

· 综 述 ·

转录组测序技术在整合网络调控中的应用现状及前景

杨 杰¹, 张 瑜²

(1. 甘肃省平凉医学高等专科学校 药理教研室, 中国甘肃 平凉 744000;
2. 重庆医药高等专科学校 中医教研室, 中国重庆 401331)

摘 要: 2005 年以来, 以 454 焦磷酸测序技术、Solexa 边合成边测序技术和 SOLiD 连续连接合成测序技术为标志的新一代测序技术相继诞生, 使得全转录组商业化测序成为可能. 针对转录因子与基因、转录因子与 miRNA、以及 miRNA 与基因所构成的整合调控网络, 转录组测序技术以高通量、高分辨率、高精度和价格低廉等优势发挥着巨大的作用. 从 miRNA 与蛋白编码基因等层次分别介绍了转录组测序技术在基础研究、临床研究等领域的应用现状及发展前景.

关键词: 转录组测序; 整合调控网络; 应用; 前景

中图分类号: Q-1

文献标识码: A

文章编号: 1007-7847(2013)06-0561-04

The Application and Perspective of Transcriptome Sequencing Approach Upon Integrated Regulatory Network

YANG Jie¹, ZHANG Yu²

(1. Department of Pharmacology, Pingliang Medical College of Gansu Province, Pingliang 744000, Gansu, China)
2. Department of Traditional Chinese Medicine, Chongqing Medical and Pharmaceutical College, Chongqing 401331, China)

Abstract: Since 2005, next-generation sequencing approaches have been emerged such as the 454 sequencing by pyrophosphoric acid, the Solexa sequencing by synthesis and the SOLiD sequencing by ligation. Then the integral transcriptomes sequencing have been commercialized. The regulatory network, which composes the transcript factors, genes and relevant miRNA, becomes the candidate objective of next-generation sequencing. Transcriptome sequencing has enormous effect due to high-flux, high resolution, and cheap cost. The application and perspective of transcriptome sequencing approaches from two layers, e. g. miRNAs, and genes, in basal research, clinical research are reviewed.

Key words: transcriptome sequencing approach; integrated regulatory network; application; perspective

(*Life Science Research*, 2013, 17(6): 561~564)

转录组是基因组特定组织细胞在特定时空点功能状态的整合表现, 乃一个细胞或细胞群所有转录 mRNA、UTRs、miRNA、ncRNAs 等总称. 转录组概念于 1997 年由 Velculescu 等^[1]首次提出, 酿酒酵母细胞转录组得以全面描述. 随后, 拟南芥、水稻以及人类常染色体转录组^[2-4]全面分析. 基因组学从基因结构研究转向功能基因组研究^[5], 急

需高通量、大规模的分析手段和方法来解决这些需求.

20 世纪 90 年代以后, 基因芯片是建立最早, 也是最为成熟的生物芯片技术, 又被称为 DNA 微阵列, 其突出的优点在于高度并行性、多样性、微型化和自动化, 但它是封闭系统, 只能检测已知序列的特征^[6]. 2005 年, 以瑞士 Roche 公司的 454

收稿日期: 2013-03-11; 修回日期: 2013-06-06

作者简介: 杨杰(1976-), 男, 四川广汉人, 讲师, 硕士, 主要从事分子生物学在中医药中的应用研究, Tel: 0933-8501528, E-mail: yangjie2012@sina.cn.

技术、美国 Illumina 公司的 Solexa 技术以及美国 ABI 公司的 SOLiD 技术为标志的新一代高通量测序技术相继诞生,使得全转录组商业化测序成为可能. 转录组测序技术,作为传统测序一次革命性的改变,其更大的优势在于它是一个开放系统,具有寻找新的信息的能力^[6]. 本文就转录组测序技术的原理及应用现状与发展前景作一综述.

1 转录组测序技术的原理

对于转录组测序实验,主要包括以下步骤. 首先,对于真核生物,用带有寡聚核苷酸 Oligo (dT)的磁珠富集 mRNA,对于原核生物,用试剂盒去除 rRNA,将所得 mRNA 随机打成片段,再用随机引物和逆转录酶从 RNA 片段合成 cDNA 片段. 然后,对 cDNA 片段进行末端修复并连接测序接头,经琼脂糖凝胶电泳获取 200 bp 左右的大小片段,并进行 PCR 扩增,构建 cDNA 文库,用 Illumina HiSeq™ 2000 进行测序. 如果从总 RNA 中只提取长度为 21~23 个碱基左右的 RNA,则得到全部的 miRNA(microRNA)转录本,相应的方法称作 miRNA-seq^[7]. 最后,对转录组测序数据进行常规转录组分析,包括数据预处理、序列拼接、数据库注释、表达分析、GO 功能分类、KEGG pathway mapping 等.

2 转录组测序技术在整合网络调控中的应用现状

转录因子、基因、miRNA 三者之间相互影响,相辅相成,共同构成了生命现象的多层次整合调控网络^[8]. 作为新一代高通量测序技术,转录组测序技术主要检测细胞内基因表达和转录组信息. 因此,我们主要从 miRNA 与蛋白编码基因等层次分别介绍转录组测序技术在基础研究、临床研究等领域中的应用现状及发展前景.

2.1 转录组测序技术与 miRNA 层次研究

miRNA,又称为 MicroRNAs,是一类内生的、长度约 20~24 个核苷酸的小分子非编码 RNA,可调节细胞内多种生理与病理过程. miRNA 与基因构成复杂的调节网络既可以通过一个 miRNA 来调控多个基因的表达,也可以通过几个 miRNA 的组合来精细调控某个基因的表达,因此 miRNA 构成了生命现象多层次网络调控中心. 转录组测序技术的应用,其中小分子 RNA 深度测序,可探测新 miRNA 和 miRNA 变异体,广泛应用于疾病诊断、各物种生长发育与中药药用部位等研究中.

2.1.1 miRNA 与疾病诊断

miRNA 能沉默靶基因的表达,而 miRNA 变异体可导致人类疾病的发生发展,因此 miRNA 特征转录谱可作为疾病诊断的生物标志物,对揭示人类重大疾病的治疗靶点有重要作用. Han 等^[9]对膀胱上皮细胞癌进行深度测序,发现 *hsa-miR-182*、*hsa-miR-183*、*hsa-miR-200a*、*hsa-miR-143* 和 *hsa-miR-195* 等 miRNA 转录本异常表达于膀胱上皮癌细胞中,并经 RT-PCR 定量验证证实,与正常对照组存在明显差异. 近年来,miRNA 也用作诊断败血症的指纹图谱. 通过 Solexa 测序平台和 RT-PCR 定量验证,进一步分析认为 *miR-15a*、*miR-16*、*miR-193b* 和 *miR-483-5p* 等 miRNA 差异表达与败血症的死亡有关^[10]. 致癌基因的选择性表达,以及小分子 RNA 的异常表达与肿瘤的形成相关. 但是 miRNA 之间相互调控的网络机制仍有待于探讨. 转录组测序有助于发现肿瘤细胞中 miRNA 转录本的异常调节. 如神经胶质瘤中 miRNA 的成熟过程,Moore LM^[11]等测序 miRNA 的成熟体与前体,认为 miRNA 生物合成通路的异常调节是神经胶质瘤发生发展的生物标志. 朊病毒病是人或动物的一种急性致死性传染病. 经小分子 RNA 深度测序,显示了朊病毒感染的神经元细胞中独特的 miRNA 信号特征,可用作朊病毒病的诊断,以及揭示其分子发病机制^[12].

2.1.2 miRNA 与生长发育

miRNA 通过沉默或降解靶基因而调控基因表达,与生物体各物种的生长发育、细胞分化等密切相关. Sun^[13]等使用 Solexa 测序平台获取了 36 周小鸡下丘脑组织的 miRNA 转录本,其中 22 条表达上调,157 条表达下调;经 KEGG 通路分析,这些高表达的 miRNA 靶基因与小鸡的代谢和发育生长相关,有助于揭示 miRNA 网络调控的分子机制. 在拟南芥^[14]、水稻^[15]等植物生长过程中,miRNA 已广泛地被转录组测序技术所鉴定. Shamimuzza-man^[16]等运用转录组测序成功地从 53 个已知的 miRNA 中鉴定了 183 个不同的靶点,其中,25 个 miRNA 与大豆胎盘母面绒毛小叶的发生、发展有关,12 个 miRNA 与大豆种皮的生长有关,剩下的 16 个 miRNA 与大豆生长发展阶段无关. 同时,也显示了大量的 miRNA 所介导的靶点为转录因子,对维持大豆成熟的生长过程是十分重要的.

2.1.3 miRNA 参与调控中药药用部位组织的形成

对于同一个中药,入味部位不同,而表现出功效的差异.功效的差异反映出中药药效成分的差异,这一问题根本的原因归结于转录组学特征的不同.比如,构成眼睛与肠道组织细胞的基因组完全相同,但因其转录组的差异,而表现出眼睛与肠道结构和功能的悬殊差别.从人参的根、茎和花中分别提取RNA,并混合经转录组测序,即使同一miRNA家族的成员也在人参不同部位表现出不同的表达方式.如,*miR6135i*^[17]在人参根中低表达,却在茎和花中适度表达;*miR6135e.2j*^[17]在人参花中高表达.这说明了*miR6135i*和*miR6135e.2j*等^[17]miRNA转录本在人参中所表现出功能的差异,为人参miRNA网络调控功能的阐释提供了重要线索.除此之外,转录组测序技术也应用于药用植物虎杖^[18]、丹参^[19]的研究中,其中虎杖根部表达的12、18、60和54个基因分属于甲羟戊酸(MVA)、2-甲基-D-赤藓糖醇4-磷酸(MEP)、莽草酸和白藜芦醇生物合成路径,它们可能参与药用活性物质蒽醌类和白藜芦醇的生物合成;在丹参根的转录组测序研究中,经过分析获得了可能参与丹参酮合成的序列27条,参与丹酚酸合成的序列29条,细胞色素P450序列70条,为中药次级代谢产物相关基因表达调控的网络分析奠定了基础.

2.2 转录组测序技术与蛋白编码基因层次研究

如上所述,miRNA成为多层次整合网络调控中心^[8].转录因子、miRNA与基因三者之间相互影响,相辅相成,导致各物种生物性状出现多样性.生物性状的多样性,归结于mRNA转录本和大量的蛋白编码基因.在这成千上万的蛋白编码基因中,转录组测序技术可大规模发现新基因,是筛查遗传病病因的一个强有力的新工具,有助于了解疾病的发病机制和开发新的治疗方法.Ho等^[20]运用转录组测序技术比较CD90⁺肝癌干细胞与CD90⁺非肿瘤肝组织的基因表达谱.经生物信息学分析,在CD90⁺肝癌干细胞中表达上调的基因与炎症、肿瘤药物耐药性和脂质代谢相关,同时发现了GPC3基因与CD90⁺肝癌干细胞具有极强的相关性,有力揭示了抗肝癌药物的治疗靶点.2013年,陈小平等^[21]分别对花生地上果针和地下果针进行从头合成转录组测序,分别有859和1068个转录本的表达显著上调,其中地上果针中,与光合作用和衰老相关的基因表达明显上调.同年,中科院遗传与发育生物学研究所^[22]用全外

显子组测序方法,在一个非综合征性白化病家系中鉴定出一个色素合成相关基因SLC24A5,是导致眼皮肤白化病的新致病基因.

3 展望与小结

真核生物基因表达调控是多水平、多层次的复杂精密调控.每一种调控方式均具有不同的时空特点.这种时空调控网络包括转录调控和转录后调控.其中,转录因子和miRNA是转录调控和转录后调控的重要调控点,它们分别影响着机体内功能基因的差异表达.首先,转录因子之间的调节受相应转录因子调节物的调控,比如,酶级联连锁反应一样.其次,miRNA既可调节最高层次表达水平的转录因子,也可调节最低层次表达水平的转录因子,而转录因子主要与启动子或增强子结合,从而影响基因的表达.实际上,miRNA最终将影响靶基因的表达,因此,miRNA构成了生命现象多层次、多水平整合网络调控中心.根据转录组测序技术的优势,利用小分子RNA深度测序技术可检测任何物种的miRNA转录本.然后,通过生物信息学软件构建各物种、各组织细胞miRNA的整合调控网络,进而从整体水平诠释生命与疾病的奥秘.比如,离子通道和连接蛋白将参与冠心病心肌梗死发病的病理生理过程^[23,24].经拓扑网络生物信息学分析,miRNA的动态调控将导致连接蛋白43基因的异常调节,并鉴定了TRPV4、SCN5A、CACNA1C和TRPC6基因的生物意义,有力揭示了冠心病心肌梗死的发病机制^[23].

转录组测序技术实质上是一种高通量测序技术,可大规模检测未知基因与小分子非编码RNA,若恰当利用生物信息学软件,可从蛋白编码基因与miRNA等层次阐释生命现象的多层次、多水平整合调控网络.

参考文献(References):

- [1] VELCULESCU V E, ZHANG L, ZHOU W, *et al.* Characterization of the yeast transcriptome[J]. Cell, 1997, 88(2): 243-251.
- [2] ZHU T, WANG X. Large-scale profiling of the *Arabidopsis* transcriptome[J]. Plant Physiology, 2000, 124(4): 1472-1476.
- [3] FITZPATRICK D R, RAMSAY J, MCGILL N I, *et al.* Transcriptome analysis of human autosomal trisomy[J]. Human Molecular genetics, 2002, 11(26): 3249-3256.
- [4] VANDEPOELE K, van de PEER Y. Exploring the plant transcriptome through phylogenetic profiling[J]. Plant Physiology, 2005, 137(1): 31-42.
- [5] KUSHWAHA H R, SINGH A K, SOPORY S K, *et al.* Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* L. Heynh and *Oryza sativa* L. reveals their developmental and stress regulation[J]. BMC genomics, 2009,

- 10: 200.
- [6] 滕晓坤, 肖华胜. 基因芯片与高通量 DNA 测序技术前景分析[J]. 中国科学(生命科学)(TENG Xiao-kun, XIAO Hua-sheng. The perspective and analysis of gene chips and high-throughput DNA sequencing approach[J]. Science in China (Life Science), 2008, 38(10): 891-899.
- [7] 王曦, 汪小我, 王立坤, 等. 新一代高通量 RNA 测序数据的处理与分析[J]. 生物化学与生物物理进展(WANG Xi, WANG Xiao-wo, WANG Li-kun, *et al.* A review on the processing and analysis of next-generation RNA-seq data[J]. Progress in Biochemistry and Biophysics, 2010, 37(8): 834-846.
- [8] CHENG C, YAN K, HWANG W, *et al.* Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data[J]. PLoS Computational Biology, 2011, 7(11): e1002190.
- [9] HAN Y, CHEN J, ZHAO X, *et al.* MicroRNA expression signatures of bladder cancer revealed by deep sequencing[J]. PLoS one, 2011, 6(3): e18286.
- [10] WANG H, ZHANG P, CHEN W, *et al.* Serum microRNA signatures identified by solexa sequencing predict sepsis patient's mortality: a prospective observational study[J]. PLoS one, 2011, 7(6): e38885.
- [11] MOORE L M, KIVINEN V, LIU Y, *et al.* Transcriptome and small RNA deep sequencing reveals deregulation of miRNA biogenesis in human glioma[J]. The Journal of Pathology, 2012, 10(9): 4109-4120.
- [12] BELLINGHAM S A, COLEMAN B M, HILL A F, *et al.* Small RNA deep sequencing reveals a distinct miRNA signature released in exosomes from prion-infected neuronal cells[J]. Nucleic Acids Research, 2012, 40(21): 10937-10949.
- [13] SUN G R, LI M, LI G X, *et al.* Identification and abundance of miRNA in chicken hypothalamus tissue determined by Solexa sequencing[J]. Genetics and Molecular Research, 2012, 11(4): 4682-4694.
- [14] ADDO-QUAYE C, ESHOO T W, BARTEL D P, *et al.* Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome[J]. Current Biology, 2008, 18(10): 758-762.
- [15] GUO W, WU G, YAN F, *et al.* Identification of novel *Oryza sativa* miRNAs in deep sequencing-based small RNA libraries of rice infected with Rice stripe virus[J]. PLoS one, 2012, 7(10): e46443.
- [16] SHAMIMUZZAMAN M, VODKIN L. Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing[J]. BMC Genomics, 2012, 13: 310.
- [17] WU B, WANG M, MA Y, *et al.* High-throughput sequencing and characterization of the small RNA transcriptome reveal features of novel and conserved microRNAs in panax ginseng[J]. PLoS one, 2012, 7(9): e44385.
- [18] 郝大程, 马培, 穆军, 等. 中药植物虎杖根的高通量转录组测序及转录组特性分析[J]. 中国科学(生命科学)(HAO Da-cheng, MA Pei, MU Jun, *et al.* The analysis of high-throughput transcriptomes sequencing and transcriptomes characteristic about hu zhang root in traditional chinese drug[J]. Science in China (Life Science)), 2012, 42(5): 398-412.
- [19] 李滢, 孙超, 罗红梅, 等. 基于高通量测序 454 GS FLX 的丹参转录组学研究[J]. 药学报(LI Ying, SUN Chao, LUO Hong-mei, *et al.* The transcriptomes research of salvia miltiorrhiza by high-throughput 454 GS FLX sequencing[J]. Pharmaceutica Sinica), 2010, 45(4): 524-529.
- [20] HO D W, YANG Z F, YI K, *et al.* Gene expression profiling of liver cancer stem cells by RNA-sequencing[J]. PLoS one, 2012, 7(5): e37159.
- [21] CHEN X, ZHU W, AZAM S, *et al.* Deep sequencing analysis of the transcriptomes of peanut aerial and subterranean young pods identifies candidate genes related to early embryo abortion[J]. Plant Biotechnology Journal, 2013, 11(1): 115-127.
- [22] WEI A H, ZANG D J, ZHANG Z, *et al.* Exome sequencing identifies SLC24A5 as a candidate gene for nonsyndromic oculocutaneous albinism[J]. The Journal of Investigative Dermatology, 2013-3-14.
- [23] ZHOU R, HANG P, ZHU W, *et al.* Whole genome network analysis of ion channels and connexins in myocardial infarction[J]. Cell Physiology Biochemistry, 2011, 27(3): 299-304.
- [24] 杨杰, 王米渠. 血瘀证分子调控网络研究历程及前瞻[J]. 中国中西医结合杂志 (YANG Jie, WANG Mi-qu. The research course and perspective of molecular regulation network and blood stasis syndrome [J]. Chinese Journal of Integrated Traditional and Western Medicine), 2012, 32(10): 1422-1424.

关于《生命科学研究》杂志网上投稿系统正式运行的公告

《生命科学研究》杂志网上投稿系统已进入正式运行阶段。请有意投稿的作者登陆本刊官网 <http://smkx.hunnu.edu.cn> 进行在线投稿。通过 E-mail 投稿的作者必须重新在网站注册投稿,稿件状态及过刊信息均可通过登陆网站查询。

特此公告!

《生命科学研究》编辑部