

网络出版时间: 2019-12-18 17:12 网络出版地址: <http://kns.cnki.net/kcms/detail/34.1065.R.20191217.1516.017.html>

人肝细胞癌中 HNF4 α 近端启动子区生物信息学分析

曹 辉¹, 许 钟², 陈燕平¹, 张玲玲², 鲁 亮¹

摘要 目的 应用生物信息学数据库挖掘数据, 预测分析肝细胞癌发生发展过程中抑癌基因肝核转录因子 4 α (HNF4 α)表达下调的转录调控机制。方法 从美国国家生物技术信息中心在线核苷酸数据库获得 HNF4 α 基因的启动子序列; 应用 PromotorScan、Patch、P-Match、AliBaba2 等在线转录因子预测软件分析 HNF4 α 转录因子结合位点; 应用在线肝脏基因数据库 liveratlas 筛查肝癌组织中表达下调的基因, 与预测的转录因子进行对比。通过 TCGA 数据库提取相关转录因子与 HNF4 α 的肝癌组织芯片表达结果, 进行相关性分析。结果 人 HNF4 α 基因在肝癌中对应转录本的启动子约 1 471 bp。应用 PromotorScan、Patch、P-Match 和 AliBaba2 等在线软件分析, 预测转录因子结合位点结果汇总后去除重复, 共涵盖 225 个转录因子。应用在线基因数据库 liveratlas 检索, 提取肝细胞癌中表达下调的基因共 2 538 个, 与上述预测的 225 个转录因子进行对比, 取交集获得目标转录因子共 17 个。相关性分析提示 HLF($r = 0.5534$)、RREB1($r = 0.4079$)、RXRA($r = 0.4247$) 等转录因子与 HNF4 α 的表达成正相关。结论 生物信息学分析提示 HLF、RREB1、RXRA 等转录因子参与 HNF4 α 在肝脏中表达的转录调控, 其中一个或多个蛋白的低表达与肝细胞癌发生发展中抑癌基因

HNF4 α 的表达下调有关。

关键词 HNF4 α ; 启动子; 转录因子; 生物信息学; 肝细胞癌

中图分类号 R 735.7

文献标志码 A **文章编号** 1000-1492(2020)01-0080-04

doi: 10.19405/j.cnki.issn1000-1492.2020.01.017

肝核转录因子 4 (hepatocyte nuclear factor 4, HNF4) 是细胞核激素受体家族中的核心成员之一, 在调控肝细胞分化和维持其生物学功能上具有重要作用, 其中 HNF4 α 与肝脏特异基因的表达密切相关^[1]。研究^[2-3]表明 HNF4 α 在肝细胞癌变中扮演着抑癌基因的角色, 在肝硬化、肝癌中其表达明显下调。作为肝细胞中的一个重要的转录因子, HNF4 α 与超过 40% 肝基因的启动子区相结合^[4], 其重要性可见一斑, 由此不难理解近年来肝脏疾病研究中 HNF4 α 受到国内外学者的重视^[5-7]。该研究拟应用生物信息学方法对 HNF4 α 基因启动子区进行序列分析, 预测该基因近端启动子区的转录因子结合位点, 进一步分析其在肝细胞癌中下调有关的转录因子, 为研究其转录调控机制、探索肝癌治疗新思路提供理论基础。

1 材料与方法

1.1 基因启动子区序列的获得 从美国国家生物技术信息中心在线核苷酸数据库(Gene Bank)获得 HNF4 α 基因序列, 基于 HNF4 α 的基因序列 ID, 下载

2019-08-26 接收

基金项目: 国家自然科学基金(编号: 81860439); 贵州省科学技术基金(编号: 黔科合 J 字[2012]2241 号)

作者单位: 贵州省人民医院¹ 肿瘤科、² 消化内科, 贵阳 550002

作者简介: 曹 辉, 女, 博士, 副主任医师;

许 钟, 男, 博士, 副主任医师, 责任作者, E-mail: xucong0567@163.com

ulated into culture dishes at $1 \times 10^6/\text{ml}$, and each group was given 3 replicate wells. The experiment was divided into two groups. Group 1: miR-25-3p up-regulated (transfected with miR-25-3p mimics), miR-25-3p down-regulated (transfected with miR-25-3p inhibitor) and NC group; Group 2: NOTCH1 up-regulated (transfected NOTCH1 mimics), NOTCH1 down-regulated (transfected with NOTCH1 inhibitor), and NC group. Real-time quantitative PCR was used to detect the expression levels of miR-25-3p and NOTCH1 in each group, transwell chamber was used to detect cell invasion ability, CCK-8 was used to detect cell proliferation, bioinformatics analysis and luciferase activity assay were used to determine whether NOTCH1 was Target gene of miR-25-3p. **Results** NOTCH1 is the target gene of miR-25-3p. Results revealed that cell invasion and proliferation increased with the overexpression of miR-25-3p by miRNA mimics and decreased with the suppression of NOTCH1. **Conclusion** miR-25-3p promotes the invasion and proliferation of human ESCC cells by targeting NOTCH1, and miR-25-3p is expected to become a new target for ESCC treatment.

Key words miR-25-3p; ESCC; NOTCH1; invasion; proliferation; bioinformatics analysis

NCBI 网站上外链接 GeneCopoeia Inc. 所提供的启动子序列。

1.2 启动子区转录因子结合位点的预测分析 通过在线启动子区域预测分析软件 PromotorScan(<http://www-bimas.cit.nih.gov/molbio/proscan/>) 分析启动子序列 通过转录因子数据库 TRANSFAC 中的在线转录因子预测软件包括 Patch、P-Match、AliB-aba2 (<http://www.gene-regulation.com/pub/programs.html>) 预测 HNF4 α 启动子序列顺式作用元件 分析其可能转录因子结合位点。

1.3 肝癌组织中相关转录因子的筛选 应用在线肝脏基因数据库 liveratlas (<http://liveratlas.hupo.org.cn/>)^[8] 筛查肝癌组织中表达下调的基因 与预测的转录因子进行对比 获得目标转录因子数据集。

1.4 相关性分析 通过 TCGA 数据库 (<https://cancergenome.nih.gov/>) 提取相关转录因子与 HNF4 α 的肝癌组织芯片表达结果, 进行相关性分析。

1.5 统计学处理 采用 GraphPad Prism 5 软件进行数据分析, 相关性分析计算相关系数 r , 以 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 人 HNF4 α 基因基本信息及启动子序列 人
HNF4 α 基因 ID: 3 172 ,定位于 20 号染色体正链
(20q13. 12) ,全长 77 045 bp 其 mRNA ID 包括: NM
_000457. 4、NM_001030003. 2、NM_001030004. 2、NM
_001258355. 1、NM _ 001287182. 1、NM _ 001287
183. 1、NM _ 001287184. 1、NM _ 175914. 4、NM _
178849. 2 和 NM_178850. 2。

通过 ACEVIEW (<https://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/>) 了解其转录本情况，选择 NM_000457.4 对应的基因启动子克隆产品 ID 为 HPRM20338, HPRM30433。

提取两者序列，应用 NCBI Blast 比对，HPRM30433 序列包含 HPRM20338 转录起始位点 (TSS) 上游的全部序列，故取 HPRM30433 序列来进一步分析，其全长 1 471 bp，位于转录起始位点 -1 371 bp 至 +99 bp 之间。

2.2 人 HNF4 α 基因启动子区转录因子结合位点预测分析结果 应用 PromotorScan 分析启动子序列, 分析结果共 35 个转录因子结合位点, 其中 Sp1 结合位点共 16 个、AP-2 共 6 个。去重后涵盖 AP-2、APRT-mouse_US、EARLY-SEQ1、GCF、HSV-tk-2nd_

distal_si、JCV_repeated_sequenc、LF-A2、PuF、SDR_RS、Sp1、T-Ag、UCE. 2 共 12 个转录因子。

应用 Patch 在线软件分析 ,参数设置 set of sites 选择 “vertebrates” ,Lower score boundary 设为 “90” , 其它采用默认设置 ,预测结果共 1 567 个转录因子结合位点 ,手工筛查物种为人的转录因子结合位点共 785 个 经汇总去重 ,共涵盖了 161 个转录因子。

应用 P-Match 在线软件分析,当 Cut-offs 使用参数为“to minimize the sum of both error rates”时,预测到 17 个转录因子结合位点(图 1),去除重复共涵盖 12 个转录因子: AP-4、ARP-4、CDP CR3、COUP-TF、CREB、c-Rel、E47、HNF-4、NF-kappaB、NRF-2、Olf-4、ZID; 当使用参数为“to minimize false negative matches”,共预测到 324 个转录因子结合位点,经汇总去重,共涵盖 27 个转录因子: AP-4、AREB6、ARP-1、BSAP、CDP、CDP CR3、COUP-TF、CREB、c-Rel、E47、Elk-4、Evi-4、HFH-4、HLF、HNF-4、myogenin/NF-4、NF-E2、NF-kappaB、NF-kappaB (p50)、NF-kappaB (p65)、Nkx2-5、NRF-2、Olf-4、RREB-4、STATx、YY1、ZID。

图 1 P-Match 预测转录因子结合位点结果图

应用 AliBaba2 在线软件分析, 取在线软件默认设置参数, 预测结果共 164 个转录因子结合位点, 经汇总去重, 共涵盖了 61 个转录因子。

以上软件预测结果汇总后去除重复,结果共涵盖了 225 个转录因子。

2.3 筛查肝癌中人 HNF4 α 启动子区可能转录因子 应用在线基因数据库 liveratlas 检索 ,选择 “Disease” ,以 “hepatocellular carcinoma” 为检索词检索 ,依次点击 “HuLDi00052” 和 “more...” ,提取肝细胞癌中表达下调的基因共 2 538 个 ,与上述预测的 225 个转录因子进行对比 ,取交集获得目标转录因子。结果预测的转录因子中共有 17 个在肝细胞癌中表达下调: Egr-1 、 GATA-3 、 HLF 、 IRF-4 、 MAZ 、 MyoD 、

Pax-5、POU1F1a、RelA、RREB1、RXR-alpha、SMAD-3、Sp1、TBP、TCF4、USF2、WT1。

2.4 表达相关性分析 从 TCGA 数据库中提取以上 17 个转录因子及 HNF4 α 在癌旁组织及肝细胞癌中表达的数据,用 GraphPad Prism 5 软件进行相关性分析及绘图,结果显示 HLF、RREB1、RXRA 等三个转录因子与 HNF4 α 的表达成正相关,计算的 Pearson 相关系数 r 分别为 0.553 4、0.407 9、0.424 7, $P < 0.0001$,差异有统计学意义(图 2~4)。

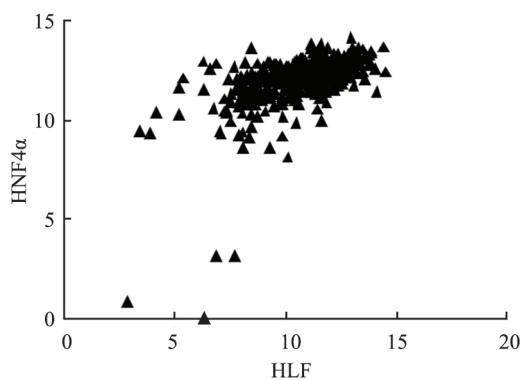


图 2 转录因子 HLF 与 HNF4 α 的相关性分析

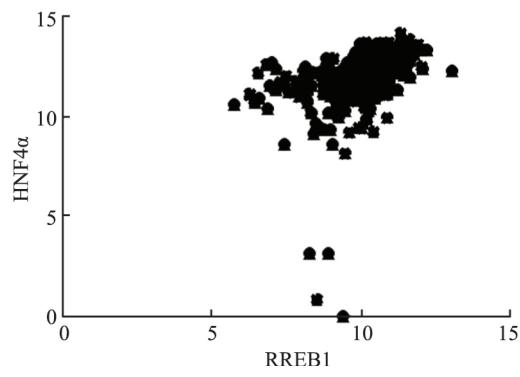


图 3 转录因子 RREB1 与 HNF4 α 的相关性分析

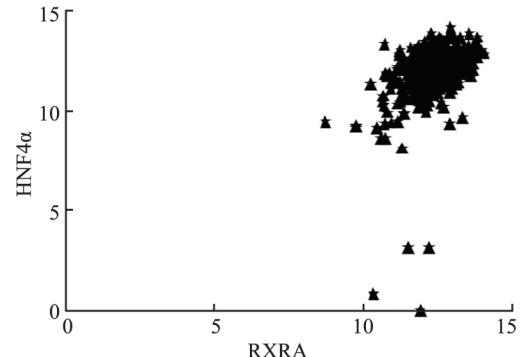


图 4 转录因子 RXRA 与 HNF4 α 的相关性分析

3 讨论

研究抑癌基因表达调控机制既有助于揭开恶性肿瘤发生发展的分子机制,也可为探索癌症治疗新策略提供新的思路。肝癌是我国和世界上最常见的恶性肿瘤之一,基于临床样本检测的研究越来越丰富^[9],多种芯片的数据也通过在线数据库公开报道。对这些数据进行深度分析,结合启动子区序列转录因子的生物信息学分析,可为抑癌基因转录调控研究提供更精准的方向。

HNF4 α 是肝癌发生中受关注的抑癌基因之一,相关的信号通路研究表明 JNK/c-Jun、Notch 和 MAPK 等信号通路可抑制 HNF4 α 的功能或抑制其表达^[10~12],PI3K/AKT 信号通路可以促进其表达^[13]。而对 HNF4 α 转录调控机制的进一步解析,可在分子水平上更充分的阐明其表达调控机制,有助于深入认识肝细胞癌变的机制。

HNF4 α 在肝癌发生中表达受抑制,与之相对应,调控其表达的转录因子在肝癌中的表达通常是下调的。本研究应用北京蛋白质组研究中心/蛋白质组国家重点实验室建立的在线肝脏基因数据库 liveratlas^[8],基于公开发表的芯片等数据提取到在肝细胞癌中表达下调的基因共 2 538 个。在 NCBI 的基因数据库中获得 HNF4 α 的启动子序列后,应用多种启动子序列分析软件来预测其顺式作用元件,合并多个软件预测结果以减少遗漏的可能,预测结果共对应 225 个转录因子。通过两组数据的交集,有效缩小了在肝癌中研究 HNF4 α 的转录调控蛋白的范围。

随后,通过 TCGA 数据库,详细地分析了潜在转录因子与 HNF4 α 表达的相关性,很幸运地从 17 个潜在目标中找到 3 个转录因子与 HNF4 α 表达成正相关,使进一步的实验研究变得更加可行。通过文献检索,未见 HLF、RREB1 与 HNF4 α 之间相互作用的报道; Tomaru et al^[14] 通过 Matrix RNAi(结合 siRNA 敲除和 QRT-PCR) 在 HepG2 细胞系中的研究结果表明,RXRA 可调控 HNF4 α 的表达,为本研究的分析结果提供了佐证,下一步将在此基础上,通过小鼠肝癌模型等体内实验来验证。

深度的数据发掘,可以借助于综合利用公开的数据库资源、结合生物信息学分析、充分应用计算机和网络资源来实现其价值。尽管预测分析结果的真

实情况有待体内外实验的评价,但通过对生物信息的获取与加工、分析与阐释,可有效缩小在肝癌发生中HNF4 α 转录调控机制探索的研究范围,节约成本和资源,提高研究效率,是一行之有效的途径,值得推广应用。

参考文献

- [1] Watt A J , Garrison W D , Duncan S A. HNF4: a central regulator of hepatocyte differentiation and function [J]. Hepatology , 2003 , 37(6) : 1249 – 53.
- [2] Frades I , Andreasson E , Mato J M , et al. Integrative genomic signatures of hepatocellular carcinoma derived from nonalcoholic fatty liver disease [J]. PLoS One 2015 ,10(5) : e0124544.
- [3] Ning B F , Ding J , Yin C , et al. Hepatocyte nuclear factor 4 alpha suppresses the development of hepatocellular carcinoma [J]. Cancer Res , 2010 ,70(19) : 7640 – 51.
- [4] Odom D T , Zilzperger N , Gordon D B , et al. Control of pancreas and liver gene expression by HNF transcription factors [J]. Science , 2004 ,303(5662) : 1378 – 81.
- [5] Fekry B , Ribas-Latre A , Baumgartner C , et al. Incompatibility of the circadian protein BMAL1 and HNF4alpha in hepatocellular carcinoma [J]. Nat Commun , 2018 ,9(1) : 4349.
- [6] Cai W Y , Lin L Y , Hao H , et al. Yes-associated protein/TEA domain family member and hepatocyte nuclear factor 4-alpha (HNF4alpha) repress reciprocally to regulate hepatocarcinogenesis in rats and mice [J]. Hepatology , 2017 ,65(4) : 1206 – 21.
- [7] 张勇,吴宁,马永,等. 肝细胞核因子4 α 抑制大鼠肝癌细胞增殖、侵袭能力和新生血管生成相关基因表达 [J]. 南京医科大学学报(自然科学版) ,2016 ,36(4) : 430 – 4.
- [8] Zhang Y , Yang C , Wang S , et al. Liver atlas: a unique integrated knowledge database for systems-level research of liver and hepatic disease [J]. Liver Int , 2013 ,33(8) : 1239 – 48.
- [9] Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma [J]. Cell , 2017 ,169(7) : 1327 – 41.
- [10] Li T , Jahan A , Chiang JY. Bile acids and cytokines inhibit the human cholesterol 7 alpha-hydroxylase gene via the JNK/c-jun pathway in human liver cells [J]. Hepatology , 2006 ,43(6) : 1202 – 10.
- [11] Tanimizu N , Miyajima A. Notch signaling controls hepatoblast differentiation by altering the expression of liver-enriched transcription factors [J]. J Cell Sci , 2004 ,117(Pt 15) : 3165 – 74.
- [12] Reddy S , Yang W , Taylor D G , et al. Mitogen-activated protein kinase regulates transcription of the ApoCIII gene. Involvement of the orphan nuclear receptor HNF4 [J]. J Biol Chem , 1999 ,274 (46) : 33050 – 6.
- [13] Magner N L , Jung Y , Wu J , et al. Insulin and IGFs enhance hepatocyte differentiation from human embryonic stem cells via the PI3K/AKT pathway [J]. Stem Cells , 2013 ,31(10) : 2095 – 103.
- [14] Tomaru Y , Nakanishi M , Miura H , et al. Identification of an inter-transcription factor regulatory network in human hepatoma cells by Matrix RNAi [J]. Nucleic Acids Res 2009 ,37(4) : 1049 – 60.

Bioinformatics analysis of the promoter region of HNF4 α gene in human hepatocellular carcinoma

Cao Hui¹ , Xu Zhong² , Chen Yanping¹ , et al

(¹*Dept of Oncology* ,²*Dept of Gastroenterology* ,Guizhou Provincial People's Hospital , Guiyang 550002)

Abstract Objective To predict and analyze the mechanism of downregulation of HNF4 α expression in hepatocellular carcinoma (HCC) by bioinformatics analysis. **Methods** The promoter sequence of HNF4 α gene obtained from NCBI GenBank was analyzed by transcription factor prediction software, including PromotorScan , Patch , P-Match and AliBaba2. Meanwhile, the down-regulated genes in human HCC were screened by Liveratlas. Then the intersected genes were analyzed based on the data extracted from TCGA database. **Results** The promoter of the corresponding transcript of human HNF4 α gene in HCC is about 1 471 bp. After removing duplication, the pool of predicting transcription factor binding sites covers 225 transcription factors by using online software. A total of 2 538 down-regulated genes in HCC were extracted from liveratlas online gene database, and 17 transcription factors were obtained by compared with the transcription factor pool predicted above. The correlation analysis indicated that HLF ($r=0.553\ 4$) , RREB1 ($r=0.407\ 9$) and RXRA ($r=0.424\ 7$) were positively correlated with the expression of HNF4 α . **Conclusion** Bioinformatics analysis indicates that HLF , RREB1 and RXRA may be involved in the transcriptional regulation of HNF4 α expression in liver. The faint expression of one or more of them may be related to the down-regulation of HNF4 α expression in the development of HCC.

Key words HNF4 α ; promoter; transcription factor; bioinformatics; HCC