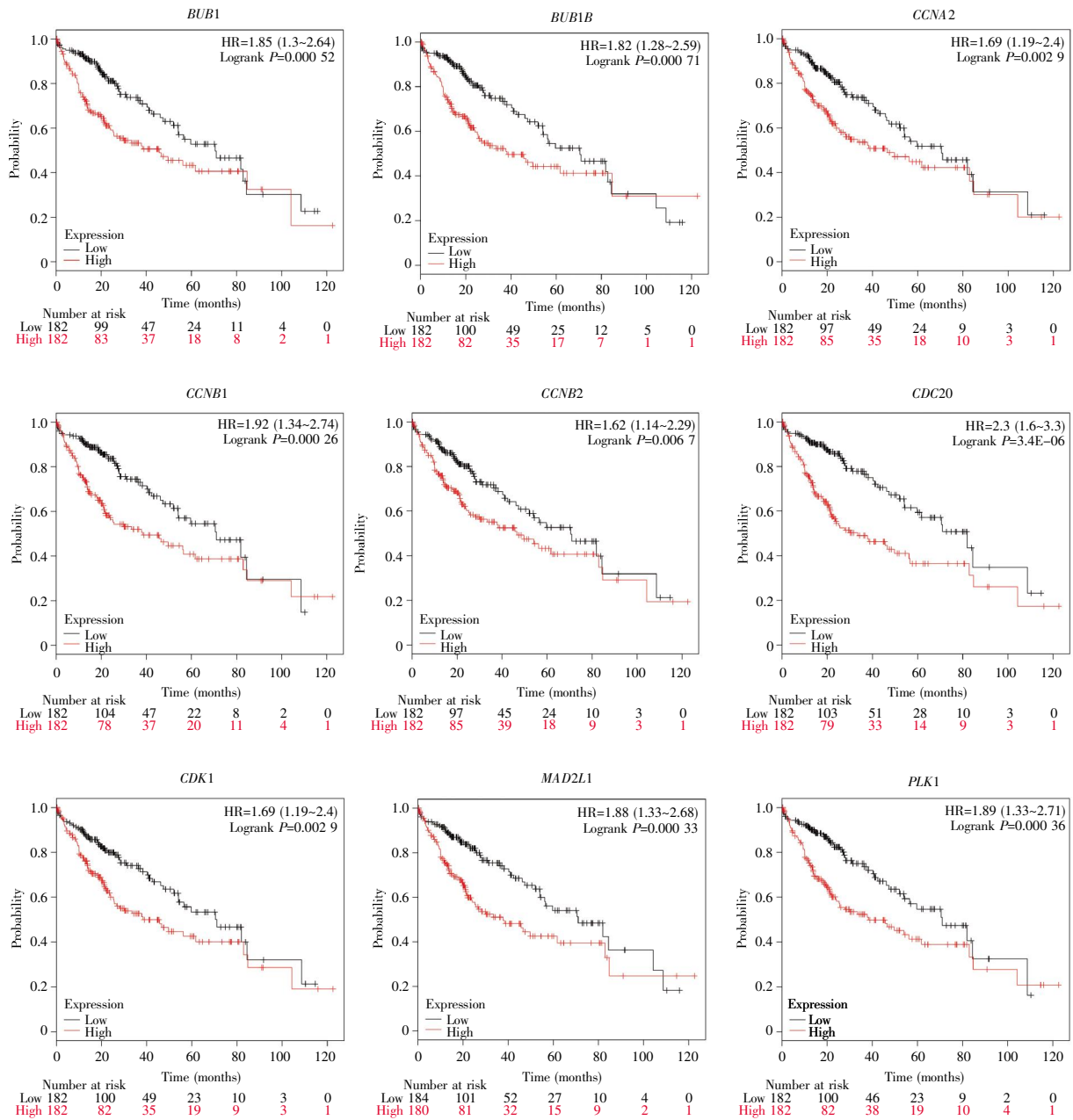


图 3 Hub genes 的总生存期分析

HR: 危险比; Logrank $P < 0.05$ 表明差异具有统计学意义。红色曲线表示高表达组, 黑色曲线代表低表达组。

Fig.3 Analysis of effects of hub genes on overall survival

HR: Hazard ratio; Logrank $P < 0.05$ stands for significant difference. Black and red lines represent low and high expression groups, respectively.

者的预后仍然较差。

本研究从 GEO 数据库检索包含肝细胞癌组织和非癌组织的基因表达谱数据 GSE84402, 利用生物信息学工具进行深入挖掘分析, 共筛选出 1 307 个差异表达基因。通过 GO 分析发现差异表达基因主要参与了有丝分裂、细胞周期、DNA 复制及物质代谢等生物学过程。KEGG 通路分析显示差异表达基因主要涉及细胞周期、DNA 复制及

物质代谢等通路。通路分析结果与 GO 富集分析结果一致, 表明细胞周期、DNA 复制及物质代谢等功能的异常, 在肝癌发生发展的过程中发挥着重要作用。进一步筛选出 *BUB1*、*BUB1B*、*CCNA2*、*CCNB1*、*CCNB2*、*CDC20*、*CDK1*、*MAD2L1*、*PLK1* 等 9 个 hub genes, 分析发现 9 个 hub genes 均与细胞周期调控相关。有研究表明细胞周期相关蛋白质的异常表达与多种肿瘤的发生和发展密切相关

关,且多是由于编码蛋白质的上游信号通路的突变或基因损伤所致^[20]。采用 KM plotter 数据库对 hub genes 进行总生存期的分析,结果显示 9 个 hub genes 在肝细胞癌患者中均为高表达基因,且与肝细胞癌的预后显著相关。

BUB1 基因编码的丝氨酸/苏氨酸激酶参与细胞有丝分裂过程^[21],其表达增加可改变有丝分裂纺锤体组装检查点通路,在细胞增殖和肿瘤进展中发挥重要的作用^[22]。Xu 等^[23]研究发现,miR-490-5p 可通过调节肝癌细胞 TGF β /Smad 通路,抑制 *BUB1* 表达,从而抑制肝癌细胞的增殖、浸润和迁移能力,进一步降低细胞存活率并促进细胞凋亡。*BUB1B* 基因定位于 15 号染色体的 15q15.1 区,*BUB1B* 蛋白是一种具有纺锤体分离调控功能的蛋白激酶^[24]。大量研究表明 *BUB1B* 参与了多种肿瘤形成过程^[25-28],但其在肝癌中的作用机制鲜有报道,仍需进一步实验研究。*CCNA2* 在细胞有丝分裂的 S/G2 期高水平表达,*CCNA2* 异常表达会导致染色体结构不稳定,进而诱导肿瘤的发生^[29]。Hung 等^[30]发现肝癌细胞中精氨酸琥珀酸裂解酶(argininosuccinate lyase, ASL)与 *CCNA2* 相互作用,并通过非酶途径促进肝癌的形成。*CCNB1* 可促进细胞从 G2 期向 M 期转变,但在肿瘤细胞中 *CCNB1* 过度表达,并与 *CDK1* 结合导致细胞无限增殖^[31]。Chai 等^[32]发现沉默肝癌细胞的 *FOXM1* 基因可显著降低 *CCNB1* 的表达水平,表明 *CCNB1* 在 *FOXM1* 诱导的肝癌细胞增殖中有重要作用,证明 *FOXM1-CCNB1* 与肝癌患者预后不良相关。*CCNB2* 在多种肿瘤组织中过表达,且与肿瘤的侵袭和临床治疗效果差相关,肝癌细胞中 *KPNA2* 可通过促进 *CCNB2/CDK1* 的表达,诱导肝癌细胞增殖,且与肝癌患者总生存时间和无病生存时间显著降低有关^[33]。*CDC20* 编码与细胞周期的 APC/C 相互作用的调节蛋白,肝细胞中 *CDC20* 高表达可加速细胞增殖,促进肝癌的发生和发展,但其分子机制及预后意义有待深入研究^[34]。*CDK1* 是细胞周期的重要调节因子^[35-36]。研究表明:肝细胞癌组织的 *CDK1* 表达水平明显高于非癌组织,*CDK1* 在凋亡蛋白诱导的肝癌细胞凋亡中起着重要的调控作用^[37];*CDK1* 的高表达与肝癌患者整体存活率低有关,利用 *CDK1* 抑制剂阻断 *CDK1/pdk1/ β -Cat* 信号转导,可提高索拉菲尼治疗肝癌临床前模型的疗效,为临床提出个性化治疗方案提供依据^[38]。*MAD2L1* 被认为是染色体控制通路的重要

介质,*MDA2L1* 表达水平异常可导致染色体不稳定,并促进肿瘤的进展,其在肝癌中的表达水平与肿瘤的大小、分期和分级密切相关^[39]。Li 等^[40]报道 *MDA2L1* 在肝癌组织和细胞中高表达,miR-200c-5p 可通过下调 *MAD2L1* 的表达,抑制肝癌细胞增殖、迁移和浸润,并促进细胞凋亡。*PLK1* 被认为是肝癌中的一个重要致癌因子^[41],其在调节细胞有丝分裂、胞质分裂、DNA 损伤应答等细胞周期过程中发挥重要作用^[42-43]。研究发现抑制 *PLK1* 表达,可抑制肝癌细胞增殖^[44];*PLK1* 可作为肝癌的预后标志物与治疗靶点,但其作用机制仍需进一步阐明^[45-46]。

综上所述,本研究利用生物信息学的方法对肝细胞癌基因表达谱数据进行挖掘和分析,得到了 1 307 个差异表达基因,其中 *BUB1*、*BUB1B*、*CCNA2*、*CCNB1*、*CCNB2*、*CDC20*、*CDK1*、*MAD2L1* 和 *PLK1* 等 9 个 hub genes 经总生存期分析发现与肝细胞癌患者预后不良有关,该结果为揭示肝细胞癌发生和发展的分子机制,筛选有效的诊断和预后相关的生物标志物及药物治疗靶点提供了参考。

参考文献(References):

- [1] TORRE L A, BRAY F, SIEGEL R L, et al. Global cancer statistics, 2012 [J]. CA: A Cancer Journal for Clinicians, 2015, 65(2): 87-108.
- [2] TANG R, LIU H, YUAN Y, et al. Genetic factors associated with risk of metabolic syndrome and hepatocellular carcinoma[J]. Oncotarget, 2017, 8(21): 35403-35411.
- [3] MOUDGIL V, REDHU D, DHANDA S, et al. A review of molecular mechanisms in the development of hepatocellular carcinoma by aflatoxin and hepatitis B and C viruses[J]. Journal of Environmental Pathology Toxicology and Oncology, 2013, 32(2): 165-175.
- [4] STEPIEN M, HUGHES D J, HYBSIER S, et al. Circulating copper and zinc levels and risk of hepatobiliary cancers in Europeans[J]. British Journal of Cancer, 2017, 116(5): 688-696.
- [5] LLOVET J M, ZUCMAN-ROSSI J, PIKARSKY E, et al. Hepatocellular carcinoma[J]. Nature Reviews Disease Primers, 2016, 2: 16018.
- [6] FARAZI P A, DEPINHO R A. Hepatocellular carcinoma pathogenesis: from genes to environment[J]. Nature Reviews Cancer, 2006, 6(9): 674-687.
- [7] SIEGEL R, MA J, ZOU Z, et al. Cancer statistics, 2014[J]. CA: A Cancer Journal for Clinicians, 2014, 64(1): 9-29.
- [8] CHEN W, ZHENG R, BAADE P D, et al. Cancer statistics in China, 2015[J]. CA: A Cancer Journal for Clinicians, 2016, 66(2): 115-132.
- [9] TAHERGORABI Z, KHAZAEI M, MOODI M, et al. From obesity to cancer: a review on proposed mechanisms[J]. Cell Biochemistry and Function, 2016, 34(8): 533-545.
- [10] DUTTA R, MAHATO R I. Recent advances in hepatocellular carcinoma therapy[J]. Pharmacology & Therapeutics, 2017, 173: 106-117.

- [11] KUDO M, TREVISANI F, ABOU-ALFA G K, *et al.* Hepatocellular carcinoma: therapeutic guidelines and medical treatment[J]. *Liver Cancer*, 2016, 6(1): 16–26.
- [12] CANALES R D, LUO Y, WILEY J C, *et al.* Evaluation of DNA microarray results with quantitative gene expression platforms[J]. *Nature Biotechnology*, 2006, 24(9): 1115–1122.
- [13] WANG H, HUO X, YANG X R, *et al.* STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4[J]. *Molecular Cancer*, 2017, 16(1): 136–151.
- [14] DAVIS S, MELTZER P S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor[J]. *Bioinformatics*, 2007, 23(14): 1846–1847.
- [15] KIM K I, VAN DE WIEL M A. Effects of dependence in high-dimensional multiple testing problems[J]. *BioMed Central Bioinformatics*, 2008, 9: 114–126.
- [16] JIAO X, SHERMAN B T, HUANG D W, *et al.* DAVID-WS: a stateful web service to facilitate gene/protein list analysis[J]. *Bioinformatics*, 2012, 28(13): 1805–1806.
- [17] SZKLARCZYK D, MORRIS J H, COOK H, *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible[J]. *Nucleic Acids Research*, 2017, 45(D1): D362–D368.
- [18] SU G, MORRIS J H, DEMCHAK B, *et al.* Biological network exploration with Cytoscape 3[J]. *Current Protocols in Bioinformatics*, 2014, 47: 8.13.1–24.
- [19] SZASZ A M, LANCZKY A, NAGY A, *et al.* Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients[J]. *Oncotarget*, 2016, 7(31): 49322–49333.
- [20] OTTO T, SICINSKI P. Cell cycle proteins as promising targets in cancer therapy[J]. *Nature Reviews Cancer*, 2017, 17(2): 93–115.
- [21] YU H, TANG Z. Bub1 multitasking in mitosis[J]. *Cell Cycle*, 2005, 4(2): 262–265.
- [22] TAKAGI K, MIKI Y, SHIBAHARA Y, *et al.* BUB1 immunolocalization in breast carcinoma: its nuclear localization as a potent prognostic factor of the patients[J]. *Hormones & Cancer*, 2013, 4(2): 92–102.
- [23] XU B, XU T, LIU H, *et al.* miR-490-5p suppresses cell proliferation and invasion by targeting BUB1 in hepatocellular carcinoma cells[J]. *Pharmacology*, 2017, 100(5–6): 269–282.
- [24] SHICHIRI M, YOSHINAGA K, HISATOMI H, *et al.* Genetic and epigenetic inactivation of mitotic checkpoint genes hBUB1 and hBUBR1 and their relationship to survival[J]. *Cancer Research*, 2002, 62(1): 13–17.
- [25] MANSOURI N, MOVAFAGH A, SAYAD A, *et al.* Targeting of BUB1b gene expression in sentinel lymph node biopsies of invasive breast cancer in iranian female patients[J]. *Asian Pacific Journal of Cancer Prevention*, 2016, 17(suppl. 3): 317–321.
- [26] HUDLER P, BRITOVSEK N K, GRAZIO S F, *et al.* Association between polymorphisms in segregation genes BUB1B and TTK and gastric cancer risk[J]. *Radiology and Oncology*, 2016, 50(3): 297–307.
- [27] HAHN M M, VREEDE L, BEMELMANS S A, *et al.* Prevalence of germline mutations in the spindle assembly checkpoint gene BUB1B in individuals with early-onset colorectal cancer[J]. *Genes Chromosomes & Cancer*, 2016, 55(11): 855–863.
- [28] MA Q, LIU Y, SHANG L, *et al.* The FOXM1/BUB1B signaling pathway is essential for the tumorigenicity and radioresistance of glioblastoma[J]. *Oncology Reports*, 2017, 38(6): 3367–3375.
- [29] KANAKKANTHARA A, JEGANATHAN K B, LIMZERWALA J F, *et al.* Cyclin A2 is an RNA binding protein that controls Mre11 mRNA translation[J]. *Science*, 2016, 353(6307): 1549–1552.
- [30] HUNG Y H, HUANG H L, CHEN W C, *et al.* Argininosuccinate lyase interacts with cyclin A2 in cytoplasm and modulates growth of liver tumor cells[J]. *Oncology Reports*, 2017, 37(2): 969–978.
- [31] PARK T J, KIM J Y, OH S P, *et al.* TIS21 negatively regulates hepatocarcinogenesis by disruption of cyclin B1-Forkhead box M1 regulation loop[J]. *Hepatology*, 2008, 47(5): 1533–1543.
- [32] CHAI N, XIE H H, YIN J P, *et al.* FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1[J]. *Biochemical and Biophysical Research Communications*, 2018, 500(4): 924–929.
- [33] GAO C L, WANG G W, YANG G Q, *et al.* Karyopherin subunit- α 2 expression accelerates cell cycle progression by upregulating CCNB2 and CDK1 in hepatocellular carcinoma[J]. *Oncology Letters*, 2018, 15(3): 2815–2820.
- [34] LI J, GAO J Z, DU J L, *et al.* Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma[J]. *International Journal of Oncology*, 2014, 45(4): 1547–1555.
- [35] SAKURIKAR N, EASTMAN A. Critical reanalysis of the methods that discriminate the activity of CDK2 from CDK1[J]. *Cell Cycle*, 2016, 15: 1184–1188.
- [36] BANYAI G, BAIDI F, COUDREUSE D, *et al.* Cdk1 activity acts as a quantitative platform for coordinating cell cycle progression with periodic transcription[J]. *Nature Communications*, 2016, 7: 11161–11172.
- [37] ZHAO J, HAN S X, MA J L, *et al.* The role of CDK1 in apoptin-induced apoptosis in hepatocellular carcinoma cells[J]. *Oncology Reports*, 2013, 30: 253–259.
- [38] WU C X, WANG X Q, CHOK S H, *et al.* Blocking CDK1/PDK1/ β -catenin signaling by CDK1 inhibitor R03306 increased the efficacy of sorafenib treatment by targeting cancer stem cells in a preclinical model of hepatocellular carcinoma[J]. *Theranostics*, 2018, 8(14): 3737–3750.
- [39] YAN H, LI Z, SHEN Q, *et al.* Aberrant expression of cell cycle and material metabolism related genes contributes to hepatocellular carcinoma occurrence[J]. *Pathology Research and Practice*, 2017, 213: 316–321.
- [40] LI Y, BAI W, ZHANG J. miR-200c-5p suppresses proliferation and metastasis of human hepatocellular carcinoma via suppressing MAD2L1[J]. *Biomedicine & Pharmacotherapy*, 2017, 92: 1038–1044.
- [41] WANG D, CHANG R, WANG G, *et al.* Polo-like kinase 1-targeting chitosan nanoparticles suppress the progression of hepatocellular carcinoma[J]. *Anti-Cancer Agents in Medicinal Chemistry*, 2017, 17(7): 948–954.
- [42] ARCHAMBAULT V, LEPINE G, KACHANER D. Understanding the polo kinase machine[J]. *Oncogene*, 2015, 34(37): 4799–4807.
- [43] IKEDA M, TANAKA K. Plk1 bound to Bub1 contributes to spindle assembly checkpoint activity during mitosis[J]. *Scientific Reports*, 2017, 7: 8794.
- [44] BOUHLAL H, OULED-HADDOU H, DEBUYSSCHER V, *et al.* RB/PLK1-dependent induced pathway by SLAMF3 expression inhibits mitosis and control hepatocarcinoma cell proliferation[J]. *Oncotarget*, 2016, 7: 9832–9843.
- [45] ZHANG L, HUANG Y, LING J, *et al.* Screening and function analysis of hub genes and pathways in hepatocellular carcinoma via bioinformatics approaches[J]. *Cancer Biomark*, 2018, 22(3): 511–521.
- [46] LIN P, WEN D Y, DANG Y W, *et al.* Comprehensive and integrative analysis reveals the diagnostic, clinicopathological and prognostic significance of polo-like kinase 1 in hepatocellular carcinoma[J]. *Cellular Physiology and Biochemistry*, 2018, 47(3): 925–947.