

基于 RNA-seq 技术对 13-三体综合征脐带血差异性表达环状 RNA 研究

林南新¹ 汤冬城^{2,3} 龚蔚蔚² 欧明林² 陈洁晶² 林华² 周献青¹ 郭辉³ 周俊³
黄锦龙¹ 戴勇^{*2,3}

(1. 广西师范大学生命科学学院, 广西 桂林 541004; 2. 中国人民解放军联勤保障部队第九二四医院肾脏科、全军器官移植与透析治疗中心、广西代谢性疾病研究重点实验室, 广西 桂林 541002; 3. 暨南大学第二临床医学院、深圳市人民医院临床医学研究中心, 广东 深圳 518020)

【摘要】 目的 本研究拟从 circRNA 为切入点, 利用 RNA-seq 技术从 13-三体综合征患者和产检正常的孕妇脐带血中筛选出差异表达的 circRNA, 为 13-三体综合征患者的诊断和治疗提供基础性研究。**方法** 选择在 2017 年 1 月 1 日至 2018 年 1 月 1 日期间深圳市人民医院确诊为 13-三体综合征患者和一名产检结果正常的孕妇作为研究对象, 使用 RNA-seq 技术检测出差异表达的 circRNA, 进行组间的差异分析, 并对 circRNA 吸附相关的微小 RNA(miRNA) 进行预测和分析。**结果** 通过与健康对照组分析可得到, 共有 8005 个差异表达的 circRNA 具有统计学上的意义, 其中表达上调的 circRNA 有 4275 个, 表达下调的 circRNA 有 3730 个, hsa_circ_0002473 是位于 13 号染色体上表达最上调的 circRNA, 与 8 种 miRNA 相互作用。**结论** 13-三体综合征患者脐带血与健康对照组的脐带血之间存在差异表达的 circRNA, hsa_circ_0002473 有可能成为 13-三体综合征诊断和治疗的生物标志物。

【关键词】 环状 RNA miRNA; 13-三体综合征; RNA-seq 生物标志物

【中图分类号】 R715.5 **【文献标识码】** A

【Abstract】 Objective This study intends to use circRNA as the starting point, and use RNA-seq technology to screen differentially expressed circRNA from the cord blood of patients with 13-trisomy syndrome and pregnant women with normal obstetric examination, so as to provide basic research for the diagnosis and treatment of patients with 13-trisomy syndrome. **Method** During the period from January, 2017 to January, 2018, patients diagnosed with 13-trisomy syndrome in Shenzhen people's hospital and a pregnant woman with normal obstetric examination were selected as the study subjects. Differentially expressed circRNA was detected using RNA-seq technique. Differences between groups were analyzed, and related micro RNA (miRNA) adsorbed by circRNA was predicted and analyzed. **Results** According to the analysis with healthy control group, a total of 8005 circRNA with differential expression were statistically significant. Among them, 4275 circRNA were up-regulated and 3730 circRNA were down-regulated. Hsa_circ_0002473 is the most up-regulated circRNA on human chromosome 13, interacting with 8 miRNA. **Conclusions** There is a differential expression in circRNA between cord blood of patients with 13-trisomy syndrome and that of healthy control group, and hsa_circ_0002473 may be a biomarker for diagnosis and treatment of 13-trisomy syndrome.

DOI: 10.13470/j.cnki.cjpd.2019.01.002

基金项目: 广西重点实验室建设项目(17-259-57)

* 通讯作者: 戴勇, E-mail: daiyong22@aliyun.com

【Key words】 circRNA miRNA; 13-trisomy syndrome; RNA-seq Biomarkers

13-三体综合征,又称为 Patau 综合征,最早在 1960 年由 Klaus Patau 所报道^[1]。13-三体综合征发生的主要原因是由于双亲之一的生殖细胞在减数分裂时不分裂所导致^[2]。13-三体综合征在新生儿中的发病率约为 1/10 000~20 000^[3],而且预后很差,90%的新生儿在 6 个月内死亡,仅有 5%的患儿能够活过 10 岁^[4]。其中,13-三体综合征的发生与孕妇的高龄密切相关。至今,13-三体综合征尚未有根治的方法,其复杂的发病机制一直未能阐明,找到可靠的生物标志物对 Patau 综合征患者及时诊断和治疗显得尤为重要。

13-三体综合征的发生和发展与非编码 RNA 的异常表达密切相关。约有 98%的人类基因组是非编码 RNA,这表明它们对生理和病理具有非常重要的作用^[5]。circRNA 是一类在哺乳动物中广泛表达的非编码 RNA,具有不编码蛋白质却能同时调控体内多个基因的功能。circRNA 有着结构稳定、序列保守性高、丰度高和组织特异性高的特点^[6]。

基因芯片一直以来都是研究转录组学最重要的工具。但是,基因芯片及技术有着固有的缺陷^[7]。近来随着高通量测序技术的迅猛发展,特别是 RNA-seq 技术出现,俨然成为了转录组学研究最具有革命性的工具。与基因芯片相比, RNA-seq 技术不仅可以获得所有 RNA 丰富的、准确性较高的转录本,而且对差异表达的基因具有定量更准确、检测范围更广、分析更加可靠的优势^[8-10]。RNA-seq 技术在肿瘤的研究中被广泛的运用,而在 13-三体综合征中研究却尚未提及。本研究拟从 circRNA 为切入点,采用 RNA-seq 技术获得 13-三体综合征患者和健康对照组之间差异性表达的 circRNA,并对差异性表达的 circRNA 进行生物信息学分析,并从 circRNA 靶向吸附的 miRNA 预测中找到参与 13-三体综合征机制的 circRNA。

1 材料和方法

1.1 研究对象 本研究获得了深圳市人民医院伦理委员会批准,并且按照赫尔辛基宣言进行。选取

2017 年 1 月 1 日至 2018 年 1 月 1 日在深圳市人民医院确诊为 13-三体综合征患者和一名产检正常的志愿者为研究对象。分别抽取患者和健康对照组的脐带血 0.5 ml,置于-80 °C 的液氮中保存。研究对象染色体核型图如下(图 1、2)

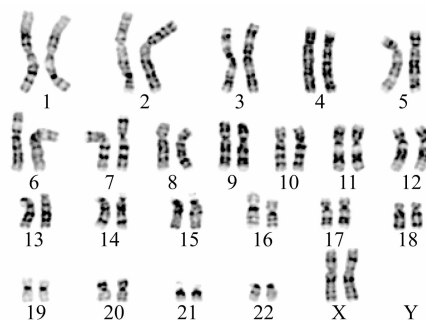


图 1 健康对照组染色体模型分析图

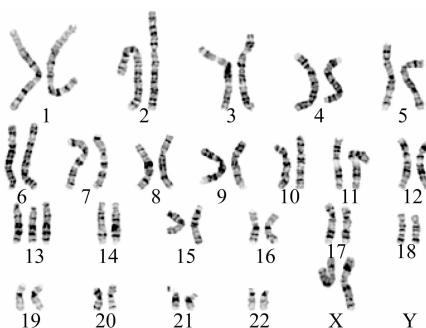


图 2 PS 患者染色体模型分析图

1.2 RNA 的提取和检测使用 Trizol 提取 RNA。用琼脂糖凝胶来分析样品 RNA 完整性以及是否存在 DNA 污染。用 NanoPhotometer Spectrophotometer 检测 RNA 纯度 (OD260/280 及 OD260/230 比值),使用 Qubit2.0 Fluorometer 对 RNA 浓度精确定量,用 Agilent 2100 bicanalyzer 精确检测侧 RNA 完整性。

1.3 文库构建及测序建库起始 RNA 为 totalRNA,总量大于 1 ug。通过 Oligo(dT)磁珠富集带有 polyA 尾的 mRNA,随后在 NEB Fragmentation Buffer 中用二价离子将得到的 mRNA 随机打断,以片段化的 mRNA 随机打断,以片段化的 mRNA 为模板,随机寡核苷酸为引物,在 M-MuLv 逆转录酶体系中合成 cDNA 第一条链,随后用 RNaseH 降

解 DNA 链,并在 DNA polymerase I 体系下,以 dNTPs 为原料合成 cDNA 第二条链,纯化后的双链 cDNA 经过末端修复,加 A 尾并连接测序接头,用 AMPureXPbeads 纯化 PCR 产物,最终获得文库。文库构建完成后,先使用 Qubit 2.0 Fluoro meter 进行初步定量,稀释文库至 $1.5 \text{ ng}/\mu\text{l}$,随后用 Agilent 2100 bioanalyzer 对文库的 insert size 进行检测,insert size 符合预期后,qRT-PCR 对文库有效浓度进行准确定量(文库有效浓度高于 2nM),以保证文库质量。文库检测合格后,使用 Illumina 高通量测序平台(HiSeq/MiSeq)进行测序。

1.4 测序数据分析由于测序获得的原始数据中包含少量带有测序接头或测序质量较低的 reads。为了保证数据分析的质量及可靠性,需要对原始数据进行过滤。主要去除带接头的 reads,去除含 N 的 reads,去除低质量 reads。同时对 clean data 进行 Q20、Q30 和 GC 含量计算。后续所有分析均是基于 clean data 进行的高质量分析。我们使用 HISAT2V.2.0.5 构建参考基因组的索引,并使用 HISAT2V.2.0.5 将配对末端 clean reads 与参照组对比。最后使用 feature Counts 用于计算映射到每个基因的读数。然后根据基因的长度计算每个基因的 FPKM,并计算映射到该基因的读数。

1.5 差异表达基因的筛选应用 DESeq2R 软件进行两个比较组合之间的差异表达分析,并用 TPM 进行表达量的归一化处理。在不同样本组差异表达 RNA 检测中,将 Fold change >2 且 $P < 0.05$ 作为筛选的标准。

2 结果与分析

2.1 差异表达的 circRNA 分析通过使用 RNA-seq 技术比较 13-三体综合征患者和健康对照组可知,共有 9039 个差异表达的 circRNA,其中有 8005 个差异表达的 circRNA 具有统计学上的意义($|\log_2 \text{FoldChange}| > 2, P < 0.05$)。表达上调的 circRNA 有 4275 个,表达下调的 circRNA 有 3730 个,差异表达 circRNA 火山图(图 3)。差异表达的 circRNA 大部分属于内含子,只有少数属于外显子。

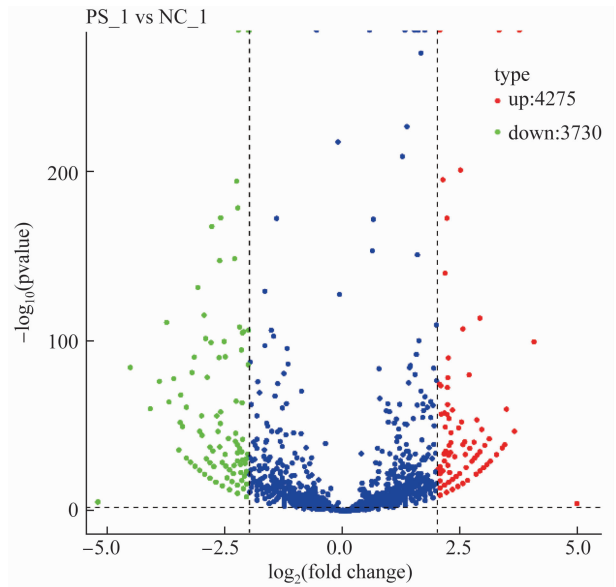


图 3 circRNA 火山图

2.2 GO 和 KEGG 富集分析对差异表达的 8005 个 circRNA 进行 GO 分析和 KEGG 分析。GO 富集分析主要在细胞代谢、代谢过程、细胞内大分子代谢生物过程中富集;在细胞内细胞器部分、细胞内、细胞器细胞组成过程中富集;而在分子结合方面,主要在 RNA 结合上富集(图 4)。KEGG 分析差异表达的 circRNA 主要在子宫内膜癌、肾细胞癌、RNA 降解、细胞周期富集(图 5)。

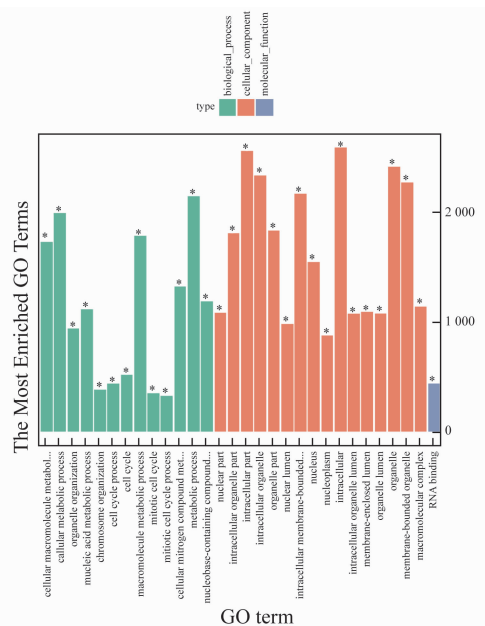


图 4 来源基因 GO 富集图

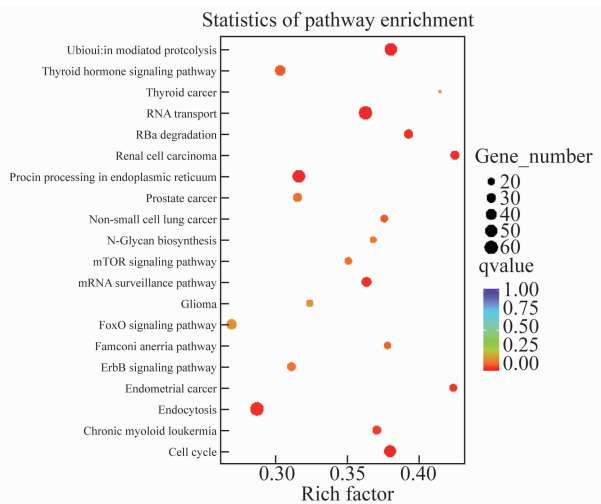


图 5 来源基因 KEGG 富集图

2.3 对差异倍数前十的 circRNA 分析从差异表达的 circRNA 筛选出表达上调前十的 circRNA 和表达下调前十的 circRNA 进行分析,发现差异表达的 circRNA 主要分布在染色体 16 上,且大部分属于内含子和外显子-内含子 circRNA (表 1/2)其中 hsa_circ_0002473 是位于 13 号染色体上表达最上调的 circRNA,和 8 种 miRNA 相互作用(has-miR-486-

3p、has-miR-130a-3p、has-miR-320C、has-miR-17-3p、has-miR-320b、has-miR-4236、has-miR-320a、has-miR-197-3p)。

3 讨论

13-三体综合征染色体变异疾病中比较常见的形式之一,它的发生和发展与非编码 RNA 的异常表达密切相关。近些年来,大量的 circRNA 的发现,引起了越来越多学者们的重视,circRNA 的功能研究也成为了新的热点。有研究报道,hsa_circ_0004018 和 hsa_circ_0005986 在 HCC(肝癌)的表达水平显著降低^[11,12]。hsa_circ_002059 在胃癌组织中的表达显著下调^[13]。Li 等^[14]发现 circHIPK3 的表达水平在膀胱癌组织中显著下调。尽管 circRNA 在肿瘤中的研究取得了一定的进展,但在 13-三体综合征的作用尚未提及。

本研究通过 RNA-seq 技术得到 13-三体综合征患者和健康对照组差异性表达的 circRNA。其中共有 9039 个 circRNA 差异性表达,8005 个 circRNA 具有统计学上的意义。表达上调的 circRNA 共有 4275 个,

表 1 差异表达量上调 10 个 circRNA

circRNA 名称	染色体定位	Fold change	circRNA 类型	基因定位
hg38_circ_0002733	chr11	14.171	exon-intron	HBG1; HBG2; HBG3
hsa_circ_0000377	chr12	10.608	intron	TAS2R14
hg38_circ_0028661	chr6	9.9333	exon-intron	TUBB2A; TUBB2B
hsa_circ_0005152	chr16	9.7604	intron	USP7
hsa_circ_0004265	chr12	8.9791	intron	KNTC1
hsa_circ_0081790	chr7	8.9791	intron	RELN
hsa_circ_0006459	chr9	8.9333	intron	GNAQ
hsa_circ_0084677	chr8	8.7864	intron	CSPP1
hsa_circ_0004344	chr16	8.734	intron	USP7
hsa_circ_00042473	chr13	8.564	intron	DNAJC3

表 2 差异表达量下调 10 个 circRNA

circRNA 名称	染色体定位	Fold change	circRNA 类型	基因定位
hg38_circ_0002752	chr11	-12.661	exon-intron	HBG2
hsa_circ_0001551	chr5	-9.1456	exon-intron	RARS
hg38_circ_0009077	chr16	-9.0461	exon-intron	HBA2; HBA1; HBM
hg38_circ_0009566	chr16	-9.0461	exon-intron	PAPD5
hsa_circ_0001953	chrY	-8.9936	intron	ZFY
hg38_circ_0008377	chr15	-8.8826	exon-intron	DENND4A
hsa_circ_0010445	chr1	-8.6982	exon-intron	HP1BP3
hsa_circ_0003715	chr9	-8.6311	exon-intron	UBQLN1
hsa_circ_0000120	chr1	-8.5607	intron	MAN1A2
hsa_circ_0009024	chrY	-8.4867	exon	-

表达下调的 circRNA 共有 3730 个,同时也对差异性表达的 circRNA 进行了 GO 和 KEGG 分析。GO 富集分析主要在细胞代谢、代谢过程、细胞内大分子代谢生物过程中富集;在细胞内细胞器部分、细胞内、细胞器细胞组成过程中富集;而在分子结合方面,主要在 RNA 结合上富集。KEGG 分析差异表达的 circRNA 主要在子宫内膜癌、肾细胞癌、RNA 降解、细胞周期富集。

同时,我们对差异倍数前十的 circRNA 进行分析。我们发现 hsa_circ_0002473 是我们研究中在 13 号染色体上表达最上调的 circRNA,与 8 种 miRNA 相互作用(has-miR-486-3p、has-miR-130a-3p、has-miR-320c、has-miR-17-3p、has-miR-320b、has-miR-4236、has-miR-320a、has-miR-197-3p)。研究证实,has-miR-486-3p 在习惯性流产的孕妇中表达显著降低^[15],我们推测这可能与 13-三体综合征流产有一定的联系。has-miR-130a-3p 在冠心病中表达显著失调^[16],has-miR-320 家族包括(has-miR-320a、has-miR-320b、has-miR-320c)均与乳腺癌、宫颈癌等女性高发疾病有关^[17,18]。我们推测 hsa_circ_0002473 和其靶向吸附相关的 miRNA 参与调控 13-三体综合征的发生和发展。目前来说,关于 circRNA 在 13-三体综合征的研究还在起步阶段,circRNA 在 13-三体综合征的功能分析尚未完善,同时由于本次研究样本量不足,仅用了一个 RNA-seq 的数据集进行分析,所研究的结果是基于生物学信息分析,研究的结论尚未得到验证,这需要我们在后期的实验中继续探索。

综上所述,本研究通过 RNA-seq 技术得到了 13-三体综合征患者脐带血差异表达的 circRNA,并推测出 hsa_circ_0002473 有可能成为 13-三体综合征新型的生物标志物,为 13-三体综合征的诊断和治疗提供了新的思路。

参考文献

- [1] Patau K, Smith DW, Therman E, et al. Multiple congenital anomaly caused by an extra autosome[J]. *Lancet*, 1960, 1: 790.
- [2] Patau K, Smith DW, Therman E, et al. Multiple congenital anomalies caused by an extra autosome[J]. *Lancet*, 1960, 1(7128):790-793.
- [3] Aypar E, Yildirim MS, Sert A, et al. A girl with metopic synostosis and trisomy 13 mosaicism; case report and view of the literature[J]. *Am J Med Genet A*, 2011, 155:638-641.
- [4] Tunca Y, Kadandate JS, Pivnick EK. Long term survival in Patau syndrome[J]. *Clin Dysmorphol*, 2001, 10(2):149-150.
- [5] Zhang Y, Xu Y, Feng L, et al. Comprehensive characterization of lncRNA-mRNA related ceRNA network across 12 major cancers[J]. *Oncotarget*, 2016, 7(39): 64148-64167.
- [6] 许妍,赵雪,堵晶晶,等.环状 RNA 研究进展[J]. *中国生物工程杂志*, 2017, 37(10):93-102.
- [7] 张建华,应安娜,俞梦琪,等.利用 RNA-seq 数据分析肝细胞癌差异表达基因及蛋白相互作用[J]. *基因组学与应用生物学*, 2018, 8:1-9.
- [8] 赵利南,赵益,孙有智.基于 RNA-seq 技术的活血化痰法治疗乳腺癌关键基因的筛选和验证[J]. *中国实验方剂学杂志*, 2019, 4:1-8.
- [9] Raz T, Kapranov P, Lipson D, et al. Protocol Dependence of Sequencing-Based Gene Expression Measurements[J]. *PLoS ONE*, 2011, 6(5):e19287.
- [10] 廖钦洪,邹勇,李洪雷,等.基于 RNA-seq 数据的生姜 NAC 转录因子家族鉴定及分析[J]. *中国中药杂志*, 2018, 43(3): 493-501.
- [11] Yao Z, Luo J, Hu K, et al. ZKSCAN1 gene and its related circular RNA(circZKSCAN1) both inhibit hepatocellular carcinoma cell growth, migration, and invasion but through different signaling pathways[J]. *Mol Oncol*, 2017, 11(4):422-437.
- [12] Fu L, Yao T, Chen Q, et al. Screening differential circular RNA expression profiles reveals hsa_circ_0004018 is associated with hepatocellular carcinoma[J]. *Oncotarget*, 2017, 8(35):58405.
- [13] Li P, Chen S, Chen H, et al. Using circular RNA as a novel type of biomarker in the screening of gastric cancer[J]. *Clinica Chimica Acta*, 2015, 444:132-136.
- [14] Li Y, Zheng F, Xiao X, et al. CircHIPK3 sponges miR-558 to suppress heparanase expression in bladder cancer cells[J]. *Embo Reports*, 2017, 18(9):1646.
- [15] Gu Y, Zhang X, Yang Q, et al. Aberrant Placental Villus Expression of miR-486-3p and miR-3074-5p in Recurrent Miscarriage Patients and Uterine Expression of These MicroRNAs during Early Pregnancy in Mice[J]. *Gynecol Obstet Invest*, 2016, 81(2):112-117.
- [16] Pan RY, Liu P, Zhou HT, et al. Circular RNAs promote TRPM3 expression by inhibiting hsa-miR-130a-3p in coronary artery disease patients[J]. *Oncotarget*, 2017, 8(36):60280-60290.
- [17] Zhang T, Zou P, Wang T, et al. Down-regulation of miR-320 associated with cancer progression and cell apoptosis via targeting Mcl-1 in cervical cancer[J]. *Tumor Biology*, 2016, 37(7):8931-8940.
- [18] Wang W, Yang J, Xiang YY, et al. Overexpression of Hsa-miR-320 Is Associated With Invasion and Metastasis of Ovarian Cancer[J]. *J Cell Biochem*, 2017, 118(11):3654-3661.

(收稿日期:2019-01-11)

编辑:宋文颖