

夏枯草的转录组测序与次生代谢产物生物合成相关基因的挖掘

朱昀昊^{1,2}, 张梦佳¹, 李璐¹, 赵乐^{1,2}, 董诚明^{1,2*}

1. 河南中医药大学药学院, 河南 郑州 450046

2. 呼吸疾病诊疗与新药研发河南省协同创新中心, 河南 郑州 450046

摘要: 目的 获得夏枯草转录组信息, 探索夏枯草次生代谢产物生物合成的分子基础。方法 利用 Illumina HiSeq 2000 高通量测序技术对夏枯草果穗、茎和叶等部位进行转录组测序, 经 Trinity 软件组装后获得 Unigene, 并进行系统的生物信息学分析。结果 共获得 77 863 条 Unigene, 平均长度为 716.72 nt。通过与 NR、Swiss-Prot、COG 等 7 个公共数据库进行 BLAST 比对, 共有 41 367 个 Unigene 获得注释, 注释率为 53.13%。在 KEGG 数据库中, 有 1 406 条 Unigene 被注释到次生代谢产物合成。通过对转录组数据深入挖掘发现, 参与夏枯草三萜类骨架合成的 Unigene 共有 60 个, 参与酚酸类合成的 Unigene 共有 24 个, 参与次生代谢后修饰的 Unigene 共有 259 个, 参与其他次生代谢合成的 Unigene 共有 118 个。结论 夏枯草转录组数据为探索夏枯草次生代谢产物的生物合成机制提供了重要的资源, 也为利用代谢工程增加夏枯草中重要次生代谢产物含量提供了基础信息。

关键词: 夏枯草; 转录组; 次生代谢; 生物合成; 测序

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2019)05-1220-07

DOI: 10.7501/j.issn.0253-2670.2019.05.028

Transcriptome analysis of *Prunella vulgaris* and identification of putative genes involved in secondary metabolism biosynthesis

ZHU Yun-hao^{1,2}, ZHANG Meng-jia¹, LI Lu¹, ZHAO Le^{1,2}, DONG Cheng-ming^{1,2}

1. School of Pharmacy, Henan University of Traditional Chinese Medicine, Zhengzhou 450046, China

2. Collaborative Innovation Center for Respiratory Disease Diagnosis and Treatment & Chinese Medicine Development of Henan Province, Zhengzhou 450046, China

Abstract: Objective To obtain the transcriptome sequence database and to explore the molecule basis of secondary metabolism biosynthesis in *Prunella vulgaris*. **Methods** The high-throughput RNA-seq technology (Illumina HiSeq 2000) was used to conduct a transcriptomic analysis of *P. vulgaris* ears, stems, and leaves. Unigenes were obtained after assembled by Trinity, and the sequencing results were analyzed with the bioinformatics analysis. **Results** A total of 77 863 Unigenes were obtained and the average length was 716.72 nt. A total of 41 367 (53.13%) unigenes were annotated by a BLAST similarity search against Nr, Swiss-Prot, COG and other four public sequence databases, and 1 406 Unigenes were assigned to secondary metabolism biosynthesis pathways by the KEGG. Based on the bioinformatic analysis, we found that 60 Unigenes were involved in the triterpenes backbone biosynthesis, 24 Unigenes catalyze synthesis of phenolic acids, and 259 Unigenes might participate in secondary metabolism post-modification. Additionally, 118 unigenes might involve in other secondary metabolism biosynthesis of *P. vulgaris*. **Conclusion** The transcriptome data of *P. vulgaris* from this study provides an important resource for understanding the biosynthesis mechanism of its secondary metabolites, and provides basic information that may aid in metabolic engineering to increase yields secondary metabolites of *P. vulgaris*.

Key words: *Prunella vulgaris* L.; transcriptome; secondary metabolism; biosynthesis

夏枯草 *Prunella vulgaris* L. 为唇形科夏枯草属 消肿散结的功效。夏枯草含有丰富的三萜类、酚酸类、黄酮类等多种次生代谢成分, 因其具有抗肿瘤、

收稿日期: 2018-07-09

基金项目: 国家自然科学基金资助项目 (81603232); 国家重点研发计划 (2017YFC1702800); 河南省重大科技专项 (171100310500); 河南中医药大学博士科研基金资助项目 (BSJJ2015-13); 中央引导地方科技发展专项资金

作者简介: 朱昀昊 (1986—), 男, 博士, 讲师, 主要从事药用植物分子生物学研究。E-mail: guxinhan123@163.com

*通信作者 董诚明, 学士, 教授。E-mail: dcm371@sohu.com

抗炎、护肝等多种生物活性而广泛地运用于临床^[1]。同时,夏枯草也是卫计委公布的药食两用中药,由于其使用的安全性和有效性一直以来都是凉茶类饮料的主要原材料之一。目前对夏枯草的研究主要集中在化学成分^[2]、药理活性^[3]、质量评价^[4]和规范化种植技术^[5-6],而对夏枯草活性次生代谢成分分子生物合成相关基因发掘和利用却鲜有报道,仅汝梅^[7]、许锋等^[8]克隆了夏枯草迷迭香酸生物合成途径中迷迭香酸合成酶、酪氨酸氨基转移酶、苯丙氨酸解氨酶等相关基因。

转录组学是研究基因转录及其调控规律的一门学科,其反映了生物组织基因组信息和外部特征的动态联系,可以在整体水平上高效地发掘特定生理功能、生源途径相关的已知或未知的功能基因,对研究特定生物学过程和植物次生代谢途径具有重要意义^[9]。第 2 代测序技术(NGS)是常用的转录组学研究方法之一,其可以同时百万级别的条数进行同时测序,又称为高通量测序技术。高通量测序技术在植物次生代谢途径研究、生物体发育等研究领域发挥着重要的作用。随着测序成本的不断下降和测序通量的不断提升,目前许多重要的药用植物,如丹参^[10]、人参^[11]、地黄^[12]、山茱萸^[13]等都已经完成了转录组测序分析,积累了一批与药用次生代谢产物合成调控相关的基因。本研究利用 Illumina HiSeq 2000 对夏枯草果穗、茎和叶进行了转录组测序,以期获得与其有效成分合成相关的基因信息,为进一步挖掘与克隆夏枯草新的功能基因奠定基础。

1 材料与方法

1.1 材料

本研究所用夏枯草采自河南省确山县夏枯草 GAP 种植基地,并经河南中医药大学董诚明教授鉴定为唇形科植物夏枯草 *Prunella vulgaris* L.。采集新鲜夏枯草叶片、果穗和茎,用水清洗干净后,滤纸吸干水分,置液氮中速冻后,放 -80 °C 超低温冰箱保存。

1.2 RNA 提取与检测

总 RNA 使用植物 RNA 提取试剂盒(康为世纪生物科技有限公司,北京)提取,利用 Agilent 2100 核酸蛋白检测仪(Agilent, Santa Clara, 美国)检测所有样品 RNA 的纯度和浓度。

1.3 RNA 文库构建与质控

文库的构建与测序在北京百迈客生物科技有

限公司完成,具体步骤参照朱昀昊等^[13]的方法。文库构建完成后,分别使用 Qubit2.0 和 Agilent 2100 对文库的浓度和插入片段大小进行检测,使用 Q-PCR 方法对文库的有效浓度进行准确定量,以保证文库质量。样品库检合格后,用 HiSeq2500 进行高通量测序,测序读长为 PE125。

1.4 转录组数据组装

为了使各样品中表达丰度较低的转录本组装得更完整,夏枯草 3 种不同组织测序样品采用合并组装,从而间接增加测序深度,使转录结果更完整,同时也有利于后续的数据分析。对原始数据进行滤过处理,去除接头、不确定碱基、低质量的序列和冗余序列。使用 Trinity 拼接干净读序得到重叠群(Contigs)。在所有的转录本中,取最长的 Contig 作为该转录本的转录本(Unigene)。

1.5 功能注释

使用 BLAST 软件将 Unigene 序列与 NR、Swiss-Prot、GO、COG、KOG、KEGG 数据库比对,预测完 Unigene 的氨基酸序列之后使用 HMMER 软件与 Pfam 数据库比对,获得 Unigene 的注释信息。

2 结果与分析

2.1 测序与组装结果统计

经过测序从夏枯草果穗、茎、叶 3 个样本中共获得 19.16 Gb clean data,各样本的 Q30 碱基比均在 92.78%及以上。说明测序结果准确度较高,可以用于后续分析。利用 Trinity 软件对 3 个夏枯草样品的测序数据进行合并组装。结果显示组装后共得到 7 830 894 个 Contigs 序列,平均长度为 45.17 nt,其中 200~300 nt 序列占总 Contigs 序列数的 99.22%,其分布特征符合 Illumina 测序的预期结果(表 1),可为后续的数据组装提供原始数据。测序数据组装共得到 185 478 条 transcript 和 77 863 条 Unigene;其中长度在 1 kb 以上的 Unigene 有 15 443 条,这些基因可作为后续实验研究的重要对象;序列长度集中于 200~2 000 nt,Unigene 平均长度为 716.72 nt。大于 2 000 nt 的序列占总 Unigene 序列的 8.09%。

2.2 基因功能注释

通过选择 BLAST 参数 E 值不大于 1×10^{-5} 和 HMMER 参数 E 值不大于 1×10^{-10} ,最终共有 41 367 个 Unigene 获得注释,注释率为 53.13%。其中有 13 081 个(16.80%)Unigene 注释到 COG 数据库中;21 563 个(27.69%)Unigene 注释到 GO 数据库中;8 882 个(21.47%)Unigene 注释到 KEGG 数据库

中; 21 531 个 (11.41%) Unigene 注释到 KOG 数据库中; 24 412 个 (31.35%) Unigene 注释到 Pfam 数据库中; 25 797 个 (33.13%) Unigene 注释到 Swiss-Prot 数据库中; 41 079 个 (52.76%) Unigene 注释到 Nr 数据库中。由于夏枯草缺乏属基因组、EST 和蛋白质序列等信息, 仍有 36 496 个 (53.13%) Unigene 不能被已有数据库注释 (表 2)。

2.3 GO 功能分类

在本研究中, 共有 21 563 条 Unigene (27.69%)

得到了 GO 注释 (图 1)。涉及分子功能的 Unigene 有 32 806 条, 其中具有催化活性 (catalytic activity) 和结合 (binding) 的 Unigene 数量最多; 与细胞组成相关的 Unigene 有 17 791 条, 其中细胞组分中细胞组分和细胞 Unigene 数量最多, 分别是 8 763 和 8 720 条; 与生物学过程相关的 Unigene 有 40 702 条, 其中代谢过程、细胞过程和单生物过程 Unigene 数量最多, 分别有 15 101、12 113 和 10 465 条。

表 1 夏枯草转录组组装结果

Table 1 Summary of Illumina transcriptome assembly for *P. vulgaris*

长度范围/bp	Contigs (占比/%)	transcript (占比/%)	Unigene (占比/%)
200~300	7 770 119 (99.22)	36 669 (19.77)	30 426 (39.08)
300~500	27 208 (0.35)	28 243 (15.23)	19 223 (24.69)
500~1 000	17 138 (0.22)	30 612 (16.50)	12 771 (16.40)
1 000~2 000	10 551 (0.13)	43 279 (23.33)	9 143 (11.74)
2 000+	5 878 (0.08)	46 675 (25.16)	6 300 (8.09)
总数	7 830 894	185 478	77 863
总长度	353 734 982	250 314 475	55 806 062
N50 长度	45	2 247	1 331
平均长度	45.17	1 349.56	716.72

表 2 基因功能注释分类

Table 2 Classification of functional annotation of *P. vulgaris* unigene

数据库	注释基因数/个	占总基因数的百分比/%	基因数 ≥ 300 nt	基因数 ≥ 1 000 nt
COG	13 081	16.80	9 812	6 073
GO	21 563	27.69	14 984	7 676
KEGG	8 882	21.47	6 498	3 343
KOG	21 531	11.41	16 143	8 798
Pfam	24 412	31.35	19 939	12 384
Swiss-Prot	25 797	33.13	20 615	11 318
Nr	41 079	52.76	29 465	14 420
All	41 367	53.13	29 574	14 438
Total	77 863	100.00	47 437	15 443

2.4 COG 和 KOG 分类

在本研究中, 共有 13 081 条 Unigene 与 COG 数据库具有同源性, 这些 Unigene 对应 COG 数据库的 25 个功能类别 (表 3)。其中一般功能预测注释结果最多, 有 3 426 条 Unigene; 其次是复制、重组和修复, 有 1 586 条; 细胞运动和核结构中的 Unigene 数量最少, 分别有 16 条和 3 条, 没有发现胞外结构类基因; 有 691 条 Unigene 被注释到次生

代谢产物合成、转运和代谢中。

在 KOG 注释中, 共有 21 531 条 Unigene 获得了注释, 得到了 25 个分类, 其中一般功能预测注释结果最多, 有 4 367 条 Unigene; 其次是翻译后修饰、蛋白质转换、伴侣蛋白信号, 有 2 359 条 Unigene; 细胞运动和核结构所占 Unigene 最少, 分别只有 9 条和 81 条; 其中有 1 203 条 Unigene 被注释到次生代谢生物合成、转运和分解中。

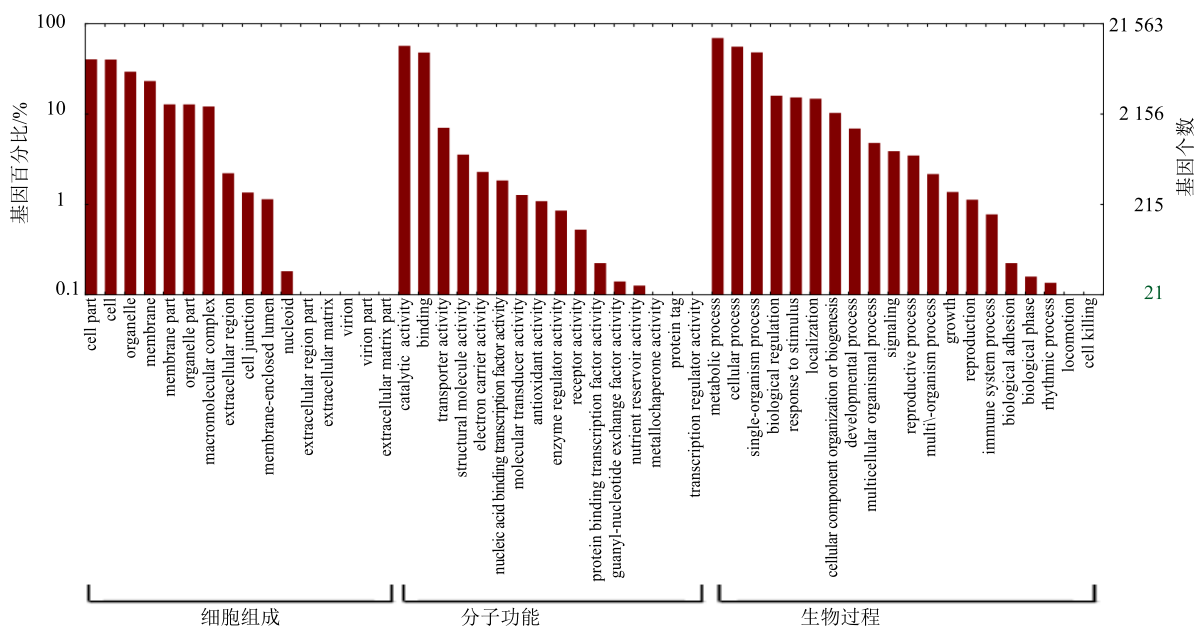


图 1 基因的 GO 功能注释

Fig. 1 Genetic GO annotation

表 3 夏枯草 unigene 的 COG 和 KOG 分类

Table 3 COG and KOG classification of *P. vulgaris* unigene

功能分类	COG	KOG
翻译、核糖体结构和生物发生	1 283	1 340
RNA 加工和修饰	1 43	760
转录	1 468	1 144
复制、重组和修复	1 586	542
染色质结构和活力	90	275
细胞周期控制、细胞分裂、染色体分区	325	492
核结构	3	81
防卫机制	278	297
信号转导机制	1 263	2 294
细胞壁膜生物发生	440	364
细胞运动	16	9
细胞骨架	232	545
胞外结构	0	238
细胞内转运、分泌和小泡运输	217	1 011
翻译后修饰、蛋白质折叠和分子伴侣	1 154	2 359
能量生成和转换	1 142	1 426
碳水化合物运输和代谢	1 084	1 337
氨基酸运输和代谢	1 110	1 039
核苷酸运输和代谢	212	239
辅酶运输和代谢	383	253
脂质运输和代谢	676	1 024
无机离子运输和代谢	700	657
次生代谢物合成、运输和代谢	691	1 203
一般功能预测	3 426	4 367
未知功能	495	991

2.5 KEGG 注释

KEGG 的注释结果描绘了 8 882 条 Unigene 所归属的 119 条代谢通路，有 2 969 条 Unigene 被注释到代谢途径 (metabolic pathways)，1 406 条 Unigene 被注释到次生代谢产物合成 (biosynthesis of secondary metabolites)，另外咖啡因代谢、油菜素生物合成、硫甙生物合成、其他类型 O 型聚糖生物合成、甜菜红碱生物合成和亚油酸代谢途径的基因最少，分别有 4、4、3、2、2 和 1 条。

2.6 夏枯草次生代谢生物合成途径相关酶的鉴定

2.6.1 三萜类

夏枯草是我国常用的中药材，其富含的三萜类、有机酸类、苯丙素类、黄酮类等多种类型的次生代谢产物可能是其发挥药理作用的物质基础^[3]。夏枯草中含有丰富的三萜类化合物，其中熊果酸和齐墩果酸的含量较高。植物中三萜类物质的生物合成的前体是由甲羟戊酸途径 (MVA) 和甲基赤藓糖醇磷酸化途径 (MEP) 共同合成的异戊烯二磷酸 (IPP) 或其异构体 3,3-二甲基稀丙基焦磷酸 (DMAPP)^[14]。在夏枯草转录组数据中，共发现 38 个 Unigene 可能编码 MVA 途径中乙酰辅酶 A 酰基转移酶 (AACT)、羟甲基戊二酰辅酶 A 合酶 (HMGS)、3-羟基-3-甲基戊二酰辅酶 A 还原酶 (HMGR) 等 6 种酶，22 个 Unigene 可能编码 MEP 途径中的 1-脱氧-D-木酮糖-5-磷酸合酶

(DXS)、1-脱氧-D-木酮糖-5-磷酸还原异构酶(DXR)等7种酶。

香叶基二磷酸、法尼基二磷酸、鲨烯、2,3-环氧鲨烯是三萜骨架生物合成中的一系列不同 C5 单元组成的中间体^[15]。催化上述中间体产生所需的酶,在夏枯草的转录组数据中均找到了编码相应酶

的候选基因,其中包括1个香叶基二磷酸合酶(GPPS)、4个法尼基二磷酸合酶(FPP)、4个鲨烯合酶(SS)、10个鲨烯环氧酶(SQE)、6个 β -香树脂合成酶(β -AS)等。

2.6.2 酚酸类 迷迭香酸具有抗病毒、抗炎、抗抑郁活性等药理活性,是夏枯草中的有效成分之一^[4],

表 4 夏枯草三萜和酚酸类生物合成相关候选基因

Table 4 Candidate genes involved in triterpenes and phenolic acids biosynthesis identified in *P. vulgaris*

简称	英文全称	Unigene 数目	酶编号
AACT	acetyl-CoA acetyltransferase	8	2.3.1.9
HMGS	3-hydroxy-3-methylglutaryl-coenzyme A synthase	11	2.3.3.10
HMGR	3-hydroxy-3-methylglutaryl-coenzyme A reductase	14	1.1.1.34
MK	mevalonate kinase	1	2.7.1.36
PMK	phosphomevalonate kinase	1	2.7.4.2
MDC	diphosphomevalonate decarboxylase	3	4.1.1.33
DXS	1-deoxy-D-xylulose-5-phosphate synthase	12	2.2.1.7
DXR	1-deoxy-D-xylulose-5-phosphate reductoisomerase	2	1.1.1.267
CMS	4-diphosphocytidyl-2-C-methyl-D-erythritol synthase		2.7.7.60
CMK	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	1	2.7.1.148
MDS	2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase	2	4.6.1.12
HDS	1-hydroxy-2-methyl-butenyl-4-diphosphate synthase	1	1.17.7.3
HDR	1-hydroxy-2-methyl-butenyl-4-diphosphate reductase 4-hydroxy-3-methylbut-2-enyl diphosphate reductase	4	1.17.1.2
IDI	isopentenyl diphosphate isomerase	2	5.3.3.2
FPS	farnesyl diphosphate synthase	4	2.5.1.10
GPPS	geranyl diphosphate synthase	1	2.5.1.1
GGPPS	geranylgeranyl pyrophosphate synthase	10	2.5.1.10
SS	squalene synthase	4	2.5.1.21
SQE	squalene monooxygenase	10	1.14.13.132
β -AS	beta-amyrin synthase	6	5.4.99.39
PAL	phenylalanine ammonialyase	8	4.3.1.5
C4H	cinnaamiaeicid-4-hydroxylase	3	1.14.13.11
4CL	hydroxycinnamate: coenzyme A ligase	9	6.2.1.12
TAT	tyrosine aminotransferase	4	2.6.1.5
HPPR	hydroxyphenylpyruvate reductase	5	1.1.1.237
RAS	hydroxycinnamoyl-CoA: hydroxyphenyllactate hydroxycinnamoyl transferase	2	2.3.1.140

也是《中国药典》2015年版中夏枯草药材的质控指标。另外,迷迭香酸也是许多其他次生代谢物的前体物质,在夏枯草植物次生代谢网络中具有重要的作用。植物中迷迭香酸的生物合成途径已经基本清晰,苯丙素代谢途径和酪氨酸代谢途径共同参与了迷迭香酸的合成和调控^[16]。在夏枯草转录组数据中,共注释得到24条Unigene可能编码迷迭香酸的生物合成途径的7种关键酶,其中包括8个苯丙氨酸解氨酶(PAL)、3个肉桂酸-4-羟化酶(C4H)、9个4-香豆素辅酶A连接酶(4CL)、4个酪氨酸氨基转移酶(TAT)5个对苯基丙酮酸还原酶(HPPR)和2个迷迭香酸合成酶(RAS)。

2.6.3 次生代谢后修饰酶 次生代谢产物骨架形成后需要经过母核的氧化、糖基化等后修饰的反应才能形成结构各异的成分,这些后修饰反应也使次生代谢产物的种类和结构更加多样。细胞色素P450(cytochrome P450, CYP450)主要参与次生代谢产物的氧化反应,糖基转移酶(glycosyltransferases, GT)催化糖基转移反应,两者在次生代谢产物的后修饰中起着重要的作用^[17]。通过搜索夏枯草转录组数据在Swiss-Prot数据库的注释结果,共找到259条Unigene被注释为CYP450,118条Unigene被注释为UGT。

2.6.4 其他次生代谢产物 夏枯草中还含有黄酮

类、苯丙素类、香豆素类、多糖类等化合物和内源性激素等生长调节物质,根据 KEGG 代谢通路分析结果,有 15 条代谢通路中 534 条 Unigene 可能参与夏枯草其他次生代谢途径。其中苯丙素类生物合成途径所占比例最大,共有 110 条 Unigene 被映射到该途径,甜菜红素和硫代葡萄糖苷生物合成途径所占比例最小,分别仅有 2 条、3 条 Unigene。

2.7 转录因子分析

转录因子可激活或抑制植物次生代谢产物生物合成途径中功能基因的表达,从而有效调控次生代谢产物合成积累。根据基因序列比对到 Pfam 数据库,参照各种转录因子的隐马氏模型文件,利用 HMMER 3.0 软件对夏枯草转录组数据进行搜索。结果显示夏枯草转录组数据预测共有 1 449 个 Unigene 被注释为转录因子,分属于 60 种转录因子类型。最多的转录因子类型是 MYB 类 (158 个),约占 10.90%;其次是 AP2/ERF 类 (共 113 个),约占 7.8%;锌指蛋白 C2H2 类 (97 个),约占 6.7%;GRAS 类 (96 个),约占 6.63%;bHLH 类 (84 个),

表 5 夏枯草其他与次生代谢产物相关的 Unigene 数量

Table 5 Number of unigene related to secondary metabolites in *P. vulgaris*

代谢途径	编号	覆盖到该通路的基因数目
甜菜红素生物合成	ko00965	2
硫代葡萄糖苷生物合成	ko00966	3
油菜素内酯生物合成	ko00905	4
黄酮、黄酮醇生物合成	ko00944	7
芪类、二芳基庚烷、姜酚生物合成	ko00945	19
二萜生物合成	ko00904	25
柠檬烯、蒎烯降解	ko00903	27
异喹啉生物碱生物合成	ko00950	29
茛菪烷、吡啶、吡啶生物碱的生物合成	ko00960	30
类黄酮生物合成	ko00941	31
玉米素生物合成	ko00908	38
类胡萝卜素生物合成	ko00906	49
N-聚糖生物合成	ko00510	73
萜类骨架生物合成	ko00900	87
苯丙素类生物合成	ko00940	110

约占 5.8%; WRKY 类 (85 个),约占 5.87%; NAC 类 (69 个),约占 4.76%和 bZIP 类 (56 个),约占 3.86%。

3 讨论

夏枯草作为著名的药食两用中药材,次生代谢

产物的安全性和复杂性是其能够被广泛应用的物质基础然而由于夏枯草富含的三萜类、酚酸类等次生代谢产物的生物合成途径极其复杂^[1,3],因此对夏枯草次生代谢生物合成调控相关功能基因的挖掘仍处于较低的水平。目前,药用植物次生代谢合成生物学研究的进展较为缓慢,主要是由于对药用植物次生代谢生物合成途径调控网络的了解较少,功能基因的克隆、表达及功能分析仍较为滞后。近年来,基于高通量测序的转录组学技术发展迅速,也为药用植物次生代谢生物合成调控基因信息发掘提供了前所未有的机遇^[18]。夏枯草三萜是其次生代谢产物的重要组成部分,而其含量和组分又主要取决于生物合成关键酶以及在细胞中的表达水平。在夏枯草转录组的 KEGG 注释中,共筛选到 60 个参与三萜合成上游部分基因,分属于 MVA 和 MEP 途径的各个环节。在夏枯草酚酸类的生物合成过程中,苯丙素代谢途径和酪氨酸代谢途径共同参与了迷迭香酸的合成和调控,通过对夏枯草转录组数据的搜索,共得到 24 条 Unigene 可能编码迷迭香酸的生物合成途径的 7 种关键酶。

次生代谢产物合成下游环节还包括母核的结构修饰以及糖基化等步骤,细胞色素 (P450) 和糖基转移酶 (UGTs) 分别是这些关键步骤的催化酶。P450s 通过氧化反应 (如羟基、酮基、羧基和双键等不饱和基团) 在不同类型次生代谢产物母核的不同位置和不同数量进行结构修饰^[19],极大地提高了所生成次生代谢产物结构的多样性、生物活性。UGT 催化糖基单元到修饰过的次生代谢产物母核上,形成苷类化合物,而使其具有水溶性^[20]。植物体内 P450 基因类及 UGT 基因家族数据庞大、功能十分复杂,因此对其功能的研究不够透彻而难以准确鉴定,在基于同源性上寻找专属反应的调控基因上仍是不小的难题,鉴定和验证这些基因需要做的工作还有很多。本研究通过对夏枯草进行转录组测序,极大的丰富了夏枯草的基因资源,为更深入的研究夏枯草次生代谢合成及调控提供基础数据,也有望推动夏枯草次生代谢相关基因的研究。

参考文献

- [1] 柏玉冰,李春,周亚敏,等. 夏枯草的化学成分及其三萜成分的抗肿瘤活性研究 [J]. 中草药, 2015, 46(24): 3623-3629.
- [2] 贾晓斌,刘光敏,陈彦,等. 基于中药组分结构理论

- 的夏枯草属药材防治肺癌物质基础研究 [J]. 中药材, 2010, 33(7): 1105-1109.
- [3] 封亮, 贾晓斌, 陈彦, 等. 夏枯草化学成分及抗肿瘤机制研究进展 [J]. 中华中医药杂志, 2008, 23(5): 428-434.
- [4] 冯伟红, 李春, 信伟梅, 等. 生物测定法用于中药质量评价的探索研究——以夏枯草抗氧化活性与总酚酸含量相关性的研究为例 [J]. 中国中药杂志, 2016, 41(14): 2660-2668.
- [5] 杨伟, 刘丽, 郭巧生, 等. 不同居群夏枯草产量构成因子与形态特征关系分析 [J]. 中国中药杂志, 2013, 38(21): 3666-3671.
- [6] 于曼曼, 刘丽, 郭巧生, 等. 氮素不同形态对比对夏枯草苗期生长及光合特性的影响 [J]. 中国中药杂志, 2011, 36(5): 530-534.
- [7] 汝梅. 夏枯草迷迭香酸生物合成途径酪氨酸支路关键酶基因的克隆与功能研究 [D]. 北京: 中国科学院教育部水土保持与生态环境研究中心, 2017.
- [8] 许锋, 曹腾, 宁迎晶, 等. 夏枯草苯丙氨酸解氨酶基因的克隆与表达分析 [J]. 华北农学报, 2012, 27(1): 39-44.
- [9] 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.
- [10] 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]. 药学学报, 2010, 45(4): 524-529.
- [11] Jayakodi M, Lee S C, Lee Y S, *et al.* Comprehensive analysis of *Panax ginseng* root transcriptomes [J]. *BMC Plant Biol*, 2015, 15(1): 138.
- [12] Wang F, Zhi J, Zhang Z, *et al.* Transcriptome analysis of salicylic acid treatment in *Rehmannia glutinosa* hairy roots using RNA-seq technique for identification of genes involved in acetoside biosynthesis [J]. *Front Plant Sci*, 2017, 8: 787.
- [13] 朱昀昊, 董诚明, 郑晓珂, 等. 基于转录组测序的山茱萸次生代谢生物合成相关基因的挖掘 [J]. 中国中药杂志, 2017, 42(2): 213-219.
- [14] Cheng A X, Lou Y G, Mao Y B, *et al.* Plant Terpenoids: Biosynthesis and Ecological Functions [J]. *J Integr Plant Biol*, 2007, 49(2): 179-186.
- [15] 张长波, 孙红霞, 巩中军, 等. 植物萜类化合物的天然合成途径及其相关合酶 [J]. 植物生理学通讯, 2007, 43(4): 779-786.
- [16] Petersen M, Abdullah Y, Benner J, *et al.* Evolution of rosmarinic acid biosynthesis [J]. *Phytochemistry*, 2009, 69(15): 1663-1679.
- [17] 李军玲, 罗晓东, 赵沛基, 等. 植物萜类生物合成中的后修饰酶 [J]. 云南植物研究, 2009, 31(5): 461-468.
- [18] Xiao M, Zhang Y, Chen X, *et al.* Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest [J]. *J Biot*, 2013, 166(3): 122.
- [19] Mizutani M, Sato F. Unusual P450 reactions in plant secondary metabolism [J]. *Arch Biochem Biophys*, 2011, 507(1): 194-203.
- [20] Gachon C M, Langlois-Meurinne M, Saindrenan P. Plant secondary metabolism glycosyltransferases: the emerging functional analysis [J]. *Trends Plant Sci*, 2005, 10(11): 542-549.