

·论著·

基于生物信息学方法分析 KIF20A 与肝癌预后的关系

曹宝全,朱洪

(昆明医科大学第二附属医院肝胆胰外科一病区,云南 昆明 650000)

摘要:目的 利用生物信息学方法分析 KIF20A 与肝癌的关系。方法 基于 GEO 数据库中的肝癌数据集 GSE121248 使用 edgeR 来筛选差异的 mRNAs,再对这些基因进行富集分析以探索其功能和潜在的作用通路;另基于 String 数据库分析后构建蛋白互作网络(PPI),经 cytoscape 可视化后得到关键基因,之后再在 GEPIA 数据库中验证。结果 共有 531 个差异表达的基因,其中上调的 mRNAs 有 152 个,下调的 mRNAs 有 379 个。DAVID 数据库分析显示,531 个差异基因的生物功能主要富集在炎症反应、氧化过程和免疫反应中;细胞成分主要集中在细胞外区域,分子功能主要集中在氧化还原酶活性、蛋白质同源二聚活性、铁离子结合和血红蛋白结合;富集的主要途径是化学致癌性、代谢和抗生素生物合成。通过 String 数据库的分析和 cytoscape 软件的结合,只筛选出一个关键 mRNA 为 KIF20A。生存曲线分析显示,KIF20A 的表达与肝癌患者的预后相关,且低表达与高生存率呈正相关($P<0.05$)。结论 KIF20A 是该调控网络的核心基因,在肝癌的发生发展中起着重要作用,其高表达与患者生存率有关,是判断肝癌预后的可靠生物标志物,可作为肝癌患者药物治疗的靶点。

关键词:肝癌;KIF20A;mRNAs

中图分类号:R735.7

文献标识码:A

DOI:10.3969/j.issn.1006-1959.2021.15.013

文章编号:1006-1959(2021)15-0047-04

Analysis of the Relationship Between KIF20A and the Prognosis of Liver Cancer Based on Bioinformatics Methods

CAO Bao-quan,ZHU Hong

(The First Ward of Hepatobiliary and Pancreatic Surgery,the Second Affiliated Hospital of Kunming Medical University, Kunming 650000,Yunnan,China)

Abstract: Objective To analyze the relationship between KIF20A and liver cancer using bioinformatics methods.**Methods** Based on the liver cancer data set GSE121248 in the GEO database, edgeR was used to screen differential mRNAs, and then these genes were enriched and analyzed to explore their functions and potential pathways;In addition, a protein interaction network (PPI) was constructed based on the analysis of the String database, and the key genes were obtained after visualization by cytoscape, and then verified in the GEPIA database.**Results** There were a total of 531 differentially expressed genes, of which 152 were up-regulated mRNAs and 379 were down-regulated mRNAs.DAVID database analysis showed that the biological functions of 531 differential genes were mainly enriched in inflammatory response, oxidation process and immune response;Cell components were mainly concentrated in the extracellular area, and molecular functions were mainly concentrated in oxidoreductase activity, protein homodimerization activity, iron ion binding and hemoglobin binding;The main ways of enrichment were chemical carcinogenicity, metabolism and antibiotic biosynthesis.Through the combination of String database analysis and cytoscape software, only one key mRNA was selected as KIF20A. Survival curve analysis showed that the expression of KIF20A was correlated with the prognosis of liver cancer patients, and low expression was positively correlated with high survival rate ($P<0.05$).**Conclusion** KIF20A is the core gene of this regulatory network and plays an important role in the occurrence and development of liver cancer.Its high expression is related to the low survival rate of patients, is a reliable biomarker for judging the prognosis of liver cancer, and can be used as a target of drug treatment for liver cancer patients.

Key words: Liver cancer;KIF20A;mRNAs

肝癌(liver cancer)属于消化系统最常见的恶性肿瘤,且每年约有 60 万患者死于肝癌^[1]。从宏观上看,肝癌可分为原发性肝癌和继发性肝癌,其中原发性肝癌的比例为 75%~90%^[2]。虽然肝部分切除和肝移植治疗早期肝癌是有效的^[3],但由于在疾病早期,患者的症状不易引起重视,导致大多数患者在就诊时已处于肝癌的中晚期,其 5 年生存率通常低于 30%^[4]。因此,寻找肝癌潜在生物标志物,对提高其诊断、治疗和预后具有重要意义。目前,鉴于现有第二代基因诊断技术的贡献,通过微阵列技术分析肝癌

的基因表现谱后,发现很多基因与肝癌的发展密切相关^[5]。因此,将基因芯片技术与生物信息学方法相结合,可以筛选出在肝癌发生发展中起关键作用的基因。本研究利用 GEO 数据库中不同肝癌样本的基因表达数据,研究肝癌与正常肝组织中差异表达的基因,现报道如下。

1 材料与方法

1.1 数据来源 肝癌基因数据集 GSE121248 的表达数据来自于 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo/>)^[6],其中有 37 个肝脏正常组织样本和 70 个肝癌组织样本。

1.2 基因筛选 利用 R 软件分析下载的数据,设定 $|\log FC| \geq 1$, $\text{adj.}P=0.05$,研究在肝癌组织和正常肝组织中差异表现的基因。

1.3 富集分析 差异基因的功能注释由 DAVID (<https://david.ncicrf.gov/>, 6.8 版)^[7]数据库分析,包括

基金项目:1.国家自然科学基金资助项目(编号:81460132);2.云南省科技厅基础研究计划(昆医联合专项)(编号:2018FE001(-224))

作者简介:曹宝全(1994.3-),男,湖北十堰人,硕士,主要从事肝胆疾病的诊治工作

通讯作者:朱洪(1970.5-),男,云南昆明人,博士,教授,博士生导师,主要从事肝胆胰脾疾病的诊治工作

生物过程、细胞成分和分子功能,然后对这些基因进行通路富集分析(keoto enclopedia of genes,KEGG), $FDR < 0.05$ 被认为差异有统计学意义。

1.4 蛋白质相互作用网络的建立及筛选关键基因 将筛选出的差异基因上传到 String(<https://string-db.org/11.0> 版本)中^[8],并从该数据库中选择“Multiple proteins”,生物种类设定为“Homo Sapiens”,minimum required interaction score 设定为“highest confidence(0.9)”进行分析。将得到的数据文件输入到 cytocape(版本 3.7.2)软件中进行可视化。然后,利用软件中 APP:cytohubba 的 DMNC 和 Degree 算法分别进行分析,得到关键基因。两种算法得到的前 10 个基因取交集后得到的共同基因被认为是关键基因。

1.5 生存分析 利用数据库 GEPIA(<http://gepia.cancer-pku.cn/>,基因表达谱分析)^[9]中肝癌患者的临床资料,绘制 K-M 曲线, $P < 0.05$ 被认为差异有统计学意义。

2 结果

2.1 基因筛选结果 通过肝癌数据集 GSE121248 获得了 531 个差异表达基因,共有 152 个高表达 mRNA 和 379 个低表达 mRNAs,见图 1。

2.2 GO 和 KEGG 富集分析 DAVID 数据库分析显示,531 差异基因的生物学功能主要富集在炎症反应、氧化过程和免疫反应中;细胞成分主要集中在细

胞外区域,分子功能主要集中在氧化还原酶活性、蛋白质同源二聚活性、铁离子结合和血红蛋白结合;富集的主要途径是化学致癌性、代谢和抗生素生物合成,见图 2。

2.3 PPI 蛋白质相互作用网络的构建和关键基因筛选 共得到 10 种差异基因,分别为 TTK、KIF4A、NUSAP1、MELK、PTTG1、NEK1、NEK1、NEK2、NEK2、PMPK2、KIKG2、CEA2、CEF2、CDK1、CCNB1、CDC20、CCNA2、BUB1B、DLP5、KIF1A、TPX2、KIF20A。在网站上(<http://bioinformatics.psb.gnt.be/webtools/Venn/>)制作韦恩图后,共同的差异基因只有 KIF20A,见图 3。

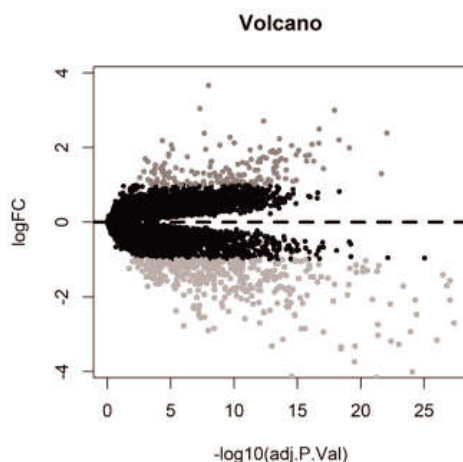
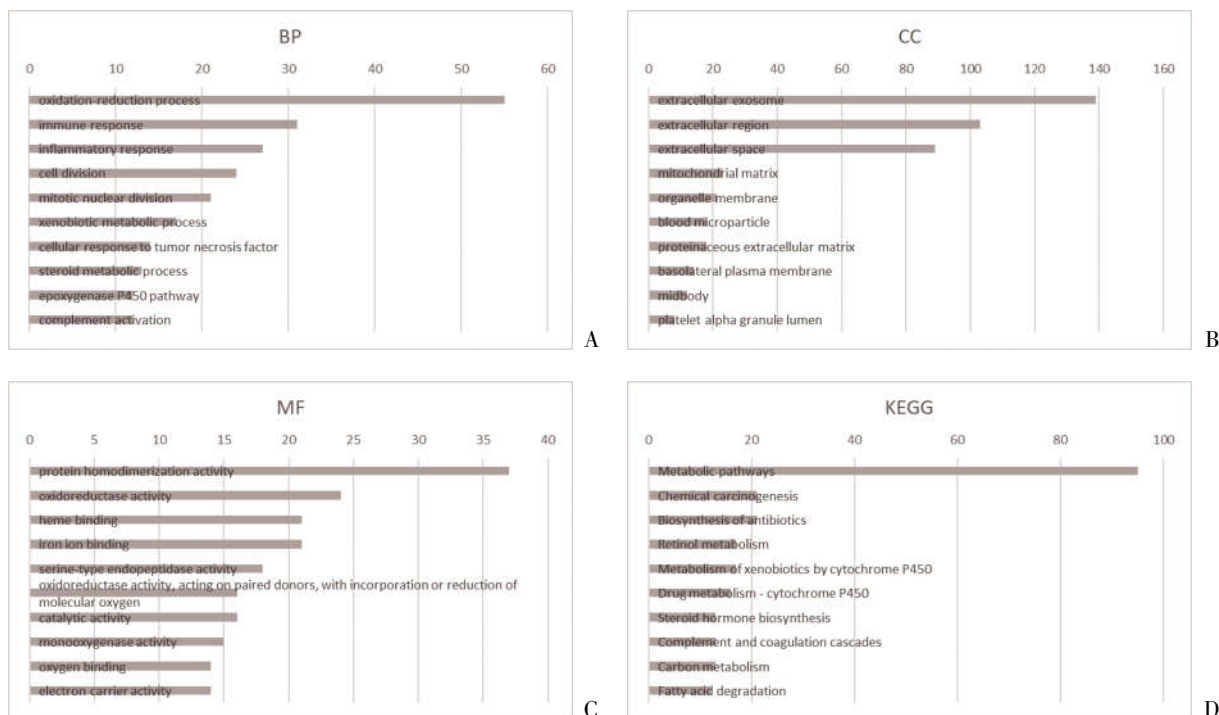
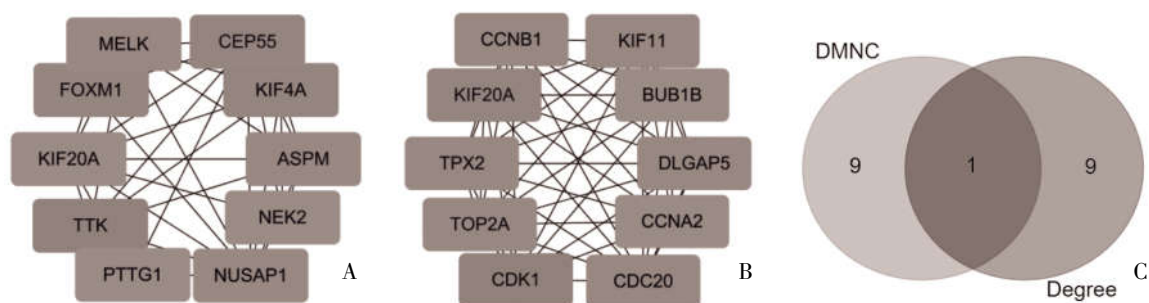


图1 筛选的差异 mRNAs 火山图



注:A:差异 RNA 富集的生物过程;B:差异 mRNA 富集的细胞成分;C:差异 mRNA 富集的分子功能;D:差异 mRNAs 丰富聚集的通路

图2 GO 和 KEGG 富集分析

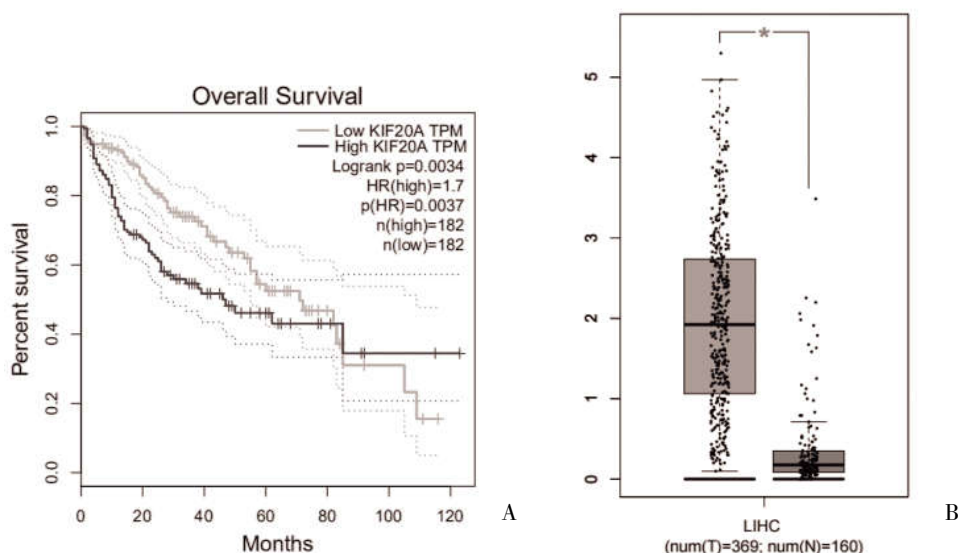


注:A:DMNC 算法得到的关键基因;B:分级算法的密钥环;C:通过两种算法获得差异基因韦恩图

图 3 关键基因筛选

2.4 生存分析 为了进一步验证关键基因 KIF20A 是否与肝癌的预后密切相关,利用 Gepia 数据库绘制了肝癌的生存分析曲线,及其在肝癌组织和正常肝组织中的表达,结果显示 KIF20A 的表达与肝癌患

者的预后相关,且低表达与高生存率呈正相关($P < 0.05$)。根据表达的箱式图分析结果,其在肝癌组织的表达量远高于肝脏的正常组织($P < 0.05$),见图 4。



注:A:肝癌中 KIF20A 的表达与患者预后关系的生存分析图;B:KIF20A 在肝癌组织和肝脏正常组织的表达箱式图

图 4 GEPIA 数据库分析结果

3 讨论

KIF20A(驱动蛋白家族成员 20A)是驱动蛋白-6 亚家族成员^[10],其表达水平与癌症的发生发展密切相关^[11]。据报道^[12,13],KIF20A 可促进膀胱癌细胞的增殖和转移,通过激活 JAK/STAT3 通路促进结肠癌的发生和发展。KIF20A 也是不同类型乳腺癌的治疗靶标和预后生物标志物^[14],是治疗卵巢癌的药物靶点^[15];同时,其可以抑制非小细胞肺癌细胞的增殖和远处转移,促进胃癌的增殖^[16]。本研究旨在探讨 KIF20A 对肝癌患者预后的影响,为肝癌靶向基因治疗提供理论依据。

本研究分析了 GEO 数据库的肝癌数据集 GSE12248 的 70 个肝癌组织样本和 37 个正常肝组织样本中的基因表达数据,得到了 531 个差异的 mRNAs,其中发挥上调作用的有 152 个,下调作用的有 379 个。GO 富集分析表明,这些差异表达基因主要通过代谢途径和化学致癌作用影响细胞外氧化

还原酶活性、免疫反应、炎症反应、血红素结合和铁结合。同时,为了进一步筛选出在肝癌发生发展中起关键作用的 mRNAs,将这 531 个差异基因上传到 String 数据库,构建由 531 个点和 935 条边组成的蛋白质相互作用网络,然后将结果输入到 Cytoscape 中进行可视化,前 10 个关键基因分别通过 DMNC 和 Degree 算法得到。经过取交集,结果发现只有 KIF20A 是肝癌发生发展的关键基因。为了进一步探讨 KIF20A 与肝癌预后的关系,本研究利用 GEPIA 数据库中肝癌患者的临床资料,绘制了其在肝癌组织和正常肝组织中表达的箱式图和 K-M 生存曲线,结果表明 KIF20A 在肝癌组织中高表达,其表达水平与肝癌患者的预后密切相关 ($P < 0.05$),提示 KIF20A 的表达量越低,肝癌患者的预后越好,生存时间越长。本研究还存在一些不足:①本研究的结果没有得到进一步细胞实验的验证;②虽然本研究中

(下转第 53 页)

(上接第 49 页)

使用的肝癌数据集样本量不小,但只有一个数据集,尚需要更多的样本研究结果来进一步证实这一结果;③肝癌的发生和发展通常是由多种因素影响,本研究未进行单因素分析和多因素分析,只从基因表达的水平上对肝癌的发展做研究。

总之,KIF20A 是该调控网络的核心基因,在肝癌的发生发展中起着重要作用,其高表达与患者生存率低有关,是判断肝癌预后的可靠生物标志物,可作为肝癌患者药物治疗的靶点。

参考文献:

- [1]Zhao C,Nguyen MH.Hepatocellular carcinoma screening and surveillance:practice guidelines and real-life practice[J].Clin Gastroenterol,2016(50):120-133.
- [2]Wong MC,Jiang JY,Goggins WB,et al.International incidence and mortality trends of liver cancer:a global profile [J].Sci Rep, 2017,3(31):7.
- [3]Dhanasekaran R,Nault JC,Roberts LR,et al.Genomic Medicine and Implications for Hepatocellular Carcinoma Prevention and Therapy[J].Gastroenterology,2019,156(2):492-509.
- [4]European Association for the Study of the Liver.EASL Clinical Practice Guidelines:Management of hepatocellular carcinoma [J].J Hepatol,2018,69(1):182-236.
- [5]Kamps R,Brandao RD,Bosch BJ,et al.Next-Generation Sequencing in Oncology:Genetic Diagnosis,Risk Prediction and Cancer Classification[J].Int J Mol Sci,2017,18(2):1-57.
- [6]Paredes J,Zabaleta J,Garai J,et al.Immune-Related gene expression and cytokine secretion is reduced among African American colon cancer patients[J].Front Oncol,2020(10):1498.
- [7]Huang da W,Sherman BT,Lempicki RA.Systematic and inte-

grative analysis of large gene lists using DAVID bioinformatics resources[J].Nat Protoc,2009,4(1):44-57.

[8]Szklarczyk D,Franceschini A,Wyder S,et al.STRING v10: protein-protein interaction networks,integrated over the tree of life[J].Nucleic Acids Res,2015(43):447-452.

[9]Tang Z,Li C,Kang B,et al.GEPIA:a web server for cancer and normal gene expression profiling and interactive analyses[J].Nucleic Acids Res,2017,45(1):98-102.

[10]Wu WD,Yu KW,Zhong N,et al.Roles and mechanisms of Kinesin-6 KIF20A in spindle organization during cell division [J].Eur J Cell Biol,2019,98(2-4):74-80.

[11]Rath O,Kozielski F.Kinesins and cancer [J].Nature Reviews Cancer,2012,12(8):527-539.

[12]Shen T,Yang L,Zhang Z,et al.KIF20A Affects the Prognosis of Bladder Cancer by Promoting the Proliferation and Metastasis of Bladder Cancer Cells[J].Dis Markers,2019(2019):4863182.

[13]Xiong M,Zhuang K,Luo Y,et al.KIF20A promotes cellular malignant behavior and enhances resistance to chemotherapy in colorectal cancer through regulation of the JAK/STAT3 signaling pathway[J].Aging(Albany NY),2019,12(16):11905-11921.

[14]Nakamura M,Takano A,Thang PM,et al.Characterization of KIF20A as a prognostic biomarker and therapeutic target for different subtypes of breast cancer[J].Int J Oncol,2020,57(1):277-288.

[15]Li Y,Guo H,Wang Z,et al.Cyclin F and KIF20A,FOXM1 target genes,increase proliferation and invasion of ovarian cancer cells[J].Exp Cell Res,2020,395(2):112212.

[16]Sheng Y,Wang W,Hong BU,et al.Upregulation of KIF20A correlates with poor prognosis in gastric cancer[J].Cancer Manag Res,2018(10):6205-6216.

收稿日期:2021-01-31;修回日期:2021-02-10

编辑/杜帆