

基于 RNA-seq 的银线草转录组分析

李依民, 胡本祥, 彭 亮, 沈 霞, 高 静, 王昌利, 颜永刚*, 张 岗*

陕西中医药大学药学院 陕西省中药基础与新药研究重点实验室, 陕西 西安 712046

摘要: 目的 获得银线草 *Chloranthus japonicus* 根茎转录组信息特征。方法 用高通量测序平台 Illumina HiSeq™ 2000 150PE 进行银线草根茎转录组测序分析。结果 转录组测序共获得 68 458 750 条高质量序列 (clean reads), Trinity *de novo* 组装获得 56 096 个 unigenes, 平均长度 801 nt。BLAST 分析显示分别有 25 773 (45.94%)、17 801 (31.73%)、16 082 (28.67%)、9 649 (17.20%) 个 unigenes 在 NR、Swiss-port、KOG、KEGG 数据库得到注释信息, 可归为 GO 分类的生物过程、细胞组分和分子功能 3 大类 40 分支, 涉及 131 个 KEGG 标准代谢通路, 其中包括 16 个次生代谢标准通路。进一步分析获得 170 个基因参与单萜、二萜、倍半萜、三萜等萜类化合物生物合成通路。同时, 转录组预测编码蛋白框序列 1 887 个; 高等植物转录因子 54 个家族。借助 MISA 软件发现 8 987 个简单重复序列 (SSRs), 二碱基重复 SSRs 数量最丰富, 有 5 948 个, 出现频率为 66.2%, 五碱基重复 SSRs 相对较少, 占 1.3%。结论 基于 RNA-seq 分析获得银线草根茎转录组信息特征及萜类合成代谢通路关键基因, 为后期银线草活性成分次生代谢途径解析及调控研究奠定基础。

关键词: 银线草; 转录组; 功能基因; 代谢通路; 简单重复序列; RNA-seq

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2018)21-4967-08

DOI: 10.7501/j.issn.0253-2670.2018.21.003

Transcriptomic analyses of *Chloranthus japonicus* by RNA-seq

LI Yi-min, HU Ben-xiang, PENG Liang, SHEN Xia, GAO Jing, WANG Chang-li, YAN Yong-gang, ZHANG Gang

Shaanxi Provincial Key Laboratory for Chinese Medicine Basis & New Drugs Research, College of Pharmacy, Shaanxi University of Chinese Medicine, Xi'an 712046, China

Abstract: Objective To obtain the transcriptome dataset of *Chloranthus japonicus*. **Methods** Using the Illumina HiSeq™ 2000 150PE, a rhizome transcriptome of *C. japonicus* was generated, followed by systemic bioinformatics analyses. **Results** A total of 68 458 750 high quality clean reads were produced by the transcriptome sequencing. Trinity *de novo* assembling resulted in a total of 56 096 unigenes with an average length of 801 nt. BLAST analysis indicated that 25 773 (45.94%), 17 801 (31.73%), 16 082 (28.67%), and 9 649 (17.20%) unigenes were successfully annotated in the NR, Swiss-port, KOG, and KEGG databases, respectively. All unigenes were classified into three major groups by GO, including biological process, cellular component, and molecular function, and then, grouped into 40 subgroups. And 131 KEGG standard pathways were designated, 16 of which were defined as the secondary metabolism. Further analysis revealed that a total of 170 unigenes were involved in the biosynthesis of mono-, di-, sesqui-, or triterpene. Meantime, 1 887 unigenes were predicted to contain protein coding sequences. Totally 54 families of transcription factors of higher plant were identified. Using MISA prediction, 8 987 simple sequence repeats (SSRs) were obtained, among which the di-nucleotide SSRs were abundant with 5 948 (66.2%), whereas the penta-nucleotide SSRs were relatively less, accounting for 1.3%. **Conclusion** The transcriptome of *C. japonicus* rhizome was generated by RNA-seq along with the identification of unigenes implicated in various terpenes biosynthesis, which will provide a fundamental basis for secondary metabolism pathway dissections and their regulatory mechanisms in this plant species.

Key words: *Chloranthus japonicus* Sieb.; transcriptome; functional gene; metabolim pathway; simple sequence repeats; RNA-seq

银线草 *Chloranthus japonicus* Sieb. 为金粟兰 毛七等, 生于山坡或山谷杂木林下荫湿处或沟边草
科金粟兰属多年生草本, 别名天王七、四块瓦、白 丛中, 主要分布于我国西北、华北和东部等省, 朝

收稿日期: 2018-04-30

基金项目: 陕西省教育厅专项 (18JK0216); 陕西中医药大学新进博士科研启动经费 (104080001); 陕西省高校首批青年杰出人才支持计划项目; 咸阳市中青年科技领军人才项目; 陕西省科技厅项目 (2016KTTSSF01-01-01); 国家中医药管理局项目 (ZYBZH-C-QIN-45)

作者简介: 李依民, 女, 博士, 讲师, 研究方向为中药资源与分子生药学。E-mail: 2051058@sntcm.edu.cn

*通信作者 张 岗, 男, 博士, 教授, 研究方向为中药资源与分子生药学。Tel/Fax: (029)38185165 E-mail: jay_gumling2003@aliyun.com
颜永刚, 男, 博士, 教授, 研究方向为中药品种、品质与资源开发。E-mail: yunfeng828@163.com

鲜、日本也有分布;根及根茎药用,始载于《神农本草经》,味辛、苦,性温,有毒,具有祛湿散寒、活血消肿、散瘀解毒的功效,主治风寒咳嗽、瘀血肿痛、毒蛇咬伤、疟疾和疮痍等^[1]。银线草药用植物资源极为丰富,药用历史悠久,在太白七药中具有重要的地位。中药化学研究表明银线草主要含有萜类、黄酮类、香豆素类、木脂素及挥发油等化合物^[2]。银线草特征性成分主要包括银线草醇、金粟兰内酯、银线草内酯、银线草呋喃醇等乌药烷型倍半萜类化合物^[2-3],药理学研究揭示其具有良好的抗菌^[4]、抗炎^[5]、抗肿瘤^[6]和抗 HIV^[7-8]等多种生物活性。此外,银线草还在生药学、组织化学定位方面有初步研究报道^[9-10]。最新研究利用二代高通量测序技术解析银线草叶绿体基因组,为该植物的系统进化及药材分子鉴定提供基础数据^[11]。

转录组测序技术即 RNA-seq,是功能基因组研究的一个重要组成部分,能够在基因组序列未知的前提下,研究特定条件下细胞中所有基因的转录本丰度、可变剪接等全局信息,结果准确、分析可靠且重复性较高,有助于从整体水平上揭示生物体生长发育、次生代谢及生理适应的转录调控规律^[12]。当前, RNA-seq 在药用植物功能基因组领域内应用广泛,已获得人参^[13]、甘草^[14]和膜荚黄芪^[15]等众多药用植物转录组数据,为阐明中药种质资源遗传基础打下基础。银线草临床功效明确、活性成分多样,具有极高的研究价值。然而,基因组相关数据匮乏,限制了对银线草资源的深入研究与开发。本研究利用 RNA-seq 进行银线草转录组测序分析,以期揭示其转录组信息特征,为该药用植物功能基因挖掘、次生代谢途径解析及调控研究提供基础数据。

1 材料与方法

1.1 材料

植物材料于 2015 年 7 月采自陕西省宝鸡市太白县黄峰山镇,经度 107°20'34.8",纬度 34°1'27.5",海拔 1 873.3 m,经陕西中医药大学张岗教授鉴定为金粟兰科金粟兰属银线草 *Chloranthus japonicus* Sieb.,取单株植株根茎液氮速冻后置于-80 °C 冰箱备用。

1.2 RNA 提取与文库构建

采用 EASYspin 植物 RNA 快速提取试剂盒(Aidlab, 中国)制备银线草根茎总 RNA,琼脂糖凝胶电泳和 NanoDrop™ 2000 分光光度计(Thermo Fisher, 美国)检测完整性。用带有 Oligo (dT) 的磁珠富集 mRNA,加入碎片化缓冲液(fragmentation buffer)将

mRNA 打断成短片段,用六碱基随机引物(random hexamers)合成 cDNA 第 1 链;然后加入缓冲液、dNTPs、RNase H 和 DNA polymerase I 合成 cDNA 第 2 链;再经过 QiaQuick PCR 试剂盒(QIAGEN, 德国)纯化并加 EB 缓冲液洗脱之后做末端修复、加 poly (A) 并连接测序接头,然后用琼脂糖凝胶电泳进行片段大小选择,最后进行 PCR 扩增构建测序文库。

1.3 转录组测序与组装

利用 Illumina HiSeq™ 2000 150PE 对银线草根茎转录组文库进行高通量测序。测序原始图像数据经碱基识别(base calling)转化为序列数据原始序列(raw reads),经数据评估、过滤除杂和冗余处理等质控得到高质量序列(clean reads),再利用 Trinity 做转录组 *de novo* 组装。Trinity 首先将具有一定长度重叠(overlap)的 reads 连成更长的片段,即得到的不含 N 的组装片段作为 unigene。

1.4 转录组功能注释

利用 BLAST 将 unigenes 与蛋白数据库 NR、Swiss-port、蛋白相邻类的聚簇(KOG)和 KEGG(京都基因与基因组百科全书)进行比对(E 值 $<1 \times 10^{-5}$),得到与相应 unigenes 具有最高序列相似性的蛋白,进而得到 unigenes 注释信息。根据 NR 注释信息,使用 Blast2GO 软件得到 unigene 的 GO (gene ontology) 注释,用 WEGO 软件对所有 unigenes 做 GO 功能分类统计,从宏观上认识该物种的基因功能分布特征。

1.5 蛋白编码框(CDS)和转录因子预测

按 NR、Swiss-Prot、KOG 和 KEGG 的优先级顺序将 unigenes 序列与以上蛋白库做 BLASTx 比对(E 值 $<1 \times 10^{-5}$)并确定该 unigene 编码区的核酸序列(序列方向 5'→3')和氨基酸序列。利用 ESTScan^[13]预测与以上数据库比对不上的 unigenes 的编码区及序列方向。将所预测的 unigenes 编码蛋白序列与植物转录因子数据库(plant TFDB)进行 hmmscan 比对,搜索转录因子家族及其成员。

1.6 简单重复序列(simple sequence repeats, SSRs)特征检测

使用 MISA 工具(MicroSAtellite identification tool)检测银线草转录组 unigenes,搜索 SSRs 并进行统计分析。

2 结果与分析

2.1 银线草转录组组装与质量分析

采用 Illumina HiSeq™ 2000 150PE 高通量测序平

台对银线草根茎进行转录组测序,共得到 68 458 750 条 raw reads,过滤产生了 66 138 234 条 clean reads,包含 6 320 552 293 个核苷酸信息,Q20(碱基量 $\geq$ 20%)、Q30(碱基量 $\geq$ 30%)分别为 98.69%、93.59%,GC 量为 45.01%,说明测序质控良好, clean reads 质量合格。Trinity 组装获得 56 096 个 unigenes,平均长度 801 nt,最长达到 14 262 nt,最短序列为 201 bp,N50 为 1 460 nt。

Unigenes 长度分布(图 1-A)显示,12 506 条 unigenes 长度超过 1 000 nt,4 786 条序列大于 2 000 nt。Reads 在 unigenes 上的覆盖情况统计(图 1-B)显示,所含 reads 数量在 11~100 的 unigenes 数量最多,为 33 677 条;其次为 reads 数量在 101~200 的 unigenes,为 4 041 条;reads 数量在 1~10、1 001~2 000、2 001~4 000 的 unigenes 分别为 2 639、3 253、2 763 条;其余 reads 分布区域对应的 unigenes 数量均相对较少。

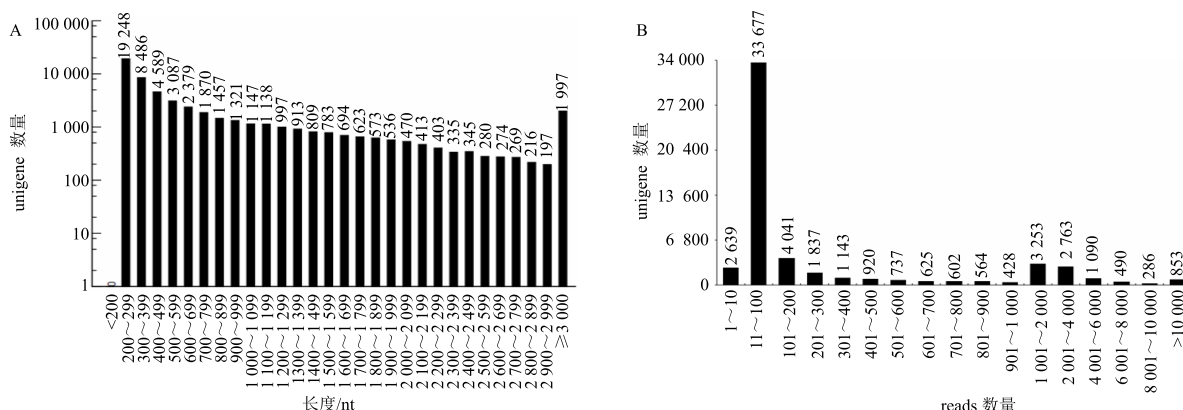


图 1 银线草转录组 unigenes 长度分布 (A) 和 reads 覆盖统计 (B)

Fig. 1 Length distribution (A) and reads coverage statistics (B) of *C. japonicus* transcriptomic unigenes

2.2 银线草转录组 unigenes 的功能注释

使用 BLAST 将所有 unigenes 与 NR、Swiss-port、KOG、KEGG 等数据库进行一致性比对分析,对各数据库注释的 unigenes 数目进行统计,进而获得银线草根茎转录组 unigenes 的功能注释信息。结果表明,25 773 条 unigenes (45.94%) 在 NR 数据库比对成功得到注释,在 Swiss-port、KOG、KEGG 等数据库获得注释的 unigenes 数目依次为 17 801 (31.73%)、16 082 (28.67%)、9 649 (17.20%)。7 408 条 unigenes 同时在所有数据库中注释,至少有 1 种数据库注释成功共 25 861 条 (46.10%),30 235 条未获得注释,见图 2。

以 NR 数据库为例进行分析, unigenes 注释同源基因的物种分布如图 2 所示,在相似序列匹配度较高的物种中,莲 *Nelumbo nucifera* Gaertn. 所占比例最高,为 7 072 条 (27.44%);其次为葡萄 *Vitis vinifera* L. 2 326 条 (9.02%),油棕 *Elaeis guineensis* Jacq. 1 270 条 (4.93%),海枣 *Phoenix dactylifera* L. 1 218 条 (4.73%),土瓶草 *Cephalotus follicularis* Labill. 714 条 (2.77%),可可 *Theobroma cacao* L. 821 条 (3.19%),其余匹配物种比例在 2.04%~4.71%,比例

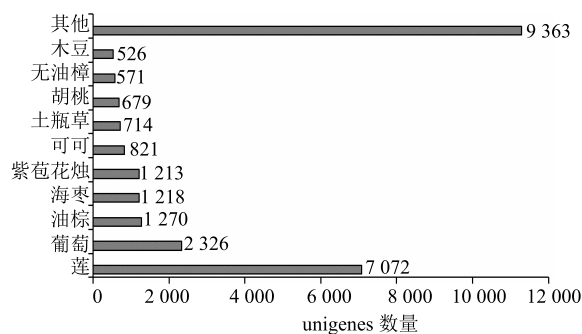


图 2 银线草转录组 unigenes 与 NR 数据库匹配物种分布
Fig. 2 Species distribution of *C. japonicus* transcriptomic unigenes against NR database

小于 1.57% 的匹配物种的 unigenes 为 9 363 条,占 36.32%。

根据 NR 注释信息得到 GO 功能分类(图 3),14 660 条 unigenes 被注释到生物过程、细胞组分和分子功能 3 个 GO 类别的 40 个小组。细胞组分中细胞 (cell) 和细胞部分 (cell part) 相关基因丰度最高,达 936 和 936 条;其次是细胞器 (organelle),有 675 条;病毒粒子 (virion)、病毒粒子组成 (virionpart) 等基因较少,为 10 条。生物过程主要聚集在代谢过程 (metabolic process) 和细胞过程

(cellular process), 涉及的基因分别有 2 036 条和 1 710 条; 应激适应 (response to stimulus)、生物调控 (biological regulation) 基因数量分别为 350、257 条。分子功能中具有催化活性 (catalytic activity) 和结合功能的基因 (binding) 数量较高, 分别为 2 152 和 1 336 条, 其他类别基因数目普遍较少。

为了进一步分析银线草转录组 unigenes 的功

能, 进行 KOG 功能分类分析 (图 4), 共得到 25 个不同的 KOG 功能类群, 种类比较全面, 包括大多数的生命活动; 一般功能预测的基因数量最多, 有 6 887 条; 翻译后修饰、蛋白反转、伴侣和信号转导机制类次之, 分别为 2 899 和 2 580 条; 加工和修饰 unigenes 数目 1 541 条; 其他种类基因丰度不尽相同。

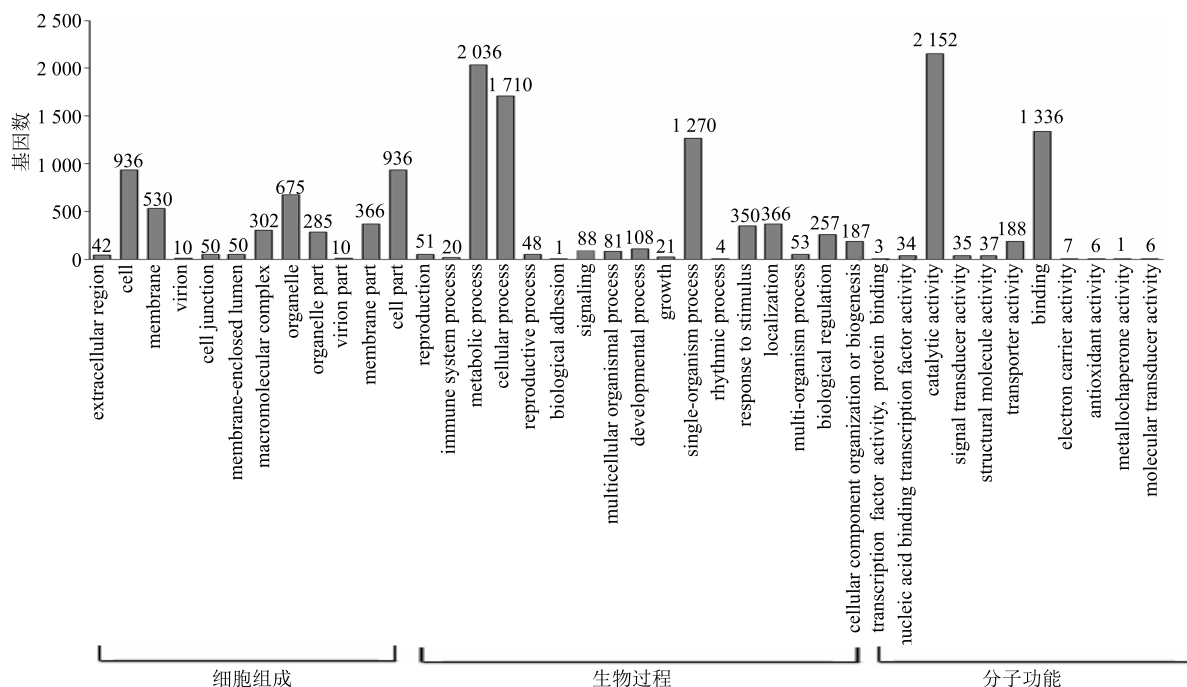


图 3 银线草根茎转录组 unigenes 的 GO 分类

Fig. 3 GO classification of *C. japonicus* transcriptomic unigenes

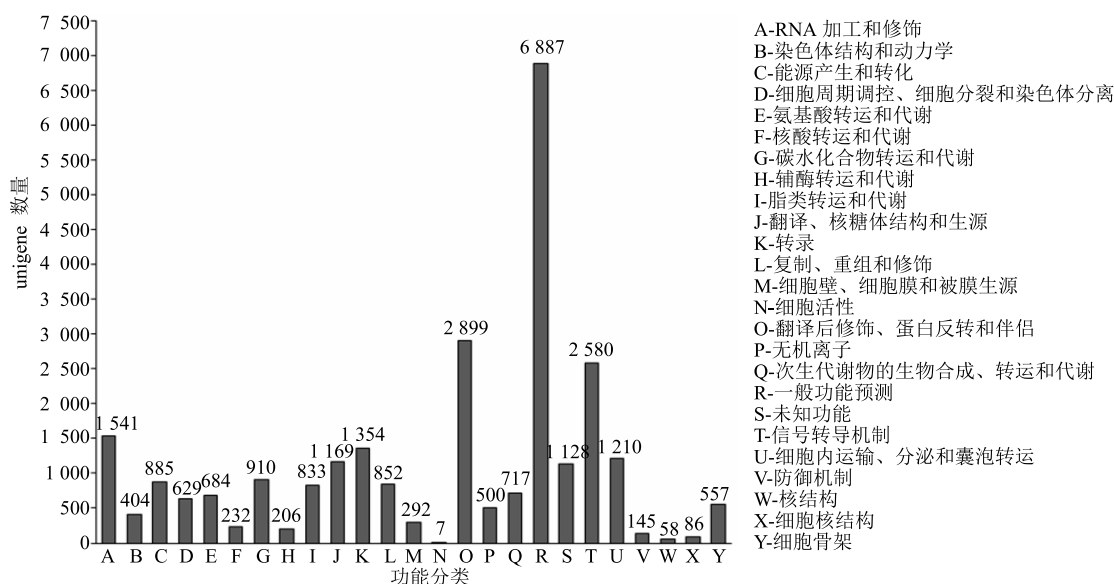


图 4 银线草转录组 unigenes 的 KOG 注释分布

Fig. 4 KOG annotation distribution of *C. japonicus* transcriptomic unigenes

银线草根茎转录组 unigenes 参与 KEGG 代谢通路分为 5 大分支: 细胞过程 (cellular processes) 460 条、环境信息处理 (environmental information processing) 308 条、遗传信息处理 (genetic information processing) 2 283 条、代谢 (metabolism)

8 510 条和有机系统 (organismal systems) 334 条。5 117 条 unigenes 获得 131 个 KEGG 标准代谢通路, 按照基因注释量大小依次排序, 选取前 11 个代谢通路信息见表 1, 这些通路包含的 unigenes 数量大于 200 条以上。

表 1 银线草转录组 unigene KEGG 通路分析统计

Table 1 KEGG pathway analysis of *C. japonicus* transcriptomic unigene

编号	代谢通路	unigenes 数量	占比/%	通路 ID
1	代谢通路	1 940	37.91	ko01100
2	次生代谢产物生物合成	1 015	19.83	ko01110
3	抗生素生物合成	465	9.08	ko01130
4	多种环境下微生物代谢	390	7.62	ko01120
5	核糖体	325	6.35	ko03010
6	植物-病原菌互作	293	5.72	ko04626
7	碳水化合物代谢	277	5.41	ko01200
8	内质网蛋白加工	248	4.85	ko04141
9	氨基酸生物合成	238	4.65	ko01230
10	剪接体	221	4.32	ko03040
11	嘌呤代谢	202	3.95	ko00230

KEGG 代谢通路分析还发现 402 条 unigenes 参与苯丙素类、萜类、生物碱及黄酮类等生物合成相关的 16 个次生代谢通路。结果表明 (表 2), 苯丙素的生物合成代谢通路 (ko00940) 基因数量最多, 为 126 个; 18 条 unigenes 与萜类、吡啶、吡啶生物碱 (ko00960) 合成相关; 花青素、咖啡因、黄酮及甜菜红色素的生物合成通路基因数较少。

银线草主要成分为萜类, 特征性成分为倍半萜类化合物, 在转录组数据中找到了大量倍半萜、单萜、二萜以及三萜类生物合成 KEGG 标准通路 (表 2)。64 条 unigenes 编码萜类化合物骨架生物合成 (ko00900); 单萜类基因 4 条; 柠檬烯和蒎烯是双环单萜的重要代表, 涉及该化合物降解途径相关基因 20 条序列; 32 条 unigenes 参与二萜类生物合成

表 2 银线草转录组 unigenes 次生代谢 KEGG 通路注释统计

Table 2 Secondary metabolism KEGG pathways annotation analysis of *C. japonicus* transcriptomic unigenes

编号	代谢通路	unigene 数量	占比/%	通路 ID
1	苯丙素的生物合成	126	2.46	ko00940
2	萜类化合物的合成及姜辣素类	31	0.61	ko00945
3	类黄酮生物合成	20	0.39	ko00941
4	萜类、吡啶、吡啶生物碱生物合成	18	0.35	ko00960
5	玉米素的生物合成	16	0.31	ko00908
6	油菜素内酯的生物合成	10	0.20	ko00905
7	花青素生物合成	4	0.08	ko00942
8	咖啡因的代谢	3	0.06	ko00232
9	黄酮和黄酮醇的生物合成	2	0.04	ko00944
10	甜菜红色素的生物合成	2	0.04	ko00965
11	萜类化合物骨架生物合成	64	1.25	ko00900
12	单萜类生物合成	4	0.08	ko00902
13	柠檬烯和蒎烯降解	20	0.39	ko00903
14	二萜类生物合成	32	0.63	ko00904
15	类胡萝卜素生物合成	39	0.76	ko00906
16	倍半萜和三萜类生物合成	11	0.21	ko00909

(ko00904); 类胡萝卜素生物合成 (ko00906) 有关的基因有 39 条; 倍半萜和三萜类生物合成 (ko00909) 基因有 11 条。其中, 倍半萜类包括无环型 (acyclic)、吉玛烯型 (germacrene), 编码基因分别为 1 和 7 条。这些基因信息为进一步发掘银线草萜类活性成分合成通路解析及其调控研究提供基础数据。

2.3 CDS 和转录因子分析

对银线草转录组所有 unigenes 的 CDS 进行分析, 通过 BLAST 比对共获得 CDS 序列 25 512 个, 利用 ESTscan 数据库分析获得 CDS 序列 1 887 个。转录因子预测发现有 54 个家族成员, 其中 bHLH、ERF、WRKY、C2H2、MYB_related、NAC、MYB、

bZIP 类占主体, 说明银线草根茎生理代谢涉及众多转录调控过程 (图 5)。

2.4 SSRs 特征分析

用 MISA 软件对转录组 unigenes 进行 SSRs 分析 (表 3), 7 074 条 unigenes 中共计 8 987 个 SSRs。其中, 二碱基重复 SSRs 数量最丰富, 有 5 948 个 (66.2%), AG/CT 类型的比例最高。三碱基重复 SSRs 数量次之, 有 2 310 个, 占 SSRs 总量的 25.7%, 其中 AAG/CTT 重复类型数量最多。四碱基和六碱基重复分别为 394、217 个, 各占 4.4%、2.4%; 五碱基重复相对较少, 仅占 1.3%。此外, 还发现 SSRs 重复单元数量也存在一定变化, 其中重复 6、7 次的比例最高, 重复 5、8 次的次之。

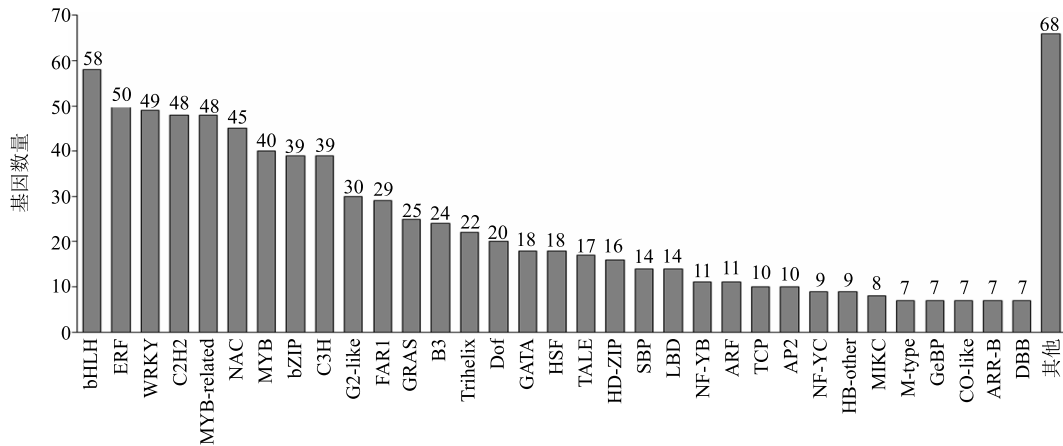


图 5 银线草转录组 unigenes 的转录因子分析

Fig. 5 Transcription factor classification of *C. japonicus* transcriptomic unigenes

表 3 银线草转录组 unigenes SSRs 分析

Table 3 SSRs analyses of *C. japonicus* transcriptomic unigenes

重复	重复单元数量												合计	占比/%
	4	5	6	7	8	9	10	11	12	13	14	≥15		
二碱基重复	0	0	1 324	1 042	882	661	547	315	87	52	221	817	5 948	66.2
三碱基重复	0	1 166	534	292	78	86	97	21	14	10	6	12	2 310	25.7
四碱基重复	309	65	16	1	3	0	0	0	0	0	0	0	394	4.4
五碱基重复	84	28	6	0	0	0	0	0	0	0	0	0	118	1.3
六碱基重复	139	62	3	10	2	1	0	0	0	0	0	0	217	2.4
合计	532	1 321	1 883	1 345	965	748	644	366	101	62	227	829	8 987	100.0

3 讨论

基于高通量测序技术的转录组测序 (RNA-seq) 通过对生物样本的总 RNA 进行测序, 在本草基因组学研究方面应用非常广泛也并取得重大进展^[12]。本研究首次采用 Illumina HiSeqTM 2000 150PE 测序平台, 进行秦岭特色中草药资源银线草的转录组测序

分析, 测序质量良好、质控严格, *de novo* 组装获得 66 138 234 条高质量 reads 参与组装, 共得到 56 096 个 unigenes, 序列长度与 reads 覆盖区域对应合理。unigenes 序列信息量庞大, 数据基本涵盖全转录组信息, 初步揭示银线草根茎的基因表达特征。

转录组测序数据通常必需借助系统的生物信

息学分析策略进行 reads 的滤过、组装, unigenes 注释及功能分类等^[12]。本研究利用 Trinity、BLAST、WEGO、ESTscan 等工具对银线草转录组序列进行注释和功能分类。基于序列一致性分析, unigenes 与 NR、Swiss-port、KOG、KEGG 比对注释成功占 46.10%, 共 25 861 条序列, 其余 30 235 条未注释, 这与人参^[13]和罗勒花^[16]等药用植物的转录组研究结果类似, 说明基于 RNA-seq 的药用植物转录组有大量序列结构特征需进一步深入挖掘分析。

GO 分类揭示银线草根茎的转录组特性与生物过程、细胞组分和分子功能相关; KOG 功能分析从基因组水平寻找直系同源体, 提高基因功能注释的准确性^[16], 本研究共得到 25 个不同的 KOG 类群, 说明银线草转录组 KOG 种类比较全面。进一步对银线草功能基因序列进行 KEGG 代谢路径注释, 发现 131 个标准 KEGG 代谢通路, 这些基因可能参与银线草水分吸收、矿质营养、光合作用和呼吸作用等生命代谢活动。还发现 16 个次生代谢 KEGG 标准通路的 unigenes 参与苯丙素类、萜类、生物碱、黄酮类等生物合成。其中, 170 个基因参与单萜、二萜、倍半萜、三萜等萜类化合物的生物合成。银线草次生代谢成分以萜类为主, 其中倍半萜类化合物又是最为重要的一类活性成分^[7]。本研究获得的转录组数据分析表明倍半萜类合成基因主要涉及无环型和吉玛烯型倍半萜类, 而主要乌药烷型倍半萜化合物合成相关基因未能得到注释, 可能由于金粟兰属该类倍半萜生物合成通路基因数据相对匮乏。尽管如此, 这些基因信息为银线草醇、金粟兰内酯、银线草内酯等活性成分生物合成通路解析及其调控研究提供重要依据。

转录本的丰度和表达调控在植物细胞代谢过程中起决定作用。转录因子 AP2/ERF、bHLH、MYB 和 WRKY 等家族在植物细胞甲羟戊酸、苯丙烷类代谢途径调控中起关键作用^[17]。本研究获得的银线草 unigenes 转录因子覆盖高等植物转录因子数据库 PlantTFDB 4.0 中 54 个家族, 与次生代谢调控密切相关的转录因子家族 unigenes 数量较多, 说明银线草初级代谢与次生代谢涉及复杂的转录调控机制。

转录组测序数据除了能够用于生物样本的基因挖掘及其表达调控研究, 还可开发大量 EST-SSR, 因此在遗传多样性、分子标记等方面应用广泛^[13]。本研究发掘银线草 7 074 条 unigenes 的 8 987 个 SSRs 位点, 重复类型以三核苷酸为主, 双核苷酸次

之, 与以三核苷酸重复类型为主的主要作物水稻、大麦或棉花等的研究结果一致^[18]。双核苷酸重复 SSRs 中 AG/CT 类型最多, 三核苷酸重复中 AAG/CTT 类型最多, 与番红花^[19]和人参^[20]等植物中双核苷酸重复的情况相同。可见大多数植物 SSRs 重复主要以双核苷酸和三核苷酸为主, 但不同物种的重复序列有差别。表明银线草基因组内具有较高丰度的 SSRs, 为银线草遗传多样性研究提供基础。

基于 RNA-seq 的银线草转录组测序初步获得大量重要数据信息, 后续通过系统分析, 以期全面解析以银线草醇、金粟兰内酯等为主要活性成分的生物合成通路及其调控机制, 揭示该药用植物的群体遗传结构与分子进化特征, 为银线草药用资源的可持续利用与控制提供理论基础。

参考文献

- [1] 南京中医药大学. 中药大辞典 [M]. 上海: 上海科学技术出版社, 2006.
- [2] Cao C M, Peng Y, Shi Q W, *et al.* Chemical constituents and bioactivities of plants of Chloranthaceae [J]. *Chem Biodiv*, 2008, 5(2): 219-238.
- [3] 林凤祥, 罗永明, 李宏焘, 等. 多穗金粟兰乌药烷型倍半萜二聚体类化学成分研究 [J]. *中草药*, 2016, 47(18): 3169-3174.
- [4] Kang T H, Lee Y M, Lee W J, *et al.* Antifungal activities of dimeric sesquiterpenes, shizukaols C and F, isolated from *Chloranthus japonicus* Sieb. [J]. *J Microb Biot*, 2017, 27(7): 1272-1275.
- [5] Zhao J J, Guo Y Q, Yang D P, *et al.* Chlojaponilactone B from *Chloranthus japonicus*: Suppression of inflammatory responses via inhibition of the NF- κ B signaling pathway [J]. *J Nat Prod*, 2016, 79(9): 2257-2263.
- [6] Tang L, Zhu H, Yang X, *et al.* Shizukaol D, a dimeric sesquiterpene isolated from *Chloranthus serratus*, represses the growth of human liver cancer cells by modulating Wnt signalling pathway [J]. *PLoS One*, 2016, 11(3): e0152012.
- [7] 杨颖, 曹颖莉, 刘海洋, 等. 银线草醇 F: 一种新结构类型 HIV-1 逆转录酶 RNaseH 活性抑制剂 [J]. *药学报*, 2012, 47(8): 1011-1016.
- [8] Yan H, Ba M Y, Li X H, *et al.* Lindenane sesquiterpenoid dimers from *Chloranthus japonicus* inhibit HIV-1 and HCV replication [J]. *Fitoterapia*, 2016, 115(1): 64-68.
- [9] 吕邵娃, 杨炳友, 毕姗姗, 等. 银线草的生药鉴定研究 [J]. *中药材*, 2009, 32(2): 203-206.
- [10] 康杰芳, 李娜, 袁琴琴, 等. 银线草营养器官结构及

- 主要成分的组织化学定位 [J]. 西北植物学报, 2010, 30(12): 2412-2416.
- [11] Sun J, Zhang G, Li Y, *et al.* The complete chloroplast genome sequence of *Chloranthus japonicus* [J]. *Mitochondrial DNA Part A*, 2015, 27(5): 3202-3204.
- [12] 陈士林, 朱孝轩, 李春芳, 等. 中药基因组学与合成生物学 [J]. 药学学报, 2012, 47(8): 1070-1078.
- [13] Chen S, Luo H, Li Y, *et al.* 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in *Panax ginseng* [J]. *Plant Cell Rep*, 2011, 30(9): 1593-1601.
- [14] Ramilowski J A, Sawai S, Seki H, *et al.* *Glycyrrhiza uralensis* transcriptome landscape and study of phytochemicals [J]. *Plant Cell Physiol*, 2013, 54(5): 697-710.
- [15] Liu X B, Ma L, Zhang A H, *et al.* High-throughput analysis and characterization of *Astragalus membranaceus* transcriptome using 454 GS FLX [J]. *PLoS One*, 2014, 9(5): e95831.
- [16] 刘 雷, 赵 欢, 冉茂中, 等. 罗勒花和叶的转录组数据组装及基因功能注释 [J]. 中草药, 2017, 48(17): 3612-3618.
- [17] Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants [J]. *Mol Plant*, 2015, 8(5): 689-708.
- [18] Varshney R K, Graner A, Sorrells M E. Genic microsatellite markers in plants: Features and applications [J]. *Trends Biotechnol*, 2005, 23(1): 48-55.
- [19] 陈国庆. 番红花 EST 资源的 SSR 信息分析 [J]. 广西植物, 2011, 31(1): 43-46.
- [20] Li C, Zhu Y, Guo X, *et al.* Transcriptome analysis reveals ginsenosides biosynthetic genes, microRNAs and simple sequence repeats in *Panax ginseng* C. A. Meyer [J]. *BMC Genomics*, 2013, 14: 245.