

• 论 著 •

# LINC01667 通过上调 FABP5 重塑肝细胞癌脂肪酸代谢\*

张凯楠<sup>1,2</sup>, 李 翔<sup>3</sup>, 伏建峰<sup>4△</sup>

1. 新疆医科大学研究生学院, 新疆乌鲁木齐 830017; 2. 新疆维吾尔自治区人民医院检验科, 新疆乌鲁木齐 830000; 3. 新疆生产建设兵团第十三师红星医院检验科, 新疆哈密 839000; 4. 新疆军区总医院全军临床检验诊断中心, 新疆乌鲁木齐 830000

**摘 要:**目的 探讨 LINC01667 对脂肪酸结合蛋白 5(FABP5)的调控作用及对肝细胞癌(HCC)细胞脂肪酸代谢的影响。方法 采用双荧光素酶报告实验检测 LINC01667 与 FABP5 启动子的结合能力,以及对启动子活性的调控作用,采用尼罗红染色和游离脂肪酸摄取实验检测过表达 LINC01667 对 HCC 细胞内脂滴水平和细胞游离脂肪酸摄取能力的影响。应用 Transwell 实验检测 LINC01667/FABP5 调控轴在 HCC 细胞迁移和侵袭中的作用效应。结果 LINC01667 可与 FABP5 启动子结合,并上调 FABP5 表达( $P<0.05$ );过表达 LINC01667 可增强 HepG2 及 HuH7 细胞中脂滴的蓄积,并且增强其对游离脂肪酸的摄取能力( $P<0.05$ );靶向抑制 FABP5 后,由 LINC01667 介导的 HCC 细胞游离脂肪酸摄取能力增强效应及其迁移、侵袭能力均被显著减弱( $P<0.01$ )。结论 LINC01667 可通过上调 FABP5 重塑 HCC 细胞脂肪酸代谢,进而促进 HCC 恶性进程。

**关键词:**肝细胞癌; LINC01667; 脂肪酸结合蛋白 5; 脂肪酸代谢

**DOI:**10.3969/j.issn.1673-4130.2025.02.001 **中图法分类号:**R735.7

**文章编号:**1673-4130(2025)02-0129-07 **文献标志码:**A

## LINC01667 reprogram fatty acid metabolism in hepatocellular carcinoma by up-regulation of FABP5\*

ZHANG Kainan<sup>1,2</sup>, LI Xiang<sup>3</sup>, FU Jianfeng<sup>4△</sup>

1. Graduate School of Xinjiang Medical University, Urumqi, Xinjiang 830054, China; 2. Department of Clinical Laboratory, Xinjiang Uygur Autonomous Region People's Hospital, Urumqi, Xinjiang 830000, China; 3. Department of Clinical Laboratory, Red Star Hospital of the 13th Division of Xinjiang Production and Construction Corps, Hami, Xinjiang 83900, China; 4. Clinical Laboratory Diagnosis Center, General Hospital of Xinjiang Military Region, Urumqi, Xinjiang 830000, China

**Abstract: Objective** To investigate the regulation of LINC01667 on fatty acid binding protein 5 (FABP5) and its effect on fatty acid metabolism in hepatocellular carcinoma (HCC) cells. **Methods** Based on the dual luciferase reporter assay, the binding ability of LINC01667 to FABP5 promoter and its regulatory effect on promoter activity were detected. In addition, the effects of overexpression of LINC01667 on the content of intracellular lipid droplets in HCC cells and on the cellular free fatty acid uptake capacity were detected by Nile red staining and free fatty acid uptake assay, respectively. The effect of LINC01667/FABP5 on the migration and invasion of HCC cells was observed by Transwell experiment. **Results** LINC01667 could bind to FABP5 promoter and up-regulate FABP5 expression ( $P<0.05$ ). Overexpression of LINC01667 could enhance the accumulation of lipid droplets in HepG2 and HuH7 cells and enhance the uptake of free fatty acids ( $P<0.05$ ). After targeted inhibition of FABP5, the enhancement of free fatty acid uptake of HCC cells mediated by LINC01667 and its migration and invasion ability were significantly weakened ( $P<0.01$ ). **Conclusion** LINC01667 could remodel the fatty acid metabolism of HCC cells through the up-regulation of FABP5, and then promote the malignant process of HCC.

**Key words:** hepatocellular carcinoma; LINC01667; fatty acid binding protein 5; fatty acid metabolism

肝细胞癌(HCC)是最常见的原发性肝癌病理类型,相关数据显示,HCC已成为全球第六大常见癌症和第三大癌症相关死亡的恶性肿瘤疾病<sup>[1-2]</sup>。近年来,随着人们饮食结构的变化,非酒精性脂肪肝

\* 基金项目:新疆维吾尔自治区自然科学基金面上项目(2022D01C245)。

作者简介:张凯楠,男,博士研究生,主要从事肝细胞癌分子机制方向研究。△ 通信作者,E-mail:dxpjf@163.com。

(NAFLD)在人群中的发病率呈逐年上升趋势,约 25%的 NAFLD 患者最终演化为肝硬化或肝癌,表明 NAFLD 正在成为 HCC 的主要危险因素<sup>[3]</sup>。已有研究证实,NAFLD 的根本原因是肝细胞中脂肪酸代谢(FAM)紊乱导致的甘油三酯异常聚集,而 NAFLD 向 HCC 演进的最基本机制则可能是 FAM 的持续失调<sup>[4]</sup>。研究表明,长链非编码 RNA(LncRNA)在 HCC 相关的 FAM 重编程中起着极其重要的作用。但是,具体作用效应和可能的分子机制仍需进一步完善。脂肪酸结合蛋白 5(FABP5)属于脂肪酸结合蛋白家族,其与长链脂肪酸的摄取、转运及代谢调节具有密切相关性<sup>[5]</sup>。本课题组前期研究发现,LINC01667 作为癌基因,可能具有调控 FABP5 蛋白表达的作用<sup>[6]</sup>。但是,FABP5 是否参与 HCC 脂肪酸代谢重编程,以及 LINC01667 在上述生物学过程中是否发挥重要作用,尚需进一步探究。

基于 HCC 的治疗现状和对 FAM 重编程在 HCC 进程中作用效应的认知,并结合课题组前期研究发现,本研究拟针对 LINC01667 及其相关分子 FABP5,对其调控 HCC 代谢重编程的作用效应及其可能的分子生物学机制进行研究,以期对 HCC 的预防和治疗提供参考。

## 1 材料与方法

**1.1 主要仪器与试剂** 荧光显微镜(日本 OLYMPUS 公司,型号:IX73);全波长酶标仪(美国 Thermo Scientific 公司,型号:Multiskan 60);流式细胞分析仪(Guava easyCyte HT 公司,型号:Millipore);萤火虫荧光素酶检测(Promega,货号:E1910);脂质荧光染色试剂盒(尼罗红法,北京索莱宝,货号:G1264);Invitrogen™ BODIPY™ FL C16(美国 Thermo Scientific 公司,D3821);Transwell 小室(美国 Corning 公司,货号:3422)。

**1.2 细胞培养和抗体** 人 HCC 细胞 HepG2 及人肝癌细胞 HuH7 购自典型培养物保藏中心细胞库(中国科学院委员会)。细胞在补充有 10%胎牛血清(美国 Gibco BRL)的 Dulbecco 改良 Eagle 培养基(美国 Gibco BRL)中,置于 5% CO<sub>2</sub>、37 °C 培养箱培养。FABP5(ab255276)及 GAPDH(ab8245)抗体购自艾博抗贸易有限公司。

## 1.3 方法

**1.3.1 双荧光素酶报告实验** 以 pGL3-basic、PCD-NA3.1 为载体分别构建 LINC01667、FABP5 启动子对照和实验质粒,按研究所需分为 4 组:pGL3-basic+PCDNA3.1 组、pGL3-basic+LINC01667-PCDNA3.1 组、FABP5-pGL3-basic+PCDNA3.1 组和 FABP5-pGL3-basic+LINC01667-PCDNA3.1 组。以海肾荧光素酶基因作为内参进行双荧光素酶报告实验检测,

对 LINC01667 可否结合 FABP5 启动子并增强其活性进行验证。

**1.3.2 蛋白质印迹(Western blot)法检测 FABP5 表达** 首先,使用 1% Triton X-100 裂解缓冲液裂解细胞,然后,使用 BCA 定量法测定细胞裂解物中的总蛋白质浓度。经 SDS-PAGE 电泳、纯化、转膜、封闭后,使用 FABP5 抗体孵育过夜(4 °C),清洗后使用辣根过氧化物酶标记二抗孵育 1 h(37 °C),对膜进行显影后采集图像,最后使用图像分析软件(Image-Pro Plus 6.0)分析图像。

**1.3.3 过表达 LINC01667 的 HCC 细胞系构建** 取生长状态良好并处于对数生长期的 HepG2 和 HUH7 细胞,计数后按照每孔  $3 \times 10^5$  /mL 于六孔板上进行铺板,加入完全培养基常规培养过夜,待转染前 2 h,吸出原有培养基,更换为新鲜 DMEM 培养基。使用过表达 LINC01667 的慢病毒对细胞进行转染,转染 24 h 后,使用磷酸缓冲液(PBS)漂洗后,更换完全培养基进行常规培养,后续使用嘌呤霉素进行稳转细胞株的筛选。干扰 FABP5 的表达采用小干扰 RNA(siRNA)进行,详细方案见文献[7]。

**1.3.4 脂质荧光染色分析** 将过表达 LINC01667 的 HepG2 和 HuH7 细胞和对照 HepG2 和 HuH7 细胞培养至六孔板内,待细胞融合度达到 80%后,撤去培养基,使用 PBS 清洗细胞 3 次,收集细胞行脂质荧光染色(尼罗红法)检测,操作严格按试剂盒说明书进行。

**1.3.5 HCC 细胞系对游离长链脂肪酸摄取能力的测定** 将过表达 LINC01667 的 HepG2 和 HuH7 细胞和对照 HepG2 和 HuH7 细胞培养于直径 6 cm 培养皿内,待细胞融合度达 80%后,撤去培养基,使用 PBS 清洗细胞 3 次,并使用 PBS 饥饿培养 2 h,向培养液中加入 200 nmol/L Invitrogen™ BODIPY™ FL C16,模拟长链脂肪酸摄取过程,上机于流式细胞分析仪,检测细胞内荧光强度。

**1.3.6 HCC 细胞迁移和侵袭能力检测** 使用 Transwell 实验检测肿瘤细胞的迁移和侵袭能力。在 Transwell 小室上层内接种对应处理后的细胞(侵袭实验需提前使用基质胶对 Transwell 小室内膜进行包被),Transwell 上室中加入无血清培养基,下室中加入含 20%血清的培养基,随后置于培养箱中,48 h 后进行染色、观察。实验以 3 个随机视野中小室外层膜上黏附的细胞作为统计量,计算迁移和侵袭的细胞数。详细方案参照文献[7]。

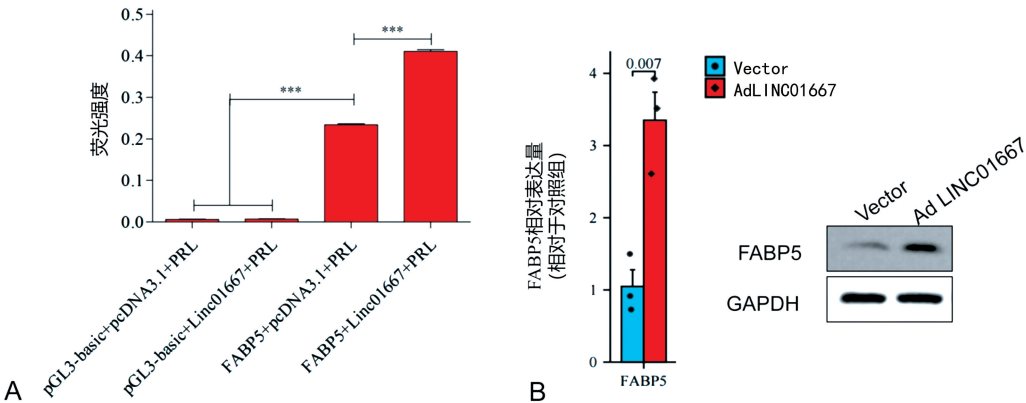
**1.4 统计学处理** 使用 SPSS22.0 进行数据统计分析,呈正态分布的计量资料以  $\bar{x} \pm s$  表示,使用独立样本 *t* 检验进行组间比较。 $P < 0.05$  差异有统计学意义。

## 2 结果

**2.1 LINC01667 可结合 FABP5 启动子并上调其表**

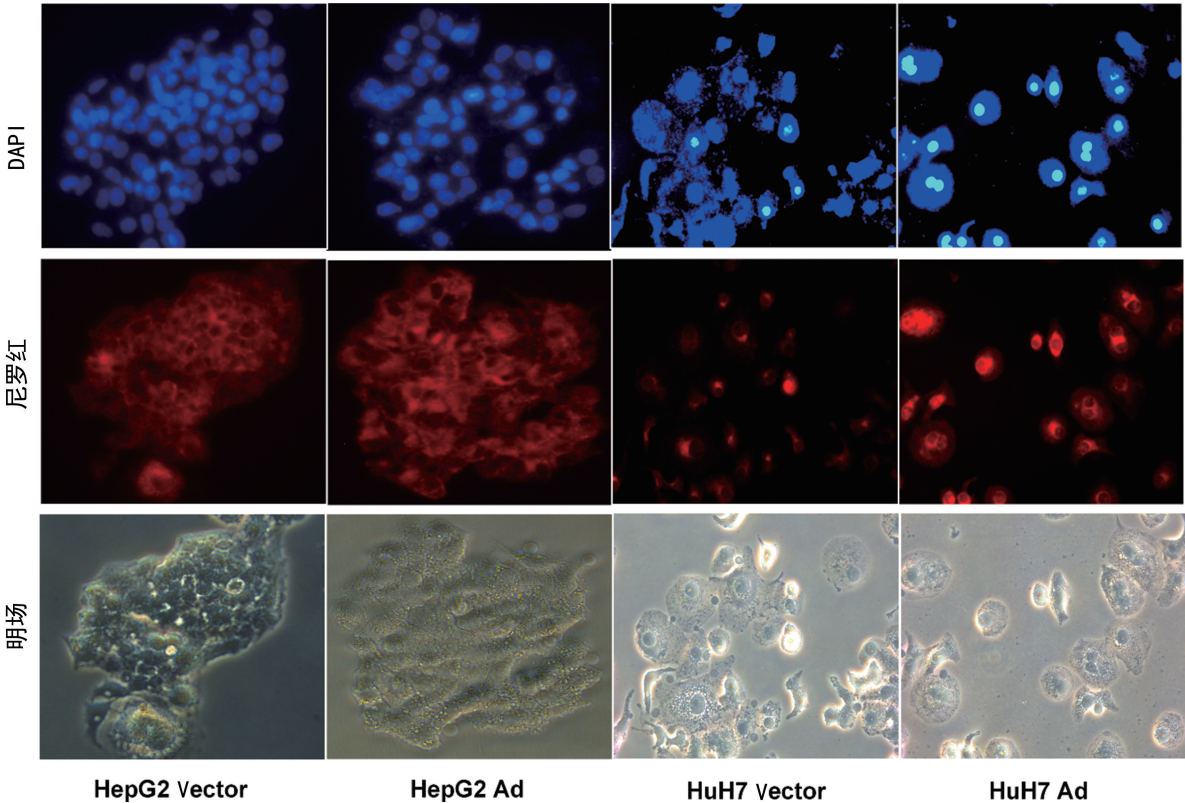
达 双荧光素酶报告实验结果显示, LINC01667 可与 FABP5 启动子区域结合, 且具有激活 FABP5 启动子活性, 进而促进其转录、翻译表达的作用效应(图 1A); Western blot 结果显示, 过表达 LINC01667 后, FABP5 的蛋白水平表达明显上调(图 1B), 差异有统计学意义( $P<0.01$ )。

2.2 过表达 LINC01667 后 HCC 细胞内的脂滴水平



注:A 为 LINC01667 与 FABP5 结合的双荧光素酶报告实验;B 为过表达 LINC01667 后 HCC 细胞 FABP5 的表达变化;Vector 代表空载体组, Ad LINC01667 代表过表达 LINC01667 组; \*  $P<0.05$ , \* \*  $P<0.001$ 。

图 1 FABP5 的启动子活性检测及表达量分析



注:蓝色代表细胞核,红色代表脂滴, HepG2 Vector 及 HuH7 Vector 代表仅转入空载体的对照细胞, HepG2 Ad 及 HuH7 Ad 代表过表达 LINC01667 的细胞。

图 2 细胞内脂滴的尼罗红染色( $\times 400$ )

2.3 LINC01667 对 HepG2 和 HuH7 细胞的脂肪酸摄取能力的影响 LINC01667 增强 HepG2 及 HuH7 细胞系对游离脂肪酸的摄取能力 以 Invitrogen™

增多 为分析过表达 LINC01667 是否会影响 HCC 细胞系的脂代谢, 采用尼罗红技术对过表达 LINC01667 的细胞及对照细胞进行染色, 结果显示: LINC01667 过表达的 HepG2 细胞和 HuH7 细胞的细胞内脂滴水平(即红色荧光)均较对照组更强, 提示 LINC01667 可促进 HCC 细胞内的脂滴蓄积。见图 2。

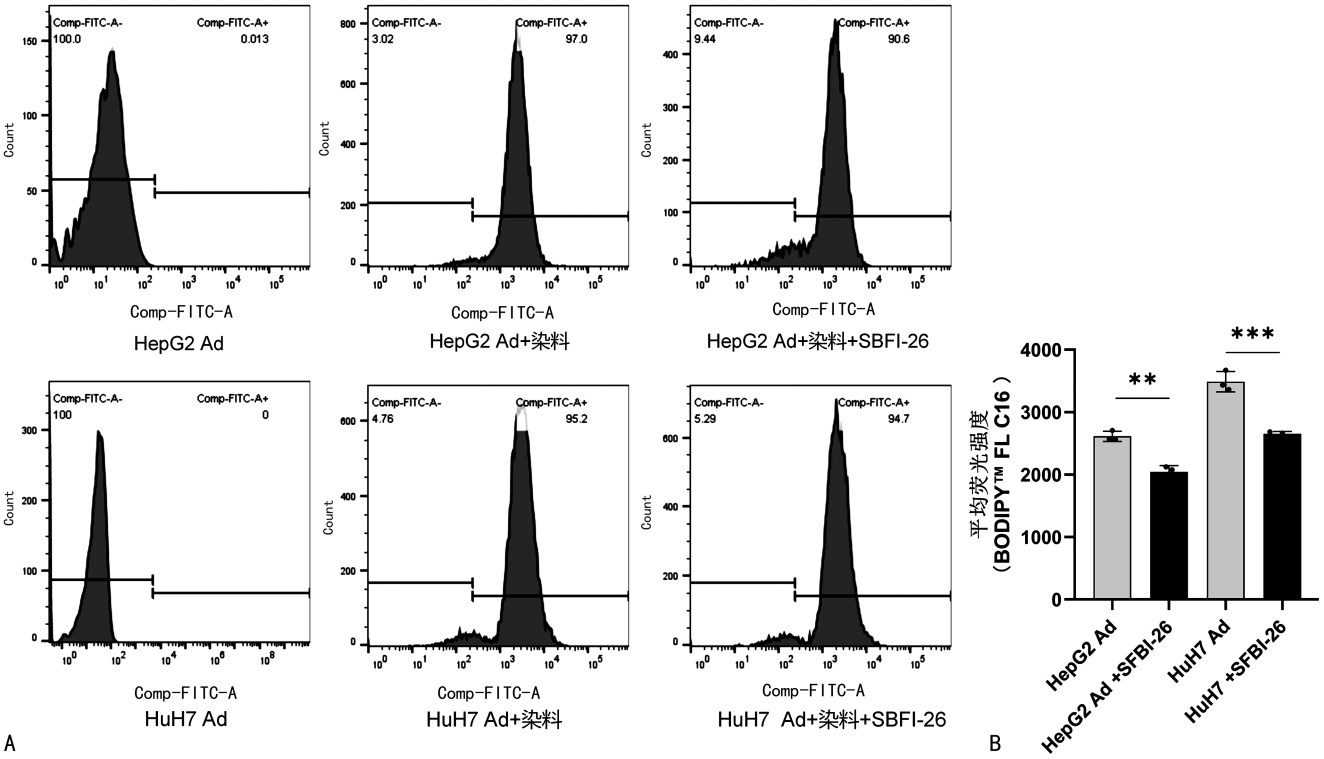
BODIPY™FL C16 模拟内环境中的游离脂肪酸, 流式细胞术检测显示:在 HepG2 和 HuH7 细胞中, 过表达 LINC01667 后, 细胞内平均荧光强度显著高于对照细



胞组,差异有统计学意义( $P<0.01$ )。结果提示,过表达 LINC01667 后,HCC 细胞系对游离脂肪酸的摄取能力显著增强。见图 3。

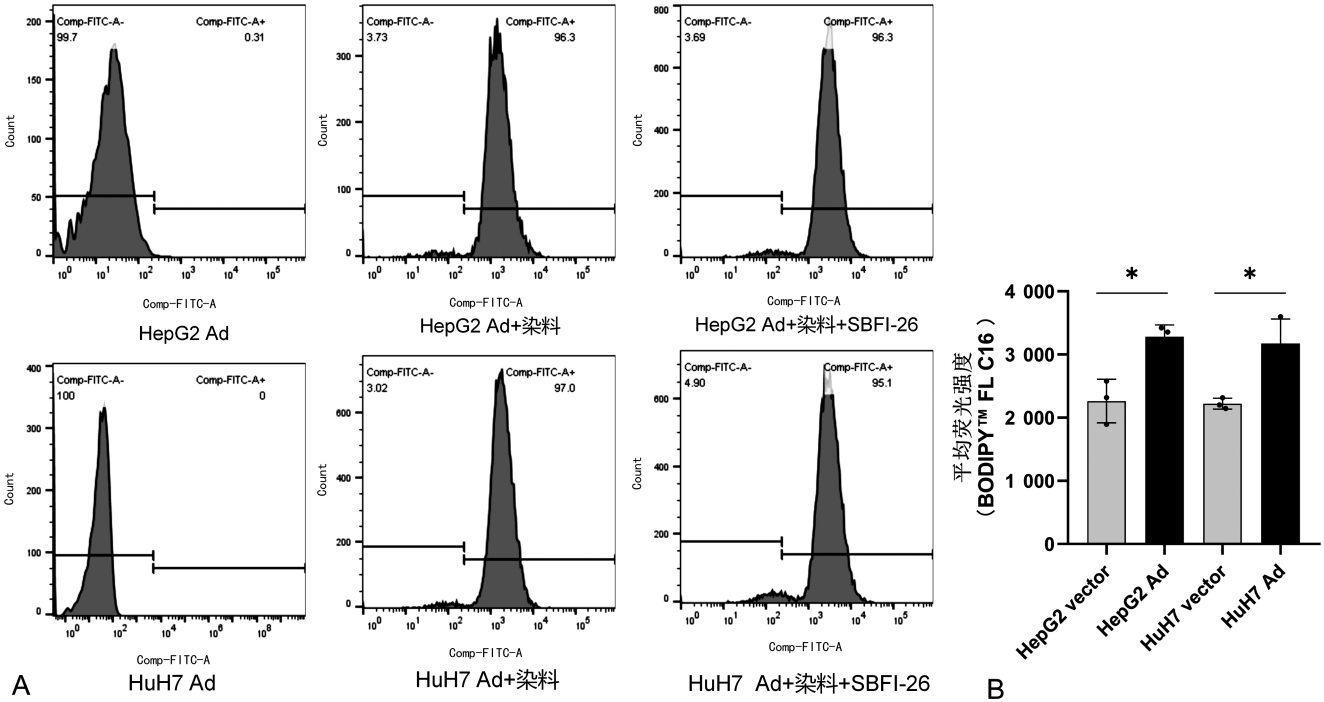
2.4 抑制 FABP5 对 LINC01667 介导的 HepG2 和 HuH7 细胞游离脂肪酸摄取能力的影响 以 FABP5

靶向抑制剂 SBFI-26 对过表达 LINC01667 后的 HepG2 和 HuH7 细胞进行体外干预,流式细胞术检测显示:靶向抑制 FABP5 后,由过表达 LINC01667 介导的脂肪酸摄取能力增强效应被显著抑制,差异有统计学意义( $P<0.01$ )。见图 4。



注:A 为流式细胞术检测结果;B 为各组统计结果;HepG2 Vector 及 HuH7 Vector 代表仅转入空载体的对照细胞,HepG2 Ad 及 HuH7 Ad 代表过表达 LINC01667 的细胞;Count 为细胞数; \*  $P<0.05$ 。

图 3 LINC01667 对 HepG2 和 HuH7 细胞脂肪酸摄取能力的影响



注:A 为流式细胞术检测结果;B 为各组统计结果;HepG2 Vector 及 HuH7 Vector 代表仅转入空载体的对照细胞,HepG2 Ad 及 HuH7 Ad 代表过表达 LINC01667 的细胞;Count 为细胞数; \*  $P<0.05$ 。

图 4 抑制 FABP5 对 LINC01667 介导的游离脂肪酸摄取能力的影响



**2.5 抑制 FABP5 对 LINC01667 介导的 HepG2 和 HuH7 细胞迁移及侵袭能力的影响** Transwell 迁移能力实验结果显示:与对照组相比,转染 ssiFABP5 后的 HCC 细胞穿透 Transwell 小室的细胞数显著下降( $P < 0.01$ );结果提示,抑制 FABP5 后,由

LINC01667 介导的 HCC 细胞迁移能力减弱;Transwell 侵袭实验结果显示,与 ssiNC 组相比,转染 ssiFABP5 后的 HCC 细胞穿透 Transwell 小室的细胞数显著下降( $P < 0.01$ );结果提示 HCC 细胞侵袭能力显著减弱。见图 5、6。

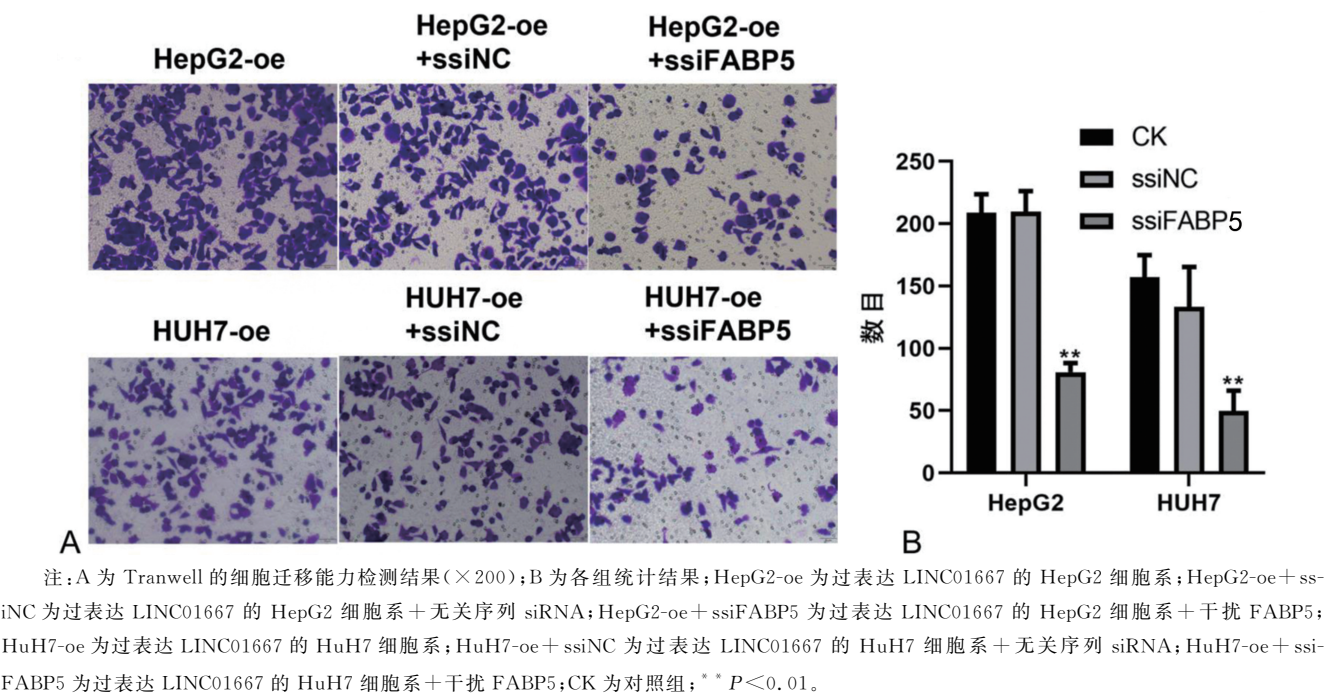


图 5 各组细胞迁移能力检测结果

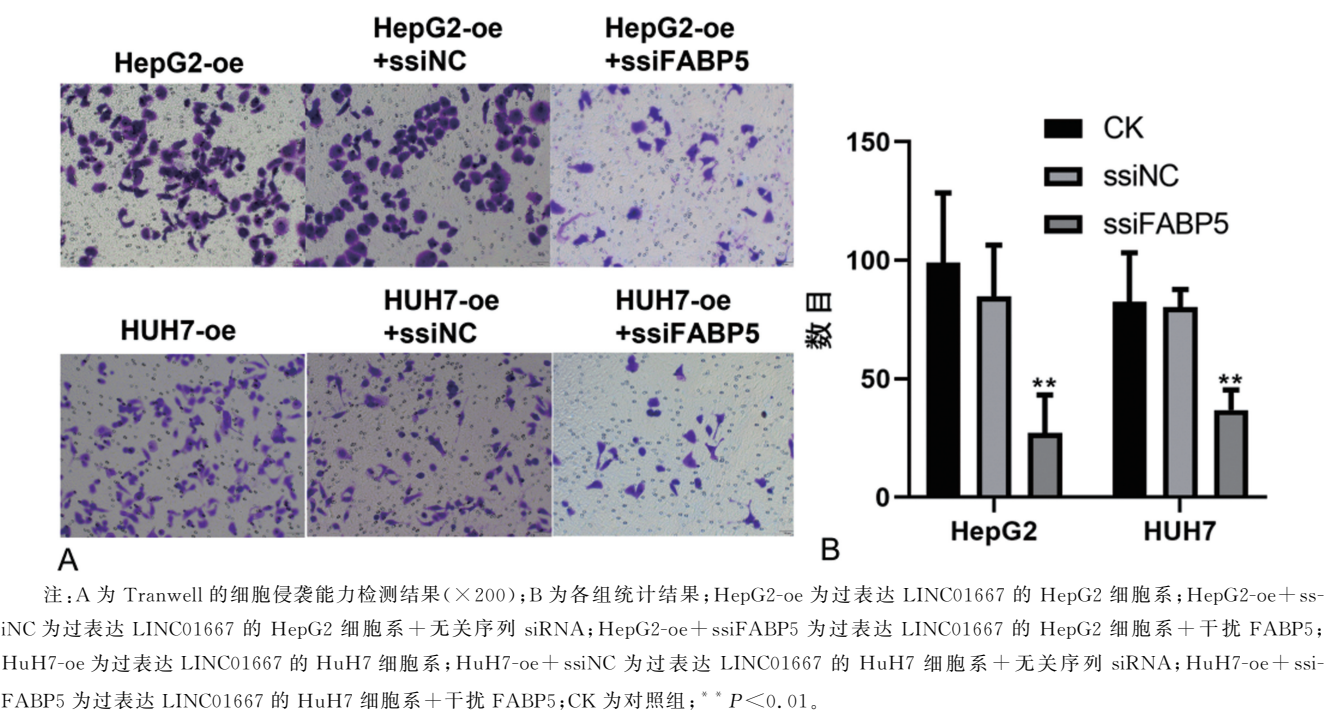


图 6 各组细胞侵袭能力检测结果

3 讨 论

HCC 起病隐匿、侵袭性强、预后不佳,其防治一直是临床医学关注的焦点和亟待解决的难题。目前,有关 HCC 的治疗手段主要以外科手术为主,辅以终

末期的放化疗。但是,肝脏是重要的消化和免疫器官,其特殊的双重血供和解剖位置导致了肝癌术后的易转移性<sup>[5-7]</sup>,进而限制了现有治疗手段作用的发挥。因此,本研究从 HCC 发生发展的分子生物学机制着

手,探寻可有效阻断 HCC 细胞增殖、侵袭进程的新型靶向药物,有可能是解决 HCC 防治难题的关键。

如前所述,NAFLD 正在成为 HCC 的主要危险因素,而 NAFLD 向 HCC 演进的最基本机制则可能是 FAM 的持续失调<sup>[4]</sup>。在本研究中,笔者验证了 LINC01667 可以结合 FABP5 DNA 启动子区域,并调控 FABP5 DNA 的转录,并且使用 Western blot 法验证了 LINC01667 可以促进 FABP5 的表达,结果初步提示 FABP5 可能是 LINC01667 的下游靶基因,且 LINC01667 通过上调 FABP5 启动子活性,进而调控其表达。但是,由于技术的限制,本研究虽然证明了 LINC01667 与 FABP5 DNA 启动子区域存在结合,但是并没有将 LINC01667 与 FABP5 DNA 启动子结合的具体碱基序列对进行解析。后续的研究工作中,笔者将尝试应用表面等离子共振技术,进一步缩小结合区域的范围,并通过构建不同突变体,以验证 LINC01667 与 FABP5 DNA 启动子结合的具体碱基序列。

FABP5 作为一种脂质伴侣,主要功能是调节细胞内脂肪酸的水平和特定的代谢途径。本研究结果显示,当 LINC01667 上调 FABP5 表达后,细胞内脂滴水平可增加,而脂滴是由甘油三酯和胆固醇酯组成,其充当脂质和能量稳态的储存细胞器<sup>[8]</sup>。在肿瘤内,细胞内脂滴经常上调,并且与肿瘤侵袭性和治疗耐药性相关<sup>[9]</sup>。例如,在卵巢癌中,遏制脂滴的形成和积累可以与贝伐珠单抗结合使用,进而抑制肿瘤的生长和增殖<sup>[10]</sup>。另有研究发现,在缺氧条件下,脂滴可在胶质母细胞瘤细胞中蓄积。抑制脂滴的产生可降低体外缺氧复氧过程中的细胞存活率,并显著阻遏体内肿瘤的增殖<sup>[11]</sup>。基于本研究结果,推测 LINC01667 可与 FABP5 DNA 启动子区域结合,并调控 FABP5 基因的转录,并且 LINC01667 上调 FABP5 表达后,可增加细胞内脂滴蓄积。由此推测,LINC01667 可能通过 FABP5 表达增强 HCC 细胞内脂滴蓄积,从而重塑 HCC 细胞的脂肪酸代谢过程,增强其肿瘤相关的恶性生物学行为。

但是,LINC01667/FABP5 调控通路如何重塑 HCC 脂肪酸代谢呢? 本研究使用 Invitrogen™ BODIPY™ FL C16 模拟体内脂肪酸摄取过程,流式细胞术检测结果显示,在 HepG2 和 HuH7 细胞中,过表达 LINC01667 后,细胞内平均荧光强度显著高于对照细胞,提示过表达 LINC01667 后,HCC 细胞系对游离脂肪酸的摄取能力显著增强。这说明 LINC01667 导致 HCC 细胞内脂滴水平增强的可能原因是其增强了细胞对外环境游离脂肪酸的摄取能力。

本研究通过 FABP5 的靶向抑制剂 SBFI-26 对过表达 LINC01667 后的 HepG2 和 HuH7 细胞系进行体外干预,流式细胞分析仪检测结果显示,当靶向抑

制 FABP5 后,由过表达 LINC01667 介导的脂肪酸摄取能力增强效应被显著遏制,表明 LINC01667 促进 HCC 细胞系脂肪酸代谢重编程的分子伴侣为 FABP5。但是目前的结果还无法证明 LINC01667 是否仅作用于 FABP5,以及该过程是否有其他关键性分子参与。

由于 LINC01667 可以通过 FABP5 重编程 HCC 脂肪酸代谢,而脂肪酸代谢又与肿瘤相关的恶性生物学行为密切相关。因此,本研究就进一步评估干预 FABP5 后对过表达 LINC01667 的 HCC 细胞恶性生物学行为的影响,应用 Transwell 实验观测 HCC 细胞系的迁移、侵袭能力,结果显示:当敲低 FABP5 后 48 h,HCC 细胞的迁移、侵袭能力均显著减弱。这说明 LINC01667 可以通过 FABP5 相关的脂肪酸代谢重编程促进 HCC 细胞系细胞恶性生物学表型。因此,基于 LINC01667/FABP5 调控轴在 HCC 细胞恶性生物学行为中的关键作用,LINC01667、FABP5 均有可能作为一种潜在的 HCC 生物标志物和(或)治疗靶点应用于临床。

综上所述,LINC01667 可以通过靶向 FABP5 启动子,上调 FABP5 表达,进而重塑 HCC 脂肪酸代谢,促进 HCC 的恶性进展。以 LINC01667/FABP5 调控轴为靶向药物作用靶点,可为抗 HCC 药物的研发提供借鉴和参考。

参考文献

[1] CHIDAMBARANATHAN-REGHUPATY S,FISHER P B,SARKAR D. Hepatocellular carcinoma (HCC): epidemiology, etiology and molecular classification[J]. Adv Cancer Res,2021,149:1-61.

[2] 陈志国,盛静,赵著洋,等. 外周血 snoRNA 单一及组合 panel 对原发性肝癌的诊断价值[J]. 国际检验医学杂志, 2024,45(10):1267-1272.

[3] LIN Y P,WANG P M,CHUANG C H,et al. Metabolic risks are increasing in non-B non-C early-stage hepatocellular carcinoma;a 10-year follow-up study[J]. Front Oncol,2022,12:816472.

[4] XU K,XIA P,CHEN X,et al. ncRNA-mediated fatty acid metabolism reprogramming in HCC[J]. Trends Endocrinol Metab,2023,34(5):278-291.

[5] PRAYUGO F B,KAO T J,ANURAGA G,et al. Expression profiles and prognostic value of fabps in colorectal adenocarcinomas[J]. Biomedicines,2021,9(10).

[6] ZHANG K,LIU H,YU M,et al. Upregulated LINC01667 expression is correlated with poor prognosis in hepatocellular carcinoma[J]. Front Oncol,2021,11:650173.

[7] 张凯楠,温雯,赵辉,等. miR-130a-3p 在 HCC 致病机制中的作用及作为预后标志物的评估[J]. 联勤军事医学, 2023,37(5):383-387.

(下转第 140 页)

• 论 著 •

# 基于检验数据的机器学习建立 2 型糖尿病患者合并冠心病的风险预测模型

顾智超, 吴昀喆, 杨帆, 陆怡德<sup>△</sup>

上海交通大学医学院附属瑞金医院检验科, 上海 200025

**摘要:**目的 分析 2 型糖尿病患者的临床指标特征, 利用机器学习筛选风险预测指标, 建立简便、有效的 2 型糖尿病合并冠心病的风险预测模型。方法 采用回顾性研究的方法, 选取 2022 年 1 月至 2023 年 11 月该院心内科住院且诊断为 2 型糖尿病合并冠心病的患者 217 例作为合并冠心病组, 同期在门诊确诊为 2 型糖尿病患者 214 例作为糖尿病组, 记录两组患者实验室常规检测数据。通过最小绝对收缩和选择算法(Lasso)筛选特征, 运用随机森林、决策树、支持向量机、极端梯度提升、Logistic 回归、K 近邻算法、人工神经网络 7 种机器学习算法建立模型, 通过受试者工作特征曲线及曲线下面积(AUC)、校准曲线、特异度、灵敏度、F1 值等指标评价不同模型的诊断效能。**结果** 通过 Lasso 回归共筛选出年龄、性别、收缩压、舒张压、心率、C 反应蛋白、血糖等 20 个关键因素。纳入模型后, 支持向量机模型具有最高的灵敏度(88.37%)、阴性预测值(82.14%)及 AUC(0.845), 随机森林模型具有最高的准确度(76.47%)、阳性预测值(76.74%)、F1 值(0.77), 而极端梯度提升算法具有较好的特异度(80.95%), 引入支持向量机模型及 SHAP 值后, 归纳得出血糖对于糖尿病合并冠心病具有正向影响。**结论** 机器学习可作为 2 型糖尿病患者合并冠心病风险评估的有效工具, 其中支持向量机、随机森林和极端梯度提升模型均有较好的预测效能, 具有一定的临床应用前景。

**关键词:**机器学习; 2 型糖尿病; 冠心病; 预测模型

**DOI:**10.3969/j.issn.1673-4130.2025.02.002

**中图法分类号:**R446.1; R587.2

**文章编号:**1673-4130(2025)02-0135-06

**文献标志码:**A

## Establishment of a risk prediction model for patients with type 2 diabetes and coronary heart disease based on machine learning of laboratory data

GU Zhichao, WU Yunzhe, YANG Fan, LU Yide<sup>△</sup>

Department of Laboratory Medicine, Ruijin Hospital, Shanghai Jiao Tong University

School of Medicine, Shanghai 200025, China

**Abstract:** **Objective** To analyze the characteristics of clinical indicators in patients with type 2 diabetes, and to establish a simple and effective risk prediction model for type 2 diabetes complicated with coronary heart disease by screening risk prediction indicators with machine learning. **Methods** A retrospective study was conducted, and 217 patients diagnosed with coronary artery disease combined with type 2 diabetes mellitus who were hospitalized in the Hospital from January 2022 to November 2023 were selected. Additionally, 214 patients diagnosed with T2DM during the same period in the outpatient department were selected as the control group. Their routine laboratory test data were recorded. The Least Absolute Shrinkage and Selection Operator (Lasso) algorithm was used to select features, and the models were built by using seven machine learning algorithms: Random Forest, Decision Tree, Support Vector Machine, eXtreme Gradient Boosting, Logistic Regression, K-Nearest Neighbor, and Artificial Neural Network. The diagnostic efficacy of different models through receiver operating characteristic curve (ROC), area under curve (AUC), calibration curve, specificity, sensitivity, F1 value, and other indicators were evaluated. **Results** Twenty key factors, including age, gender, systolic blood pressure, diastolic blood pressure, heart rate, C-reactive protein and blood glucose were selected using Lasso regression. When incorporated into various models, the SVM model exhibited the highest sensitivity (88.37%), negative predictive value (82.14%), and area under curve (0.845). The Random Forest model had the highest accuracy (76.47%), positive predictive value (76.74%), and F1 score (0.77). Meanwhile, the XGBoost algorithm demonstrated relatively good specificity (80.95%). After introducing the SHAP model, it was inferred that blood glucose had a significant positive impact on the occurrence of coronary heart disease in