

北苍术转录组分析及倍半萜类生物合成关键基因挖掘

赵建华, 赵重阳, 孙成振, 史凤玉, 陈丽娜, 郑金双*

河北科技师范学院农学与生物科技学院, 河北秦皇岛066004

摘要: 倍半萜类化合物是北苍术主要活性成分及质量控制指标成分。为研究北苍术倍半萜类化合物生物合成通路, 利用Illumina HiSeq™ 2000 150PE高通量测序平台对北苍术根茎进行转录组测序, 得到8.86 G数据, 59 065 569条高质量reads。Trinity组装获得50 330条条带, 平均长度1 214 nt。所有条带在NR、NT、Pfam、KOG、Swiss-prot、KEGG、GO数据库中得到注释, 按GO功能分类将其归为3大类53个分支; KEGG分析发现共有11 235条条带参与130个KEGG标准代谢通路, 421条条带参与21个次生代谢标准通路, 其中14条条带参与MVA途径关键酶的生物合成, 12条条带参与MEP途径关键酶的生物合成。实时荧光定量PCR (qRT-PCR)揭示倍半萜类合成关键基因*FPPS*、*HMGR*和*DXS*在北苍术不同器官中均有表达, 在茎、根茎、花和越冬芽中相对表达量较高, 且*HMGR*、*DXS*基因的相对表达量高于*FPPS*基因。本文为北苍术倍半萜类生物合成途径和分子调控机制的研究提供参考。

关键词: 北苍术; 转录组; 代谢通路; 倍半萜; qRT-PCR

北苍术(*Atractylodes chinensis*)为菊科苍术属多年生草本植物, 其干燥根茎是药材苍术的重要来源, 被《神农本草经》列为中品, 性温, 味辛、苦。现代药理研究表明, 北苍术的主要化学成分为苍术素、茅术醇、 β -桉油醇等倍半萜类化合物(武子敬等2010; 李万娟等2016), 具燥湿健脾、祛风散寒、明目等功效。

目前, 关于北苍术分子生物学方面的研究鲜见报道, 仅见利用DNA条形码和分子标记技术进行苍术类药材鉴别及其亲缘关系研究(邹小兴等2009; 邵婧等2015)。转录组测序技术是功能基因组研究的一个重要组成部分, 是无参考基因组植物分子机理研究的重要手段, 能够快速、准确地提供大量转录本信息, 获得有效的基因编码序列, 有助于从整体水平上揭示植物生长发育、次生代谢及生理适应的转录调控规律(陈士林等2012)。开展无参考基因组药用植物转录组研究, 对阐明药用植物发育机制、SSR分子标记开发、次生代谢产物生物合成途径的基础理论研究具有重要意义。目前已获得西洋参(Sun等2010)、人参(Chen等2011)、甘草(Ramilowski等2013)等药用植物的转录组数据, 并在药用植物自身发育机制研究(闻静等2019)、关键基因挖掘(李欢等2018)和SSR分子标记开发(贺润丽等2018)等方面广泛应用。研究表明, 植物中萜类化合物的合成途径有两条: (1)

甲羟戊酸(mevalonic acid, MVA)途径, 位于细胞质中, 羟甲基戊二酰辅酶A还原酶(3-hydroxy-3-methylglutaryl coenzyme A reductase, HMGR)是第一个限速酶(Lichtenthaler等1997), 法尼基焦磷酸合酶(farnesyl pyrophosphate synthase, FPPS)是催化生成倍半萜类的关键酶(Ahmed等2016), 催化生成法尼基二磷酸(farnesyl diphosphate, FPP) (Vranová等2013), 为倍半萜类化合物提供C15骨架; (2) 2-C-甲基-D-赤藓糖醇-4-磷酸(2-C-methyl-D-erythritol-4-phosphate, MEP)途径, 位于质体中, 1-脱氧-D-木酮糖-5-磷酸合酶(1-deoxy-D-xylose-5-phosphate synthetase, DXS)是第一个关键限速酶(Estévez等2001; Schwender等2001; Cordoba等2009)。

倍半萜类化合物含量的提高对苍术药材质量具有重要意义, 欲提高北苍术根茎中倍半萜类化合物含量, 主要依赖于对其生物合成分子机理的解析。由于北苍术基因组和转录组信息缺乏, 阻碍了倍半萜类化合物生物合成分子机理的研究进程, 严重制约了北苍术种质资源利用与育种工作。本研究利用高通量测序技术开展三年生北苍术根茎

收稿 2019-12-19 修定 2020-05-27

资助 河北省高等学校科学技术研究项目(QN2019162)、河北省自然科学基金(H201907047)和河北省重点研发计划(19226354D)。

* 通讯作者(jinshuangk@163.com)。

转录组分析, 以期揭示北苍术转录组的表达特征, 挖掘倍半萜类化合物生物合成关键基因, 探究其时空表达情况, 为北苍术倍半萜类生物合成途径和分子调控机制研究提供参考。

1 材料与方法

1.1 材料

2014年10月, 本试验用北苍术种子来源于3个不同产地: 产地1 (秦皇岛市同盛医药有限公司基地)、产地2 (承德市围场县)、产地3 (承德市宽城县), 经南京农业大学郭巧生教授鉴定为菊科植物北苍术[*Atractylodes chinensis* (DC.) Koidz]的种子。种子于2015年4月初播于河北科技师范学院试验农场, 至2017年11月取产地1三年生北苍术根茎, 液氮速冻后, 置于-80℃超低温冰箱保存, 用于转录组测序。

1.2 转录本组装与分析

三年生北苍术根茎转录组高通量测序与转录本组装参考李依民等(2018)的方法。组装获得的高质量条带的功能注释、转录因子预测、简单重复序列标记(simple sequence repeats, SSR)分析参考熊阳阳等(2019)的方法。

1.3 倍半萜类化合物生物合成基因挖掘

根据文献报道的植物倍半萜类化合物生物合成途径, 结合KEGG注释结果和数据库中已知的基因信息, 利用BLAST进行比对, 确定北苍术转录组数据库中与倍半萜类化合物生物合成相关的基因(彭亮等2018)。

1.4 倍半萜类生物合成关键基因表达情况分析

分别取3个不同产地北苍术的三年生植株的茎、叶、花、根茎、越冬芽, 置于液氮中, 提取RNA, 经反转录获得cDNA, 依据HMGR、FPPS (桑晓华等2017)和DXS (邓娟等2017)基因序列设计引物进

行实时荧光定量PCR (real time quantitative PCR, qRT-PCR), 内参基因为UBQ2 (桑晓华等2017), 反应体系和程序参照桑晓华等(2017)的方法。基因表达情况采用相对定量法, 参照2^{-ΔΔCt}法进行分析。引物序列见表1。

2 实验结果

2.1 北苍术转录组组装与质量分析

三年生北苍术根茎高通量转录组测序共产生59 065 569条高质量reads, 包含8.86 G核苷酸信息, Q₂₀ (碱基量≥20%)和Q₃₀ (碱基量≥30%)分别为97.09%和92.14%, GC含量为45.55%, 测序质控良好, 高质量条带质量合格。Trinity组装得到50 330条条带, 最长的15 540 nt, 最短的301 nt, 平均1 214 nt, N₅₀ (拼接条带由长到短排列, 相加之和占总基因组大小50%时的条带长度)为1 748 bp, N₉₀ (拼接条带由长到短排列, 相加之和占总基因组大小90%时的条带长度)为544 bp。长度为501~1 000 nt的条带数量最多, 有49 202条; 其次是1 001~2 000 nt长度的条带, 有48 602条; 301~500 nt长度的条带有42 135条; 长度高于2 000 nt的条带有31 447条。

2.2 北苍术转录组中条带的功能注释

获得注释信息最多的是NR数据库, 有24 294条(48.26%)条带, 注释信息最少的是KOG数据库, 有6 779条(13.46%), 有3 270条条带在7个数据库同时得到注释, 占全部条带的6.49% (表2)。

在NR数据库中, 条带注释的物种分布如图1所示, 相似条带匹配度较高的物种中, 刺苞菜蓟(*Cynara cardunculus*)所占比例最高为37.50%, 其次为莴苣(*Lactuca sativa*) 29.30%, 向日葵(*Helianthus annuus*) 17.40%, 甜菜(*Beta vulgaris*)和葡萄(*Vitis vinifera*)均占0.90%。

表1 qRT-PCR所用引物序列
Table 1 Primer sequences for qRT-PCR

基因	正向引物序列(5'→3')	反向引物序列(5'→3')
UBQ2	GGTTGAGGGGAGGAATGC	AGACGAAGGACAAGGTGA
HMGR	AAACAACATAAATCCGGCGAACT	AGGCGTCGATGGCTTTTG
DXS	CTCTTTGATTGGGTCTGGG	GTTATGGCACCATTTCAGTCT
FPPS	CTGCCCTTGGTTGGTGCGTT	TTCTTGGGACATGGTTGCGA

表2 北苍术转录组条带功能注释情况统计

Table 2 Summary of functional annotation of *A. chinensis* transcriptomic unigenes

数据库	注释成功条带数量/条	注释成功基因比例/%
NR数据库注释	24 294	48.26
NT数据库注释	17 225	34.22
KO数据库注释	11 235	22.32
SwissProt数据库注释	20 682	41.09
PFAM数据库注释	18 047	35.85
GO数据库注释	18 047	35.85
KOG数据库注释	6 779	13.46
在所有数据库均被注释	3 270	6.49
在至少一个数据库被注释	31 160	61.91

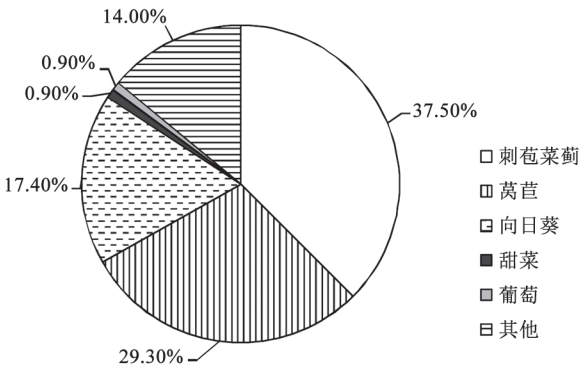


图1 北苍术转录组条带与NR数据库匹配物种分布
Fig.1 Distribution of the matching species of *A. chinensis* transcriptome in NR database

对北苍术条带进行GO数据库基因功能分类, 共有18 047条条带注释到生物过程、细胞组分和

分子功能3个类别53个功能组(图2)。分子功能中丰度较高的是结合, 有9 819条; 其次是催化活性, 有7 267条; 细胞组分中细胞和细胞要素均有4 522条; 其次是细胞器有2 987条; 生物过程组分中细胞过程丰度最高, 有9 407条; 其次是代谢过程, 有8 824条。

从图3可以发现, 共有6 779条条带被注释到25种KOG分类中。注释为翻译后修饰、蛋白转换和分子伴侣的条带最多, 有1 021条; 注释为细胞运动性的条带数量最少, 仅2条; 其他类型的基因表达丰度各不相同。

通过KEGG对北苍术条带的代谢途径及其功能进行分析发现, 共有11 235条条带得到功能注释, 涉及5条主通路19条子通路, 以翻译的条带数量最多有1 020条; 其次为折叠、分类和降解有879条;

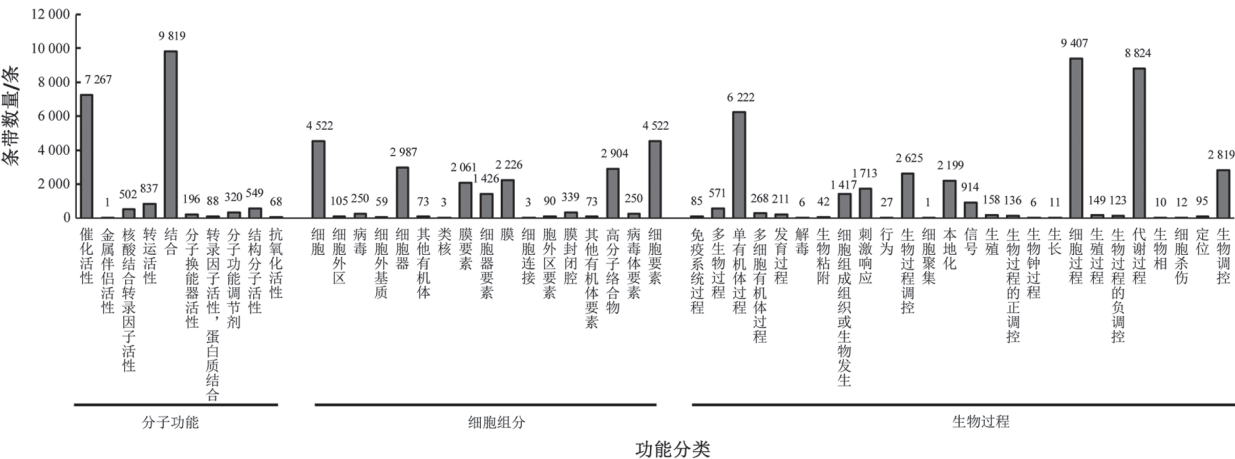


图2 北苍术转录组的GO分类
Fig.2 GO classification of *A. chinensis* transcriptomic unigenes

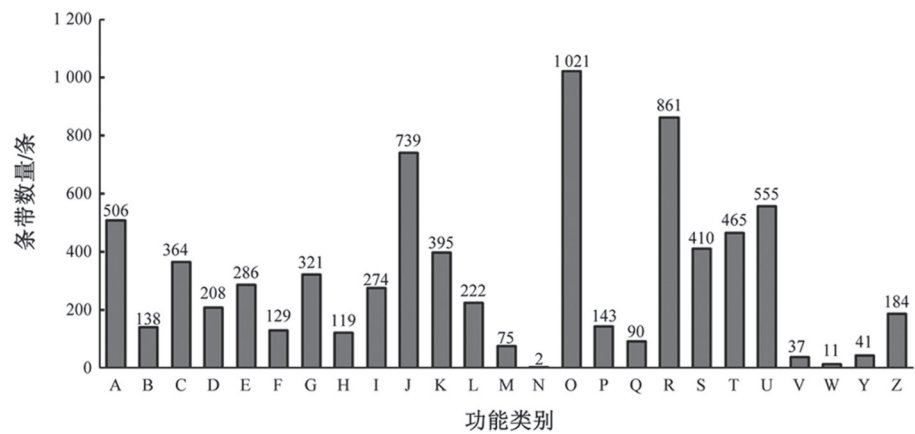


图3 北苍术转录组的KOG分类
Fig.3 KOG classification of *A. chinensis* transcriptomic unigenes

A: RNA加工与修饰; B: 染色体结构与动态; C: 能量产生与转换; D: 细胞周期控制、细胞分裂、染色体分区; E: 氨基酸运输与代谢; F: 核苷酸运输与代谢; G: 糖类运输与代谢; H: 辅酶运输与代谢; I: 脂类运输与代谢; J: 翻译、核糖体结构和生物起源; K: 转录; L: 复制、重组和修复; M: 细胞壁/膜/囊膜生物起源; N: 细胞运动性; O: 翻译后修饰、蛋白转运、分子伴侣; P: 无机离子运输与代谢; Q: 次生代谢产物生物合成; R: 一般功能预测; S: 功能未知; T: 信号传导机制; U: 胞内运输、分泌、囊泡运输; V: 防御机制; W: 胞外结构; Y: 核结构; Z: 细胞骨架。

碳水化合物代谢的有877条; 膜运输的数量最少, 仅56条(图4)。KEGG代谢通路分析表明, 有11 235条条带涉及130个KEGG标准代谢通路, 其中条带数量大于150条的代谢通路有18条(表3)。

2.3 编码区和转录因子分析
对北苍术转录组条带的编码区(CDS)进行分析, 通过Blast比对获得25 835个CDS序列, 利用ESTscan分析获得23 791个序列。转录因子预测分

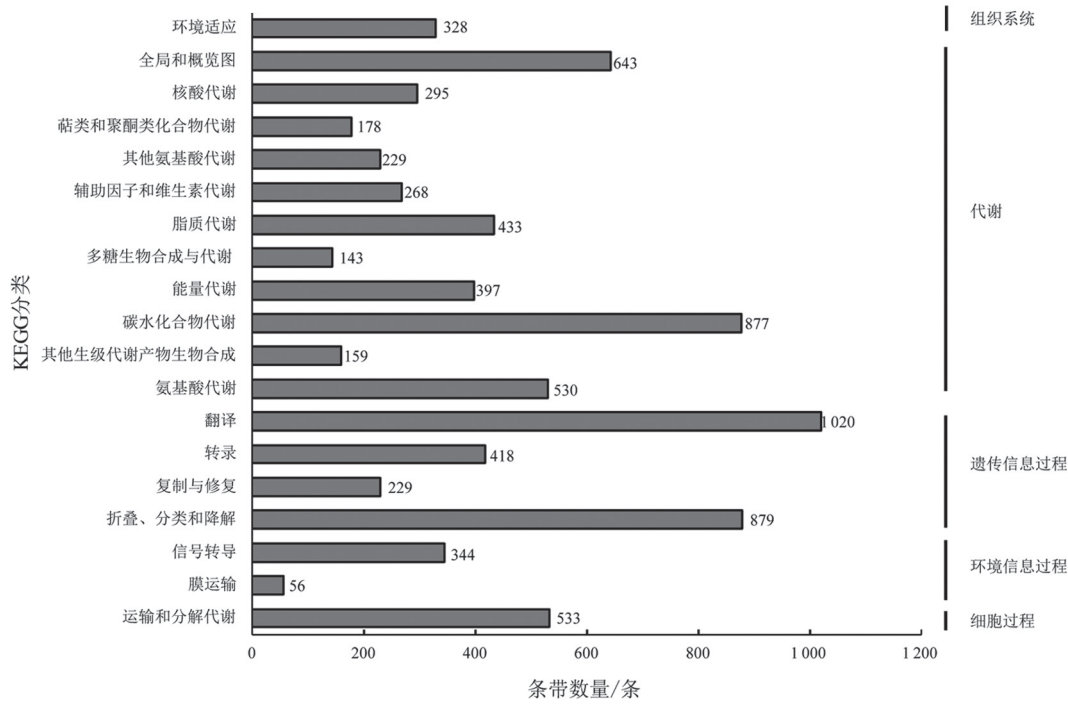


图4 北苍术转录组的KEGG分类
Fig.4 KEGG classification of *A. chinensis* transcriptomic unigenes

表3 北苍术转录组条带 KEGG通路分析

Table 3 KEGG pathway analysis of *A. chinensis* transcriptomic unigenes

编号	代谢通路	条带数量/条	比例/%	通路ID
1	核糖体	453	4.03	ko03010
2	碳代谢	383	3.41	ko01200
3	内质网蛋白质加工	337	3.00	ko04141
4	氨基酸合成	324	2.88	ko01230
5	剪接体	305	2.71	ko03040
6	内吞作用	292	2.60	ko04144
7	植物病原体互作	278	2.47	ko04626
8	植物激素信号转导	245	2.18	ko04075
9	RNA转运	245	2.18	ko03013
10	嘌呤代谢	239	2.12	ko00230
11	淀粉和蔗糖代谢	234	2.08	ko00500
12	泛素介导的蛋白水解	199	1.77	ko04120
13	RNA降解	185	1.65	ko03018
14	mRNA监测通路	183	1.63	ko03015
15	糖酵解/糖异生	180	1.60	ko00010
16	氧化磷酸化	173	1.54	ko00190
17	氨基酸和核苷酸糖代谢	162	1.44	ko00520
18	嘧啶代谢	152	1.35	ko00240

析见图5,发现包含76个家族成员,以AP2-EREBP、MYB、C3H和WRKY占主体。

2.4 SSRs特征分析

对北苍术转录组50 330条条带进行SSRs分析,共检测到13 096个SSRs (表4)。结果表明,单碱基重复SSRs数量最丰富有6 329个(48.33%),其中A/T类的数量最多,有2 715个;其次为二碱基重复,有4 326个(33.03%),其中AG/CT类的数量最多,有605个;三碱基重复有2 008个(15.33%),其中ACC/CCT类的数量最多,有279个;四碱基重复、五碱基重

复和六碱基重复相对较少,分别为146个(1.11%)、156个(1.19%)和131个(1.00%),重复单元数高于9的数量极少。

2.5 倍半萜类化合物生物合成途径关键基因的鉴定

KEGG代谢通路分析发现421条条带参与黄酮类、生物碱类、萜类等生物合成的21个次生代谢标准通路(表5)。其中苯丙烷生物合成代谢通路条带数量最多(87条);其次为萜类骨架生物合成(65条)。北苍术根茎中主要活性成分为倍半萜类化合物,在北苍术转录组中与萜类骨架生物合成相关的

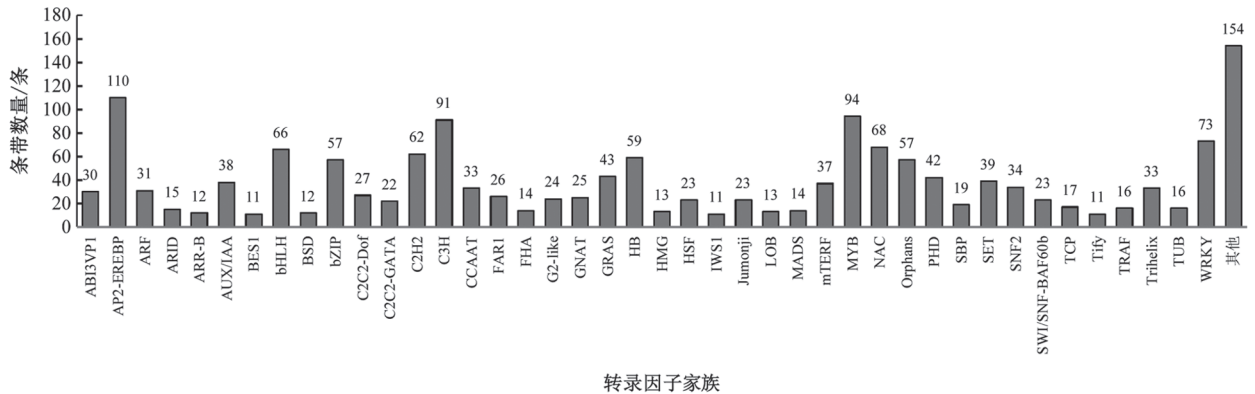


图5 北苍术转录组条带的转录因子分析

Fig.5 Transcription factor classification of *A. chinensis* transcriptomic unigenes

表4 北苍术转录组中条带的SSRs分布

Table 4 SSRs distribution of *A. chinensis* transcriptomic unigenes

重复	重复单位数量/个											合计	比例/%
	5	6	7	8	9	10	11	12	13	14	≥15		
单碱基重复	0	0	0	0	0	2 764	1 229	701	416	294	925	6 329	48.33
二碱基重复	0	1 185	732	541	380	267	194	189	129	136	573	4 326	33.03
三碱基重复	1 172	460	197	81	41	28	6	3	6	4	10	2 008	15.33
四碱基重复	105	27	12	1	0	0	1	0	0	0	0	146	1.11
五碱基重复	111	28	8	3	0	0	0	0	0	0	0	156	1.19
六碱基重复	104	12	7	4	3	0	0	0	0	0	0	131	1.00
合计	1 492	1 712	956	630	424	3 059	1 430	893	551	434	1 508	13 096	100.00

表5 北苍术转录组条带次生代谢KEGG通路分析

Table 5 Secondary metabolism KEGG pathway annotation analysis of *A. chinensis* transcriptomic unigenes

编号	代谢通路	条带数量/条	比例/%	通路ID
1	苯丙烷生物合成	87	20.67	ko00940
2	萜类骨架生物合成	65	15.44	ko00900
3	类胡萝卜素生物合成	49	11.64	ko00906
4	莨菪烷、哌啶和吡啶生物碱的生物合成	30	7.13	ko00960
5	玉米素生物合成	27	6.41	ko00908
6	单菌素生物合成	23	5.46	ko00261
7	倍半萜和三萜生物合成	21	4.99	ko00909
8	异喹啉生物碱生物合成	20	4.75	ko00950
9	黄酮和黄酮醇生物合成	19	4.51	ko00941
10	二苯乙烯类、二芳基庚烷类和姜辣素生物合成	15	3.56	ko00945
11	二萜生物合成	15	3.56	ko00904
12	黄酮和黄酮醇的生物合成	9	2.14	ko00944
13	油菜素内酯生物合成	9	2.14	ko00905
14	咖啡因生物合成	8	1.90	ko00232
15	柠檬烯与蒎烯降解	7	1.66	ko00903
16	单萜类生物合成	6	1.43	Ko00902
17	单环β-内酰胺生物合成	4	0.95	ko00965
18	葡萄糖甙生物合成	4	0.95	ko00966
19	花青素生物合成	1	0.24	ko00942
20	吲哚生物碱生物合成	1	0.24	ko00901
21	异黄酮生物合成	1	0.24	ko00943

条带数量排第2位, 有65条; 与倍半萜和三萜生物合成代谢通路相关的条带数量排第7位, 有21条。

在北苍术转录组数据库中, 发现有14条条带可能编码MVA途径中6种酶, 其中注释为HMGS和MK的条带有3条, 注释为其他4种酶的均有2条; 12条条带可能编码MEP途径中的7种酶, 其中注释为DXS的条带有5条, DXR有2条, 注释为其他5种酶的均有1条(表6)。

2.6 北苍术不同器官倍半萜类生物合成关键基因的时空表达

利用qRT-PCR技术, 检测了北苍术3个不同产地、5个不同器官中倍半萜类生物合成关键基因 *FPPS*、*HMGR*和*DXS*的相对表达量(图6), 发现在北苍术不同产地、不同器官中3个基因相对表达量差异显著, 总体表达趋势相似, 均为叶中相对表达量最低, 而在茎、根茎、花、越冬芽中相对表

表6 北苍术转录组中与倍半萜类生物合成相关条带

Table 6 Unigenes related to sesquiterpenes biosynthesis in *A. chinensis* transcriptomic unigenes

途径	酶全称	简称	条带数量/条
MVA	酰基转移酶(acetyl-CoA acetyltransferase)	AACT	2
	羟甲基戊二酰辅酶A合成酶(3-hydroxy-3-methylglutaryl CoA synthase)	HMGS	3
	羟甲基戊二酰辅酶A还原酶(3-hydroxy-3-methylglutaryl CoA reductase)	HMGR	2
	甲羟戊酸激酶(mevalonate kinase)	MK	3
	磷酸甲羟戊酸激酶(phosphomevalonate kinase)	PMK	2
	甲羟戊酸焦磷酸脱羧酶(mevalonate-5-pyrophosphate decarboxylase)	MDC	2
MEP	1-脱氧-D-木酮糖-5-磷酸合酶(1-deoxy-D-xylulose-5-phosphate synthase)	DXS	5
	1-脱氧-D-木酮糖-5-磷酸还原异构酶(1-deoxy-D-xylulose-5-phosphate reductoisomerase)	DXR	2
	4-二磷酸胞苷-2-C-甲基赤藓糖醇合酶(4-diphosphocytidyl-2-C-methylerythritol synthase)	CMS	1
	4-二磷酸胞苷基-2-C-甲基赤藓糖醇激酶(4-diphosphocytidyl-2-C-methylerythritol kinase)	CMK	1
	2-C-甲基-D-赤藓糖醇2,4-环二磷酸合成酶(2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase)	MECPS	1
	4-羟基-3-甲基丁-2-烯基二磷酸合成酶(4-hydroxy-3-methylbut-2-enyl-diphosphate synthase)	HDS	1
	4-羟基-3-甲基丁-2-烯基二磷酸还原酶[4-hydroxy-3-methylbut-2-enyl diphosphate reductase]	HDR	1

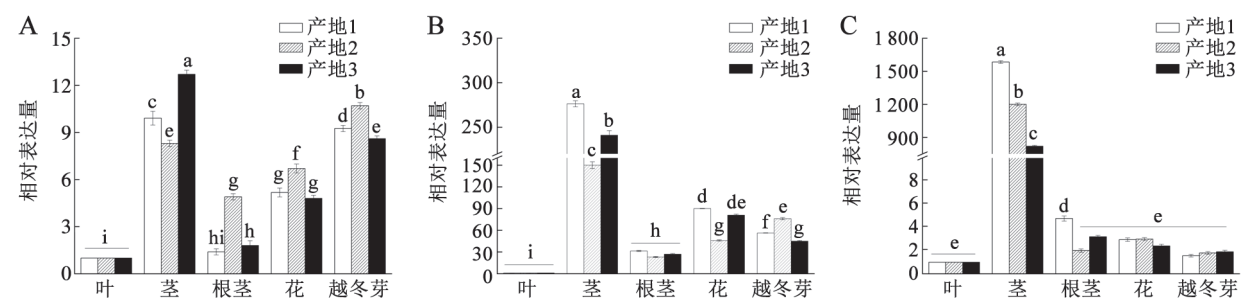


图6 北苍术不同产地和器官中FPPS、HMGR和DXS基因表达分析
Fig.6 mRNA expression levels of FPPS, HMGR and DXS in different habitats and organs of *A. chinensis*
A: FPPS; B: HMGR; C: DXS. 不同小写字母代表差异显著($P<0.05$).

达量相对较高, 且HMGR和DXS基因的相对表达量高于FPPS基因。

3 讨论

第二代高通量转录组测序技术在药用植物分子生物学研究方面得到广泛应用。本文首次以北苍术根茎为材料, 利用Illumina HiSeq™ 2000 150PE高通量测序平台建立北苍术转录组数据库, 测序质量良好、质控严格, 获得50 330条条带, 在NR、NT、Pfam、KOG、Swiss-prot、KEGG和GO数据库中被全部注释成功, 说明基于第二代高通量测序平台的北苍术转录组分析可行, 为北苍术的分子生物学研究提供了丰富的序列资源。

倍半萜类化合物是北苍术根茎中主要活性成分, 其生物合成途径复杂, 目前尚未见北苍术倍半

萜类化合物生物合成分子机理相关的研究报道。本文通过研究北苍术根茎转录组特征, 挖掘到65条与萜类骨架合成相关的条带以及21条与倍半萜和三萜生物合成相关的条带。Ahmed等(2016)研究茅苍术(*Atractylodes lancea*)根茎转录组特征, 获得62 352条条带, 挖掘到依赖MVA途径和MEP途径合成的5种未知的倍半萜类物质。Huang等(2016)对茅苍术叶片和根茎进行转录组分析, 获得92 366条条带, 比较发现, 叶片和根茎差异表达基因中有39条条带与次生代谢途径相关, 其中18条与萜类和聚酮化合物代谢相关。杨国等(2019)通过白术(*Atractylodes macrocephala*)转录组高通量测序, 鉴定到32个与MVA途径关键酶合成相关的条带, 30个与MEP途径关键酶合成相关的条带。本研究挖掘到14条条带参与MVA途径和12条条带参与MEP

途径中关键酶的合成。通过调控关键酶基因的表达量可有效的调控萜类化合物的生物合成, 这些条带的发现为揭示北苍术倍半萜类化合物生物合成通路分析及其调控机理提供重要线索。

qRT-PCR技术是研究植物代谢途径中关键酶相对表达分析的重要手段。Huang等(2016)转录组分析发现, 茅苍术叶和根茎的差异表达基因主要涉及信号转导、初级代谢和次级代谢。Yang等(2013)研究丹参(*Salvia miltiorrhiza*)叶与根转录组差异表达基因, 发现丹参酮生物合成途径中的编码关键酶SmCPS、SmKSL和CYP76AH1的基因在根中相对表达量较高。本研究发现北苍术*FPFS*、*HMGR*和*DXS*基因在不同器官中表达量不同, 但不同产地间差异不大, 其中茎、根茎、花、越冬芽中表达量相对较高, 叶中最低, 提示影响北苍术倍半萜类化合物合成的关键基因的表达可能集中在茎、根茎、花、越冬芽等部位。

参考文献(References)

- Ahmed S, Zhan CS, Yang YY, et al (2016). The transcript profile of a traditional Chinese medicine, *Atractylodes lancea*, revealing its sesquiterpenoid biosynthesis of the major active components. PLOS One, 11: e0151975
- Chen S, Luo H, Li Y, et al (2011). 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in *Panax ginseng*. Plant Cell Rep, 30: 1593–1601
- Chen SL, Zhu XX, Li CF, et al (2012). Genomics and synthetic biology of traditional Chinese medicine. Acta Pharm Sin, 47: 1070–1078 (in Chinese with English abstract) [陈士林, 朱孝轩, 李春芳等(2012). 中药基因组学与合成生物学. 药学学报, 47: 1070–1078]
- Cordoba E, Salmi M, Leon P (2009). Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. J Exp Bot, 60: 2933–2943
- Deng J, Wan QY, Gong L, et al (2017). Cloning and analysis of *DXS* gene from *Atractylodes lancea*. Chin J Exp Trad Med Form, 23: 39–44 (in Chinese with English abstract) [邓娟, 万倩芸, 龚玲等(2017). 茅苍术*DXS*基因的克隆与分析. 中国实验方剂学杂志, 23: 39–44]
- Estévez JM, Cantero A, Reindl A, et al (2001). 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. J Biol Chem, 276: 22901–22909
- He RL, Han YL, Wang F, et al (2018). Analysis on SSR information in transcriptome of *Astragalus membranaceus* var. *mongolicus* and its polymorphism. Chin J Chin Mat Med, 43: 1838–1843 (in Chinese with English abstract) [贺润丽, 韩毅丽, 王芳等(2018). 蒙古黄芪转录组SSR标记开发及遗传多样性研究. 中国中药杂志, 43: 1838–1843]
- Huang QQ, Huang X, Deng J, et al (2016). Differential gene expression between leaf and rhizome in *Atractylodes lancea*: a comparative transcriptome analysis. Front Plant Sci, 7: 348
- Li H, Zhang N, Li YM, et al (2018). High-throughput transcriptomic sequencing of *Rheum palmatum* L. seedlings and elucidation of genes in anthraquinone biosynthesis. Acta Pharm Sin, 53: 1908–1917 (in Chinese with English abstract) [李欢, 张娜, 李依民等(2018). 利用转录组测序挖掘掌叶大黄蒽醌类生物合成相关基因. 药学学报, 53: 1908–1917]
- Li WJ, Guo YL, Shang CL, et al (2016). Analysis of chemical constituents of *Atractylodes chinensis* rhizome by GC-MS. Chin J Exp Trad Med Form, 22: 66–70 (in Chinese with English abstract) [李万娟, 郭艳玲, 商春丽等(2016). 北苍术化学成分的GC-MS分析. 中国实验方剂学杂志, 22: 66–70]
- Li YM, Hu BX, Peng L, et al (2018). Transcriptomic analyses of *Chloranthus japonicus* hy RNA-seq. Chin Trad Herb Drug, 49: 4967–4974 (in Chinese with English abstract) [李依民, 胡本祥, 彭亮等(2018). 基于RNA-seq的银钱草转录组分析. 中草药, 49: 4967–4974]
- Lichtenthaler HK, Rohmer M, Schwender J (1997). Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. Physiol Plant, 101: 643–652
- Peng L, Zhang G, Yan YG, et al (2018). Analysis of transcriptomes and exploring of anthraquinones biosynthetic pathway genes in *Rubia cordifolia* L. Chin Trad Herb Drug, 49: 1890–1896 (in Chinese with English abstract) [彭亮, 张岗, 颜永刚等(2018). 茜草转录组分析及蒽醌类化合物合成相关基因的挖掘. 中草药, 49: 1890–1896]
- Ramilowski JA, Sawai S, Seki H, et al (2013). *Glycyrrhiza uralensis* transcriptome landscape and study of phytochemicals. Plant Cell Physiol, 54: 679–710
- Sang XH, Gu W, Chao JG, et al (2017). Selection of reference genes of *Atractylodes lancea* and its application in biosynthesis of active ingredients. Plant Physiol J, 53: 1680–1686 (in Chinese with English abstract) [桑晓华, 谷巍, 巢建国等(2017). 茅苍术内参基因筛选及其在活性成分生物合成研究中的应用. 植物生理学报, 53: 1680–1686]
- Schwender J, Gemünden C, Lichtenthaler HK (2001). Chlorophyta exclusively use the 1-deoxyxylulose 5-phosphate/2-C-methylerythritol 4-phosphate pathway for the biosynthesis of isoprenoids. Planta, 212: 416–423
- Shao J, Gu W, Chao JG, et al (2015). Molecular identification of *Atractylodes lancea* and its closely related species based on ITS2 sequence. Chin Trad Herb Drug, 46:

- 1209–1215 (in Chinese with English abstract) [邵婧, 谷巍, 巢建国等(2015). 基于ITS2 序列的茅苍术及其近缘种DNA分子鉴定. 中草药, 46: 1209–1215]
- Sun C, Li Y, Wu Q, et al (2010). *De novo* sequencing and analysis of the American ginseng root transcriptome using a GS FLX Titanium platform to discover putative genes involved in ginsenoside biosynthesis. *BMC Genomics*, 11: 262
- Vranová E, Coman D, Gruissem W (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol*, 64: 665–700
- Wen J, Zhao G, Yang YQ, et al (2019). Analysis of comparative transcriptome in the development of *Tetragastria hemsleyanum* tuber. *Mol Plant Breeding*, (17): 5660–5667 (in Chinese with English abstract) [闻静, 赵刚, 杨玉琪等 (2019). 三叶青块根发育比较转录组分析. 分子植物育种, (17): 5660–5667]
- Wu ZJ, Lin YZ, Liu JG, et al (2010). Study of the components of volatile oil of *Atractylodes chinensis* by GC-MC. *Ginseng Res*, (3): 13–15 (in Chinese with English abstract) [武子敬, 林跃中, 刘金刚等(2010). 北苍术挥发油成分GC-MS分析. 人参研究, (3): 13–15]
- Xiong YY, Wang CM, Wang J (2019). Analysis of transcriptome differences of clematis single and double petals. *Mol Plant Breeding*, 17: 3859–3864 (in Chinese with English abstract) [熊阳阳, 王昌命, 王锦(2019). 基于RNA-seq的铁线莲转录组信息分析. 分子植物育种, 17: 3859–3864]
- Yang G, Li HH, Jin YF, et al (2019). Analysis of genes related to biosynthesis of sesquiterpene in *Atractylodes macrocephala* by transcriptome. *Plant Physiol J*, 55: 1827–1838 (in Chinese with English abstract) [杨国, 李鸿慧, 金叶飞等(2019). 基于转录组分析白术倍半萜生物合成的相关基因. 植物生理学报, 55: 1827–1838]
- Yang L, Ding G, Lin H, et al (2013). Transcriptome analysis of medicinal plant *Salvia miltiorrhiza* and identification of genes related to tanshinone biosynthesis. *PLOS One*, 8: e80464
- Zou XX, Huang LQ, Cui GH, et al (2009). Genetic relationships of *Atractylodes* plants. *Acta Pharm Sin*, 44: 680–686 (in Chinese with English abstract) [邹小兴, 黄璐琦, 崔光红等(2009). 苍术属植物的遗传关系研究. 药学报, 44: 680–686]

Transcriptomic analysis of *Atractylodes chinensis* and elucidation of genes in sesquiterpenes biosynthesis

ZHAO Jianhua, ZHAO Chongyang, SUN Chengzhen, SHI Fengyu, CHEN Lina, ZHENG Jinshuang*
College of Agronomy and Biotechnology, Hebei Normal University of Science & Technology, Qinhuangdao, Hebei 066004, China

Abstract: Sesquiterpenes are the main active constituents and the index components for the quality control of *Atractylodes chinensis*. To study the sesquiterpenes biosynthesis, rhizomes of *A. chinensis* were subjected to a high-throughput transcriptomic sequencing analysis by Illumina HiSeq™ 2000 150PE. A total of 8.86 G clean data resulting in 59 065 569 clean reads were generated. Trinity assembly yielded 50 330 unigenes, with an average of 1 214 nt. Functional annotation revealed that all unigenes were successfully annotated in the NR, NT, Pfam, KOG, Swiss-prot, KEGG and GO databases. GO classification contained 3 major groups with 53 subgroups. KEGG analysis indicated that 11 235 unigenes were implicated in 130 KEGG standard metabolic pathways, and 421 unigenes were related to 21 standard secondary metabolic pathways, among which 14 and 12 unigenes were involved in the biosynthesis of key enzymes in MVA and MEP pathways respectively. qRT-PCR analyze showed that *FPPS*, *HMGR* and *DXS* genes were expressed in different organs of *A. chinensis*, especially in stems, rhizomes, flowers and hibernacula, meanwhile, the relative expression levels of *HMGR* and *DXS* genes were higher than *FPPS* gene. This study provide a fundamental basis for sesquiterpenes biosynthesis and molecular regulatory mechanism in *A. chinensis*.

Key words: *Atractylodes chinensis*; transcriptome; metabolic pathway; sesquiterpenes; qRT-PCR

Received 2019-12-19 Accepted 2020-05-27

This work was supported by the Science and Technology Research Projects of Hebei Higher Education Institutions (QN2019162), Hebei Natural Science Foundation (H201907047) and the Key Research and Development Projects of Hebei Province (19226354D).

*Corresponding author (jinshuangk@163.com).