

DOI:10.11937/bfyy.201712043

# 转录组的测序方法及应用研究概述

姚娜<sup>1,2</sup>, 刘秀明<sup>1,2</sup>, 董园园<sup>1</sup>, 王南<sup>1</sup>, 孟璐璐<sup>1</sup>, 李海燕<sup>1,2</sup>

(1. 吉林农业大学 生物反应器与药物开发教育部工程研究中心, 吉林 长春 130118;

2. 吉林农业大学 生命科学学院, 吉林 长春 130118)

**摘 要:**转录组研究是基因的结构和功能研究的基础, 转录组测序技术在近年来取得了很大的进展, 被应用到多种动植物中。利用其能够获得大量未知序列的优势, 已经被广泛应用到未知基因挖掘、分子调控及相关代谢途径中。该研究主要介绍了转录组研究的方法、转录组测序技术的发展及其在动植物中的重要应用, 并讨论其存在的问题及发展趋势。

**关键词:**转录组; 测序; 454 测序; Solexa; SOLiD

**中图分类号:**Q 944 **文献标识码:**A **文章编号:**1001-0009(2017)12-0192-07

转录组(transcriptome)是指某一特定的生理条件下, 细胞、组织或生物体内所有转录产物的集合, 即转录后所有 mRNA 的总称, 是研究细胞表型和功能的一个重要手段。转录组这个概念最初是由 VELCULESCU 等<sup>[1]</sup>在研究酵母基因组进行基因表达时提出的, 任何细胞的转录组在研究细胞生物学和生物化学方面都能够提供一些新的有用的信息。转录组研究是基因的结构和功能研究的基础, 是发掘新的功能基因的重要途径<sup>[2]</sup>。

转录组学(transcriptomics), 同蛋白质组学和代谢组学, 均属于功能基因组学研究范畴<sup>[3]</sup>, 是一门从整体水平上研究基因表达和转录调控的学科。转录组学作为一种新的研究方法, 利用全部基因的表达调控、蛋白质功能等信息来解决生物学问题, 将基因组学研究带入了一个高速发展的时代<sup>[4]</sup>。转录组学的研究目的不仅是不同转录组样本中每个基因的表达水平的变化, 也包括转录组的定位和注释及每个基因在基因组中的功能结构的测定<sup>[5-7]</sup>。

## 1 转录组的研究方法

研究转录组的基本方法有基因芯片技术(gene chip technology)、基因表达系列分析(serial analysis of gene expression, SAGE)<sup>[8]</sup>、转录组测序(RNA-Seq)<sup>[9]</sup>。

### 1.1 基因芯片技术

基因芯片技术是随着人类基因组计划而发展起来的基于杂交技术的多学科交叉的新型生物技术<sup>[10]</sup>。基因芯片是生物芯片的一种, 又称 DNA 芯片(DNA chips)<sup>[11]</sup>, 是把大量 DNA 分子片段按照特定的排列方式而固定在硅片、塑料、玻璃及尼龙膜等载体上, 从而形成致密、有序的 DNA 分子点阵。其基本原理是分子生物学中的核酸分子原位杂交技术, 即利用碱基互补配对的原则, 通过各种分子技术手段将核苷酸固定到支持物上, 然后将处理好的样品与其进行杂交, 通过检测每个探针分子的杂交信号而获取样品的分子数量和信息<sup>[12]</sup>。其制备过程包括: 试验设计、数据规范化、数据分析、探针和目的片段的关系及数据保存<sup>[13]</sup>, 基因芯片的主要的应用表现在基因表达分析方面, 能够在很短的时间内测定出不同功能状态、不同组织、部位等基因的差异表达, 从而得到特异的基因表达谱, 因此大大提高了对相关基因研究的效率。基因芯片的产生为分子生物学基础理论研究提供了一个巨大的技术平台, 最早的用途之一就是基因组测序, 能够测定重复序列或较长序列的 DNA 序列及进行不同基因组之间的同源区域的序列比较。此外, 基因芯片还被用于肿瘤

**第一作者简介:**姚娜(1980-), 女, 硕士, 实验师, 现主要从事生物反应器等研究工作。E-mail:52730603@qq.com.

**责任作者:**李海燕(1971-), 女, 博士, 教授, 现主要从事植物抗逆工程与分子生物学等研究工作。E-mail:hyli99@163.com.

**基金项目:**国家自然科学基金资助项目(31501366, 31101172); 国家“863”高技术研究发展计划资助项目(2011AA100606); 吉林省科技厅资助项目(20150623024TC-11); 大学生创新创业训练计划资助项目(201510193033)。

**收稿日期:**2016-12-15

的诊疗、药物研发及种苗的培养和筛选等方面,利用基因芯片技术可同时获取肿瘤细胞在各个生长时期与肿瘤生长相关基因的表达模式,将在其生长阶段的各个环节中共同表达的众多基因作为一个基因簇来进行研究,从而可以更好地筛选出与肿瘤相关的起作用的分子和可能起关键性调控作用的基因。基因芯片在控制药物使用方面也发挥着重要的作用,可以通过检测药物作用对生物体内基因的表达水平的影响,进而研究药物对基因表达调控的影响而了解该药物的分子机制及生物作用。但基因芯片的发展也有一定的局限性,如试验成本高、样品的制备和操作繁杂、特异性低等,也在某种程度上限制了该技术的发展。

## 1.2 基因表达系列分析

基因表达系列分析(SAGE),是基于 Sanger 测序法的一种基因表达差异分析技术,最早是在 1995 年被提出<sup>[14]</sup>,各相关学科中 SAGE 技术是一种数字化的序列表达技术,是分子生物学与生物信息学有机结合的产物<sup>[15-17]</sup>。该技术能够分析特定细胞或组织表达的基因,或比较不同组织、不同时空条件下基因表达的差异性<sup>[18]</sup>。该技术的主要优点在于假阳性率低、灵敏性很高、试验周期相对较短、可重复性强,可以检测到表达水平很低的基因,适用于寻找生物中的一些没有完成基因组测序的新基因,或为鉴定难以检测的开放阅读框和新基因提供依据<sup>[19]</sup>。基因表达系列分析的主要原理<sup>[20]</sup>(图 1)是先将 3'端一个短的寡核苷酸序列确定为转录本的标签,然后将这

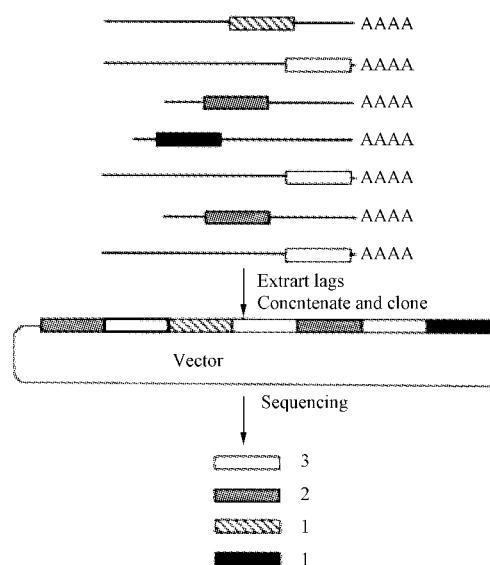


图 1 SAGE 的基本原理<sup>[20]</sup>

Fig. 1 Principle of SAGE

些标签串联到一起,形成大量的串联体,克隆到载体上进行测序获得代表转录本信息的标签序列,将获得的串联体的测序信息与已知的标签数据库进行对比分析,可以确定某一特殊条件或生理状态下生物体内基因的表达模式,并根据测序的标签出现的频率来确定基因的表达丰度。鲍忠赞等<sup>[19]</sup>利用基因序列分析技术研究了高温处理前后家蚕的基因表达差异,获得了不同处理的基因表达谱,构建了 2 个家蚕 SAGE 文库,得到了大量差异表达基因,为家蚕抗高温基因的鉴定以及探究基因调控的网络关系奠定基础。

## 1.3 转录组测序

转录组测序,即 RNA-Seq,是最近发展起来的基于新一代测序技术(next-generation sequencing)的利用深度测序技术进行转录组分析的技术<sup>[21]</sup>。自 1964 年首个完整基因的核苷酸序列发表以来,测序技术对分子生物学的发展起到了巨大的推动作用,RNA-seq 技术开始逐步取代基因芯片和基因表达系列分析技术<sup>[22]</sup>。转录组测序采用数字化信号,将细胞中所有的转录产物反转录为 cDNA 文库,然后将 cDNA 文库中的 DNA 随机剪切成小片段,在 cDNA 两端加上接头利用新一代高通量测序仪测序,直到获得足够的序列,所得序列通过比对或从头组装拼接,形成全基因组范围的转录谱(图 2)<sup>[23]</sup>。RNA-Seq 因其通量高、灵敏度高、分辨率高及不受物种限制等优点,在转录组学研究中占有重要地位,被认为是一种在转录水平上更为精确的测定分析方法<sup>[24]</sup>,该技

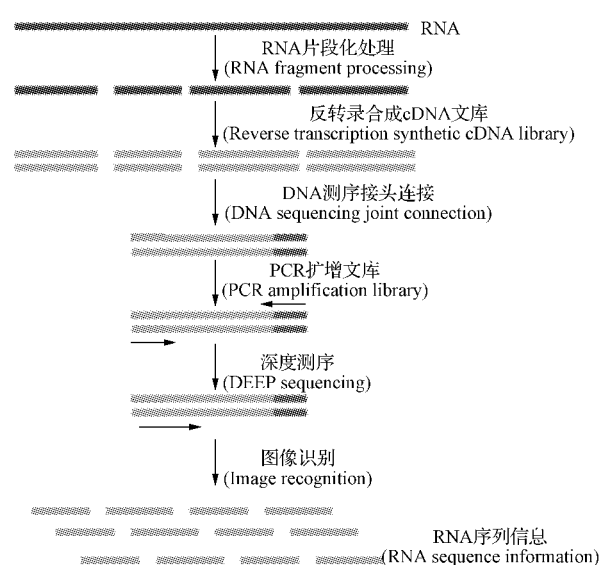


图 2 RNA-Seq 试验流程<sup>[23]</sup>

Fig. 2 Progress of RNA-Seq

术已经在动植物等研究领域开展了广泛的应用。

## 2 转录组测序研究的发展

测序技术在最近 30 年内取得了很大的进展<sup>[25]</sup>,即 Sanger 测序的传统的测序方法,是将 cDNA 片段克隆到细菌载体上,然后利用载体上的引物进行测序,这种方法适用于 cDNA 文库的构建<sup>[26]</sup>,随着毛细管测序的出现,该方法被称为表达序列标签(expressed sequence tags, ESTs)的高通量测序,被广泛地应用在动植物及微生物的基因发现、基因注释、单核苷酸的多态性(single nucleotide polymorphisms, SNP)发现、分子标记及表达分析等方面<sup>[27]</sup>。自 2005 年以来,随着测序技术的发展,新的测序技术迅速发展起来并有取代传统测序技术的优势,被称为新一代测序技术,包括 454 测序(罗氏生物公司)<sup>[28]</sup>、Solexa 技术(Illumina 公司)<sup>[29]</sup>及 SOLiD 技术(ABI 公司)<sup>[30]</sup>,可以说是对传统测序方法的一次变革。新一代测序技术又被称作深度测序或高通量测序,相对于传统的 Sanger 测序而言,其优点在于测序通量高,测序时间短且测序成本显著下降。该技术能在单核苷酸水平上对任意物种的整体转录活动检测,不仅能够分析转录本的结构和表达水平,同时还可以发现未知的转录本和稀有转录本,能够更加全面精确地获得转录组信息。

### 2.1 454 测序

454 测序技术是基于焦磷酸测序法并结合单分子乳液 PCR(emulsion PCR)和微流体技术的新型超高通量测序技术,它能够提供一个更高的转录组覆盖深度,特别适合一些未知基因组序列的物种<sup>[31]</sup>。该测序平台一次测序工作可产生 100 万条序列,序列的平均长度约为 400 bp,数据总量约 500 Mb<sup>[32]</sup>。454 测序技术的基本原理是依靠荧光信号的生物发光而进行的 DNA 序列分析技术。首先建立一个微反应器,用特殊设计的 DNA 捕获磁珠与短的 DNA 片段结合,并被扩增试剂乳化,从而形成油包水的混合物,这样就建立了只包含 1 个磁珠和 1 个独特片段的微反应器,每个独特的片段就会在自己的微反应器里进行独立的扩增。接下来,乳液的混合物被打破,而扩增的片段却仍然结合在磁珠上。然后将携带短片段的 PCR 产物的捕获磁珠放入只能容纳 1 个磁珠的 PTP 板中进行测序。在 4 个单独试剂瓶里的 4 种碱基(A、T、C、G),依照 T、A、C、G 的顺序依次循环进入 PTP 板,且每次只能进入 1 个碱基。如果发生碱基互补配对,就会释放 1 个焦磷酸。

而释放的焦磷酸会在 ATP 硫酸化酶和荧光素酶的作用下,释放出光信号并实时地被仪器所配置的高灵敏度 CCD 捕获到。因此,当有 1 个碱基和测序模板进行配对时,就会捕捉到一分子的光信号,以此类推,就能够准确、快速地确定待测模板的碱基序列。454 测序技术的主要优点是快速且准确率高、产生片段的数量多、不需要克隆和进行电泳等。

目前,454 测序技术已经被应用到多种物种中。ALAGNA 等<sup>[33]</sup>对 2 个橄榄品种花后 45 d 和 135 d 的 4 个不同样本 454 测序,获得 261 485 reads,总计 58 Mb,平均长度 217~224 nt,并发现了一些在果实成熟期和果实代谢调节及酚含量的相关差异表达基因。HE 等<sup>[34]</sup>利用 454 测序技术对金银花的花和芽进行测序,得到了 1 000 000 用于分析的 reads,平均长度 400 bp,并在数据库中获得了与细胞色素相关的 188 个基因、与糖基转移酶基因相关的 245 个基因。足叶草的 454 测序结果中产生了 1 503 232 个高质量的 reads,平均长度 138 bp<sup>[35]</sup>。应用该测序技术,DUARTE 等<sup>[36]</sup>对 8 个有代表性基因型的豌豆进行 454 测序,结果获得了 3 800 000 reads,并从中鉴定了 1 538 SNPs。WANG 等<sup>[37]</sup>对地毯草进行 454 测序,获得了 53 193 raw reads,1 942 SSR 引物并利用 24 个物种鉴定了 100 对引物。454 测序技术也被应用到亚麻<sup>[38]</sup>、兰花<sup>[39]</sup>、长春花<sup>[40]</sup>等。

### 2.2 Solexa 测序

Solexa 测序,即 Illumina 测序,是目前应用较为广泛的新一代测序系统,具有高灵敏度、高准确度、高通量和低成本等优点<sup>[41]</sup>,是目前应用在遗传分析和功能基因组等研究领域的主要技术。该技术主要利用单分子阵列方法,从细胞中提取 DNA,然后打断变小的片段,再将接头连接到片段上,经 PCR 扩增后制成文库,测序读取的长度一般为 120 bp。Solexa 测序技术在动物 microRNAs (miRNAs)研究中应用的较多,XU 等<sup>[42]</sup>通过 Solexa 测序对鹅卵巢的 miRNA 转录组进行测序,获得 21 241 283 clean reads 并对 1 328 个已知的 miRNAs 和 22 个新发现的 miRNAs 进行鉴定。MI 等<sup>[43]</sup>利用 Solexa 测序建立了雄性和雌性 2 个 miRNA 文库,并从 2 个文库中鉴定出 60 个已知和 124 个未知的 miRNAs。罗非鱼 2 个样本的 Solexa 测序结果共获得了 9 804 081 个 Unigenes,并通过生物信息学分析鉴定了 764 个成熟期 miRNA,209 个 miRNA-5p 和 202 个 miRNA-3p<sup>[44]</sup>。此外,该测序平台也被广泛地应用到其它物种中,如在木耳<sup>[45]</sup>中应用 Solexa 测序产生了大量的

基因组信息用于基因的挖掘和分子标记研究。WU等<sup>[46]</sup>对感染黑穗病的甘蔗进行 Solexa 测序,将获得的数据库在 NCBI 上比对,获得了 2 015 个差异表达基因,其中有 1 125 个上调基因和 890 个下调基因。目前,该测序技术越来越受到研究者的青睐,相关研究者为了获得更多的相关转录组数据,也被应用在棉铃虫<sup>[47]</sup>、武昌鱼<sup>[48]</sup>、里氏木霉<sup>[49]</sup>、大白菜<sup>[50]</sup>、肉鸡<sup>[51]</sup>、山羊<sup>[52]</sup>、珠子参<sup>[53]</sup>、辣椒<sup>[54]</sup>等研究上。

### 2.3 SOLiD 测序

SOLiD 测序技术的基本原理,是将接头与 DNA 片段结合后进行乳液 PCR,主要的特点在于每一步测序都是通过连接反应完成的,低成本、高通量及高准确度是其主要优势,但 SOLiD 测序因其测序序列较短,一般为 50 bp 左右,因此需要坚实的生物信息学基础分析转录组数据。MITRA 等<sup>[55]</sup>第一次应用 SOLiD 技术对人类样本进行测序,表明该技术在分析复杂的微生物学方面是可行的,且成本较低。尽管 SOLiD 测序每个基因都会被测 2 遍,错配率较低,但相对于 454 测序和 Solexa 测序而言,应用上还是受到了一定的限制。但对于模式植物来说,Solexa 测序和 SOLiD 测序相对而言都是较好的选择,因为二者的成本远远低于 454 测序<sup>[56]</sup>。

## 3 高通量测序技术的重要应用

高通量测序技术又称为第二代测序技术,也被称为深度测序、高能量测序,由于其高通量和低成本的优势,越来越受到研究者的重视,现已经被广泛地应用到缺乏基因组研究的各种动植物物种中<sup>[57]</sup>。

### 3.1 未知基因挖掘

对于一些未知基因组的物种,要深入研究相关基因有一定的困难。高通量测序技术能够找到一些重要的代谢途径中的基因,根据转录组测序结果进行下一步的分析。YAN 等<sup>[58]</sup>对月季的不同开花时期进行转录组测序,发现了与香味代谢合成途径中的一些相关基因,并鉴定了 5 个与之相关的基因。GIORDANO 等<sup>[59]</sup>利用高通量测序技术对毛花雀稗牧草的根、茎、叶和花进行了转录组测序,发现了 2 339 个 SSRs 标记的引物并鉴定了 96 个 SSRs,发现其中的 34% 在有性生殖和无性生殖系中表现出多态性。说明基因和基因组信息的发展有助于基因发现和表达的研究,与农艺性状相关的一些功能注释对分子育种和种质资源的开发将起到巨大的推动作用。从中国小鲵<sup>[60]</sup>的转录组测序结果中找到并鉴定了与免疫、性别和繁殖相关的一些基因,并获得了 31 982 个 SSRs 和 460 923 个 SNPs。从转录

组中获得一些推定的基因,在多项研究中都被证实了测序数据的准确性和可靠性,因此,对于分子育种及表达分析等方面具有重要的作用<sup>[61-63]</sup>。

### 3.2 分子标记

分子标记中最常用的 2 类是简单重复序列 (SSR) 和单核苷酸多态性 (SNP),根据转录组测序,研究者已经从多种物种中发现并检测了大量的 SSRs 和 SNPs。KAYA 等<sup>[64]</sup>对土耳其橄榄树进行 illumina 测序,发现了大量的 SNPs,并结合 AFLP (amplified fragment length polymorphism) 和 SSR 在 96 个基因型的橄榄中进行鉴定基因变异情况。在牧草植物龙须草<sup>[65]</sup>转录组测序中,获得了 26 438 832 个可用于分析的 reads,在同源基因中进行比对检测到 6 681 SSRs 和 47 177 SNPs,设计出 5 723 对 SSR 引物。在辣椒<sup>[66]</sup>、鱼腥草<sup>[67]</sup>、苜蓿<sup>[68]</sup>、洋葱<sup>[69]</sup>等其它物种高通量测序中也发现了大量的 SSRs 或 SNPs 作为分子标记进行研究。

### 3.3 相关代谢途径

转录组测序还有一个重要的作用就是在测序结果中,根据测得数据的拼接结果找到该物种基因所参与的代谢途径。WANG 等<sup>[70]</sup>为了研究萝卜中芥子油苷代谢途径和在萝卜根茎形成期中营养和风味相关的生物合成代谢的分子机制,利用转录组测序获得了长度为 55.73 Mb、66 110 000 对 reads 的数据库。LI 等<sup>[71]</sup>首次利用转录组测序在河蟹中研究了渗透调节代谢途径中的相关基因的表达分析,通过荧光定量 PCR 鉴定了 15 个差异表达基因。

## 4 结论

近年来,高通量测序技术在各个领域应用的比较广泛,转录组测序技术的开展对于新物种的发现、新基因的挖掘、分子标记及代谢调控的研究具有重要的理论意义,特别是针对一些资源信息薄弱的物种,具有重要的应用价值。尽管转录组测序技术能够获得不同条件下的大量的序列信息,但是对于研究者而言,对于获得的转录组数据,一般只针对自己的研究领域挖掘一些新的基因或序列,而对于其它的可供研究的序列还可以继续探索,从而达到有效利用转录组数据的目的。

在现代生物技术飞速发展的今天,只凭借传统的研究方法去开发和探索一些未知信息较为困难且研究缓慢,转录组测序凭借获得的大量未知序列及基因,可以针对不同的研究目标,从而寻找与之相关的基因和序列,进行基因表达、转录调控研究、基因

诊断和治疗、基因检测、基因表达谱分析等<sup>[72]</sup>其它方面的研究,特别是针对一些未知基因组的物种,利用转录组测序技术将是一个新的探索。

### 参考文献

- [1] VELCULESCU V E, ZHANG L, ZHOU W, et al. Characterization of the yeast transcriptome[J]. Cell, 1997, 88(2): 243-251.
- [2] 周华, 张新, 刘腾云, 等. 高通量转录组测序的数据分析与基因发掘[J]. 江西科学, 2012, 30(5): 607-611.
- [3] 吴琼, 孙超, 陈士林, 等. 转录组学在药用植物研究中的应用[J]. 世界科学技术(中医药现代化), 2010, 12(3): 457-461.
- [4] 薛建江, 邱景富. 病原菌感染宿主的转录组学研究进展[J]. 河北北方学院学报(医学版), 2007, 24(5): 63-66.
- [5] COSTA V, ANGELINI C, de FEIS I, et al. Uncovering the complexity of transcriptomes with RNA-Seq[J]. J Biomed Biotechnol, 2010, 6(27): 853-916.
- [6] RUAN Y, LE BER P, NG H H, et al. Interrogating the transcriptome[J]. Trends Biotechnol, 2004(22): 23-30.
- [7] WANG Z, GERSTEIN M, SNYDER M. RNA-Seq: A revolutionary tool for transcriptomics[J]. Nat Rev Genet, 2009(10): 57-63.
- [8] VELCULESCU V E, ZHANG L, VOGELSTEIN B, et al. Serial analysis of gene expression[J]. Science, 1995, 270(5235): 484-487.
- [9] 岳桂东, 高强, 罗龙海, 等. 高通量测序技术在动植物研究领域中的应用[J]. 中国科学: 生命科学, 2012, 42(2): 107-124.
- [10] 刘思言, 沈毓武, 王丕武. 基因芯片技术在植物中的应用研究进展[J]. 广东农业科学, 2013(8): 136-138.
- [11] RAMSAY G. DNA chips: State of the art[J]. Nature Biotechnology, 1998, 16(1): 40-44.
- [12] 张骞, 盛军. 基因芯片技术的发展和应[J]. 中国医学科学学报, 2008, 30(3): 344-347.
- [13] 姜庆利. 检测柯萨奇病毒、副流感病毒基因芯片的建立[D]. 长春: 吉林大学, 2013.
- [14] MARK D. Serial analysis of gene expression: ESTs get smaller[J]. Bioessays, 1996, 18(4): 261-262.
- [15] NIELSEN K L, HØGH A L, EMMERSEN J. Deep SAGE-digital transcriptomics with high sensitivity, simple experimental protocol and multiplexing of samples[J]. Nucleic Acids Res, 2006, 34: 133.
- [16] MOLINA C, ZAMAN-ALLAH M, KHAN F. The salt-responsive transcriptome of chickpea roots and nodules via deep Super SAGE[J]. BMC Plant Biol, 2011(11): 31.
- [17] CHUM W W, KWAN H S, AU C H, et al. Cataloging and profiling genes expressed in *Lentinula edodes* fruiting body by massive cDNA pyrosequencing and Long SAGE[J]. Fungal Genet Biol, 2011, 48: 359-369.
- [18] 王大力, 黄大鹏. 基因表达系列分析技术研究进展[J]. 动物医学进展, 2009, 30(4): 69-72.
- [19] 鲍忠赞, 张彩霞, 徐世清, 等. 应用基因表达系列分析(SAGE)技术研究高温处理前后家蚕的基因表达差异[J]. 蚕业科学, 2012, 38(3): 456-467.
- [20] 邓永键, 郝飞. 基因表达系列分析及其应用前景[J]. 免疫学杂志, 2003, 19(4): 315-321.
- [21] 刘红亮, 郑丽明, 刘青青, 等. 非模式生物转录组研究[J]. 遗传, 2013, 35(8): 955-970.
- [22] 井赵斌, 魏琳, 俞靓, 等. 转录组测序及其在牧草基因资源发掘中的应用前景[J]. 草业科学, 2011, 28(7): 1364-1369.
- [23] 祁云霞, 刘永斌, 荣威恒. 转录组研究新技术: RNA-Seq 及其应用[J]. 遗传, 2011, 33(11): 1191-1202.
- [24] 张春兰, 秦孜娟, 王桂芝, 等. 转录组与 RNA-Seq 技术[J]. 生物技术通报, 2012(12): 51-56.
- [25] WALL P K, LEEBENS-MACK J, CHANDERBALI A S, et al. Comparison of next generation sequencing technologies for transcriptome characterization[J]. BMC Genomics, 2009(10): 347.
- [26] ADAMS M D, KELLEY J M, GOCAYNE J D, et al. Complementary DNA sequencing: expressed sequence tags and human genome project[J]. Science, 1991, 252(5013): 1651-1656.
- [27] BOUCK A, VISION T. The molecular ecologist's guide to expressed sequence tags[J]. Mol Ecol, 2007, 16(5): 907-924.
- [28] MARGULIES M, EGHOLM M, ALTMAN W E, et al. Genome sequencing in microfabricated high-density picolitre reactors[J]. Nature, 2005, 437(7057): 376-380.
- [29] PORRECA G J, ZHANG K, LI J B, et al. Multiplex amplification of large sets of human exons[J]. Nat Methods, 2007, 4(11): 931-936.
- [30] 张全芳, 李军, 范仲学, 等. 高通量测序技术在农业研究中的应用[J]. 山东农业科学, 2013, 45(1): 137-140.
- [31] WANG Z W, JIANG C, WEN Q, et al. Deep sequencing of the *Camellia chekiangoleosa* transcriptome revealed candidate genes for anthocyan in biosynthesis[J]. Gene, 2014, 538(1): 1-7.
- [32] 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, 2010, 45(4): 524-529.
- [33] ALAGNA F, AGOSTINO N D, TORCHIA L, et al. Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development[J]. BMC Genomics, 2009(10): 399.
- [34] HE L, XU X, LI Y, et al. Transcriptome analysis of buds and leaves using 454 pyrosequencing to discover genes associated with the biosynthesis of active ingredients in *Lonicera japonica* Thunb. [J]. PLoS One, 2013, 8(4): e62922.
- [35] BHATTACHARYA D, SINHA R, HAZRA S, et al. De novo transcriptome analysis using 454 pyrosequencing of the Himalayan Mayapple, *Podophyllum hexandrum* [J]. BMC Genomics, 2013(14): 748.
- [36] DUARTE J, RIVIERE N, BARANGER A, et al. Transcriptome sequencing for high throughput SNP development and genetic mapping in pea[J]. BMC Genomics, 2014, 15(1): 126.
- [37] WANG XL, LI Y, LIAO L, et al. Isolation and characterization of microsatellite markers for *Axonopus compressus* (Sw.) Beauv. (Poaceae) using 454 sequencing technology[J]. Genet Mol Res, 2015, 14(2): 4696-4702.
- [38] WANG F, CHEN H, LI X, et al. Mining and identification of polyunsaturated fatty acid synthesis genes active during camelina seed development using 454 pyrosequencing[J]. BMC Plant Biol, 2015(15): 147.
- [39] ZHU G, YANG F, SHI S, et al. Transcriptome characterization of cymbidium sinense 'Dharma' using 454 pyrosequencing and its application in the identification of genes associated with leaf color variation[J]. PLoS One, 2015, 10(6): e0128592.
- [40] MOELLER S, WÖHRMANN T, HUETTEL B, et al. Development of 18 polymorphic microsatellite markers for *Vinca minor* (Apocynaceae) via 454 pyrosequencing [J]. Appl Plant Sci, 2015, 3(5):

1500015.

[41] 陈新. 榛子花芽转录组文库的 Solexa 测序及冷调节基因的表达谱分析[D]. 北京: 中国林业科学研究院, 2011.

[42] XU Q, ZHANG Y, CHEN Y, et al. Identification and differential expression of microRNAs in ovaries of laying and broody geese (*Anser cygnoides*) by Solexa sequencing[J]. PLoS One, 2014, 9(2): e87920.

[43] MI X, WEI Z, ZHOU Z, et al. Identification and profiling of sex-biased microRNAs from sea urchin *Strongylocentrotus nudus* gonad by Solexa deep sequencing[J]. Comp Biochem Physiol Part D Genomics Proteomics, 2014, 10C: 1-8.

[44] XIAO J, ZHONG H, ZHOU Y, et al. Identification and characterization of microRNAs in ovary and testis of Nile tilapia (*Oreochromis niloticus*) by using solexa sequencing technology[J]. PLoS One, 2014, 9(1): e86821.

[45] ZHOU Y, CHEN L, FAN X, et al. *De novo* assembly of auricularia polytricha transcriptome using illumina sequencing for gene discovery and ssr marker identification[J]. PLoS One, 2014, 9(3): e91740.

[46] WU Q, XU L, GUO J, et al. Transcriptome profile analysis of sugarcane responses to *Sporisorium scitaminea* infection using Solexa sequencing technology[J]. Biomed Res Int, 2013; 298920.

[47] ZHAO Z, WU G, WANG J, et al. Next-generation sequencing-based transcriptome analysis of *Helicoverpa armigera* larvae immune-primed with *Photobacterium luminescens* TT01[J]. PLoS One, 2013, 8(11): e80146.

[48] YI S, GAO Z X, ZHAO H, et al. Identification and characterization of microRNAs involved in growth of blunt snout bream (*Megalobrama amblycephala*) by Solexa sequencing[J]. BMC Genomics, 2013(14): 754.

[49] KANG K, ZHONG J, JIANG L, et al. Identification of microRNA-like RNAs in the filamentous fungus *Trichoderma reesei* by Solexa sequencing[J]. PLoS One, 2013, 8(10): e76288.

[50] WANG F, LI H, ZHANG Y, et al. MicroRNA expression analysis of rosette and folding leaves in Chinese cabbage using high-throughput Solexa sequencing[J]. Gene, 2013, 32(2): 222-229.

[51] PAN Y Z, WU S G, DAI H C, et al. Solexa sequencing of microRNAs on chromium metabolism in broiler chicks[J]. Nutrigenet Nutrigenomics, 2013, 6(3): 137-153.

[52] YUAN C, WANG X, GENG R, et al. Discovery of cashmere goat (*Capra hircus*) microRNAs in skin and hair follicles by Solexa sequencing[J]. BMC Genomics, 2013(14): 511.

[53] 张绍鹏, 金健, 胡炳雄, 等. 珍稀药用植物珠子参的转录组测序及分析[J]. 中国中药杂志, 2015, 40(1): 2084-2089.

[54] RAVEENDAR S, NA Y W, LEE J R, et al. The complete chloroplast genome of *Capsicum annuum* var. *glabriusculum* using illumina sequencing[J]. Molecules, 2015, 20(7): 13080-13088.

[55] MITRA S, FÖRSTER-FRÖMME K, DAMMS-MACHADO A, et al. Analysis of the intestinal microbiota using SOLiD 16S rRNA gene sequencing and SOLiD shotgun sequencing[J]. BMC Genomics, 2013, 5: S16.

[56] 李小白, 向林, 罗洁, 等. 转录组测序(RNA-seq)策略及其数据在分子标记开发上的应用[J]. 中国细胞生物学学报, 2013, 35(5): 720-

726.

[57] WANG L, LI P, BRUTNELL T P. Exploring plant transcriptomes using ultra high-throughput sequencing[J]. Brief Funct Genomics, 2010, 9(2): 118-128.

[58] YAN H, ZHANG H, CHEN M, et al. Transcriptome and gene expression analysis during flower blooming in *Rosa chinensis* 'Pallida' [J]. Gene, 2013, 540(1): 96-103.

[59] GIORDANO A, COGAN NO, KAUR S, et al. Gene discovery and molecular marker development, based on high-throughput transcript sequencing of *Paspalum dilatatum* Poir [J]. PLoS One, 2014, 9(2): e85050.

[60] CHE R, SUN Y, WANG R, et al. Transcriptomic analysis of endangered Chinese salamander: identification of immune, sex and reproduction-related genes and genetic markers [J]. PLoS One, 2014, 9(1): e87940.

[61] NG D W, SHI X, NAH G, et al. High-throughput RNA-seq for allelic or locus-specific expression analysis in *Arabidopsis*-related species, hybrids, and allotetraploids[J]. Methods Mol Biol, 2014, 1112: 33-48.

[62] LIU R, DONG Y, FAN G, et al. Discovery of genes related to witches broom disease in *Paulownia tomentosa* × *Paulownia fortunei* by a *de novo* assembled transcriptome[J]. PLoS One, 2013, 8(11): e80238.

[63] HSU J C, CHIEN T Y, HU C C, et al. Discovery of genes related to insecticide resistance in *Bactrocera dorsalis* by functional genomic analysis of a *de novo* assembled transcriptome[J]. PLoS One, 2012, 7(8): e40950.

[64] KAYA H B, CETIN O, KAYA H, et al. SNP discovery by illumina-based transcriptome sequencing of the olive and the genetic characterization of *Turkish olive* genotypes revealed by AFLP, SSR and SNP markers[J]. PLoS One, 2013, 8(9): e73674.

[65] ZOU D, CHEN X, ZOU D. Sequencing, *de novo* assembly, annotation and SSR and SNP detection of sabaigrass (*Eulaliopsis binata*) transcriptome[J]. Genomics, 2013, 102(1): 57-62.

[66] NICOLAI M, PISANI C, BOUCHET J P, et al. Discovery of a large set of SNP and SSR genetic markers by high-throughput sequencing of pepper (*Capsicum annuum*) [J]. Net Mol Res, 2012, 11(3): 2295-2300.

[67] WEI L, LI S, LIU S, et al. Transcriptome analysis of *Houttuynia cordata* Thunb. by Illumina paired-end RNA sequencing and SSR marker discovery[J]. PLoS One, 2014, 9(1): e84105.

[68] LIU Z, CHEN T, MA L, et al. Global transcriptome sequencing using the illumina platform and the development of EST-SSR markers in autotetraploid alfalfa[J]. PLoS One, 2013, 8(12): e83549.

[69] 李满堂, 张仕林, 邓鹏, 等. 洋葱转录组 SSR 信息分析及其多态性研究, 园艺学报, 2015, 42(6): 1103-1111.

[70] WANG Y, PAN Y, LIU Z, et al. *De novo* transcriptome sequencing of radish (*Raphanus sativus* L.) and analysis of major genes involved in glucosinolate metabolism[J]. BMC Genomics, 2013, 14: 836.

[71] LI E, WANG S, LI C, et al. Transcriptome sequencing revealed the genes and pathways involved in salinity stress of Chinese mitten crab, *Eriocheir sinensis* [J]. Physiol Genomics, 2014, 46(5): 177-190.

[72] 刘英华, 陈瑛. 高通量测序技术的最新研究进展[J]. 中国妇幼保健, 2013, 28(13): 1990-1992.

DOI:10.11937/bfyy.201712044

# 天麻种子萌发菌研究进展

王彩云<sup>1</sup>, 侯俊<sup>2</sup>, 王永<sup>1</sup>, 张翔宇<sup>1</sup>, 阮培均<sup>1</sup>

(1. 毕节市中药研究所, 贵州 毕节 551700; 2. 大方县扶贫开发办公室, 贵州 毕节 551700)

**摘要:**天麻是一种高度退化的兰科植物, 种子必须在紫萁小菇等小菇属真菌的帮助下才能完成萌发。紫萁小菇等天麻种子萌发菌的发现及应用, 是近年来天麻科研工作的一项新突破, 目前国内外对该菌的研究仅处于起步阶段。为了摸清天麻种子萌发菌生长与环境的关系, 找到适宜的培养条件和培养方法, 筛选出各天麻主产区的适宜萌发菌菌株, 提高天麻种子萌发率及天麻产量, 也为了充分利用萌发菌, 开拓萌发菌后熟产品, 缓解市场对野生天麻的巨大需求, 更为了进一步深入研究和大面积推广萌发菌, 该文章综述了天麻萌发菌的发现进程、生物学特性、分离纯化及提纯复壮技术、培养条件及菌株制作技术、化学成分及药理学研究进展, 以期为萌发菌的充分利用及深入研究提供参考依据。

**关键词:**天麻; 紫萁小菇; 石斛小菇; 有性繁殖; 胞内多糖

**中图分类号:**S 567.23<sup>+</sup>9 **文献标识码:**A **文章编号:**1001-0009(2017)12-0198-05

天麻(*Gastrodia elata*)属兰科(Orchidaceae)植物天麻属(*Gastrodia*)多年生草本植物, 又称明天麻、赤

箭、木浦、白龙皮、定风草等, 可用于治疗肢体麻木、小儿惊风、头痛眩晕、癫痫抽搐、破伤风等疾病<sup>[1-2]</sup>。近年来发现天麻对治疗抑郁症、老年性痴呆以及糖尿病有一定的疗效<sup>[3-5]</sup>, 导致国内外需求量剧增。

天麻是一种高度退化的兰科植物, 其种子必须在小菇属真菌的帮助下才能萌发, 而后必须与蜜环菌共生, 方能正常生长繁殖<sup>[6]</sup>。天麻种子萌发的营养源自萌发菌, 自1980年学者们从天麻种子发芽的原球茎中首次分离出与天麻种子具有共生作用的“京陕807-02号”等菌株以来, 研究者通过菌种分离试验, 先后成功地筛选出12种天麻种子萌发菌, 均

**第一作者简介:**王彩云(1989-), 女, 硕士, 助理研究员, 现主要从事药用植物资源评价与利用等研究工作。E-mail: wang-caiyun0716@126.com.

**责任作者:**阮培均(1963-), 男, 本科, 研究员, 现主要从事药用植物遗传与育种等研究工作。E-mail: rpj3819@126.com.

**基金项目:**国家“十二五”科技支撑计划资助项目(2015BAI05B03); 毕节市成果转化资助项目(毕科成字(2014)8号); 毕节市科技支撑计划资助项目(毕科合(2016)48号)。

**收稿日期:**2016-12-05

## Advances in Application and Sequencing Methods of Transcriptome

YAO Na<sup>1,2</sup>, LIU Xiuming<sup>1,2</sup>, DONG Yuanyuan<sup>1</sup>, WANG Nan<sup>1</sup>, MENG Lulu<sup>1</sup>, LI Haiyan<sup>1,2</sup>

(1. Ministry of Education Engineering Research Center of Bioreactor and Pharmaceutical Development, Jilin Agricultural University, Changchun, Jilin 130118; 2. College of Life Sciences, Jilin Agricultural University, Changchun, Jilin 130118)

**Abstract:** Transcriptome is the basis of the research of structure and function of genes, transcriptome sequencing technology has made great progress in recent years and been applied to many plants and animals species. It has been widely applied to unknown gene mining, molecular regulation and related metabolic pathways for its lots of known sequence. This study reviewed methods of transcriptome, development of transcriptome sequencing and its application in plants and animals, we discussed the problem of transcriptome and the development trend in the future.

**Keywords:** transcriptome; sequencing; 454 sequencing; Solexa; SOLiD