

技术与方法

环状RNAs翻译功能研究方法的建立

李哲 肖丙秀*

(宁波大学医学院, 浙江省病理生理学技术研究重点实验室, 宁波 315211)

摘要 环状RNAs(circular RNAs, circRNAs)翻译功能的发现改变了circRNAs作为非编码RNA的传统认识, 研究circRNAs的翻译功能是circRNAs生物学功能研究领域中的一个新方向。该文旨在建立一种circRNAs翻译功能的研究方法。该研究基于胃癌患者癌组织和血浆circRNAs芯片检测结果, 筛选同步变化的circRNAs, 利用开放阅读框查找器(open reading frame finder, ORF finder)和编码潜力评估工具(coding-potential assessment tool, CPAT)分别检测上述circRNAs是否存在开放阅读框架(open reading frame, ORF)和翻译的可行性; 随后使用内部核糖体进入位点查找器(internal ribosome entry site finder, IRES finder)和circRNADb数据库分析circRNAs是否存在内部核糖体进入位点(internal ribosome entry site, IRES), 进而采用基于序列特征的多物种m6A强大预测(robust prediction of N6-methyladenosine sites with sequence-based features in multiple species, M6AMRFS)和基于序列的RNA腺苷甲基化位点预测器(sequence-based RNA adenosine methylation site predictor, SRAMP)等工具筛选circRNAs是否含有m6A修饰位点; 最后通过交集分析, 筛选出可能具有翻译功能的circRNAs。结果显示, 7种circRNAs可能具有翻译功能, 其中hsa_circ_0001479同时具有IRES和m6A介导的翻译功能; hsa_circ_0001278、hsa_circ_0001439和hsa_circ_0000220可能通过IRES介导翻译, 而hsa_circ_0001017、hsa_circ_0008351和hsa_circ_0004406可能通过m6A介导翻译。该研究为探讨circRNAs是否具有翻译功能提供了参考依据, 对所选择的分析方法的可行性通过构建过表达质粒和Western blot加以验证。结果表明, 该研究建立了一套具有可行性的circRNAs翻译功能研究方法。

关键词 环状RNAs; 翻译; 生物学功能; 实验技术

Establishment of Methods for Studying Circular RNAs' Translational Function

LI Zhe, XIAO Bingxiu*

(School of Medicine, Ningbo University, Zhejiang Key Laboratory of Pathophysiology, Ningbo 315211, China)

Abstract The discovery of the translation function of circRNAs (circular RNAs) has changed the traditional understanding of circRNAs as a non-coding RNAs and has become an important direction in the functional research field of circRNAs. The purpose of this paper is to establish a method for studying the translation function of circRNAs. In this study, based on the results of circRNA microarray detection in cancer tissue and plasma of patients with gastric cancer, synchronous changes of circRNAs were screened by ORF finder (open reading frame

收稿日期: 2020-12-30 接受日期: 2021-03-11

宁波市科技创新团队项目(批准号: 2017C110019)资助的课题

*通讯作者。Tel: 0574-87609605, E-mail: xiaobingxiu@nbu.edu.cn

Received: December 30, 2020 Accepted: March 11, 2021

This work was supported by the Scientific Innovation Team Project of Ningbo (Grant No.2017C110019)

*Corresponding author. Tel: +86-574-87609605, E-mail: xiaobingxiu@nbu.edu.cn

URL: <http://www.cjcb.org/arts.asp?id=5536>

finder) and CPAT (coding-potential assessment tool), to detect the feasibility of ORF and translation of the above circRNAs, respectively. IRES finder (internal ribosome entry site finder) and circRNADb database were used to analyze whether there were IRES (internal ribosome entry site) in circRNAs, and then M6AMRFS (robust prediction of N6-methyladenosine sites with sequence-based features in multiple species) tool and SRAMP (sequence-based RNA adenosine methylation site predictor) tool were used to screen whether circRNAs contained M6A modification sites. Finally, the circRNAs with translation function were screened out by intersection analysis. The results show that seven kinds of circRNAs may have translation function, among which hsa_circ_0001479 has both IRES and M6A-mediated translation function; hsa_circ_0001278, hsa_circ_0001439 and hsa_circ_0000220 may mediate translation through IRES, while hsa_circ_0001017, hsa_circ_0008351 and hsa_circ_0004406 may mediate translation through M6A. This study provides a reference for exploring whether circRNAs has translation function. The feasibility of the selected analysis method is verified by constructing overexpression plasmid and Western blot. The results show that a set of feasible methods for the study of the translation function of circRNAs has been successfully established.

Keywords circular RNA; translation; biological function; experimental technology

环状RNAs(circular RNAs, circRNAs)是一类环形RNA分子,可通过套索驱动、内含子配对或RNA结合蛋白(RNA binding protein, RBP)驱动环化形成^[1]。circRNAs根据来源的不同可分为三种类型:外显子circRNA(exonic circRNA, EcRNA)、内含子circRNA(circular intronic RNA, CiRNA)和外显子-内含子circRNA(exon-intronic circRNA, EICRNA)等^[2]。CircRNAs主要分布于细胞质中,但也有部分CiRNA和EICRNA存在于细胞核内;它们具有组织和时序特异性,与生理功能调节和多种疾病发生密切相关^[3-4], circRNAs以共价键形成环状结构;与线性RNA相比, circRNAs不易被RNase R所降解,因此,具有半衰期长和成为新型诊断标记物的优势^[5-6]。

迄今为止, circRNAs的生物学功能主要有三个方面^[7]。(1) 作为微小RNA(microRNA, miRNA)“海绵”,调控下游靶基因的表达。(2) 与RBP相互作用,影响其功能发挥。(3) 翻译产生蛋白质。

尽管circRNAs长期被归类为非编码RNAs^[4],但现今发现的大多数circRNAs来源于基因外显子,且许多circRNAs包含起始密码子AUG和开放阅读框(open reading frame, ORF),也可以与多聚核糖体结合^[3,8]。近期有研究发现,有些circRNAs中高度保守的ORF在体内和体外均能以独立于5'帽结构的方式编码功能肽或蛋白质,例如:它们可通过内部核糖体进入位点(internal ribosome entry site, IRES)诱导、腺苷甲基化(N6-methyladenosine, m6A)基团修饰和滚动循环扩增等方式翻译产生蛋白质^[9]。此外,由于

circRNAs具有独特的共价闭合结构,因此,如果它们的ORF跨剪接位点进行循环,合成的蛋白质肽链长度甚至就会超过其本身序列翻译出来肽链长度^[10]。

CircRNAs的翻译功能越来越引起学者的关注,但由于高通量测序后筛选得到的circRNAs较多,迫切需要一种高速、简便鉴定具有翻译功能circRNAs的方法。为此,本研究先利用胃癌患者癌组织和血浆circRNAs芯片检测结果筛选出同步变化的circRNAs,然后采用多种算法和工具筛选具有翻译功能的circRNAs。本实验方法将为鉴定具有翻译功能的circRNAs提供参考。

1 材料与方法

1.1 胃癌相关circRNAs筛选流程

1.1.1 筛选血浆和癌组织同步变化的胃癌相关circRNAs 利用课题组前期胃癌相关circRNAs芯片检测结果^[11-12],从胃癌组织circRNAs表达谱GSE89143^[11]和胃癌患者血浆circRNAs表达谱GSE93541^[12]分别筛选出高表达和低表达中分别排列在前100位的circRNAs,并把两者结果进行交集分析,得到在胃癌患者血浆和癌组织中同步变化的circRNAs作为候选circRNAs。随后,通过circBase数据库(<http://www.circbase.org/>)检索,下载这些circRNAs的序列及相关信息。

1.1.2 识别circRNAs中的ORF ORF是从ATG(RNA序列中为AUG)开始到终止密码子的核苷酸序列。针对上述筛选出的circRNAs,先采用开放阅读框查

找器(open reading frame finder, ORF finder; http://www.bioinformatics.org/sms2/orf_find.html)检测其是否存在ORF; 再采用编码潜力评估工具(coding-potential assessment tool, CPAT; <http://rna-cpat.sourceforge.net/>), 依据ORF coverage、fickett score、hexamer score、coding probability等4个方面的情况分析这些circRNAs是否有存在ORF、序列保守性和翻译的可能性。

1.1.3 预测IRES IRES是一种RNA调控元件, 可募集核糖体, 实现核糖体组装和协助蛋白质翻译, 并启动独立于5'帽结构的翻译过程^[13]。针对1.1.2步骤筛选得到的circRNAs, 先使用内部核糖体进入位点查找器(internal ribosome entry site finder, IRES finder; <https://github.com/xiaofengsong/IRESfinder>)检索其是否具有IRES; 并通过circRNADb数据库(<https://omictools.com/circrnadb-tool>)检索其序列是否在数据库中存在, 选择*R*值大于1.5的circRNAs; 将两者进行交集分析, 最后得出目标circRNAs。

1.1.4 预测m6A修饰 由于非依赖5'帽结构启动翻译的另一种机制是通过m6A修饰位点进行, 为此, 针对1.1.2步骤筛选得到的circRNAs, 先后采用基于序列特征的多物种m6A强大预测器(robust prediction of N6-methyladenosine sites with sequence-based features in multiple species, M6AMRFS)工具(<http://server.malab.cn/M6AMRFS/>)和基于序列的RNA腺苷甲基化位点预测(sequence-based RNA adenosine methylation site predictor, SRAMP)工具(<http://www.cuilab.cn/sramp/>)

分别检测是否含有m6A修饰位点; 最后将两种方法的结果进行交集分析, 得到目标circRNAs(图1)。

1.2 主要材料和试剂

工具细胞人胚肾细胞株HEK293T由宁波大学医学院龚朝辉教授提供。Hsa_circ_000220过表达载体pLC5-ciR和相应的空载体pLC5购自广州吉赛生物科技有限公司。RPMI 1640培养液购自美国Corning公司。转染试剂Lipofectamine 2000、TRIzol购自美国Invitrogen公司。GoScript Reverse 逆转录试剂盒和qPCR Master Mix荧光定量PCR试剂盒购自美国Promega公司。抗体购自美国Sigma-Aldrich公司。

1.3 实验方法

1.3.1 细胞培养和转染 人胚肾细胞株HEK293T使用含10%胎牛血清的RPMI 1640培养液, 37 °C、5% CO₂条件下培养。按照转染说明书对细胞进行转染, 转染48 h后收获细胞。

1.3.2 载体构建 CircRNA过表达载体pLC5-ciR购自广州吉赛生物技术有限公司, 再将上下游环化框架插入到pLC5-ciR质粒载体中。上游环化框架包含内源侧翼基因组序列和EcoR I酶切位点; 下游环化框架包含部分倒置的上游序列和BamH I酶切位点。在HEK293T细胞中扩增hsa_circ_000220序列, 然后将扩增片段连接到载体的上下游环化框架反向重复(inverted repeat, IR)序列之间。

1.3.3 过表达验证 细胞转染培养48 h后, 使用

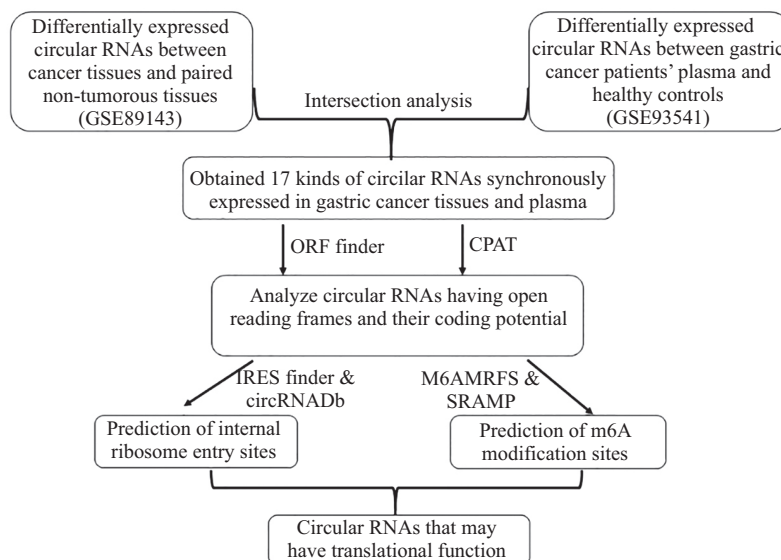


图1 预测具有翻译功能的环状RNA

Fig.1 Predict circRNAs that have function of translation

TRIzol提取总RNA; 酶标仪测定其浓度和纯度, 然后按照GoScript Reverse说明书合成cDNA, 随后使用针对hsa_circ_0000220的特异性引物在Applied Biosystems™ QuantStudio™ 3实时定量PCR仪上进行PCR。Hsa_circ_0000220的正反向引物序列分别为5'-AGA GTG CTC AGG AAT GAA AC-3'和5'-GCC TGA ACA GGT GCA ATA AC-3', 甘油醛-3-磷酸脱氢酶(glyceraldehyde-3-phosphate dehydrogenase, *GAPDH*)的正反向引物序列分别为5'-CCT GAC CTG CGT GTG GAC T-3'和5'-GCT GTG GAT GGG GAG GTG TC-3'。反应条件为: 95 °C预变性10 min; 95 °C 10 s, 60 °C 34 s, 共40个循环。用 $2^{-\Delta\Delta Cq}$ 法计算表达水平, 实验重复3次。将qRT-PCR产物进行测序, 并把测序结果在基因库中采用BLAST进行序列对比分析, 确定其准确性和一条引物是否跨过剪切位点。

1.3.4 Western blot检测 细胞转染培养48 h后, 提取细胞总蛋白质。取10 μ g样品经12%聚丙烯酰胺凝胶电泳分离后电转至聚偏二氟乙烯(polyvinylidene fluoride, PVDF)膜上, 将膜放入封闭液中37 °C摇床上封闭2 h。封闭完成后, 加一抗(抗Flag抗体) 4 °C孵育过夜, 后加二抗(重组抗小鼠IgG抗体) 37 °C孵育2 h, 电化学发光(electrochemiluminescence, ECL)

显色, 然后使用化学发光成像仪进行处理分析。

1.4 统计分析

采用SPSS 18.0进行统计分析, 两样本均数比较采用 t 检验, $P < 0.05$ 为差异具有统计学意义。

2 结果

2.1 基于circRNAs芯片数据筛选候circRNAs

血液是临床化验的最常用材料, 采集外周血也是易于接受的非创性方法。为提高circRNAs筛选的准确性, 降低结果的假阳性, 对课题组前期获得的胃癌患者癌组织和癌旁组织circRNAs差异表达谱GSE89143与胃癌患者血浆与健康人血浆circRNAs差异表达谱GSE93541进行了交集分析, 共发现17种circRNAs在胃癌组织和血浆中同步变化。接下来, 把上述circRNAs用于筛选具有翻译功能的circRNAs。

2.2 含有开放阅读框的circRNAs

先将筛选出的circRNAs通过circBase数据库检索其具体信息, 然后下载其序列于本地计算机。分别将circRNAs序列输入到ORF finder和CPAT工具分析其翻译潜能及其序列的保守性^[14-15]。结果发现, 其中16种circRNAs具有ORF(表1)。

表1 具有开放阅读框的环状RNA
Table 1 The circRNAs containing ORF

环状RNA名称	基因位置	宿主基因符号	开放阅读框	开放阅读框大小	Fickett得分	六分体得分	编码概率
CircRNA ID	Gene position	Host gene symbol	ORF	ORF size	Fickett score	Hexamer score	Coding probability
hsa_circ_0005556	chr2:15693549-15698758	<i>NBAS</i>	Yes	129	0.540 1	-0.107 3	0.005 4
hsa_circ_0001013	chr2:61339656-61345251	<i>KIAA1841</i>	Yes	417	0.695 3	-0.005 7	0.195 2
hsa_circ_0001479	chr5:43122140-43162033	<i>ZNF131</i>	Yes	951	0.949 2	0.199 7	0.999 7
hsa_circ_0001190	chr21:38792600-38794168	<i>DYRK1A</i>	Yes	273	0.659 7	-0.204 7	0.017 9
hsa_circ_0000220	chr10:15863654-15889942	<i>FAM188A</i>	Yes	282	1.007 9	-0.096 4	0.071 7
hsa_circ_0001811	chr8:74585341-74601048	<i>STAU2</i>	Yes	294	0.873 3	0.067 8	0.282 7
hsa_circ_0001017	chr2:61749745-61761038	<i>XPO1</i>	Yes	300	0.664 3	0.011 8	0.183 9
hsa_circ_0007707	chr10:98667021-98667504	<i>LCOR</i>	Yes	27	0.808 9	-0.660 6	0.000 1
hsa_circ_0008351	chr3:197592293-197602646	<i>LRCH3</i>	Yes	489	1.129 9	0.000 5	0.682 0
hsa_circ_0001897	chr9:134381500-134381840	<i>POMT1</i>	Yes	162	0.556 9	-0.040 9	0.012 4
hsa_circ_0004406	chr1:1737913-1756938	<i>GNB1</i>	Yes	282	0.989 1	0.364 3	0.710 0
hsa_circ_0001278	chr3:31617887-31621588	<i>STT3B</i>	Yes	273	0.974 5	-0.271 8	0.240 6
hsa_circ_0001498	chr5:76758919-76760634	<i>WDR41</i>	Yes	27	0.743 3	-0.666 2	0.000 1
hsa_circ_0001439	chr4:129913321-129925031	<i>SCLT1</i>	Yes	273	1.117 4	0.115 1	0.999 4
hsa_circ_0006936	chr6:150092297-150094305	<i>PCMT1</i>	Yes	99	0.483 5	-0.032 2	0.007 8
hsa_circ_0001955	chr15:64495280-64508912	<i>CSNK1G1</i>	Yes	762	0.708 9	0.160 4	0.995 3

表2 具有IRES和m6A位点的环状RNA

Table 2 The circRNAs containing IRES and m6A sites

环状RNA名称 CircRNA ID	CircRNADb数据集 CircRNADb		IRES查找器 IRES finder		M6AMRFS工具 M6AMRFS tool	SRAMP工具 SRAMP tool	
	R得分 R score	IRES位置 IRES site	得分 Score	IRES位置 IRES site	得分 Score	得分 Score	结论 Decision
hsa_circ_0005556	No search results		0.80	1-174	No search results	No search results	
hsa_circ_0001013	No search results		0.98	601-774	0.67	0.659	High confidence
hsa_circ_0001479	1.66	28-166	0.82	51-224	0.97	0.803	Very high confidence
hsa_circ_0001190	No search results		0.98	551-724	0.71	0.727	Very high confidence
hsa_circ_0001811	1.57	9-132	0.63	101-274	0.99	0.607	High confidence
hsa_circ_0001017	1.57	25-16	No search results		0.99	0.710	Very high confidence
hsa_circ_0008351	No search results		0.96	3 751-3 924	0.98	0.726	Very high confidence
hsa_circ_0001897	No search results		No search results		No search results	No search results	
hsa_circ_0004406	No search results		No search results		0.99	0.716	Very high confidence
hsa_circ_0001278	1.57	78-229	0.92	151-324	No search results	0.726	Very high confidence
hsa_circ_0001439	1.56	174-28	0.84	223-396	0.77	0.649	High confidence
hsa_circ_0006936	No search results		No search results		No search results	No search results	
hsa_circ_0001955	1.51	697-784	0.74	51-224	0.27	0.870	Very high confidence
hsa_circ_0001498	1.53	31-175	0.65	8-181	0.63	0.687	Very high confidence
hsa_circ_0000220	1.63	136-279	0.92	301-474	0.99	0.601	High confidence
hsa_circ_0007707	No search results		No search results		0.27	0.571	High confidence

2.3 具有非依赖5'帽的翻译起始机制的circRNAs

在筛选出具有ORF的circRNAs基础上, 进一步鉴定它们是否拥有IRES介导或m6A介导翻译的功能机制。

2.3.1 含有IRES的circRNAs 为了分析circRNAs是否具有IRES, 先运用IRES查找器寻找可能存在的IRES位点, 并列出生得分最高的IRES位点(以Score>0.8为标准^[16])及其序列区间; 然后运用circRNADb数据库检索IRES活性^[17]。经对两种结果进行交集分析, 发现存在IRES的circRNAs有hsa_circ_0001479、hsa_circ_0001278、hsa_circ_0001439和hsa_circ_0000220等4种(表2)。

2.3.2 具有m6A修饰的circRNAs m6A可直接结合真核起始因子3(eIF3)以招募43S复合体并绕过m7G, 进而启动翻译^[18-19]。迄今为止, 研究者开发了多种用于预测RNA m6A修饰位点的工具, 例如DeepM6aseq、iRNA-Methyl、iRNA-3typeA、M6AMRFS、SRAM和m6Acomet等。本研究选取可直接采用网页分析的M6AMRFS和SRAM工具用于预测m6A位点。先后运用M6AMRFS工具(以Score>0.8为标准^[20])和SRAMP工具(以Very high为标准^[21])筛选m6A修饰位点, 再将两个结果进

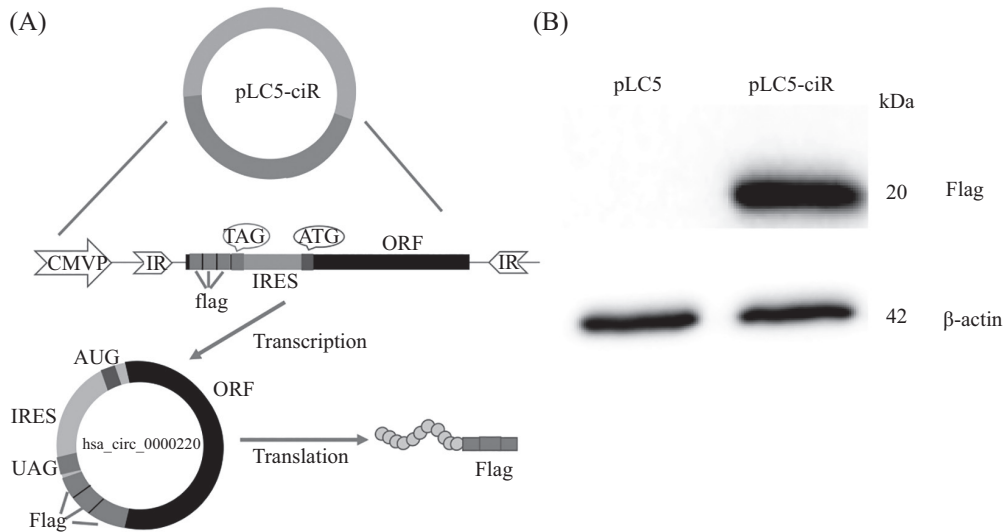
行交集分析。结果发现, hsa_circ_0001479、hsa_circ_0001017、hsa_circ_0008351和hsa_circ_0004406等4种circRNAs具有m6A修饰位点(表2)。

综合分析IRES和m6A修饰位点的筛选结果, 发现7种具有翻译功能的circRNAs, 其中hsa_circ_0001479兼有IRES和m6A介导翻译的功能; hsa_circ_0001278、hsa_circ_0001439和hsa_circ_0000220只通过IRES介导翻译, 而hsa_circ_0001017、hsa_circ_0008351和hsa_circ_0004406只通过m6A介导翻译。

2.4 验证circRNAs具有翻译功能

为了更好地说明预测结果的准确性, 本研究选择了在circRNADb数据库和IRES查找器预测结果(表2)中IRES得分均最高的hsa_circ_0000220作为研究对象, 验证其翻译功能。

先成功构建含有Flag标签的hsa_circ_0000220过表达质粒(图2A)。为增加实验的可靠性, 在构建pLC5-ciR质粒时, 把3×Flag插入至hsa_circ_0000220序列ORF的终止密码子之前。将过表达质粒pLC5-ciR和空白对照质粒pLC5分别转入HEK293T细胞株中, 48 h后提取总蛋白和总RNA分别进行Western blot和qRT-PCR。Western blot结果显示, 过表达组有阳性蛋白——Flag标签蛋白的表达(图2B), 这说明

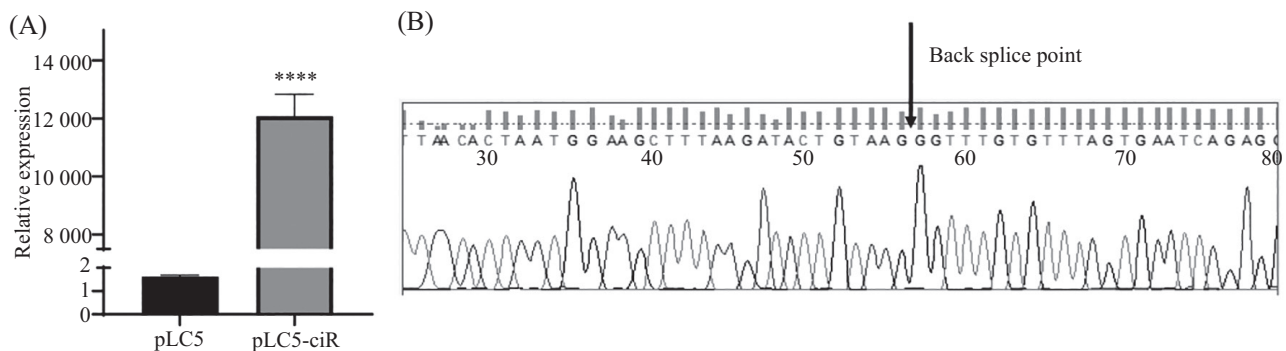


A: 过表达hsa_circ_0000220质粒pLC5-ciR示意图; B: Western blot验证转染HEK293T细胞后翻译生成的含有Flag的融合蛋白。CMVP: 巨细胞病毒启动子; IR: 反向重复; ATG和AUG分别为质粒和circRNA中的起始密码子; TAG和UAG分别为质粒和circRNA中的终止密码子; ORF: 开放阅读框; pLC5: 空载质粒; pLC5-ciR: 带有Flag标签的hsa_circ_0000220过表达质粒。

A: schematic diagram of hsa_circ_0000220 overexpression plasmid pLC5-ciR; B: Western blot confirmation results of encoding fusion protein with positive Flag protein in HEK293T cells transfected with pLC5-ciR. CMVP: cytomegalovirus promoter; IR: inverted repeat; ATG and AUG are the start codon in plasmid and circRNA, respectively; TAG and UAG are the stop codon in plasmid and circRNA, respectively; ORF: open reading frame; pLC5: empty vector; pLC5-ciR: hsa_circ_0000220 expressing vector with Flag.

图2 过表达质粒示意图和Western blot验证结果

Fig.2 Schematic diagram of overexpressed plasmid and results of Western blot verification



A: HEK293T细胞中hsa_circ_0000220的相对表达; B: hsa_circ_0000220的qRT-PCR产物测序结果。pLC5: 空载质粒; pLC5-ciR: 带有Flag标签的hsa_circ_0000220过表达质粒。**** $P < 0.0001$, $n=3$ 。

A: the relative expression of hsa_circ_0000220 in HEK293T cells; B: sequencing result of qRT-PCR product of hsa_circ_0000220. pLC5: empty vector; pLC5-ciR: hsa_circ_0000220 expressing vector with Flag. **** $P < 0.0001$, $n=3$.

图3 qRT-PCR检测hsa_circ_0000220相对表达及其测序结果

Fig.3 qRT-PCR detection of hsa_circ_0000220 relative expression and sequencing results

hsa_circ_0000220序列中包含翻译元件启动翻译机制并有效翻译出Flag的融合蛋白(图2A)。qRT-PCR结果显示, 细胞中hsa_circ_0000220表达水平明显上调(图3A); 而测序结果证实了扩增产物的准确性(图3B)。

3 讨论

尽管目前发现, circRNAs具有调控细胞功能并参

与疾病发生过程的能力, 但由于它们没有5'帽结构和3'多聚A尾, circRNAs曾一度被认为是非编码RNA^[1-3]。但最近越来越多的研究结果表明, circRNAs可以翻译产生短肽或蛋白质^[19,22]。因此, 根据是否具有翻译功能, 作者把circRNAs分为编码circRNAs(coding circRNAs)和非编码circRNAs(non-coding circRNAs)^[4]。

对于circRNAs翻译功能的探索将开启一个

表3 已报道的具有翻译功能的环状RNA

Table 3 Reported coding circRNAs

环状RNA名称	开放阅读框	开放阅读框大小	Fickett	六分体	编码	CircRNADb数据集		IRES查找器		M6AMRFS工具		SRAMP工具	
			得分	得分	概率	CircRNADb		IRES finder		M6AMRFS tool		SRAMP tool	
CircRNA ID	ORF	ORF size	Fickett score	Hexamer score	Coding probability	R得分	IRES位置	得分	IRES位置	得分	得分	得分	决定
						R score	IRES site	Score	IRES site	Score	Score	Score	Decision
hsa_circ_0000423	Yes	219	0.691 7	-0.243 6	0.009 3	1.32	101-236	0.94	51-224	0.97		0.803	Very high confidence
hsa_circ_0001451	Yes	582	1.138 5	0.247 4	0.980 3	1.59	71-217	0.97	201-374	1.00		0.870	Very high confidence
hsa_circ_0017250	Yes	524	1.013 0	0.037 8	0.999 7	1.61	118-246	0.78	301-474	0.44		0.687	High confidence

隐藏的人类蛋白质组学新方向,并增强人们对circRNAs生物学功能的了解。最近有不少研究发现,一些circRNAs不仅可以充当miRNA分子海绵发挥作用,还可通过其编码的肽和蛋白质发挥生物学功能^[22-25]。例如,circ-SHPRH可以编码SHPRH-146aa肽来抑制神经胶质瘤的发生^[22],另外,它还可以作为miRNA海绵抑制肝细胞肝癌的进展^[24]。circ-ZNF609不仅可以作为miRNA海绵来促进乳腺癌的进展,而且还翻译出蛋白质来发挥作用^[25]。为证实本方法的可靠性,作者对最近报道的几种具有翻译功能的circRNAs^[26-28]进行了翻译机制分析,结果发现它们均可通过IRES或m6A介导翻译(表3)。

CircRNAs具体翻译功能的发现,在疾病诊断和治疗中也有重要意义和价值。CircRNAs编码的肽或蛋白质具有成为肿瘤新型治疗靶点和肿瘤新型生物标志物的潜力。CircRNAs编码的肽或蛋白质将成为抗肿瘤药物筛选的新资源;circRNAs结构的稳定性,说明其半衰期长;其翻译产物在疾病标志物方面具有巨大的潜力,circRNAs将是一类新型生物标志物。

本研究中使用的生物信息学分析工具主要有IRES查找器、M6AMRFS和SRAMP等。其中IRES查找器是基于Logit模型的预测工具,参数精度可达到80%以上,准确度达到73%,可明显提高预测IRES的效率;另外,IRES查找器是基于Python语言的独立脚本,可应用于高通量筛选。M6AMRFS是一种新的基于机器学习的m6A预测工具,具有强大而广谱的预测能力。SRAMP预测m6A时仅需RNA序列,而无需加载外部组学数据,但SRAMP仅利用来自5个组织的高通量测序数据集,具有一定的指向性。所以,本研究将M6AMRFS和SRAMP联用,以提高筛选m6A的准确性。

另外,本研究为验证筛选策略的可靠性,以Flag标签蛋白作为阳性对照,在构建pLC5-ciR质粒

时,把标签基因序列插入至hsa_circ_0000220序列ORF的终止密码子之前,而IRES位于起始密码子之前(图2A)。为说明内源性circRNAs在细胞内可以真正表达出蛋白质,下一步可以通过质谱鉴定,而做IRES活性验证实验和m6A活性验证实验可以分别说明IRES元件和m6A修饰启动了circRNA蛋白质翻译。

CircRNAs的翻译功能越来越引起学者的关注,但由于高通量测序后得到的候选circRNAs数目庞大,建立高效、简便的筛选具有翻译功能circRNAs的方法显得非常重要。本研究利用胃癌患者癌组织和血浆circRNAs芯片检测结果筛选出同步变化的circRNAs,然后综合采用多种算法和工具筛选具有翻译功能的circRNAs(图1)。在精选具有ORF的circRNAs基础上(表1),采用2种工具分别针对IRES和m6A进行筛选(表2),可以更加快速地从庞大的高通量circRNAs芯片和测序结果中筛选出可能具有翻译功能的circRNAs,提高了筛选目标circRNAs的效率,降低了筛选具有翻译功能的circRNAs的技术难度,具有推广应用价值。

参考文献 (References)

- [1] 杨平,倪超,陈思,等.环状RNA及其在胃癌中作用的研究进展[J].中国细胞生物学学报(YANG P, NI C, CHEN S, et al. Advances in Circular RNAs and their roles in gastric cancer [J]. Chinese Journal of Cell Biology), 2017, 39(12): 1633-41.
- [2] 郭健民,陈熙,邹军.环状RNA的特性、生物合成和主要功能[J].中国细胞生物学学报(GUO J M, CHEN X, ZOU J. The feature, biosynthesis and main function of circular RNAs [J]. Chinese Journal of Cell Biology), 2016, 38(3): 335-40.
- [3] RUAN Y, LI Z, SHEN Y, et al. Functions of circular RNAs and their potential applications in gastric cancer [J]. Expert Rev Gastroenterol Hepatol, 2020, 14(2): 85-92.
- [4] LI Z, RUAN Y, ZHANG H, et al. Tumor-suppressive circular RNAs: mechanisms underlying their suppression of tumor occurrence and use as therapeutic targets [J]. Cancer Sci, 2019, 110(12): 3630-8.
- [5] XIE Y, SHAO Y, SUN W, et al. Downregulated expression of

- hsa_circ_0074362 in gastric cancer and its potential diagnostic values [J]. *Biomark Med*, 2018, 12(1): 11-20.
- [6] 王敏, 杨睿, 钟镇宇, 等. 环状RNA circTCF25通过miR-103a-3p/miR-107调控CDK6的表达促进膀胱癌的增殖和迁移[J]. *中国细胞生物学学报*(WANG M, YANG R, ZHONG Z Y, et al. CircRNA circTCF25 regulates the expression of CDK6 by miR-103a-3p and miR-107 to promote the proliferation and migration of bladder cancer [J]. *Chinese Journal of Cell Biology*), 2018, 40(6): 929-36.
- [7] FU L, JIANG Z, LI T, et al. Circular RNAs in hepatocellular carcinoma: functions and implications [J]. *Cancer Med*, 2018, 7(7): 3101-9.
- [8] MO Y, WANG Y, XIONG F, et al. Proteomic analysis of the molecular mechanism of lovastatin inhibiting the growth of nasopharyngeal carcinoma cells [J]. *J Cancer*, 2019, 10(10): 2342-9.
- [9] LEI M, ZHENG G, NING Q, et al. Translation and functional roles of circular RNAs in human cancer [J]. *Mol Cancer*, 2020, 19(1): 30.
- [10] KONG S, TAO M, SHEN X, et al. Translatable circRNAs and lncRNAs: driving mechanisms and functions of their translation products [J]. *Cancer Lett*, 2020, 483: 59-65.
- [11] SHAO Y, LI J, LU R, et al. Global circular RNA expression profile of human gastric cancer and its clinical significance [J]. *Cancer Med*, 2017, 6(6): 1173-80.
- [12] LI T, SHAO Y, FU L, et al. Plasma circular RNA profiling of patients with gastric cancer and their droplet digital RT-PCR detection [J]. *J Mol Med (Berl)*, 2018, 96(1): 85-96.
- [13] KING H A, COBBOLD L C, WILLIS A E. The role of IRES trans-acting factors in regulating translation initiation [J]. *Biochem Soc Trans*, 2010, 38(6): 1581-6.
- [14] ROMBEL I T, SYKES K F, RAYNER S, et al. ORF-FINDER: a vector for high-throughput gene identification [J]. *Gene*, 2002, 282(1/2): 33-41.
- [15] WANG L, PARK H J, DASARI S, et al. CPAT: coding-potential assessment tool using an alignment-free logistic regression model [J]. *Nucleic Acids Res*, 2013, 41(6): e74.
- [16] ZHAO J, WU J, XU T, et al. IRESfinder: identifying RNA internal ribosome entry site in eukaryotic cell using framed k-mer features [J]. *J Genet Genomics*, 2018, 45(7): 403-6.
- [17] CHEN X, HAN P, ZHOU T, et al. circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations [J]. *Sci Rep*, 2016, 6: 34985.
- [18] RIES R J, ZACCARA S, KLEIN P, et al. m6A enhances the phase separation potential of Mrna [J]. *Nature*, 2019, 571(7765): 424-8.
- [19] WU P, MO Y, PENG M, et al. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA [J]. *Mol Cancer*, 2020, 19(1): 22.
- [20] QIANG X, CHEN H, YE X, et al. M6AMRFS: robust prediction of N6-methyladenosine sites with sequence-based features in multiple species [J]. *Front Genet*, 2018, 9: 495.
- [21] ZHOU Y, ZENG P, LI Y H, et al. SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features [J]. *Nucleic Acids Res*, 2016, 44(10): e91.
- [22] ZHANG M, HUANG N, YANG X, et al. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis [J]. *Oncogene*, 2018, 37(13): 1805-14.
- [23] ASHWAL-FLUSS R, MEYER M, PAMUDURTI N R, et al. CircRNA biogenesis competes with pre-mRNA splicing [J]. *Mol Cell*, 2014, 56(1): 55-66.
- [24] QIN M, LIU G, HUO X, et al. Hsa_circ_0001649: a circular RNA and potential novel biomarker for hepatocellular carcinoma [J]. *Cancer Biomark*, 2016, 16(1): 161-9.
- [25] LEGNINI I, DI TIMOTEO G, ROSSI F, et al. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis [J]. *Mol Cell*, 2017, 66(1): 22-37.e29.
- [26] ZHENG X, CHEN L J, ZHOU Y, et al. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling [J]. *Mol Cancer*, 2019, 18(1): 47.
- [27] YANG Y, GAO X, ZHANG M, et al. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis [J]. *J Natl Cancer Inst*, 2018, 110(3): 304-15.
- [28] XIA X, LI X, LI F, et al. A novel tumor suppressor protein encoded by circular AKT3 RNA inhibits glioblastoma tumorigenicity by competing with active phosphoinositide-dependent kinase-1 [J]. *Mol Cancer*, 2019, 18(1): 131.