

基于生物信息学方法筛选和验证肝癌预后标志物

米宁宁¹, 白明圳¹, 高 龙¹, 马海东¹, 付文康¹, 林延延^{1,2}, 孟文勃^{1,2*}

(1. 兰州大学第一临床医学院, 中国甘肃 兰州 730099; 2. 兰州大学第一医院 普外科, 中国甘肃 兰州 730099)

摘要: 运用生物信息学方法探究肝细胞癌(hepatocellular carcinoma, HCC)发生发展的核心基因及预后标志物。下载 GEO (Gene Expression Omnibus)数据库中的 GSE112790 芯片数据及癌症基因组图谱(The Cancer Genome Atlas, TCGA)数据库中的肝癌数据, 分析得到 151 个差异表达基因(differentially expressed gene, DEG)并筛选出 10 个核心基因。生存分析表明, *BUB1B*、*CDC20*、*ASPM* 和 *DLGAP5* 基因高表达显著影响患者预后。Oncomine 数据库分析结果证实, *BUB1B*、*CDC20* 和 *DLGAP5* 的表达水平与肿瘤分级和血管浸润明显相关。HPA 数据库及肝癌组织芯片的免疫组织化学实验结果均显示, 相对于正常肝组织, 肝癌组织中 *CDC20* 和 *DLGAP5* 蛋白高表达。Cox 分析结果提示, *CDC20* 和 *DLGAP5* 可作为肝癌患者预后的独立危险因素。此外, *CDC20* 甲基化水平是影响其表达水平的重要因素, 并且和多种免疫细胞的表达相关。上述研究结果表明, *CDC20* 可作为肝癌患者预后评估的潜在生物标志物或治疗靶点。

关键词: 肝细胞癌(HCC); 核心基因; 预后; 生物信息学; 免疫细胞

中图分类号: Q811.4, R735.7

文献标识码: A

文章编号: 1007-7847(2022)06-0538-11

Identification and Validation of Prognostic Markers in Hepatocellular Carcinoma Based on Bioinformatics Analysis

MI Ning-ning¹, BAI Ming-zhen¹, GAO Long¹, MA Hai-dong¹, FU Wen-kang¹,
LIN Yan-yan^{1,2}, MENG Wen-bo^{1,2*}

(1. The First School of Clinical Medicine, Lanzhou University, Lanzhou 730099, Gansu, China; 2. Department of General Surgery, the First Hospital of Lanzhou University, Lanzhou 730099, Gansu, China)

Abstract: In order to identify the hub genes and biomarkers of hepatocellular carcinoma (HCC) by bioinformatics methods, the GSE112790 dataset from the Gene Expression Omnibus (GEO) and HCC data from The Cancer Genome Atlas (TCGA) database were screened for differential expression analysis. A total of 151 differentially expressed genes (DEGs) were identified and 10 hub genes were screened. Survival analysis indicated that upregulated expression of *BUB1B*, *CDC20*, *ASPM* and *DLGAP5* genes significantly affected the prognosis of patients. Analysis of Oncomine database showed that expression levels of *BUB1B*, *CDC20* and *DLGAP5* were obviously related to tumor grade and vascular invasion. From HPA database and immunohistochemistry of liver cancer tissue microarray, it was found that expression levels of *CDC20* and *DLGAP5* in HCC tissues were higher than those in normal tissues. Cox analysis showed that the high expression of *CDC20* and *DLGAP5* could be used as an independent risk factor for the prognosis of patients with HCC. Moreover, the *CDC20* methylation level was also found to be an important factor affecting its expression level and a variety of immune cell contents. In conclusion, *CDC20* may be a potential biomarker for HCC prognosis evaluation and therapeutic target.

Key words: hepatocellular carcinoma (HCC); hub gene; prognosis; bioinformatics; immune cell

(*Life Science Research*, 2022, 26(6): 538-548)

收稿日期: 2021-11-03; 修回日期: 2022-01-05; 网络首发日期: 2022-11-09

基金项目: 国家自然科学基金资助项目(82060551, 32160255); 甘肃省自然科学基金项目(20JR10RA676, 0JR10RA674); 兰州市城关区项目(2019JSCX0092, 2019RCCX0038, 2019SHFZ0033)

作者简介: 米宁宁(1995—), 男, 甘肃天水人, 博士研究生; *通信作者: 孟文勃(1978—), 男, 河北景州人, 博士, 主任医师, 教授, 硕/博士研究生导师, 主要从事肝胆胰外科、消化道肿瘤、内镜外科研究, Tel: 0931-8356947, E-mail: mengwb@lzu.edu.cn.

肝癌是常见的消化系统恶性肿瘤,致病因素包括慢性病毒性肝炎、长期酒精摄入、微量元素缺乏、性激素调节异常、亚硝胺类物质与黄曲霉素的直接毒性作用等^[1]。肝细胞癌(hepatocellular carcinoma, HCC)是最常见的原发性肝脏恶性肿瘤,占肝脏肿瘤的90%以上^[2]。HCC发病率逐年升高,临床上早期无明显特征性症状及体征,虽然血清甲胎蛋白(alpha-fetoprotein, AFP)是当前诊断肝癌常用而又重要的临床检测指标,但是仍有约30%的肝癌病人AFP水平正常,所以肝癌的早期诊断和治疗仍值得深入研究,尤其在发展中国家,医疗卫生条件较差,肝癌患者预后往往更差^[3]。探究HCC发生发展的分子机制对于疾病的诊断和治疗非常重要,微阵列和高通量测序技术已经应用于大多数恶性肿瘤分子标志物的筛选及分子机制的研究。本研究通过生物信息学方法,利用癌症基因组图谱(The Cancer Genome Atlas, TCGA)数据库的HCC基因表达谱数据和GEO(Gene Expression Omnibus)数据库的GSE112790芯片数据,分析得到HCC与正常组织之间的差异表达基因(differentially expressed gene, DEG),并通过功能富集分析、蛋白质-蛋白质相互作用(protein-protein interaction, PPI)分析筛选得到与HCC发生发展相关的核心基因,随后经进一步分析得到与肿瘤患者临床特征和预后密切相关的基因,最后通过免疫组织化学(immunohistochemistry, IHC)染色验证其在蛋白质水平的表达,从而获得HCC的预后标志物。

1 材料与方法

1.1 数据的来源

从GEO数据库(<http://www.ncbi.nlm.nih.gov/geo/>; 2020-07-09)下载基于GPL570平台(Affymetrix Human Genome U133 Plus 2.0 Array)的GSE-112790^[4],该数据集包含15例正常肝组织和183例肿瘤组织的测序结果;同时,下载TCGA数据库(<https://cancergenome.nih.gov/>; 2020-07-09)的374例HCC组织和50例正常肝组织的表达谱数据。

1.2 差异表达基因的筛选

使用R软件(版本3.5.2)的“limma”包(版本3.36.5)对GSE112790芯片数据做差异分析,使用RMA(Robust Multi-array Average)法完成芯片数据的背景校正及标准化^[5];使用“edgeR”包(版本3.24.3)完成TCGA数据的差异分析^[6]。得到的差异

表达基因需满足: adjusted $P < 0.01$ 并且 $|\log_2(\text{fold change})| > 2.0$ 。

1.3 差异表达基因的功能富集分析

差异表达基因的功能富集分析主要包括基因本体论(Gene Ontology, GO)与京都基因和基因组数据库(Kyoto Encyclopedia of Genes and Genomes, KEGG)信号通路富集分析,通过DAVID(Database for Annotation, Visualization and Integrated Discovery; <https://david.ncifcrf.gov/>; 2020-07-12)在线数据库和R语言的“clusterProfiler”包完成。DAVID拥有较为全面的生物信息数据资源和分析工具,旨在提供基因或蛋白质的功能注释^[7]。 $P < 0.05$ 为富集结果具有统计学意义。

1.4 蛋白质互作网络的构建

利用STRING(Search Tool for the Retrieval of Interacting Genes; <https://string-db.org/>; 2020-07-12)在线数据库构建蛋白质-蛋白质相互作用网络,设置交互作用的综合评分 > 0.400 。使用Cytoscape软件(版本3.6.0)对网路进行可视化。Cytoscape的插件MCODE(Molecular Complex Detection; 版本1.5.1)可以完成网络的聚类,从而得到拓扑学意义上连接最紧密的结构,利用该插件可以挖掘与肝癌显著相关的核心网络,并进一步研究该网络的生物学功能。筛选标准如下: degree cutoff=2; node score cutoff=0.2; maximum depth=100; k-score=2。

1.5 核心基因的筛选和分析

以蛋白质互作网络的连接度作为核心基因的筛选标准,得到连接度最高的前10个基因。通过cBioPortal在线数据库(<http://www.cbioportal.org/>; 2020-08-03)得到核心基因表达水平与患者总生存期(overall survival, OS)和无病生存期(disease-free survival, DFS)的关系^[8-9]。使用OncoPrint数据库(<https://www.oncoprint.org/>; 2020-08-03)分析核心基因的表达水平与肿瘤分级和血管浸润之间的关系。利用HPA(Human Protein Atlas; <https://www.proteinatlas.org/>; 2020-08-07)数据库^[10]提供的IHC信息,比较核心基因在肝癌组织和正常组织中的蛋白质表达差异。利用TCGA的数据,通过单因素和多因素Cox风险模型分析核心基因的临床参数与HCC预后的关系,并计算风险比(hazard ratio, HR)与95%置信区间(95% confidence interval, 95% CI)。

1.6 免疫组织化学染色

为了进一步验证核心基因在蛋白质水平的表

达情况,我们通过购自上海芯超生物科技有限公司的 HCC 组织芯片(HLivH180Su15, HLivH028P-G01)完成对与肝癌发生发展密切相关的两个核心基因的验证。利用二甲苯和无水乙醇梯度试剂进行组织样品脱蜡,采用 0.01 mol/L 柠檬酸盐缓冲液进行组织抗原修复,向组织切片滴加内源性过氧化物酶阻断反应后,用山羊血清进行封闭;分别加入 DLGAP5 抗体(DF6920;江苏亲科生物研究中心有限公司)、CDC20 抗体(10252-1-AP;武汉三鹰生物技术有限公司)并置于 4 °C 冰箱过夜,随后用辣根酶标羊抗小鼠/兔 IgG 聚合物(福州迈新生物技术开发有限公司)孵育 1 h;向组织切片滴加新鲜配制的 DAB 显色液显色,经过复染、脱水后,利用中性树胶封片。由两名病理科医生采用盲法判读结果。IHC 评分根据染色强度和阳性细胞比例两项指标综合判定。

1.7 核心基因表达与甲基化及免疫细胞的相关性分析

通过 UCSC 数据库(<http://xena.ucsc.edu/>; 2021-09-26)下载肝癌的甲基化数据,分析核心基因的甲基化水平,并通过 Spearman 法得到甲基化与基因表达是否存在相关性。使用 R 软件 survival 包(版本 3.2-11)分析甲基化位点与 HCC 患者 OS 和无进展生存期(progression-free survival, PFS)的关系。使用 TIMER 数据库分析核心基因表达与免疫细胞含量之间的关系。

2 结果

2.1 差异表达基因筛选结果

GSE112790 芯片包含 15 例正常肝组织和 183 例肿瘤组织,通过差异分析得到满足筛选条件的 264 个差异基因,其中上调的差异基因 143

个,下调的差异基因 121 个。TCGA 数据包含 374 例 HCC 样品和 50 例正常肝组织样品,经分析得到 3 297 个上调的差异基因和 300 个下调的差异基因。整合分析后得到 151 个共有的差异基因,包含 85 个上调基因和 66 个下调基因(表 1)。

2.2 差异表达基因的功能富集

GO 富集分析由生物过程、细胞组分和分子功能 3 个部分组成。分析结果显示,差异表达基因与 53 个生物过程、21 个细胞组分、15 个分子功能有关。在生物过程层面,差异基因主要与细胞分裂、染色体分离、DNA 复制、环氧化酶 P450 途径以及负性生长调控等有关;在细胞组分层面,这些基因大多参与染色体、动粒、纺锤体微管等的组成;在分子功能层面,其主要集中在蛋白激酶结合、微管结合、组蛋白激酶活性等(排名前 10 位的条目见图 1A)。KEGG 结果表明,差异基因主要富集在 P53 信号通路、AMPK 信号通路、细胞周期、细胞沉默等信号通路(图 1B)。

2.3 蛋白质互作网络和核心模块的构建

使用 STRING 在线数据库构建差异表达基因的蛋白质互作网络,并使用 Cytoscape 完成图形可视化。结果如图 2A 所示,网络由 112 个节点和 1 157 条边组成。进一步通过 MCODE 插件得到网络的核心模块(图 2B),并对构成核心模块的基因完成 KEGG 分析,结果主要富集在 P53 信号通路、病毒致癌机制、细胞周期等信号通路。

2.4 核心基因的筛选与验证

按照蛋白质互作网络中的连接度筛选出前 10 个基因作为核心基因,分别为 *CCNB1*、*CDK1*、*BUB1B*、*CDC20*、*CCNA2*、*KIF11*、*DLGAP5*、*DTL*、*NUSAP1* 和 *ASPM*。生存分析结果表明, *BUB1B*、*ASPM*、*CDC20* 和 *DLGAP5* 跟患者的 OS 和 DFS

表 1 TCGA 和 GEO 数据库共有的 151 个差异基因
Table 1 A total of 151 DEGs from both TCGA and GEO databases

DEG	Gene
Upregulated	<i>CCNB1</i> , <i>MNS1</i> , <i>GINS1</i> , <i>ZIC2</i> , <i>ANLN</i> , <i>ST8SIA6-AS1</i> , <i>FOXM1</i> , <i>CDK1</i> , <i>RACGAP1</i> , <i>C12orf75</i> , <i>SULT1C2</i> , <i>KIF14</i> , <i>IRX3</i> , <i>CENPL</i> , <i>KIF4A</i> , <i>SPP1</i> , <i>MELK</i> , <i>ZWINT</i> , <i>NDC80</i> , <i>CCNA2</i> , <i>NUF2</i> , <i>PTTG1</i> , <i>TRIM16</i> , <i>UBE2T</i> , <i>CRNDE</i> , <i>STK39</i> , <i>PDZK1IP1</i> , <i>PRC1</i> , <i>CENPW</i> , <i>G6PD</i> , <i>MCM2</i> , <i>DLGAP5</i> , <i>RAD51AP1</i> , <i>CKAP2L</i> , <i>GPC3</i> , <i>PLVAP</i> , <i>NEK2</i> , <i>ACSL4</i> , <i>E2F8</i> , <i>SPINK1</i> , <i>S100P</i> , <i>DEPDC1B</i> , <i>LINC01419</i> , <i>FLVCR1</i> , <i>AKR1B10</i> , <i>COL15A1</i> , <i>EZH2</i> , <i>RBM24</i> , <i>KIF11</i> , <i>COX7B2</i> , <i>PAGE4</i> , <i>ROBO1</i> , <i>CDC20</i> , <i>PBK</i> , <i>TRIP13</i> , <i>CDC7</i> , <i>SHCBP1</i> , <i>ASPM</i> , <i>CD24</i> , <i>UBE2C</i> , <i>NQO1</i> , <i>PHYHIP1</i> , <i>RRM2</i> , <i>MDK</i> , <i>TOP2A</i> , <i>MAGEA6</i> , <i>TRIM71</i> , <i>KIF15</i> , <i>BUB1B</i> , <i>HJURP</i> , <i>CAP2</i> , <i>DTL</i> , <i>FAM83D</i> , <i>HMMR</i> , <i>GMNN</i> , <i>KIF20A</i> , <i>LCN2</i> , <i>SP5</i> , <i>PEG10</i> , <i>CTHRC1</i> , <i>UHRF1</i> , <i>TTK</i> , <i>CDKN3</i> , <i>NCAFG</i> , <i>NUSAP1</i>
Downregulated	<i>MT1G</i> , <i>CYP26A1</i> , <i>SERPINE1</i> , <i>DIRAS3</i> , <i>LINC01093</i> , <i>CXCL14</i> , <i>TSLP</i> , <i>SLC22A1</i> , <i>HAO2</i> , <i>MT1F</i> , <i>SLC25A47</i> , <i>FITM1</i> , <i>KCNN2</i> , <i>SLCO1B3</i> , <i>MYOM2</i> , <i>CYP1A2</i> , <i>MT1E</i> , <i>CNDP1</i> , <i>BCO2</i> , <i>FCN3</i> , <i>ACSM3</i> , <i>CYP2C19</i> , <i>GBA3</i> , <i>TTC36</i> , <i>CLEC4G</i> , <i>CYP2B6</i> , <i>GYS2</i> , <i>PCK1</i> , <i>FOLH1B</i> , <i>LPA</i> , <i>GHR</i> , <i>CLEC1B</i> , <i>FOSB</i> , <i>MARCO</i> , <i>MT1H</i> , <i>CLEC4M</i> , <i>NPY1R</i> , <i>LYVE1</i> , <i>ASPG</i> , <i>FOS</i> , <i>PLAC8</i> , <i>HAMP</i> , <i>CFP</i> , <i>NAT2</i> , <i>SLITRK3</i> , <i>RND3</i> , <i>IL1RAP</i> , <i>RDH16</i> , <i>CXCL12</i> , <i>CRHBP</i> , <i>AVPR1A</i> , <i>CXCR2</i> , <i>CYP4A11</i> , <i>MT1X</i> , <i>IGFALS</i> , <i>MT2A</i> , <i>PZP</i> , <i>FBP1</i> , <i>AADAT</i> , <i>OIT3</i> , <i>MT1M</i> , <i>CETP</i> , <i>APOF</i> , <i>MT1HL1</i> , <i>C9</i> , <i>FCN2</i>

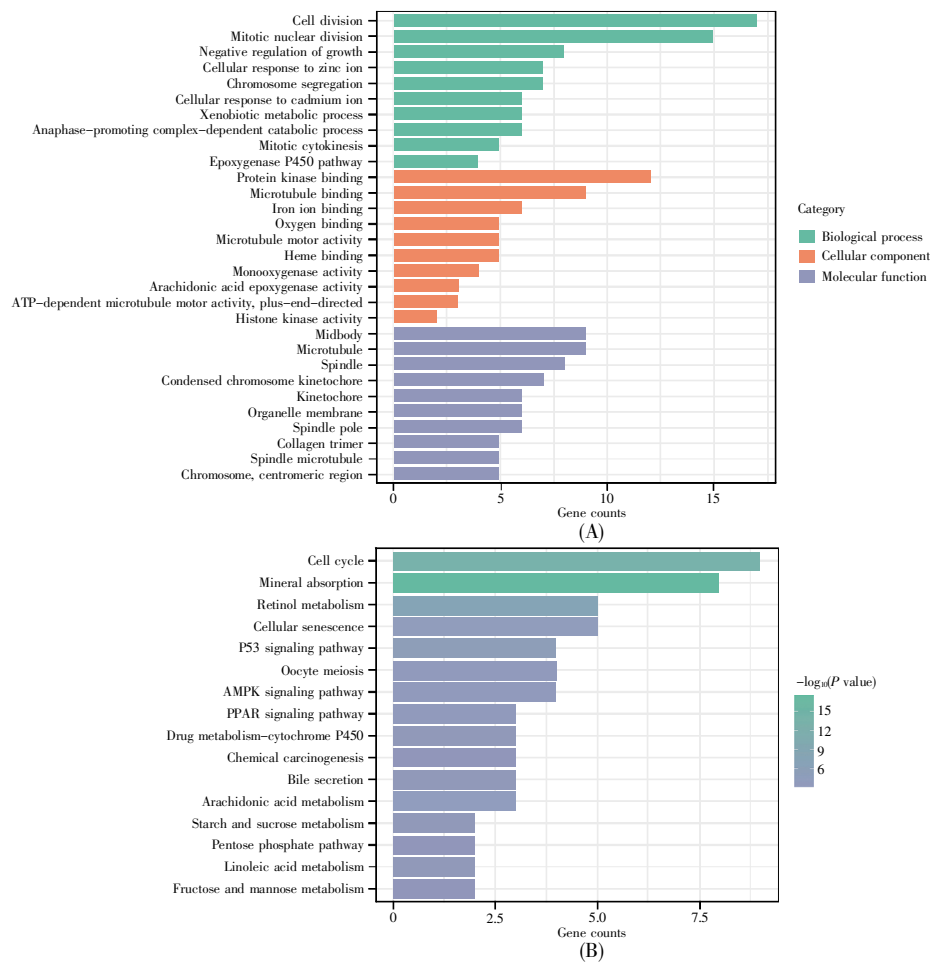


图 1 差异表达基因的 GO 和 KEGG 富集分析

(A) GO 功能富集分析的柱状图(前 10 条); (B) KEGG 通路分析的柱状图。

Fig.1 Functional enrichment analysis of DEGs

(A) GO enrichment analysis (top 10); (B) KEGG pathway analysis.

显著相关(图 3)。在 Oncomine 数据库的 Wurmbach 数据集^[11]中, *BUB1B*、*CDC20* 和 *DLGAP5* 的 mRNA 表达水平与肿瘤分级和血管浸润明显相关(图 4)。进一步在 HPA 数据库比较核心基因在蛋白质水平的表达情况, 结果显示, 在正常肝组织中未见 *CDC20* 和 *DLGAP5* 染色阳性, 在 HCC 组织中可见 *CDC20* 和 *DLGAP5* 有不同程度的染色阳性(图 5A)。肝癌组织芯片中 *CDC20* 和 *DLGAP5* 的表达水平如图 5B 所示, *CDC20* 表达定位在肝癌细胞的细胞质, *DLGAP5* 主要见于细胞质, 也表达于细胞核。此外, 在 14 例肝癌组织中, 11 例存在 *CD-20* 高表达, 3 例低表达; 在 14 例癌旁组织中, 3 例存在 *CDC20* 高表达, 11 例低表达, 差异有统计学意义($P < 0.05$)。在 90 例肝癌组织中, 71 例存在 *DLGAP5* 高表达, 19 例低表达; 在 90 例癌旁组织中, 16 例存在 *DLGAP5* 高表达, 74 例低表达, 差异有统计学意义($P < 0.05$)。单因素和多因素 Cox 回

归分析结果表明, *CDC20* 和 *DLGAP5* 可作为肝癌患者预后的独立危险因素(表 2、表 3)。

2.5 *CDC20* 甲基化分析

研究报道 *DLGAP5* 在 HCC 中高表达, 其表达水平可能受甲基化调节^[12]。但是 *CDC20* 甲基化与其表达水平对 HCC 的预后和免疫微环境的影响缺乏足够的实验数据。因此, 为了进一步探究 *CDC20* 甲基化与其表达水平的关系, 我们通过 UCSC 数据库中的肝癌甲基化数据进行了分析。图 6 显示的是 *CDC20* 甲基化位点与甲基化水平的关系, 由左到右表示甲基化水平逐渐降低。从图 7A 可知, *CDC20* 基因的甲基化水平与其表达水平呈负相关; 甲基化位点 cg06373377 和 cg05-525368 与基因表达量呈负相关, 其余位点无明显相关性。生存分析表明, 甲基化位点 cg06373377 与 HCC 患者 OS 和 PFS 相关, cg05525368 与患者 PFS 相关; 这两个位点的甲基化水平越高, 患者预

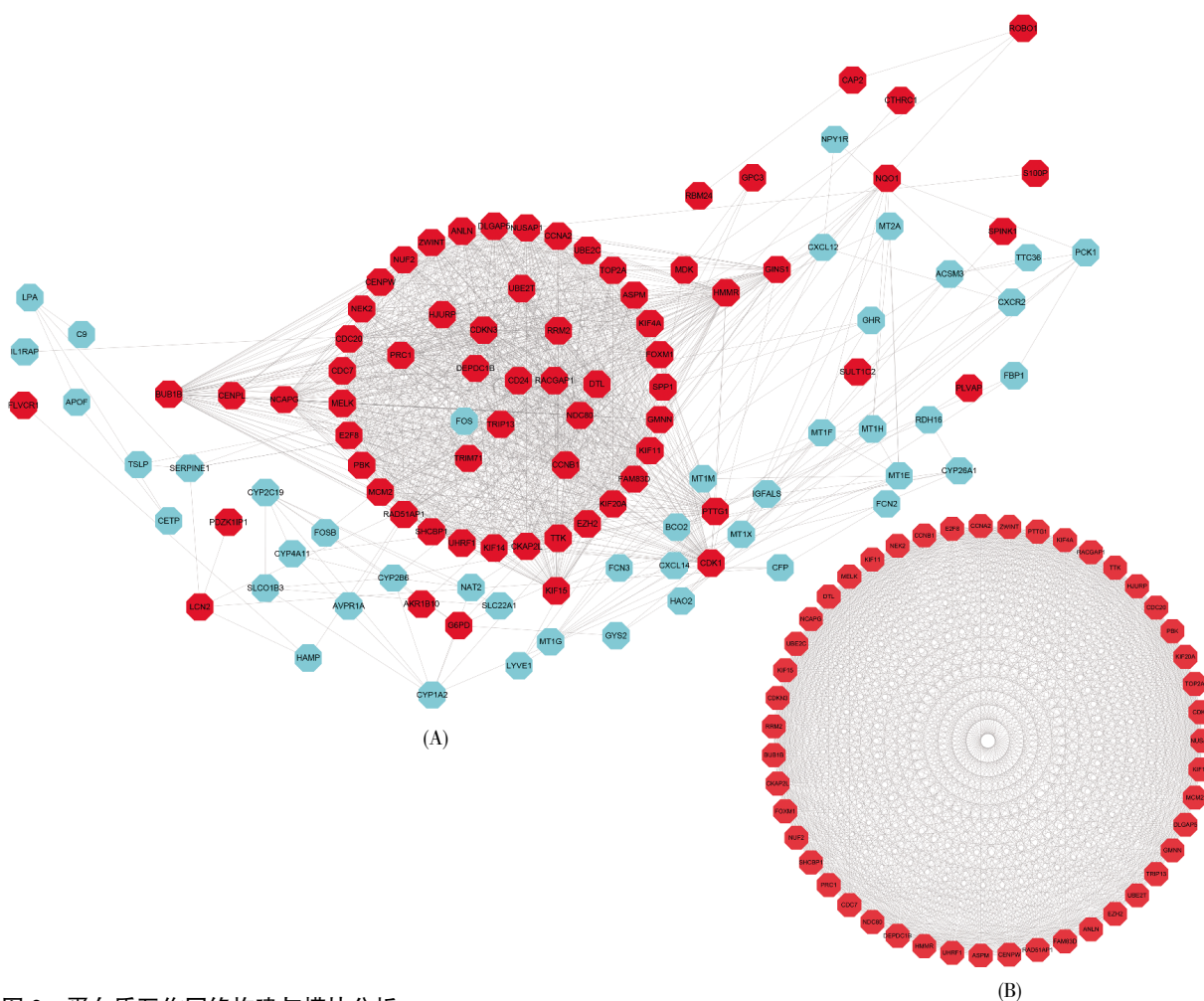


图 2 蛋白质互作网络构建与模块分析

(A) 差异基因的蛋白质互作网络, 红色和蓝色图形分别代表上调和下调的基因; (B) 网络的核心模块。

Fig.2 PPI network construction and module analysis

(A) PPI network of DEGs. The red and blue balls represent upregulated and downregulated genes, respectively; (B) The most significant module in PPI network.

后越好(图 7B)。

2.6 *CDC20* 与免疫细胞的相关性

研究发现, 肿瘤微环境中存在大量的免疫细胞, 与肿瘤的发生发展和免疫治疗密切相关^[13]。基于此, 我们分析了 *CDC20* 与免疫细胞含量的关系, 结果表明: *CDC20* 的表达水平与 CD4⁺ T 细胞、B 细胞、巨噬细胞、中性粒细胞等免疫细胞的含量相关(图 8)。

3 讨论

近年来, 随着高通量测序技术的发展, 越来越多的研究通过生物信息学方法筛选与肿瘤发生发展相关的生物标记物及治疗靶点。本研究通过下载 GEO 数据库中 GSE112790 芯片及 TCGA 数据库中 HCC 的基因表达谱数据, 使用生物信息学方法筛选到 151 个共同存在于上述两个数据集中

的 HCC 组织与正常组织的差异表达基因, 包含 66 个下调基因和 85 个上调基因。GO 分析结果显示, 这些共同差异表达的基因主要富集在细胞分裂、染色体分离、DNA 复制、环氧化酶 P450 途径以及负性生长调控等生物过程; 可能参与染色体、动粒、纺锤体微管等的组成; 其分子功能主要集中在蛋白激酶结合、微管结合、组蛋白激酶活性等。KEGG 分析提示, 差异基因主要富集在 P53 信号通路、AMPK 信号通路、细胞周期、细胞沉默等信号通路。通过构建蛋白质互作网络, 获得 10 个连接度最高的核心基因: *CCNA2*、*CDK1*、*BUB1B*、*CDC20*、*ASPM*、*KIF11*、*DLGAP5*、*CCNB1*、*DTL* 和 *N-USAP1*。生存分析结果表明, *DLGAP5*、*ASPM*、*CD-C20* 和 *BUB1B* 基因高表达与 HCC 患者不良预后相关。Oncomine 数据库的分析结果显示, *BUB1B*、*CDC20* 和 *DLGAP5* 的表达水平与肿瘤分级和血

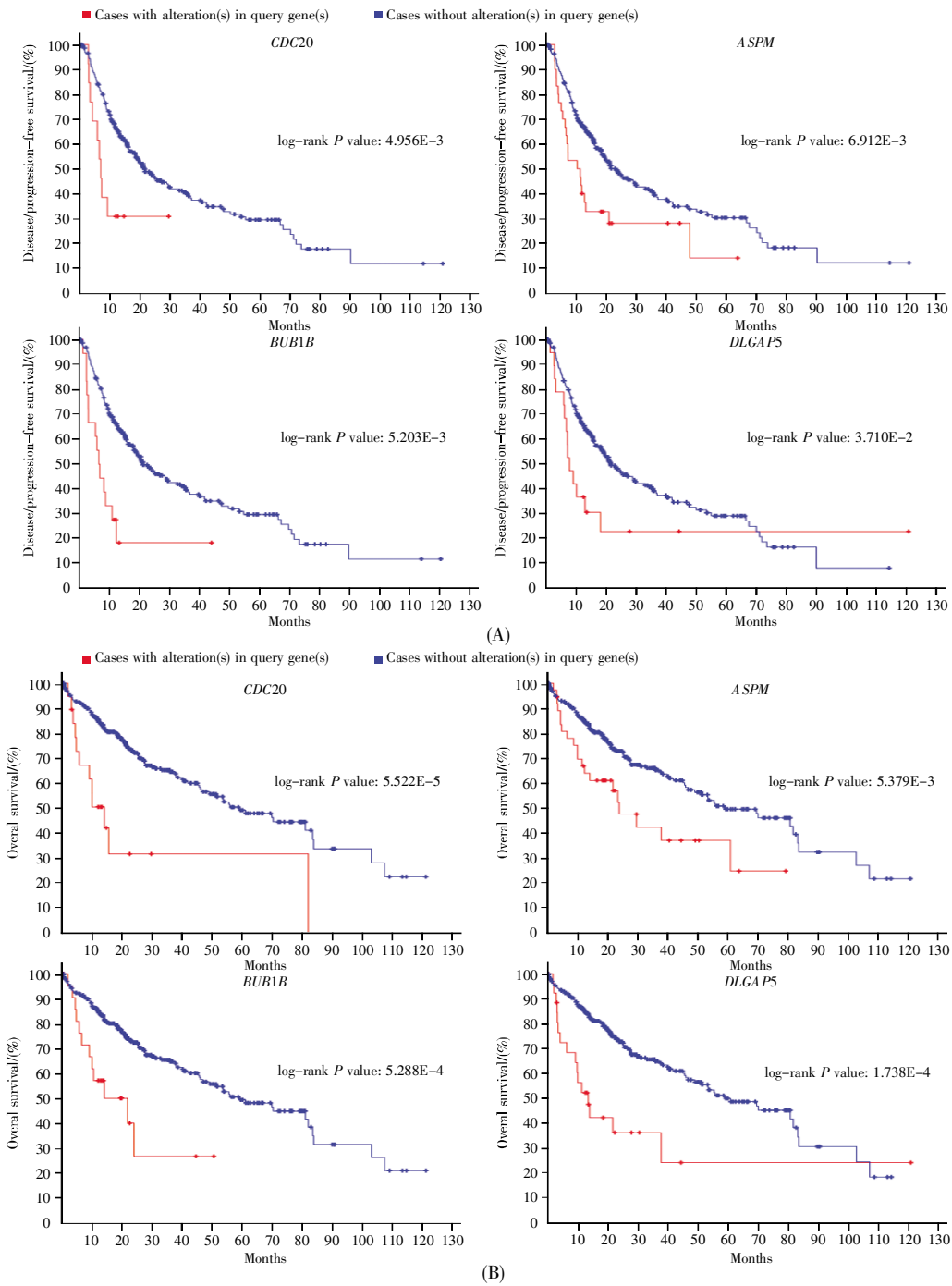


图 3 核心基因对 HCC 患者无病生存期(A)和总生存期(B)的影响分析
 Fig.3 Analyses of DFS (A) and OS (B) in patients with HCC of hub genes

管浸润明显相关。IHC 结果显示, CDC20 和 DLGAP5 蛋白在肝癌组织中高表达。单因素及多因素 Cox 回归分析结果表明, CDC20 和 DLGAP5 可以作为肝癌预后的独立危险因素。肝癌甲基化数据的分析结果提示, CDC20 的甲基化水平可以影响其表达, 并与多种免疫细胞含量正相关。上述结果表明, CDC20 可作为 HCC 患者预后评估的生物标志物或潜在的治疗靶点。

CDC20 (*cell division cycle 20*)能够参与细胞周期调控, 编码在细胞周期中与后期促进复合体/环状体相互作用的调节蛋白^[14], 其高表达与前列腺癌^[15]、乳腺癌^[16]和结肠癌^[17]的不良预后有关。Zhao 等^[18]发现 CDC20 在肝癌组织中高表达, 可能通过 Bcl-2/Bax 途径调节 P53 突变肝癌细胞的增殖和放射敏感性, 是 HCC 患者不良预后的标志物。本研究发现, HCC 肿瘤组织中 CDC20 在 mRNA

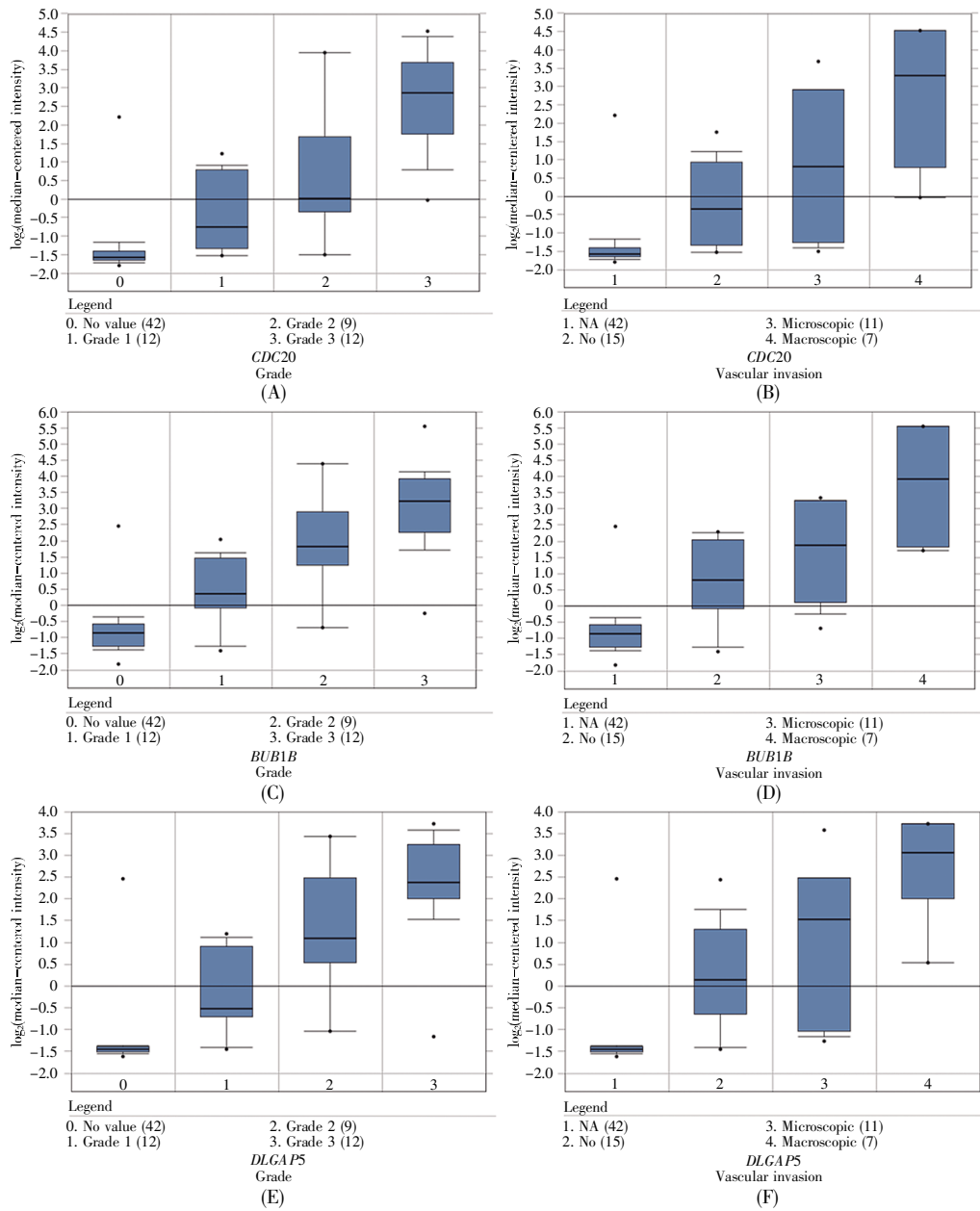


图 4 Wurmbach 数据集中 *BUB1B*、*CDC20* 和 *DLGAP5* 表达水平与 HCC 肿瘤分级(A、C、E)和血管浸润(B、D、F)的关系
 Fig.4 Associations between *BUB1B*, *CDC20* and *DLGAP5* expression and tumor grade (A, C and E), and between their expression and vascular invasion (B, D and F) from the Wurmbach liver dataset

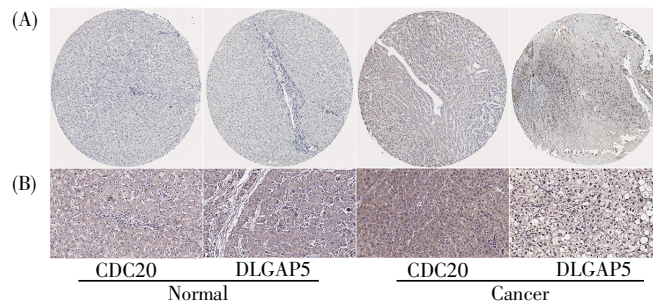


图 5 *CDC20* 和 *DLGAP5* 在肝癌组织和正常组织中的表达

(A) HPA 数据库图像; (B) 组织芯片图像($\times 400$)。

Fig.5 Expression of *CDC20* and *DLGAP5* in HCC and normal tissues

(A) Immunohistochemistry images from HPA database; (B) Immunohistochemistry images from tissue microarray ($\times 400$).

表 2 *CDC20* mRNA 表达水平与患者总生存期的单因素和多因素回归分析
Table 2 Univariate and multivariate analyses of *CDC20* mRNA expression and patient OS

Parameter	Univariate analysis			Multivariate analysis		
	HR	95% CI	P	HR	95% CI	P
Age	1.005	0.987~1.023	0.591	1.012	0.992~1.033	0.227
Gender	0.780	0.487~1.249	0.301	1.101	0.652~1.859	0.720
Grade	1.017	0.746~1.387	0.914	0.953	0.680~1.337	0.782
Stage	1.865	1.456~2.388	<0.001	1.016	0.392~2.634	0.974
Tumor	1.804	1.434~2.270	<0.001	1.554	0.666~3.627	0.308
Metastasis	3.850	1.207~12.281	0.023	2.383	0.600~9.460	0.217
Node	2.022	0.494~8.276	0.328	1.719	0.248~11.903	0.583
<i>CDC20</i>	1.440	1.251~1.657	<0.001	1.392	1.198~1.619	<0.001

表 3 *DLGAP5* mRNA 表达水平与患者总生存期的单因素和多因素回归分析
Table 3 Univariate and multivariate analyses of *DLGAP5* mRNA expression and patient OS

Parameter	Univariate analysis			Multivariate analysis		
	HR	95% CI	P	HR	95% CI	P
Age	1.005	0.987~1.023	0.591	1.014	0.994~1.034	0.186
Gender	0.780	0.487~1.249	0.301	1.176	0.689~2.005	0.553
Grade	1.017	0.746~1.387	0.914	0.993	0.707~1.395	0.968
Stage	1.865	1.456~2.388	<0.001	1.004	1.395~2.659	0.993
Tumor	1.804	1.434~2.270	<0.001	1.544	0.645~3.697	0.330
Metastasis	3.850	1.207~12.281	0.023	2.402	0.601~9.594	0.215
Node	2.022	0.494~8.276	0.328	1.686	0.256~11.087	0.587
<i>DLGAP5</i>	1.438	1.244~1.663	<0.001	1.375	1.175~1.609	<0.001

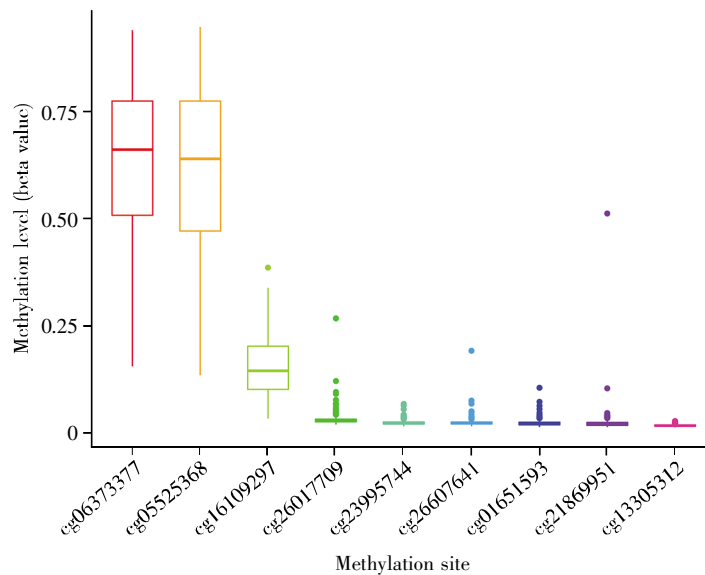


图 6 *CDC20* 基因的甲基化位点与甲基化水平的关系
Fig.6 The relationship between methylation site and methylation level of *CDC20*

水平和蛋白质水平高表达,与 HCC 多个病理指标有关,可以通过多种途径促进肿瘤的发生发展,进而影响患者的临床预后,为 HCC 的分子机制探索和治疗提供了新的证据。Zhang 等^[19]研究表明,*CDC20* 在前列腺癌中的高表达常提示预后不良。Cheng 等^[20]研究表明,*CDC20* 与肿瘤转移过程相关,可作为肿瘤潜在的治疗靶点,人们可以通过研制其靶向抑制剂来抑制肿瘤的发生发展及转移

过程。*BUB1B* 作为纺锤体检测点蛋白质参与细胞有丝分裂的正常过程^[21]。周岩等^[22]的研究指出,*BUB1B* 基因在 HCC 患者中高表达并且与肿瘤细胞的增殖和侵袭有关,可以作为 HCC 患者预后的分子标志物。此外,*BUB1B* 在许多类型的人类恶性肿瘤中表达上调,与肿瘤的形成及侵袭有关^[23-26]。但是,也有研究报道,*BUB1B* 低表达可能与结直肠癌的发展和转移有关^[27]。*DLGAP5* (*disks large*-

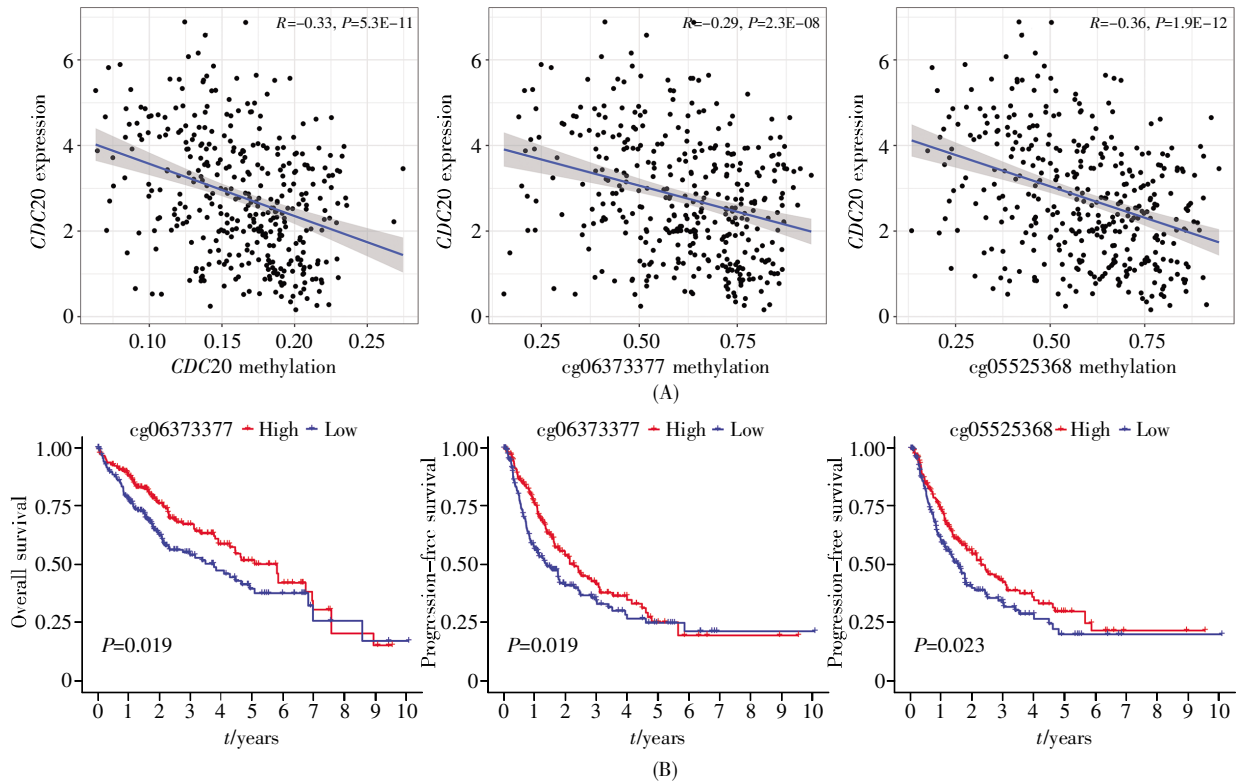


图7 CDC20 基因甲基化与表达水平(A)和患者预后(B)的关系

Fig.7 Associations between methylation level of CDC20 and its expression (A), and between methylation level of CDC20 and HCC prognosis (B)

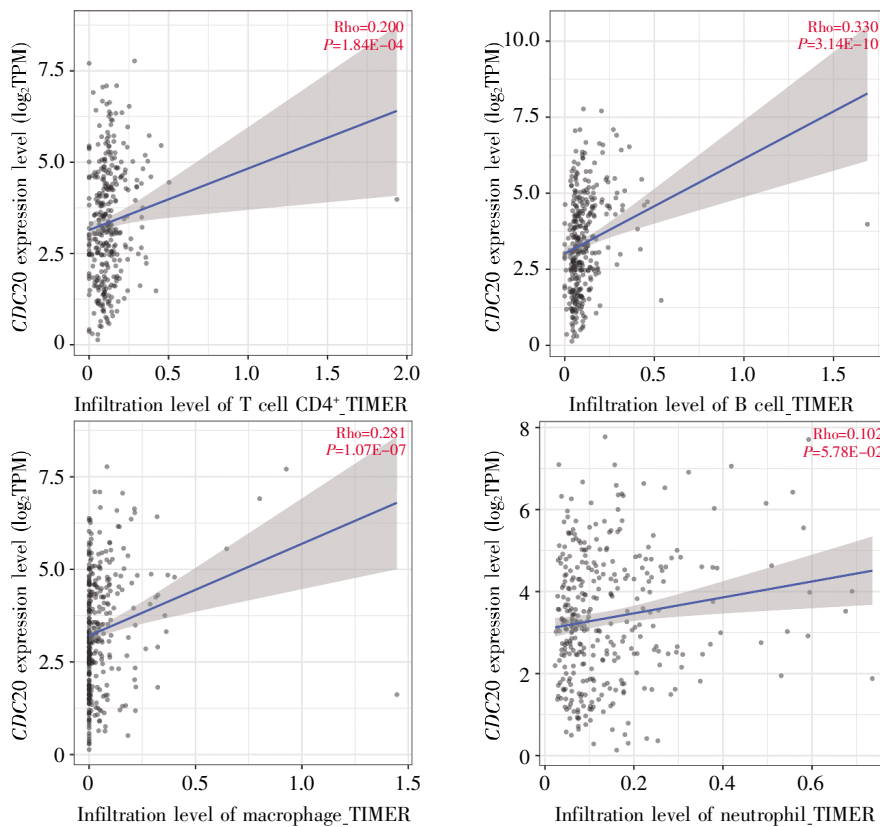


图8 CDC20 表达与免疫细胞含量的相关性

Fig.8 Associations between CDC20 gene expression and immune cell contents

associated protein 5)编码微管相关蛋白,作为一种潜在的细胞周期调节因子,可导致癌细胞的发生和发展^[28]。Liao 等^[12]的研究表明, *DLGAP5* 的表达受甲基化调节,并且其高表达促进肝癌细胞的增殖、迁移和侵袭。Chen 等^[29]的一项研究认为, *DTL* (*denticleless protein homolog*)作为细胞周期的重要调节因子,其 mRNA 水平高表达跟肝癌患者的分级和生存率显著相关; *DTL* 作为潜在的治疗靶点通过诱导肿瘤细胞衰老来治疗肝癌。有研究表明, *CCNA2* (*cyclin-A2*)属于高度保守的细胞周期蛋白,在乳腺癌和胰腺癌中高表达^[30-31],其过表达与肿瘤的进展密切相关。*NUSAP1* (*nucleolar and spindle-associated protein 1*)是一种微管相关蛋白,在细胞有丝分裂过程中起着重要的作用^[32]。研究表明, *NUSAP1* 高表达与前列腺癌、结肠癌和肝癌有关^[33-35]。周期蛋白依赖性激酶 1 (*cyclin-dependent kinase 1, CDK1*)是细胞周期的重要调控因子,通过调节中心体周期以及有丝分裂过程,在真核细胞周期调节中起着关键作用,其表达异常会导致人体的正常生理功能发生紊乱^[36]。*CDK1* 在结肠癌中已被确定为有效的临床预后标志物之一^[37]。*CCNB1* (*G2/mitotic-specific cyclin-B1*)是一种参与有丝分裂的调节蛋白,主要通过活化 *CDC2* 并与之形成复合物,参与 *G2/M* 期的转变过程^[38], *CCNB1* 的异常表达会破坏细胞周期的调控机制,可能引起细胞畸形生长及肿瘤的发生。既往研究表明, *CCNB1* 高表达与 HCC 患者预后不良相关^[39]。Chen 等^[40]研究表明,大多数驱动蛋白超家族蛋白 (*kinesin superfamily protein, KIF*)的异常高表达与 HCC 的发生及预后显著相关。*KIF11* 作为 *KIF* 家族成员之一,研究其与 HCC 的关系具有重要的临床意义。Lin 等^[41]研究发现, *ASPM* (*abnormal spindle-like microcephaly-associated protein*)可能是引起肝癌侵袭、转移能力增强的潜在分子标志物,与肿瘤的早期复发和预后有关。

综上所述,本研究通过对 HCC 和正常组织进行基因差异表达分析,得到了与 HCC 发生发展相关的多个核心基因,其中 *CDC20* 与 HCC 的发生、发展及预后有关,有望成为肝癌预后的临床指标及潜在的治疗靶点。

参考文献(References):

[1] BRAY F, FERLAY J, SOERJOMATARAM I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and

- mortality worldwide for 36 cancers in 185 countries[J]. *CA: A Cancer Journal for Clinicians*, 2018, 68(6): 394-424.
- [2] ANWANWAN D, SINGH S K, SINGH S, et al. Challenges in liver cancer and possible treatment approaches[J]. *Biochimica et Biophysica Acta-Reviews on Cancer*, 2020, 1873(1): 188314.
- [3] DIMITROULIS D, DAMASKOS C, VALSAMI S, et al. From diagnosis to treatment of hepatocellular carcinoma: an epidemic problem for both developed and developing world[J]. *World Journal of Gastroenterology*, 2017, 23(29): 5282-5294.
- [4] SHIMADA S, MOGUSHI K, AKIYAMA Y, et al. Comprehensive molecular and immunological characterization of hepatocellular carcinoma[J]. *EBioMedicine*, 2019, 40: 457-470.
- [5] RITCHIE M E, PHIPSON B, WU D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies[J]. *Nucleic Acids Research*, 2015, 43(7): e47.
- [6] ROBINSON M D, MCCARTHY D J, SMYTH G K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data[J]. *Bioinformatics*, 2010, 26(1): 139-140.
- [7] HUANG D W, SHERMAN B T, LEMPICKI R A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources[J]. *Nature Protocols*, 2009, 4(1): 44-57.
- [8] GAO J J, AKSOY B A, DOGRUSOZ U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal[J]. *Science Signaling*, 2013, 6(269): p11.
- [9] CERAMI E, GAO J J, DOGRUSOZ U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data[J]. *Cancer Discovery*, 2012, 2(5): 401-404.
- [10] UHLEN M, ZHANG C, LEE S, et al. A pathology atlas of the human cancer transcriptome[J]. *Science*, 2017, 357(6352): eaan2507.
- [11] WURMBACH E, CHEN Y B, KHITROV G, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma[J]. *Hepatology*, 2007, 45(4): 938-947.
- [12] LIAO W J, LIU W L, YUAN Q, et al. Silencing of *DLGAP5* by siRNA significantly inhibits the proliferation and invasion of hepatocellular carcinoma cells[J]. *PLoS One*, 2013, 8(12): e80789.
- [13] GALLI F, AGUILERA J V, PALERMO B, et al. Relevance of immune cell and tumor microenvironment imaging in the new era of immunotherapy[J]. *Journal of Experimental & Clinical Cancer Research*, 2020, 39: 89.
- [14] JEONG S M, BUI Q T, KWAK M, et al. Targeting *CDC20* for cancer therapy[J]. *Biochimica et Biophysica Acta-Reviews on Cancer*, 2022, 1877(6): 188824.
- [15] ZHANG Q, HUANG H, LIU A, et al. Cell division cycle 20 (*CDC20*) drives prostate cancer progression via stabilization of β -catenin in cancer stem-like cells[J]. *EBioMedicine*, 2019, 42: 397-407.
- [16] KARRA H, REPO H, AHONEN I, et al. *CDC20* and securin overexpression predict short-term breast cancer survival[J]. *British Journal of Cancer*, 2014, 110(12): 2905-2913.
- [17] WU W J, HU K S, WANG D S, et al. *CDC20* overexpression predicts a poor prognosis for patients with colorectal cancer[J]. *Journal of Translational Medicine*, 2013, 11: 142.
- [18] ZHAO S, ZHANG Y C, LU X Q, et al. *CDC20* regulates the cell proliferation and radiosensitivity of P53 mutant HCC cells through the *Bcl-2/Bax* pathway[J]. *International Journal of Biological Sciences*, 2021, 17(13): 3608-3621.
- [19] ZHANG Q, HUANG H, LIU A, et al. Cell division cycle 20 (*CDC20*) drives prostate cancer progression via stabilization of β -catenin in cancer stem-like cells[J]. *EBioMedicine*, 2019, 42: 397-407.

- [20] CHENG S J, CASTILLO V, SLIVA D. CDC20 associated with cancer metastasis and novel mushroom derived CDC20 inhibitors with antimetastatic activity[J]. *International Journal of Oncology*, 2019, 54(6): 2250–2256.
- [21] SHIN H J, BAEK K H, JEON A H, *et al.* Dual roles of human BubR1, a mitotic checkpoint kinase, in the monitoring of chromosomal instability[J]. *Cancer Cell*, 2003, 4(6): 483–497.
- [22] 周岩, 母汉友, 王淑芬, 等. BUB1B 在原发性肝癌中的表达及对肝癌细胞增殖和侵袭的影响[J]. *新医学*(ZHOU Yan, MU Han-you, WANG Shu-fen, *et al.* Clinical significance of BUB1B in hepatocellular carcinoma and its effect on proliferation and invasion of hepatoma cells[J]. *Journal of New Medicine*), 2019, 50(4): 272–277.
- [23] QIU J N, ZHANG S P, WANG P, *et al.* BUB1B promotes hepatocellular carcinoma progression via activation of the mTORC1 signaling pathway[J]. *Cancer Medicine*, 2020, 9(21): 8159–8172.
- [24] JIAO C Y, FENG Q C, LI C X, *et al.* BUB1B promotes extrahepatic cholangiocarcinoma progression via JNK/c-Jun pathways[J]. *Cell Death & Disease*, 2021, 12(1): 63.
- [25] KOYUNCU D, SHARMA U, GOKA E T, *et al.* Spindle assembly checkpoint gene BUB1B is essential in breast cancer cell survival[J]. *Breast Cancer Research and Treatment*, 2021, 185(2): 331–341.
- [26] CHEN J, LIAO Y, FAN X M. Prognostic and clinicopathological value of BUB1B expression in patients with lung adenocarcinoma: a meta-analysis[J]. *Expert Review of Anticancer Therapy*, 2021, 21(7): 795–803.
- [27] 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.
- [28] LAPRISE P, VIEL A, RIVARD N. Human homolog of disc-large is required for adherens junction assembly and differentiation of human intestinal epithelial cells[J]. *The Journal of Biological Chemistry*, 2004, 279(11): 10157–10166.
- [29] CHEN Y C, CHEN I S, HUANG G J, *et al.* Targeting DTL induces cell cycle arrest and senescence and suppresses cell growth and colony formation through TPX2 inhibition in human hepatocellular carcinoma cells[J]. *OncoTargets and Therapy*, 2018, 11: 1601–1616.
- [30] GAO T, HAN Y, YU L, *et al.* CCNA2 is a prognostic biomarker for ER+ breast cancer and tamoxifen resistance[J]. *PLoS One*, 2014, 9(3): e91771.
- [31] DONG S, HUANG F, ZHANG H, *et al.* Overexpression of BUB1B, CCNA2, CDC20, and CDK1 in tumor tissues predicts poor survival in pancreatic ductal adenocarcinoma[J]. *Bioscience Reports*, 2019, 39(2): BSR20182306.
- [32] HUSSAIN S, BENAVENTE S B, NASCIMENTO E, *et al.* The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability[J]. *The Journal of Cell Biology*, 2009, 186(1): 27–40.
- [33] GORDON C A, GONG X, GANESH D, *et al.* NUSAP1 promotes invasion and metastasis of prostate cancer[J]. *Oncotarget*, 2017, 8(18): 29935–29950.
- [34] LIU Z X, GUAN C Q, LU C H, *et al.* High NUSAP1 expression predicts poor prognosis in colon cancer[J]. *Pathology Research and Practice*, 2018, 214(7): 968–973.
- [35] ZHOU L, DU Y Y, KONG L Q, *et al.* Identification of molecular target genes and key pathways in hepatocellular carcinoma by bioinformatics analysis[J]. *OncoTargets and Therapy*, 2018, 11: 1861–1869.
- [36] FOUREST-LIEUVIN A, PERIS L, GACHE V, *et al.* Microtubule regulation in mitosis: tubulin phosphorylation by the cyclin-dependent kinase Cdk1[J]. *Molecular Biology of the Cell*, 2006, 17(3): 1041–1050.
- [37] SUNG W W, LIN Y M, WU P R, *et al.* High nuclear/cytoplasmic ratio of Cdk1 expression predicts poor prognosis in colorectal cancer patients[J]. *BMC Cancer*, 2014, 14: 951.
- [38] BROWN N R, LOWE E D, PETRI E, *et al.* Cyclin B and cyclin A confer different substrate recognition properties on CDK2[J]. *Cell Cycle*, 2007, 6(11): 1350–1359.
- [39] GU J S, LIU X R, LI J, *et al.* MicroRNA-144 inhibits cell proliferation, migration and invasion in human hepatocellular carcinoma by targeting CCNB1[J]. *Cancer Cell International*, 2019, 19(1): 15.
- [40] CHEN J L, LI S, ZHOU S, *et al.* Kinesin superfamily protein expression and its association with progression and prognosis in hepatocellular carcinoma[J]. *Journal of Cancer Research and Therapeutics*, 2017, 13(4): 651–659.
- [41] LIN S Y, PAN H W, LIU S H, *et al.* ASPM is a novel marker for vascular invasion, early recurrence, and poor prognosis of hepatocellular carcinoma[J]. *Clinical Cancer Research*, 2008, 14(15): 4814–4820.