

消化系统疾病

DOI: 10.13406/j.cnki.cyxh.001988

基于数据挖掘分析 *KIF2C* 在肝细胞癌的表达及临床意义张越美¹, 赵洪波², 朱亚玲¹, 蒋保三¹, 王 刚¹, 何春艳¹

(1. 华侨大学医学院分子和遗传流行病学课题组, 泉州 362000;

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, 昆明 650500)

【摘要】目的:通过数据挖掘分析肝细胞癌中 *KIF2C* 基因的表达并探讨其临床意义。**方法:**使用 Oncomine 和 GEPIA 数据库分析 *KIF2C* 在肝细胞癌组织中的表达情况;通过 GEPIA 数据库进行 *KIF2C* 的表达程度与病理分级及肝细胞癌生存期的相关性分析;最后利用 STRING 数据库分析与 *KIF2C* 相互作用的蛋白。**结果:**通过 Oncomine 和 GEPIA 数据库分析显示,在 mRNA 水平上,*KIF2C* 在肝细胞癌中显著高表达($P=0.000$),表达水平与其预后为负相关($P=0.000$),且肝癌病理分级越高,*KIF2C* 表达水平越高。通过 STRING 数据库分析显示与 *KIF2C* 相关的蛋白有 PLK1、AURKB、CENPA 等。**结论:**通过对肿瘤基因数据库进行挖掘发现,*KIF2C* 在肝细胞癌组织中明显高表达且与患者预后相关,为肝细胞癌的致病机制研究和抗肿瘤靶向治疗提供了理论支持。

【关键词】肝细胞癌;*KIF2C* 基因;数据库分析**【中图分类号】**R735.7**【文献标志码】**A**【收稿日期】**2018-07-21

Expression and clinical significance of *KIF2C* gene in hepatocellular carcinoma: An analysis based on data mining

Zhang Yuemei¹, Zhao Hongbo², Zhu Yaling¹, Jiang Baosan¹, Wang Gang¹, He Chunyan¹

(1. Research Group of Molecular and Genetic Epidemiology, Huaqiao University School of Medicine;

2. Institute of Molecular Clinical Medicine, Kunming Medical University and Key Laboratory of
Stem Cell and Regenerative Medicine of Yunnan Province)

【Abstract】Objective: To investigate the expression and clinical significance of the *KIF2C* gene in hepatocellular carcinoma by data mining. **Methods:** Oncomine and GEPIA databases were used to analyze the expression of the *KIF2C* gene in hepatocellular carcinoma. GEPIA database was used to investigate the correlation of *KIF2C* expression with pathological classification and survival time. STRING database was used to determine the proteins interacting with *KIF2C*. **Results:** The analysis using Oncomine and GEPIA databases showed high mRNA expression of *KIF2C* in hepatocellular carcinoma tissue ($P=0.000$), and its expression level was negatively correlated with the prognosis of hepatocellular carcinoma ($P=0.000$); the expression level of *KIF2C* increased with the increase in pathological classification. The analysis using STRING database showed that PLK1, AURKB, and CENPA proteins were associated with *KIF2C*. **Conclusion:** Tumor gene database mining shows that *KIF2C* is highly expressed in hepatocellular carcinoma tissue and is associated with patient prognosis, which provides a theoretical basis for research on the pathogenesis of hepatocellular carcinoma and anti-tumor targeted therapy.

【Key words】hepatocellular carcinoma; *KIF2C* gene; database analysis

肝细胞癌(hepatocellular carcinoma, HCC)是常见的 5 种恶性肿瘤之一,居全球疾病发病率第 5 位、

死亡率第 3 位,且发病率呈逐年上升趋势^[1]。我国为 HCC 高发地区,每年约有 14 万人死于 HCC,居恶性肿瘤疾病致死率第 2 位^[2]。近年来, HCC 的诊治研究虽取得了显著进展,但由于疾病确诊通常在晚期,早已错过了最佳治疗时期,故 HCC 总体疗效无明显改观且预后性极差,患者 5 年生存率仅 15% 左右^[3]。血清甲胎蛋白(alpha fetoprotein, AFP)是目前诊断 HCC 最重要的肿瘤标志物,但由于 AFP 的敏感性和特异

作者介绍:张越美, Email: 1611316029@hqu.edu.cn,

研究方向:生物化学与分子生物学。

通信作者:何春艳, Email: hecy@hqu.edu.cn。

基金项目:福建省第四批创业创新人才引进计划资助项目(编号:Z16X0079)。

优先出版: <http://kns.cnki.net/kcms/detail/50.1046.R.20190530.1555.008.html>
(2019-05-31)

性较差,在疾病的早期诊断时,其假阳性或假阴性率可达 40%,难以对 HCC 提供精确诊断^[4]。HCC 与肝癌细胞不受生长调控系统的控制和不间断的增殖和分裂息息相关^[5-6],其中着丝粒的正常分裂在肝癌细胞的增殖过程中起重要作用。因此,探索与着丝粒相关的靶基因及其表达情况,成为深入研究 HCC 亟待解决的问题。

有丝分裂着丝粒相关驱动蛋白 *KIF2C* 位于染色体 1p34,是驱动蛋白家族-13(kinesin-13)中最具代表性的成员之一,参与染色体分离、二极纺锤体和微管解聚的形成、细胞周期的调节和有丝分裂等生命活动^[7]。*KIF2C* 不仅在纺锤体装配中起重要作用,还可确保染色体在有丝分裂过程中进行适当的分离,提示 *KIF2C* 的异常表达可能在癌症的发生发展中发挥重要作用。Chen 等^[8]认为驱动蛋白家族(kinesin superfamily proteins, KIFs)的若干成员与 HCC 的总生存时间相关。此外,还有研究报道表明 *KIF2C* 是舌癌、脑胶质瘤和非小细胞肺癌的标记分子^[9-11]。但目前尚未有文献报道阐述 *KIF2C* 与 HCC 的表达及预后情况的作用。因此,本研究通过已有的生物信息数据库对 HCC 中 *KIF2C* 基因的表达进行深入分析并探讨其临床意义,以期对 HCC 的致病机制研究和抗肿瘤靶向治疗提供线索和依据。

1 材料与方法

1.1 数据库分析方法

1.1.1 *KIF2C* 基因的表达情况 Oncomine 数据库是癌症基因芯片数据库,意在整合和挖掘癌症基因信息,可用于比较癌症和正常组织的差异表达分析^[12]。设定检索条件为① Data type:mRNA;② Analysis type:hepatocellular carcinoma vs. normal analysis;③ 设定条件:Over-expression,选择 *KIF2C* 基因后,设定条件为 Gene:*KIF2C*,其余同前条件。由北京大学研制开发的基于 TCGA 和 GTEx 数据库的 GEPIA 数据库(<http://gepia.cancer-pku.cn/>)可用于分析基因在癌症和正常组织的差异表达,可得出标准 RNA 测序数据,其中包含 33 种恶性肿瘤,8 587 个正常组织样本和 9 736 个肿瘤样本^[13],通过 GEPIA 分析 *KIF2C* 的表达情况来验证 Oncomine 数据库的结果。

1.1.2 *KIF2C* 基因与病理分级和生存期的相关性 通过 GEPIA 数据库进行 *KIF2C* 的表达水平与病理分级与 HCC 总生存期和无病生存期的相关性分析。选择 GEPIA 数据库中 TCGA 数据集,*KIF2C* 的表达水平与病理分级的分析,参数设置为 Gene:*KIF2C*;Use major stage:yes;Datasets:

LIHC;Log Scale:yes。*KIF2C* 的表达水平与肝细胞癌总生存期和无病生存期的相关性分析,参数设置为 Gene:*KIF2C*;Methods:Overall Survival/Disease Free Survival;Group Cutoff:Median;95% Confidence Interval:yes;Axis Units:Months;Datasets:LIHC。

1.1.3 *KIF2C* 相关蛋白分析 STRING 10.5 数据库(<http://string-db.org/>)是以已知蛋白质和预测蛋白质之间相互作用的数据组成的数据库,包括直接的相互作用和间接的相互作用^[14]。为了探索该基因的分子机制,用 STRING 10.5 分析 *KIF2C* 的蛋白-蛋白相互作用(protein-protein interaction, PPI),输入“*KIF2C*”,物种类型选择人类,置信度默认>0.4 (medium confidence 0.400)。

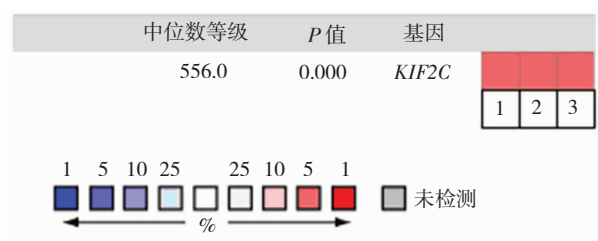
1.2 统计学分析

所有数据分析均采用在线数据库完成。Oncomine 数据库中,差异表达基因用 *t* 检验方法进行统计学分析。GEPIA 中,以 log2(fold change)展示 *KIF2C* 在肿瘤组织和正常对照组织中高通量数据的表达差异;通过单因素方差分析进行 *KIF2C* 与病理分级的关系;通过 log-rank 检验进行 *KIF2C* 与 HCC 总生存期和无病生存期的相关性分析。检验水准 $\alpha=0.05$ (双侧)。

2 结果

2.1 Oncomine 数据库 *KIF2C* 表达水平

在 Oncomine 数据库中设置的检索条件下,得到 3 个与 HCC 相关的数据集。分析结果显示,*KIF2C* 在 HCC 中明显高表达($P=0.000$),如图 1 所示。



注:1. Roessler Liver 数据集:HCC 组织 vs. 正常组织,2010;
2. Roessler Liver 2 数据集,HCC 组织 vs. 正常组织,2010;
3. Wurnbach Liver 数据集:HCC 组织 vs. 正常组织,2007
(红色表示上调,灰色表示未检测,白色表示与正常组织相当)

图 1 *KIF2C* 在肝细胞癌中表达上调

在 mRNA 水平上,*KIF2C* 在 3 个数据集中均在 HCC 组织中高表达,差异具有统计学意义(图 2)。

使用 GEPIA 分析 TCGA 中 419 个肝组织样本,其中正常肝组织 50 例(灰色),HCC 组织 369 例(红色)。*KIF2C* 在 HCC 中的表达量明显较正常肝组织高,差异具有统计学意义($P<0.01$),与 Oncomine 数据库分析结果一致(图 3)。

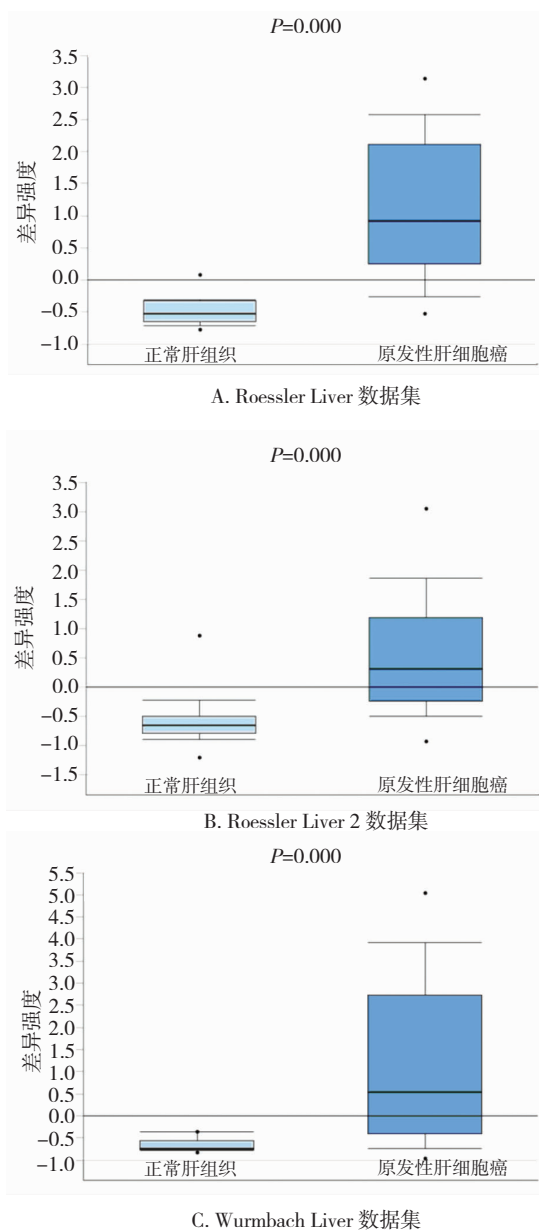


图2 Oncomine 各子数据集中 *KIF2C* 在正常肝组织和 HCC 中的表达情况

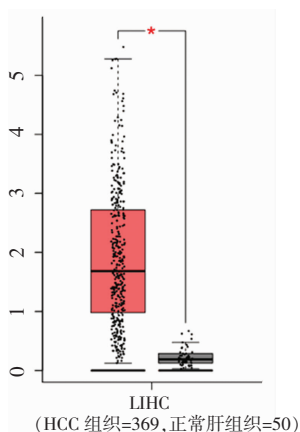


图3 *KIF2C* 在 419 个样本中的表达

2.2 *KIF2C* 表达水平与 HCC 病理分级的关系(方差分析)

选择 TCGA 的 LIHC (HCC) 数据库, 在 GEPIA 中分析 *KIF2C* 表达与 HCC 患者病理分级的关系, 病理分级 I、II、III 分别对应肝癌细胞高、中、低分化, 其分化程度越低恶性程度越高, *KIF2C* 表达水平也越高(图 4)。

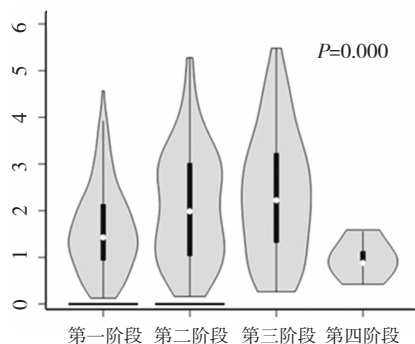
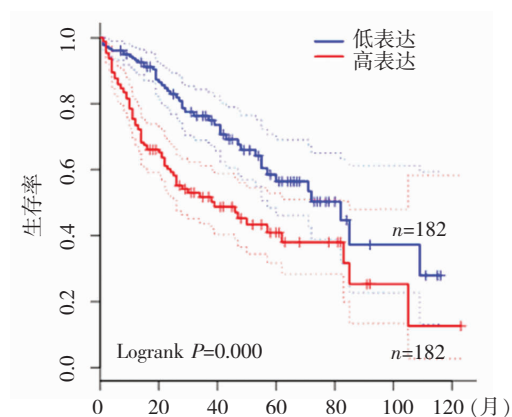


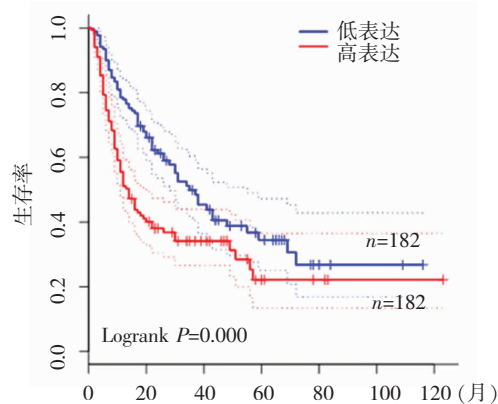
图4 病理分级中的 *KIF2C* 表达强度

2.3 *KIF2C* 表达水平与 HCC 患者生存期的关系

选择 TCGA 的 LIHC (肝细胞癌) 数据库, 在 GEPIA 中分析 *KIF2C* 表达与 HCC 患者生存期的关系, 得 *KIF2C* 高、低表达患者各 182 例的生存期曲线, 发现 *KIF2C* 表达量越高者, 其预后越差(图 5)。



A. *KIF2C* 表达与 HCC 总生存期的相关性



B. *KIF2C* 表达与 HCC 无病生存期的相关性

图5 *KIF2C* 表达水平与 HCC 患者生存期的关系

2.4 与 *KIF2C* 相关蛋白分析

利用 STRING 数据库分析 *KIF2C* 蛋白的相互作用网络, 节点数为 11 个, 如图 6 所示。其中以 *KIF2C* 为中心有多个相互作用的蛋白质, 多数蛋白质涉及细胞周期。

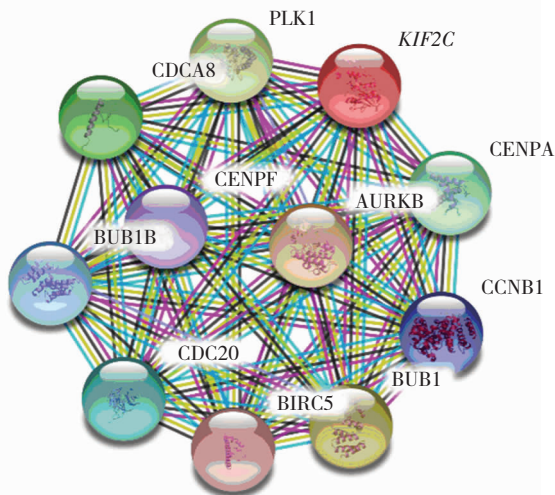


图 6 与 *KIF2C* 相关的蛋白网络图

3 讨论

Oncomine 和 GEPIA 数据库分析结果显示, 在 mRNA 水平上, 与正常肝组织相比 *KIF2C* 在 HCC 中明显高表达 ($P=0.000$), 表达水平与其预后呈负相关 ($P=0.000$); HCC 恶性程度越高, *KIF2C* 表达水平越高 ($P=0.000$); STRING 数据库分析结果显示与 *KIF2C* 相关的蛋白有 PLK1、AURKB、CENPA 等。结合文献报道, 蛋白 PLK1 在人类大多数肿瘤中高表达, 并与肿瘤细胞的增殖及患者的预后密切相关^[5], AURKB 激酶在卵母细胞减数分裂中起重要的调节作用, 与细胞增殖相关^[16]。这提示 *KIF2C* 可能在肿瘤细胞增殖通路中起重要作用。

综上所述, 本研究通过肿瘤基因数据库信息挖掘, 发现 *KIF2C* 在 HCC 组织中呈高表达, 与患者预后呈负相关, 涉及细胞周期, 与肿瘤细胞增殖密切相关, 为 HCC 致病机制研究和抗肿瘤靶向治疗提供线索和依据。但本研究所得结论均由已发表的基因芯片数据分析而来, 并未对该基因在 mRNA 和蛋白质水平上的表达量进行确定, 致使研究结果存在一定局限性。故在此思路下, 后期可进一步通过实验验证, 探索 *KIF2C* 与 HCC 发生发展机制的关系,

并将 *KIF2C* 作为抗肿瘤靶点之一进行相关性研究。

参 考 文 献

- [1] 魏矿荣, 彭侠彪, 梁智恒, 等. 全球肝癌流行概况[J]. 中国肿瘤, 2015, 24(8): 621-630.
- [2] Chen W, Zheng R, Zeng H, et al. Annual report on status of cancer in China, 2011[J]. Chin J Cancer Res, 2015, 27(1): 2-12.
- [3] 张 玥. 全球与中国肝癌疾病负担分析及肝癌早期诊断标志物的研究[D]. 北京: 北京协和医学院, 2015.
- [4] 谢 延, 赵云鹏, 张胜男, 等. 肝细胞癌的肿瘤标记物研究进展[J]. 实用预防医学, 2010, 17(6): 1243-1247.
- [5] Forner A, Reig M, Varela M, et al. Diagnosis and treatment of hepatocellular carcinoma. Update consensus document from the AEEH, SEOM, SERAM, SERVEI and SETH[J]. Med Clin (Barc), 2016, 146(11): 511.e1-22.
- [6] Roy D, Sheng GY, Herve S, et al. Interplay between cancer cell cycle and metabolism: challenges, targets and therapeutic opportunities[J]. Biomed Pharmacother, 2017, 89: 288-296.
- [7] Zhang L, Shao H, Huang Y, et al. PLK1 phosphorylates mitotic centromere-associated kinesin and promotes its depolymerase activity[J]. J Biol Chem, 2011, 286(4): 3033-3046.
- [8] Chen J, Li S, Zhou S, et al. Kinesin superfamily protein expression and its association with progression and prognosis in hepatocellular carcinoma[J]. J Cancer Res Ther, 2017, 13(4): 651-659.
- [9] Bie L, Zhao G, Wang YP, et al. Kinesin family member 2C (*KIF2C*/MCAK) is a novel marker for prognosis in human gliomas[J]. Clin Neurol Neurosurg, 2012, 114(4): 356-360.
- [10] Wang CQ, Xiang FG, Li YJ, et al. Relation between the expression of mitotic centromere-associated kinesin and the progression of squamous cell carcinoma of the tongue[J]. Oral Surg Oral Med Oral Pathol Oral Radiol, 2014, 117(3): 353-360.
- [11] 宋世龙, 毕明宏, 王洪亚, 等. *Kif2c* 在非小细胞肺癌组织中的表达及其临床意义[J]. 临床肿瘤学杂志, 2016, 21(6): 481-484.
- [12] 马小安, 吴 涛, 郭 卉, 等. 基于 Oncomine, TCGA 数据库挖掘 p57, CKS1 及 SKP2 在肝细胞肝癌中的表达及预后意义[J]. 临床医学研究与实践, 2016, 1(25): 1-4.
- [13] 齐 闯, 赵 虎, 陈 展, 等. 基于数据挖掘分析酪氨酸蛋白激酶受体 TYRO3 在膀胱癌中的表达及临床意义[J]. 现代肿瘤医学, 2018, 26(1): 57-61.
- [14] 陈 展, 崔 强, 齐 闯, 等. 基于数据挖掘分析 PEG3 基因在肾透明细胞癌中的表达及预后意义[J]. 现代肿瘤医学, 2018, 26(4): 548-553.
- [15] 周 燕, 唐运莲, 甘润良, 等. PLK1 在肿瘤发生中的研究进展[J]. 医学研究杂志, 2015, 44(3): 158-161.
- [16] 许 琳, 刘 彤, 韩 峰, 等. AURKB 与 MAPK 参与调控小鼠受精卵早期发育[J]. 中国科学(生命科学), 2012, 42(1): 37-47.

(责任编辑: 冉明会)