

转录组测序技术在药用植物研究中的应用

慧 芳, 刘秀岩, 李宗谕, 刘福顺, 杨世海*

吉林农业大学中药材学院, 吉林 长春 130118

摘要: 转录组测序技术是一种新发展的转录组学研究方法。在物种基因组信息未知的情况下, 进行测序得到遗传信息, 转录组测序技术是很重要的分子生物学方法, 已经被广泛应用于各个研究领域。而药用植物遗传信息匮乏, 对药用植物的转录组进行测序在基因组学是比较活跃的领域, 可以了解植物的基因表达并分析其功能和调控机制。对药用植物的转录组学研究有助于解决遗传育种、筛选优良抗性基因等问题。介绍了转录组测序的发展, 并从功能基因挖掘和次生代谢产物途径探索等方面综述了近年来转录组测序在药用植物研究中的应用, 为药用植物的研究提供了更多基础数据。

关键词: 药用植物; 转录组测序; 功能基因; 次生代谢产物; 遗传育种; 筛选优良抗性基因

中图分类号: R282.1

文献标志码: A

文章编号: 0253-2670(2019)24-6149-07

DOI: 10.7501/j.issn.0253-2670.2019.24.033

Application of transcriptome sequencing in study of medicinal plants

HUI Fang, LIU Xiu-yan, LI Zong-yu, LIU Fu-shun, YANG Shi-hai

College of Chinese Medicinal Materials, Jilin Agricultural University, Changchun 130118, China

Abstract: Transcriptome sequencing technology is a newly developed transcriptomics method. In the absence of the genome information of the species, sequencing is carried out to obtain genetic information. Transcriptome sequencing technology is an important molecular biology method with wide application in various research fields. Due to the deficiency of the genetic information of medicinal plants, and the transcriptome sequencing research of medicinal plants is an active field in genomics. Researchers can investigate plant gene expression and analyze its function and regulatory mechanisms. Transcriptomics research on medicinal plants can help to solve problems such as genetic breeding, excellent resistance genes screening, and etc. This paper introduces the development of transcriptome sequencing, and reviews the application of transcriptome sequencing in medicinal plants in recent years from functional gene mining and secondary metabolite pathway exploration, which provide more basic data for medicinal plant studies.

Key words: medicinal plant; transcriptome sequencing technology; functional genes; secondary metabolite; genetic breeding; screening for excellent resistance genes

转录组测序(transcriptome sequencing)是对某一物种的 mRNA 进行高通量测序, 而转录组测序的结果反映了特定条件和特定时间点的基因表达情况。随着时间或外部的环境变化, 转录组也会随之变化, 除了转录衰减和 mRNA 降解现象, 转录组包括了细胞内所有能转录出的 mRNA, 而转录组测序的结果反映了特定条件和特定时间点的基因表达情况^[1]。在参考基因组序列存在的情况下, 转录组测序可以实现序列的变异鉴定^[2-4]、空间和时间表达谱测定^[5-6]、基因模型结构的验证和鉴定^[7-8]。在没有参考基因组的情况下, 转录组测序也可以挖掘基

因^[9-12], 进行对比分析^[13], 计算其表达丰度^[14-15]。

转录组研究的方法比较多, 如基因表达序列分析(serial analysis of gene expression)、表达序列标签(expressed sequence tag, EST)、基因芯片、高通量测序等^[16-17]。而现阶段转录组学研究主要依靠高通量测序技术。但是, 对于非模式生物来说, 大多数植物的基因组信息较为匮乏, 而高通量测序的优势在于无需参考基因组, 在没有获得物种的基因组信息的情况下仍可以进行转录组测序。另外, 高通量测序获得的数据可以覆盖全部转录本, 这对于挖掘功能基因是十分有利的。

收稿日期: 2019-03-01

基金项目: 国家中医药管理局中药标准化项目: 甘草等 7 种中药饮片标准化建设(ZYBZH-Y-JL-25)

作者简介: 慧 芳(1994—), 女, 在读硕士, 研究方向为中药资源学。Tel: 15568871395 E-mail: hf29559363@163.com

*通信作者 杨世海, 教授, 博士生导师, 研究方向为中药生物技术、中药育种、药用植物栽培。Tel: 15568877267 E-mail: jlyangs@163.com

1 转录组测序技术

1.1 第 1 代测序技术

沃森和克里克在 1953 年发现了双螺旋结构^[18]，了解核酸的空间结构后，人们开始研究对应的序列信息，测序技术也开始发展。1964 年，Holley 完成了对酵母丙氨酸转移 tRNA 的测序^[19]，自此核酸测序才正式出现。随后测序技术迅速发展。1977 年，生物化学家 Sanger 发明了链终止测序法，一种基于 T4 DNA 聚合酶和大肠杆菌 DNA 聚合酶的快速 DNA 测序方法^[20-21]。同年，Gilbert 发明了化学降解测序法^[22]。这 2 种测序技术被认为是第 1 代测序技术。而后 Sanger 测序法渐渐占据了主要地位，现在普遍认为第 1 代测序技术为 Sanger 测序法。

1.2 第 2 代测序技术

人类基因组的测序工作结束后，人们想通过测序的方法来获得其他物种的基因信息^[23-24]。而传统的 Sanger 测序法比较费时费力、低通量且成本很高，不能满足日益增长的测序需求。很快人们开发出了第 2 代测序技术（next-generation sequencing, NGS）。第 2 代测序技术成本低、速度快，而且覆

盖度较深，可以同时对数百万个 DNA 进行测序，核酸测序技术进入自动化时代^[25-27]，渐渐取代了 Sanger 测序法。高通量测序目前可在 5 个主要商业平台上获得：Roche (454)、ABI-SOLi D、Illumina 基因组分析仪、Ab I3730xl 基因组分析仪以及 Heli Scope。现今应用较多的为前 3 者。

1.3 第 3 代测序技术

随着时间的推移，第 2 代测序技术的各种问题渐渐凸显，随后诞生了第 3 代测序技术，比较有代表性的测序技术：纳米孔单分子测序技术（Oxford Nanopore Technologies 公司）和单分子实时测序（single-molecular real-time, SMRT, PacBio 公司）。第 3 代测序技术使用的是单分子测序技术，通量高和测序读数长。近年来应用较为广泛的是 PacBio RSII 测序平台，无需进行 PCR 扩增即可完成测序。平均读长可以达到 10~15 kb，而最大读长可达 64.5 kb^[28]。而第 3 代测序的准确率明显比前 2 代低，PacBio 公司的 SMRT 测序技术准确率只能达到 85% 左右^[29]。这样一来，为了控制错误率无形中增加了测序的技术难度和成本^[30]。

表 1 3 代测序技术比较

Table 1 Comparison of three generations of sequencing technologies

测序技术	公司	测序仪	测序方法	测序长度	数据量
第 1 代	Applied Biosystems	ABI 3100/3130 ABI 3700 ABI 3730	Sanger 测序	700~1 000 bp	56~96 kb ^[31]
第 2 代	Roche	Genome Sequencer FLX	焦磷酸测序、边合成边测序	300~600 bp	400~600 Mb
	Applied Biosystems	SOLID 5 500 xl	边合成边测序	250 bp/300 bp	300 Gb
	Illumina	Illumina Hieq 2000	边合成边测序	101 bp	600 Gb
第 3 代	Oxford Nanopore Technologies	纳米孔单分子测序	电信号测序	2~5 kb	200 Mb
	PacBio	SMRT、PacBio RSII	边合成边测序	10~15 kb	350 Gb~1 Tb ^[32]

2 转录组测序技术在药用植物中的应用

药用植物合成了无数的次级代谢产物，而这些次生代谢产物是研发新药的重要来源，如青蒿素（artemisinin）、人参皂苷（ginsenoside）、紫杉醇（taxol）等，但是它们的含量并不是很高。所以造成了药用植物的无序开发，资源日益减少。通过克隆关键酶基因及代谢工程生产药用植物药效成分已经成为药物生产和新药开发的主要研究方法^[33]。

相比已经完成基因组测序的模式生物，药用植物缺少基因组数据，遗传背景不清晰，发展缓慢。

药用植物高通量转录组测序技术的出现，可以挖掘药用植物有效成分生物合成的功能基因并分析其表达规律，探索有效成分的生物合成途径及其调控机制，从而提高药用植物有效成分的含量。药用植物高通量测序为挖掘功能基因、探索药用植物有效成分的生物合成途径与调控机制、探索药材道地性分子机制、阐明基因功能提供了新工具。通过测序得到的信息，还可以挖掘参与植物生长发育、抗病抗逆等优良性状的基因。对于药用植物的保护具有极其重大的意义^[33-37]。

2.1 探索次生代谢产物生物合成途径及功能基因

研究药用植物的初生代谢产物或次生代谢产物的生物合成途径,就需要鉴定途径中的相关酶和基因;在有些途径的分节点往往存在对整条途径起重要的关键酶(限速酶)。因此,阐明药用植物初生代谢产物或次生代谢产物的生物合成途径,挖掘相关基因是非常必要的^[38]。

博落回 *Macleaya cordata* (Willd.) R. Br. 是罂粟属的一种传统中药,富含血根碱和白屈菜红碱,二者具有抗菌及促进生长等作用,对博落回进行转录组测序^[39],从测序结果中鉴定了 16 个参与血根碱和白屈菜红碱合成的候选基因,并对 14 个基因进行了体外功能验证。白及 *Bletilla striata* (Thunb.) Reichb.f. 含有多糖、联苕、二氢菲、甾体、三萜和挥发油等多种有效成分,白及胶质的主要成分为白及多糖,主要由葡萄糖和甘露聚糖组成。对白及全株进行转录组测序^[40],得到了 243 410 条 Unigenes 参与了半乳糖、果糖、糖胺多糖和甘露糖等多种糖类代谢,筛选到己糖激酶(62 个 Unigenes)、甘露糖变位酶(1 个 Unigene)、果糖激酶(61 个 Unigenes)、果糖-1,6-二磷酸酶(1 个 Unigenes)、二磷酸果糖激酶(9 个 Unigenes)、6-磷酸果糖激酶(12 个 Unigenes)基因。对芦荟 *Aloe vera* (Linn.) N. L. Burman var. *chinensis* (Haw.) Berg. 根和叶进行转录组测序^[41],分别得到 113 063 和 141 310 个 Unigenes,筛选到 16 个与皂苷、木质素、蒽醌和类胡萝卜素等次生代谢产物生物合成有关的基因。对毛鸡骨草 *Abrus mollis* Hance 的转录组测序^[42]得到 53 743 个 Unigenes,其中 7 572 个 Unigenes 注释到 128 条不同的 KEGG 途径。注释到苯丙氨酸、酪氨酸和色氨酸生物合成途径中的有 72 个 Unigenes,在苯丙烷类生物合成途径中注释到的 Unigenes 有 204 个,黄酮类生物合成途径中注释了 23 个 Unigenes,22 个 Unigenes 注释到了异黄酮生物合成途径。倍半萜类化合物是洋甘菊 *Matricaria chamomilla* L. 中主要的活性物质,对洋甘菊的花、茎、根和叶子进行测序^[43],获得 83 741 个 Unigenes,其中参与倍半萜类化合物合成的 Unigenes 有 61 个。黄精 *Polygonatum sibiricum* Delar. ex Redoute 中的黄精多糖具有提高免疫力的功能,在黄精的转录组测序中^[44],筛选到 138 个 Unigenes,编码 20 种与黄精多糖生物合成相关的酶。对大戟 *Euphorbia peginensis* Rupr. 的根、茎、

叶进行转录组测序^[45],得到了 157 491 个 Unigenes,其中选取 26 个 Unigenes 作为二萜合酶候选基因,通过 qRT-PCR 进一步分析,选择了 10 个参与二萜生物合成基因进行了验证。Liu 等^[46]对长春花 *Catharanthus roseus* (Linn.) G. Don 进行转录组测序,在测序数据中筛选出了 1 个属于 WRKY 家族的转录因子,对其进行克隆和功能验证,进一步确定该基因在萜类吲哚生物碱(TIAs)生物合成中的功能,并阐明了长春花萜类吲哚生物碱生物合成通路中转录因子的调控网络。

冬凌草中的冬凌草甲素属于贝壳杉烷型四环二萜类化合物,可以减缓肿瘤生长。对冬凌草进行转录组测序^[47],共得到 44 626 个 Unigenes,选择与萜类骨架和二萜类生物合成相关的候选基因 1-脱氧-D-木酮糖-5-磷酸还原异构酶(DXR)、1-脱氧-D-木酮糖-5-磷酸合酶(DXS)、内根-贝壳杉烯合成酶(KS)、贝壳杉烯氧化酶(KO),通过 qRT-PCR 分析这些候选基因在根、叶和茎中的表达模式。半夏 *Pinellia ternata* (Thunb.) Berit. 中主要的生物活性成分为生物碱,如生物合成由 L-苯丙氨酸开始的麻黄碱。半夏经测序后得到 89 068 个 Unigenes^[48],在测序数据中鉴定了 6 个注释为苯丙氨酸氨裂解酶(PAL)的 Unigenes 和 52 个参与麻黄碱和苯甲酸生物合成的 Unigenes。齐墩果烷型三萜皂苷是桔梗的主要化学成分,桔梗皂苷 D 是主要的生物活性成分,经过转录组测序^[49],获得 34 053 个 Unigenes,筛选到了 19 个参与三萜皂苷生物合成途径的候基因。

2.2 SSR 分子标记开发

分子标记可以反映物种在基因水平上的遗传多样性。简单重复序列(simple sequence repeat, SSR)技术广泛应用于基因定位、辅助育种^[50]、遗传图谱、分子标记筛选、品种纯度鉴定^[51]等方面的研究,是目前应用最广泛的 DNA 分子标记技术之一^[52]。SSR 可分为 2 类:基因组 SSR 和表达序列标签(EST-SSR)。以随机基因组序列为背景开发基因组 SSR 较为困难^[53-54]。但是在转录的 RNA 序列中鉴定 EST-SSR 比非编码序列更加容易。开发大量有价值的 EST 序列对于基因的注释、发现^[55-56]、分子标记的发展^[57-58]尤其重要。

对心籽绞股蓝 *Gynostemma cardiospermum* Cogn. ex Oliv.和绞股蓝 *Gynostemma pentaphyllum* (Thunb.) Makino 进行转录组测序^[51],得到 71 607 个 Unigenes,开发并验证了大量 EST-SSR 标记,筛选

到 3 891 个 SSR 序列, 设计了 2 596 对引物, 随机选择其中 360 对引物进行验证, 最终得到 15 个 SSR 标记可被用于评估该物种的遗传多样性。在七叶一枝花 *Paris polyphylla* Smith. 转录组测序数据^[59]中鉴定了 3 853 个 EST-SSR 分子标记, 随机设计引物, 其中 39 对引物扩增性良好, 开发了 9 个 SSR 标记, 测试了 11 个种质的 55 个七叶一枝花样品的基因型。

丛琨^[60]对味连 *Coptis chinensis* Franch. 和云连 *C. teeta* Wall. 的叶片和根进行转录组测序, 将大量转录组数据与公共数据库序列进行比对、功能注释。筛选出大量与异喹啉生物碱生物合成相关的基因和功能性 SSR。张金渝等^[61]利用转录组测序技术开发 SSR, 利用 SSR 标记对 6 个文山三七 *Panax notoginseng* (Burk.) F. H. Chen 居群的遗传多样性及其遗传结构进行分析, 发现三七具有丰富的遗传多样性。黄海燕等^[62]建立了适合杜仲 *Eucommia ulmoides* Oliver 的 SSR-PCR 反应体系, 共筛选出 13 对具有多态性的引物, 为杜仲的遗传多样性、群体遗传学、亲缘关系等研究奠定了基础。邓科君等^[63]开发了 83 对 SSR 引物, 对 13 个不同居群的丹参 *Salvia miltiorrhiza* Bunge 样品以及 10 种鼠尾草属物种进行了扩增、多态性及通用性检测。郭林林^[64]对丹参品系 ZH23 进行简化基因组测序, 利用得到的序列信息查找开发 SSR, 发现可以用于设计引物的 SSR 共 6 058 个, 并设计出 6 058 对特异性的 SSR 引物, 选取 665 对 SSR 引物进行筛选, 有 568 对引物产生了清晰的条带, 共有 356 对 SSR 引物在亲本间具有多态性。杨雯等^[65]对虎耳草属植物混合样本

进行基因组测序, 利用 MISA 程序识别和定位 SSR 位点, 随机合成 120 对引物进行多态性引物筛选, 再从多态性引物中选取易扩增、多态性高的引物, 共扩增出 2 687 个多态性条带, 每对引物平均扩增出了 158 个多态性条带。

2.3 次生代谢产物生物合成途径挖掘

三萜皂苷是远志 *Polygala tenuifolia* Willd. 重要的药用成分之一。对不同年份的远志进行转录组测序^[66], 有 176 个 Unigenes 参与萜类化合物骨架生物合成, 参与其他次级代谢产物的生物合成的 Unigenes 有 522 个。杠柳 *Periploca sepium* Bunge 中主要的生物活性物质 C₂₁ 类固醇和杠柳毒苷均来自类固醇合成途径, 对杠柳叶、根、不定根及愈伤组织进行转录组测序^[67], 筛选到 24 个参与类固醇生物合成途径的基因。贯叶金丝桃 *Hypericum perforatum* L. 中的金丝桃素和褪黑激素是其主要的活性成分, 具有抗抑郁、抗炎、抗病毒、抗癌和抗菌作用, 测序结果^[12]显示, 2 359 个 Unigenes 与次生代谢产物通路有关, 与金丝桃素和褪黑激素合成有关的 Unigenes 有 260 个。对虎杖 *Polygonum cuspidatum* Sieb. et Zucc. 根部进行转录组测序^[68], 获得 86 418 个 Unigenes, 总共有 22 572 个 Unigenes 注释到了 119 个 KEGG 通路。分别有 12、18、60、54 个 Unigenes 分别注释到甲羟戊酸 (MVA)、甲基-D-赤藓糖醇-4-磷酸 (MEP)、莽草酸和白藜芦醇生物合成途径。

转录组测序技术应用于药用植物的研究情况见表 2。

表 2 药用植物转录组测序情况汇总

Table 2 Summary of transcriptome sequencing of medicinal plants

植物名称	植物种属	测序平台	有效成分	Unigene 数量
毛地黄	毛地黄属	454	强心苷	23 532
灯盏细辛	飞蓬属	GAIIX	灯盏乙素	73 092
冬凌草	香茶属	HiSeq2000	冬凌草甲素	44 626
虎眼万年青	虎眼万年青属	HiSeq2000	甾醇糖苷 OSW-1	10 418
桔梗	桔梗属	HiSeq2000	桔梗皂苷 D	34 053
短小蛇根草	蛇根草属	HiSeq2000	喜树碱	35 608
黄精	黄精属	HiSeq2500	黄精多糖	74 130
甘草	甘草属	HiSeq2000	萜类、黄酮类	94 828、305 100
半夏	半夏属	HiSeq2000	麻黄碱	89 068
大戟	大戟属	HiSeq2000	二萜类化合物	157 491
钩藤	钩藤属	HiSeq2000	钩藤碱、异钩藤碱	100 940
滇重楼	重楼属	HiSeq2000	甾体	87 577
丹参	鼠尾草属	454	丹参酮	64 139

续表 2

植物名称	植物种属	测序平台	有效成分	Unigene 数量
苦参	槐属	HiSeq2000	生物碱、黄酮类	83 325
白及	白芨属	HiSeq4000	白及胶质	243 410
远志	远志属	HiSeq2000	三萜皂苷	39 625
杠柳	杠柳属	Hiseq 2500	C ₂₁ 类固醇、杠柳毒苷	71 629
芦荟	芦荟属	HiSeq2000	皂苷、木质素等	113 063、141 310
毛鸡骨草	相思子属	HiSeq2000	黄酮类	53 743
贯叶金丝桃	金丝桃属	HiSeq2000	金丝桃素、褪黑激素	59 184
虎杖	虎杖属	HiSeq2000	白藜芦醇	86 418
川续断	川续断属	GAllx	三萜皂苷	43 243
菰蓝	菰蓝属	HiSeq2000	生物碱	33 238
番泻叶	山扁豆属	Illumina Miseq	番泻苷	42 230、37 174
洋甘菊	菊科	Hiseq 2500	萜类	83 741
党参	党参属	HiSeq2000	党参多糖	45 511
山苍子	木姜子属	HiSeq2000	单萜、倍半萜	68 648
小叶买麻藤	买麻藤属	HiSeq2000	黄酮类、二苯乙烯类	94 816
金钗石斛	石斛属	HiSeq4000	石斛碱、酚类、多糖	61 998

3 结语

本文综述了转录组测序技术在药用植物中的研究进展, 转录组测序技术是一种转录组学的研究方法, 遗传信息可以通过测序获得, 在物种基因组信息未知的情况下, 通过测序可得到遗传信息。由于药用植物遗传信息的匮乏, 转录组测序技术逐渐应用于药用植物转录组的研究中。为了保护重要的药用植物及濒危药用植物的多样性及其可持续利用, 在转录组水平上进行测序是非常必要的。

本文从功能基因挖掘、SSR 分子标记开发和次生代谢产物生物合成途径探索 3 个方面阐述了转录组测序技术在药用植物中的应用。除了以上应用以外, 转录组学的应用推进了药用植物中天然产物发现。植物可合成无数次级代谢产物, 而这些次生代谢产物可以作为药物开发的新来源, 研究药用植物有效成分的生物合成及其遗传机制有助于大规模生产这些有效成分。转录组测序技术还可以应用于某些有毒中药的减毒研究, 它的出现有助于选育品质产量皆优的药用植物, 阐明药用植物遗传信息及调控网络, 探索中药防病治病的分子机制。转录组测序技术在药用植物研究领域必将拥有广阔的应用前景。

参考文献

[1] 王丽鸾. 基于 EST 数据库和转录组测序的茶树 DNA 分子标记开发与应用研究 [D]. 北京: 中国农业科学院,

2011.

- [2] Hansey C N, Vaillancourt B, Sekhon R S, *et al.* Maize (*Zea mays* L.) genome diversity as revealed by RNA-sequencing [J]. *PLoS One*, 2012, 7(3): e33071.
- [3] Lai K, Duran C, Berkman P J, *et al.* Single nucleotide polymorphism discovery from wheat next generation sequence data [J]. *Plant Biotechnol J*, 2012, 10(6): 743-749.
- [4] Grit H, Thomas S, Michael S, *et al.* From RNA-seq to large-scale genotyping-genomics resources for rye (*Secale cereale* L.) [J]. *BMC Plant Biol*, 2011, 11(1): 131-143.
- [5] Massa A N, Childs K L, Lin H, *et al.* The transcriptome of the reference potato genome *Solanum tuberosum* group *Phureja* clone DM1-3 516R44 [J]. *PLoS One*, 2011, 6(10): e26801.
- [6] Dugas D V, Monaco, M K, *et al.* Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid [J]. *BMC Genom*, 2011, 12(1): 514-534.
- [7] Filichkin S A, Priest H D, Givan S A, *et al.* Genome-wide mapping of alternative splicing in *Arabidopsis thaliana* [J]. *Genome Res*, 2010, 20(1): 45-58.
- [8] Zhang G, Guo G, Hu X, *et al.* Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome [J]. *Genome Res*, 2010, 20(5): 646-654.
- [9] Giddings L A, Liscombe D K, Hamilton J P, *et al.* A stereoselective hydroxylation step of alkaloid

- biosynthesis by a unique cytochrome P450 in *Catharanthus roseus* [J]. *J Biol Chem*, 2011, 286(19): 16751-16757.
- [10] Wang Y, Alonso A P, Wilkerson C G, *et al.* Deep EST profiling of developing fenugreek endosperm to investigate galactomannan biosynthesis and its regulation [J]. *Plant Mol Biol*, 2012, 79(3): 243-258.
- [11] Jensen J K, Schultink A, Keegstra K, *et al.* RNA-Seq analysis of developing nasturtium seeds (*Tropaeolum majus*): Identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis [J]. *Mol Plant*, 2012, 5(5): 984-992.
- [12] He M, Wang Y, Hua W, *et al.* De novo, sequencing of *Hypericum perforatum*, transcriptome to identify potential genes involved in the biosynthesis of active metabolites [J]. *PLoS One*, 2012, 7(7): e42081.
- [13] Wang Y, Zeng X, Iyer N J, *et al.* Exploring the switchgrass transcriptome using second-generation sequencing technology [J]. *PLoS One*, 2012, 7(3): e34225.
- [14] Guo S, Liu J, Yi Z, *et al.* Characterization of transcriptome dynamics during watermelon fruit development: Sequencing, assembly, annotation and gene expression profiles [J]. *BMC Genomics*, 2011, 12(1): 454-466.
- [15] Lu J, Du Z X, Kong J, *et al.* Transcriptome analysis of *Nicotiana tabacum* infected by cucumber mosaic virus during systemic symptom development [J]. *PLoS One*, 2012, 7(8): e43447.
- [16] 吴春颖, 宋经元, 陈士林. 表达序列标签在药用植物研究中的应用 [J]. 中草药, 2008, 39(5): 778-782.
- [17] 胡松年. 基因表达序列标签 (EST) 数据 [M]. 杭州: 浙江大学出版社, 2005.
- [18] Watson J D, Crick F H. Molecular structure of nucleic acid [J]. *Nature*, 1953, 171(4356): 737-738.
- [19] Holley R W, Everett G A, Madison J T. Nucleotide sequences in the yeast alanine transfer ribonucleic acid [J]. *J Biol Chem*, 1965, 240(5): 2122-2128.
- [20] Sanger F, Nicklen S, Coulson A R. DNA sequencing with chain-termination inhibitors [J]. *Proceed Nat Acad Sci USA*, 1977, 74(12): 5463-5467.
- [21] Sanger F, Coulson A R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase [J]. *J Mol Biol*, 1975, 94(3): 441-448.
- [22] Maxam A M, Gilbert W. A new method for sequencing DNA [J]. *Proceed Nat Acad Sci USA*, 1977, 74(2): 560-564.
- [23] Aparicio S, Chapman J, Stupka E, *et al.* Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes* [J]. *Science*, 2002, 297(5585): 1301-1310.
- [24] Yu J, Hu S, Wang J, *et al.* A draft sequence of the rice genome (*Oryza sativa* L. spp. *indica*) [J]. *Science*, 2002, 296(5565): 79-92.
- [25] Huang L L, Yang X, Sun P. *et al.* The first illumina-based de novo transcriptome sequencing and analysis of safflower flowers [J]. *PLoS One*, 2012, 7(6): e38653.
- [26] 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, 2012, 42(5): 398-412.
- [27] 宋利璞. 药用植物金银花及药用真菌蝙蝠蛾拟青霉的转录组学研究 [D]. 北京: 中国科学院北京基因组研究所, 2014.
- [28] Rhoads A, Au K F. PacBio sequencing and its applications [J]. *Genom Proteom Bioinform*, 2015, 13(5): 278-289.
- [29] 张欣欣, 王铭杰. 新一代测序技术在 HBV 变异研究中的应用 [J]. 临床肝胆病杂志, 2015, 31(4): 514-519.
- [30] 卢 鹏, 金静静, 李泽锋, 等. 基于第三代测序技术的基因组组装方法及其在烟草中的应用 [J]. 烟草科技, 2018, 51(2): 87-94.
- [31] Mitreva M, Mardis E R. *Large-Scale Sequencing and Analytical Processing of ESTs* [M]. New Jersey: Humana Press, 2009.
- [32] Xin J, Zhang R C, Wang L, *et al.* Researches on transcriptome sequencing in the study of traditional Chinese medicine [J]. *Evid Based Compl Alt*, 2017, 2017(4): 1-9.
- [33] 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): 457-462.
- [34] 郭 淑. 基于转录组测序的石斛生物碱和人参皂苷生物合成相关基因的发掘、克隆及鉴定 [D]. 北京: 北京协和医学院, 2013.
- [35] 王海静, 严铭铭, 邵 帅, 等. 人参三七皂苷化学成分及药理作用对比研究 [J]. 人参研究, 2008, 20(1): 2-11.
- [36] 岳跃冲, 范燕萍. 植物萜类合成酶及其代谢调控的研究进展 [J]. 园艺学报, 2011, 38(2): 379-388.
- [37] 张 婕. 川贝母产地生态适宜性分析和转录组学研究 [D]. 北京: 北京中医药大学, 2018.
- [38] 郑夏生. 岗梅转录组及三萜皂苷生物合成相关酶基因的挖掘 [D]. 广州: 广州中医药大学, 2014.
- [39] 张红艳, 罗 杰. 中国科学家完成博落回全基因组测序并解析茛菪基异奎宁类生物碱途径 [J]. 植物学报, 2018, 53(3): 289-292.
- [40] 姜福星, 魏丕伟, 魏帼英, 等. 白芨转录组特性分析 [J]. 分子植物育种, 2018, 16(7): 2155-2165.
- [41] Choudhri P, Rani M, Sangwan R S, *et al.* De novo, sequencing, assembly and characterisation of *Aloe vera*, transcriptome and analysis of expression profiles of genes related to saponin and anthraquinone metabolism [J]. *BMC Genomics*, 2018, 19(1): 427-447.
- [42] Yuan X, Li K, Huo W, *et al.* De novo transcriptome sequencing and analysis to identify genes involved in the biosynthesis of flavonoids in *Abrus mollis*, leaves [J].

- Russ J Plant Physiol, 2018, 65(3): 333-344.
- [43] Zhang W, Tao T, Liu X, et al. De novo assembly and comparative transcriptome analysis: Novel insights into sesquiterpenoid biosynthesis in *Matricaria chamomilla* L. [J]. *Acta Physiol Plant*, 2018, 40(7): 129-142.
- [44] Wang S, Wang B, Hua W, et al. De novo assembly and analysis of *Polygonatum sibiricum* transcriptome and identification of genes involved in polysaccharide biosynthesis[J]. *Int J Mol Sci*, 2017, 18(9): 1950-1966.
- [45] Cao X, Zhang F, Yuan B, et al. De novo, transcriptome sequencing and analysis of *Euphorbia pekinensis* Rupr. and identification of genes involved in diterpenoid biosynthesis [J]. *Plant Gene*, 2017, 12: 33-42.
- [46] Liu J Q, Cai J, Wang R, et al. Transcriptional regulation and transport of terpenoid indole alkaloid in *Catharanthus roseus*: Exploration of new research directions [J]. *Int J Mol Sci*, 2017, 18(1): 53.
- [47] Su X, Li Q, Chen S, et al. Analysis of the transcriptome of *Isodon rubescens*, and key enzymes involved in terpenoid biosynthesis [J]. *Biotechnol Biotechnol Equip*, 2016, 30(3): 592-601.
- [48] Zhang G, Jiang N, Song W, et al. De novo sequencing and transcriptome analysis of *Pinellia ternata* identify the candidate genes involved in the biosynthesis of benzoic acid and ephedrine [J]. *Front Plant Sci*, 2016, doi: 10.3389/fpls.2016.01209.
- [49] Ma C H, Gao Z J, Zhang J J, et al. Candidate genes involved in the biosynthesis of triterpenoid saponins in *Platycodon grandiflorum* identified by transcriptome analysis [J]. *Front Plant Sci*, 2016, doi: 10.3389/fpls.2016.00673.
- [50] Kantartzi S K, Ulloa M, Sacks E, et al. Assessing genetic diversity in *Gossypium arboreum* L. cultivars using genomic and EST-derived microsatellites [J]. *Genetica*, 2009, 136(1): 141-147.
- [51] Zhao Y M, Zhou T, Li Z H, et al. Characterization of global transcriptome using illumina paired-end sequencing and development of EST-SSR markers in two species of *Gynostemma* (Cucurbitaceae) [J]. *Molecules*, 2015, 20(12): 21214-21231.
- [52] 刘列钊, 林呐. 油菜简单重复序列 SSR (simple sequence repeat) 研究进展 [J]. *生命科学*, 2004, 16(3): 173-176.
- [53] Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: A review [J]. *Mol Ecol*, 2002, 11(1): 1-16.
- [54] Squirrell J, Hollingsworth P M, Woodhead M, et al. How much effort is required to isolate nuclear microsatellites from plants? [J]. *Mol Ecol*, 2003, 12(6): 1339-1348.
- [55] Bouck A, Vision T. The molecular ecologist's guide to expressed sequence tags [J]. *Mol Ecol*, 2007, 16(5): 907-924.
- [56] Emrich S J, Barbazuk W B, Li L, et al. Gene discovery and annotation using LCM-454 transcriptome sequencing [J]. *Genome Res*, 2007, 17(1): 69-73.
- [57] Barbazuk W B, Emrich S J, Chen H D, et al. SNP discovery via 454 transcriptome sequencing [J]. *Plant J* 2007, 51: 910-918.
- [58] Novaes E, Drost D R, Farmerie W G, et al. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome [J]. *BMC Genom*, 2008, 9(1): 312-325.
- [59] Wang L, Yang Y, Zhao Y, et al. De novo characterization of the root transcriptome and development of EST-SSR markers in *Paris polyphylla* Smith var. *yunnannensis*, an endangered medical plant [J]. *J Agric Sci Technol*, 2016, 18(2): 437-452.
- [60] 丛 琨. 转录组分析发现黄连异喹啉生物碱合成相关基因及 SSR 标记 [D]. 昆明: 云南农业大学, 2017.
- [61] 张金渝, 杨维泽, 崔秀明. 三七栽培居群遗传多样性的 EST-SSR 分析 [J]. *植物遗传资源学报*, 2011, 12(2): 249-254.
- [62] 黄海燕, 杜红岩, 乌云塔娜, 等. 基于 SSR 分子标记的杜仲遗传多样性体系建立 [J]. *林业科学研究*, 2013, 26(6): 795-799.
- [63] 邓科君, 张 勇, 熊丙全. 药用植物丹参 EST-SSR 标记的鉴定 [J]. *药学报*, 2009, 44(10): 1165-1172.
- [64] 郭林林. 丹参基因组 SSR 标记的开发及其在连锁图谱构建的应用 [D]. 济南: 山东农业大学, 2016.
- [65] 杨 雯, 蒋 伟, 钟国跃, 等. 虎耳草属植物 SSR 分子标记的开发及应用 [J]. *中国中药杂志*, 2018, 43(10): 2057-2066.
- [66] Tian H, Xu X, Zhang F, et al. Analysis of *Polygala tenuifolia* transcriptome and description of secondary metabolite biosynthetic pathways by illumina sequencing [J]. *Int J Genom*, 2015, 2015(2/3): 782635.
- [67] Li X, Zhang J, Lu F, et al. De novo sequencing and transcriptome analysis reveal key genes regulating steroid metabolism in leaves, roots, adventitious roots and calli of *Periploca sepium* Bunge [J]. *Front Plant Sci*, 2017, 8(594): 1-15.
- [68] Hao D C, Ma P, Mu J, et al. De novo characterization of the root transcriptome of a traditional Chinese medicinal plant *Polygonum cuspidatum* [J]. *Sci China Life Sci*, 2012, 55(5): 452-466.