

茅苍术二倍体和同源四倍体叶片转录组比较

张子璇^a, 谭明谱^b, 王 将^a, 张成才^a, 罗丽娜^a, 向增旭^{a, ①}

(南京农业大学: a. 园艺学院, b. 生命科学学院, 江苏 南京 210095)

摘要: 采用转录组测序技术对茅苍术 [*Atractylodes lancea* (Thunb.) DC.] 二倍体和同源四倍体叶片进行无参转录组测序,并在前期研究基础上筛选叶片中的高可信度差异表达 unigenes。结果显示:共获得 495 797 个 unigenes, GC 含量为 46.54%, 平均长度为 402.8 bp, N50 和 N90 分别为 413 和 225 bp。与 GO、NR、String、SwissProt、KEGG 和 Pfam 数据库的比对结果表明:共注释到 259 841 个 unigenes, 占 unigenes 总数的 52.41%。NR 数据库比对结果显示:茅苍术与同科植物 *Cynara cardunculus* var. *scolymus* (Linn.) Fiori 和向日葵 (*Helianthus annuus* Linn.) 的 unigenes 序列同源性明显高于其他种类。通过表达量比较,共筛选出 923 个差异表达 unigenes。GO 功能分类结果表明:这些差异表达 unigenes 的功能被分成 3 个大类 61 个小类。KEGG 富集分析结果表明:这些差异表达 unigenes 被注释到 5 个大类中,其中,注释到代谢中的差异表达 unigenes 最多(75),且富集在次生代谢物生物合成、淀粉和蔗糖代谢及脂肪酸延伸等 8 条代谢通路中。茅苍术二倍体和同源四倍体叶片和全株的共有高可信度差异表达 unigenes 有 298 个,其中,热激蛋白基因 (*DnaK/Hsp70*)、ACC 氧化酶基因 (*ACO*)、脂肪酸脱氢酶基因 (*FAD2*) 和长链-3-羟酰基-CoA 脱水酶基因 (*PHS1/PAS2*) 的表达量在同源四倍体中上调,而乙烯受体基因 (*ETR/ERS*)、丙二烯氧化物合酶基因 (*AOS*)、脂氧合酶基因 (*LOX*) 和 14-3-3 结构域蛋白基因 (*YWHAE*) 的表达量在同源四倍体中下调。研究结果显示:茅苍术二倍体和同源四倍体叶片在细胞发育、代谢、能量转化、物质积累及信号转导上存在差异,筛选的差异表达 unigenes 可能参与茅苍术叶片发育调控,筛选的高可信度差异表达基因可作为茅苍术倍性育种的关键候选基因。

关键词: 茅苍术; 二倍体; 同源四倍体; 差异表达 unigene; 功能注释

中图分类号: Q946-33; S567.21⁺1 文献标志码: A 文章编号: 1674-7895(2021)04-0041-09

DOI: 10.3969/j.issn.1674-7895.2021.04.05

Comparison on transcriptomes from leaves of diploid and autotetraploid of *Atractylodes lancea* ZHANG Zixuan^a, TAN Mingpu^b, WANG Jiang^a, ZHANG Chengcai^a, LUO Lina^a, XIANG Zengxu^{a, ①} (Nanjing Agricultural University: a. College of Horticulture, b. College of Life Sciences, Nanjing 210095, China), *J. Plant Resour. & Environ.*, 2021, 30(4): 41-49

Abstract: *De novo* transcriptome sequencing and analysis were conducted for leaves of diploid and autotetraploid of *Atractylodes lancea* (Thunb.) DC. by using transcriptome sequencing technology, and high confidence differentially expressed unigenes in leaves were screened on the basis of preliminary research. The results show that 495 797 unigenes are obtained in total, the GC content is 46.54%, the average length is 402.8 bp, and the N50 and N90 are 413 and 225 bp, respectively. The comparison results with GO, NR, String, SwissProt, KEGG, and Pfam databases show that there are 259 841 unigenes annotated in total, accounting for 52.41% of total unigenes. The comparison result in NR database shows that the unigene sequence homology of *A. lancea* with the same family species *Cynara cardunculus* var. *scolymus* (Linn.) Fiori and *Helianthus annuus* Linn. are evidently higher than that with other species. 923 differentially expressed unigenes are obtained through expression level comparison. GO functional classification result shows that these differentially expressed unigenes are classified into

收稿日期: 2021-01-16

基金项目: 财政部和农业农村部国家现代农业产业技术体系资助项目(CARS-21)

作者简介: 张子璇(1996—),女,河北张家口人,硕士研究生,主要从事药用植物栽培研究。

①通信作者 E-mail: zxxiang@njau.edu.cn

引用格式: 张子璇, 谭明谱, 王 将, 等. 茅苍术二倍体和同源四倍体叶片转录组比较[J]. 植物资源与环境学报, 2021, 30(4): 41-49.

3 categories and 61 subcategories. KEGG enrichment analysis result shows that these differentially expressed unigenes are annotated to 5 categories, in which, the differentially expressed unigenes annotated to metabolism is the most (75) and enriched in 8 pathways, including biosynthesis of secondary metabolites, starch and sucrose metabolism, and fatty acid elongation, etc. There are 298 common high confidence differentially expressed genes in leaves and whole plants of diploid and autotetraploid of *A. lancea*, in which, the expression levels of heat shock protein gene (*DnaK/Hsp70*), ACC oxidase gene (*ACO*), fatty acid dehydrogenase gene (*FAD2*), and long chain-3-hydroxyacyl-CoA dehydratase gene (*PHS1/PAS2*) are up-regulated in autotetraploid, while those of ethylene receptor gene (*ETR/ERS*), allene oxide synthase gene (*AOS*), lipoxygenase gene (*LOX*), and 14-3-3 domain-containing protein gene (*YWHAE*) are down-regulated in autotetraploid. It is suggested that there are differences in cellular development, metabolism, energy conversion, material accumulation, and signal transduction of diploid and autotetraploid leaves of *A. lancea*, the obtained differentially expressed unigenes may be involved in leaf development regulation of *A. lancea*, and the obtained high confidence differentially expressed genes can be served as key candidate genes for polyploid breeding of *A. lancea*.

Key words: *Atractylodes lancea* (Thunb.) DC.; diploid; autotetraploid; differentially expressed unigene; functional annotation

多倍体植物具有基因剂量效应,其体内重复基因的表达发生变化,导致植物体在生长发育、形态、生理及环境胁迫耐性等方面表现出优势性状^[1]。多倍体育种是目前药用植物选育新品种的重要手段,很多药用植物已成功诱导出多倍体且获得的多倍体具有活性成分含量高、抗逆性强等特点。例如:穿心莲 [*Andrographis paniculata* (Burm. f.) Nees] 同源四倍体叶和茎中穿心莲内酯含量以及茎中脱水穿心莲内脂含量均高于二倍体^[2];甜叶菊 (*Stevia rebaudiana* Bert.) 四倍体的逆境适应能力较二倍体更强,能有效利用弱光和低浓度 CO₂^[3]。与二倍体相比,同源四倍体的叶片发生明显变化,如 *Platycodon grandiflorum* A. De Candolle 同源四倍体的叶片长度和宽度、气孔长度和宽度、超氧化物歧化酶和过氧化物酶活性、可溶性糖和蛋白质含量及总叶绿素含量均升高^[4]。

茅苍术 [*Atractylodes lancea* (Thunb.) DC.] 隶属于菊科 (Compositae) 苍术属 (*Atractylodes* DC.), 为多年生草本植物,其干燥根茎可入药,具有燥湿健脾、祛风散寒和明目的功效^[5],主要产地为江苏、湖北及安徽等省份,道地产区为江苏镇江茅山^[6]。近年来,由于掠夺性采挖和生境变化等原因,茅苍术野生资源濒临枯竭,开展茅苍术育种和品质改良研究迫在眉睫。

转录组测序技术能为植物基因表达研究提供数据资料,是研究未完成基因组测序物种的基因及基因功能挖掘的重要手段^[7-8]。利用转录组测序技术系统分析茅苍术二倍体和同源四倍体叶片的转录组信息,能进一步明确茅苍术染色体加倍后一系列性状变异的分子机制,为挖掘茅苍术的功能基因提供基础资

料。本项目组已经初步研究了不同倍性茅苍术组培苗间的表型差异及生理特征和转录组差异,结果表明:与二倍体相比,茅苍术同源四倍体叶片发生明显的表型变化,包括叶面积指数、叶长、叶宽、叶厚、叶绿素含量、气孔的横径和纵径以及保卫细胞中的叶绿体数显著增加,叶片的可溶性糖、可溶性蛋白质及可溶性淀粉含量升高^[9-10]。

鉴于此,以茅苍术二倍体和同源四倍体试管苗为研究对象,采用转录组测序技术对其叶片进行无参转录组测序分析,并在前期研究^[10]基础上筛选叶片中的高可信度差异表达 unigenes,以揭示茅苍术同源四倍体叶片硕大和抗逆性强的分子机制,并挖掘叶片内直接参与同源四倍体叶片性状变异和抗逆性的相关候选基因。

1 材料和方法

1.1 材料

供试材料为南京农业大学园艺学院中药生物技术实验室前期培养的茅苍术二倍体和同源四倍体试管苗。选取实验室内继代 60 d 且生长健壮的茅苍术二倍体和同源四倍体试管苗,在生根培养基 (添加 0.5 mg · L⁻¹ NAA 的 1/2MS 培养基) 中培养 20 d 后,分别选取长势基本一致的二倍体和同源四倍体试管苗各 5 株,每株采集基生叶片 2 枚,相同倍性的试管苗叶片混匀。按照上述方法重复取样 3 次。采集的叶片经液氮速冻后置于 -80 °C 冰箱中保存、备用。

1.2 方法

1.2.1 总 RNA 的提取和纯化 采用 TRIzol[®] Reagent (美国 Invitrogen 公司) 分别提取茅苍术二倍体和同源四倍体叶片的总 RNA; 使用 NanoDrop 2000 超微量分光光度计(美国 Thermo 公司) 检测总 RNA 的浓度和纯度; 采用 DNase I [天根生化科技(北京)有限公司] 纯化总 RNA, 去除总 RNA 中残留的基因组 DNA; 获得的总 RNA 经 Agilent 2100 生物分析仪(美国 Agilent 公司) 检测合格后用于后续分析。选取质量合格的总 RNA 样品(质量在 5 μg 及以上, 质量浓度在 200 $\text{ng} \cdot \mu\text{L}^{-1}$ 及以上, $\text{OD}_{260}/\text{OD}_{280}$ 为 1.8~2.2) 构建 cDNA 文库。

1.2.2 cDNA 文库构建和转录组测序 采用磁珠法^[11] 分离总 RNA 中的 mRNA; 采用 Fragmentation buffer(美国 Invitrogen 公司) 将获得的 mRNA 随机断裂成长度约 200 bp 的小片段; 以 mRNA 为模板, 采用 SuperScript II Reverse Transcriptase 试剂盒(美国 Agilent 公司) 反转录合成 cDNA 第 1 链, 并使用 $10\times$ DNA Polymerase buffer 和 DNA Polymerase I 试剂(美国 Promega 公司) 合成 cDNA 第 2 链。使用 End Repair Mix(美国 Enzymatics 公司) 将 cDNA 的粘性末端补成平末端, 随后在 3' 末端加上 1 个 A 碱基并连接测序接头; 采用 Illumina HiSeq 2000 高通量测序平台(美国 Illumina 公司) 测序, 获得原始测序数据。

1.2.3 测序数据过滤及组装 使用 Cutadapt v1.16 软件过滤原始测序数据, 除去污染、含 N 比高于 10% 和低质量的 reads, 获得高质量测序数据(clean data)。由于茅苍术无参考基因组序列, 使用 Trinity v2.6.6 软件对所有 clean data 进行从头组装, 去冗余后得到 unigenes。

1.2.4 基因功能注释和差异表达基因分析 使用 BLASTx v2.2.25 软件将得到的 unigenes 比对到 NR、String、SwissProt 和 KEGG 数据库, 分别进行基因功能注释; 使用 Blast2GO 软件(<http://www.blast2go.com>) 进行 GO 注释; 使用 HMMER v3.2.1 软件进行 Pfam 注释。

使用 edgeR v3.24 软件分析茅苍术二倍体和同源四倍体叶片基因的表达量差异, 将 $|\log_2\text{FC}| \geq 1$ 且 $\text{FDR} \leq 0.05$ 作为筛选差异表达基因的标准, 其中, FC 为差异倍数, FDR 为误判率。对筛选出的 unigenes 进行 GO 功能分类统计和 KEGG 富集分析; 比较项目组前期对茅苍术二倍体和同源四倍体试管苗全株^[10] 及本研究中茅苍术二倍体和同源四倍体试管苗叶片的基因差异表达情况, 据此筛选出叶片中的高可信度差异表达 unigenes。

2 结果和分析

2.1 转录组测序数据和组装结果分析

茅苍术二倍体和同源四倍体叶片转录组的测序分析结果(表 1) 表明: 原始 reads 有 448 453 428 条, clean reads 有 447 891 390 条, 所有样本的 Q20 均大于 98%, Q30 均大于 94%, GC 含量在 46% 以上。

对获得的 clean reads 进行组装和拼接, 共获得 750 700 个转录本数据, 总碱基数 364 718 682 bp, 最长 39 024 bp, 最短 201 bp, 平均长度 485.8 bp, N50 和 N90 分别为 585 和 239 bp; 去冗余后获得 495 797 个 unigenes, 这些 unigenes 的总碱基数为 199 717 026 bp, GC 含量 46.54%, 平均长度 402.8 bp, N50 和 N90 分别为 413 和 225 bp(表 2)。

表 1 茅苍术二倍体和同源四倍体叶片转录组测序分析

Table 1 Sequencing analysis on transcriptomes from leaves of diploid and autotetraploid of *Atractylodes lancea* (Thunb.) DC.

样本 ¹⁾ Sample ¹⁾	原始 reads 数 Number of original reads	Clean reads 数 Number of clean reads	过滤后碱基数/bp Base number after filtration	Q20/%	Q30/%	GC 含量/% GC content
S1	91 333 112	91 235 388	13 608 302 939	98.19	94.58	46.21
S2	87 022 318	86 917 894	12 968 986 207	98.22	94.57	46.69
S3	77 160 790	77 095 004	11 489 084 164	98.35	94.96	46.12
S4	77 178 726	77 079 860	11 466 591 535	98.34	95.03	46.59
S5	67 990 854	67 835 418	10 096 882 596	98.01	94.21	46.54
S6	47 767 628	47 727 826	7 113 735 712	98.15	94.46	46.79
总计 Total	448 453 428	447 891 390				

¹⁾ S1, S2, S3: 二倍体的 3 个生物学重复 Three biological replicates of diploid; S4, S5, S6: 同源四倍体的 3 个生物学重复 Three biological replicates of autotetraploid.

表2 茅苍术叶片转录本和 unigene 分析

Table 2 Analysis on transcript and unigene from leaves of *Atractylodes lancea* (Thunb.) DC.

项目 Item	总数 Total number	总碱基数/bp Total base number	GC 含量/% GC content	最长长度/bp The longest length	最短长度/bp The shortest length	平均长度/bp Average length	N50/bp	N90/bp
转录本 Transcript	750 700	364 718 682	44.84	39 024	201	485.8	585	239
Unigene	495 797	199 717 026	46.54	39 024	201	402.8	413	225

2.2 基因功能注释分析

将茅苍术转录组去冗余后获得的 unigenes 分别与 GO、NR、String、SwissProt、KEGG 和 Pfam 数据库进行比对,结果(表3)显示:共注释到 259 841 个 unigenes,占 unigenes 总数的 52.41%。其中,注释到 String 数据库的 unigenes 数最多(180 360),占 unigenes 总数的 36.38%;注释到 Pfam 数据库的 unigenes 数最少(66 311),占 unigenes 总数的 13.37%。

表3 茅苍术转录组 unigenes 与 6 个数据库的比对结果

Table 3 Comparison results of unigenes in transcriptomes of *Atractylodes lancea* (Thunb.) DC. with 6 databases

数据库 Database	注释 unigenes 数 Number of annotated unigenes	占 unigenes 总数的 比例/% Percentage of total unigenes
GO	153 547	30.97
NR	137 750	27.78
String	180 360	36.38
SwissProt	155 290	31.32
KEGG	80 662	16.27
Pfam	66 311	13.37
总计 Total	259 841	52.41

NR 数据库中茅苍术与其他植物转录组 unigenes 序列的比对结果(图1)显示:茅苍术与 *Cynara cardunculus* var. *scolymus* (Linn.) Fiori 的 unigenes 序列同源性最高,共有 28 614 个 unigenes 序列相似,占该数据库注释 unigenes 数的 20.77%;与向日葵 (*Helianthus annuus* Linn.) 的 unigenes 序列同源性较高,有 21 191 个 unigenes 序列相似,占该数据库注释 unigenes 数的 15.38%;与大麦 (*Hordeum vulgare* subsp. *vulgare* Linn.)、克里藻 [*Klebsormidium nitens* (Kützing) Lokhorst]、黄石斛 (*Dendrobium catenatum* Lindl.)、旋蒴苣苔 [*Boea hygrometrica* (Bunge) R. Br.]、衣藻 (*Chlamydomonas eustigma*)、小立碗藓 [*Physcomitrella patens* (Hedw.) Bruch et Schimp.]、蓖麻 (*Ricinus communis* Linn.) 和胶球藻 C-169 (*Coccomyxa subellipsoidea* C-169) 的 unigenes 序列同

源性较低,有 2 012~5 475 个 unigenes 序列相似,均占该数据库注释 unigenes 数的 4% 以下。另外,茅苍术与白梨 (*Pyrus bretschneideri* Rehd.)、甜菜 (*Beta vulgaris* subsp. *vulgaris* Linn.) 和绿藻 (*Ostreococcus tauri*) 等植物的 unigenes 序列同源性更低,均占该数据库注释 unigenes 数的 1.8% 以下。

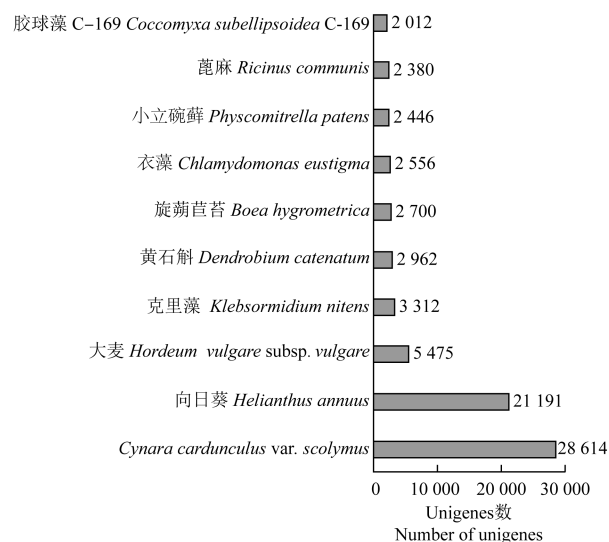


图1 NR 数据库中茅苍术与部分植物转录组 unigenes 序列的比对结果

Fig. 1 Comparison results of unigene sequences in transcriptomes of *Atractylodes lancea* (Thunb.) DC. and some plants in NR database

2.3 差异表达 unigenes 分析

比较茅苍术二倍体和同源四倍体叶片 unigenes 的表达量,筛选出 923 个差异表达 unigenes。与二倍体相比,同源四倍体中表达量上调的 unigenes 有 317 个,表达量下调的 unigenes 有 606 个。

2.3.1 GO 功能分类结果 GO 功能分类结果(表4)表明:茅苍术二倍体和同源四倍体叶片差异表达 unigenes 的功能被分成生物过程、细胞组分和分子功能 3 个大类 61 个小类。在生物过程大类中,注释为细胞过程的差异表达 unigenes 最多,共有 342 个;注释为代谢过程、应激反应、生物调节和生物过程调节

的差异表达 unigenes 较多, 分别有 307、249、219 和 201 个。在细胞组分大类中, 注释为细胞、细胞部分、细胞器、细胞器部分和细胞膜的差异表达 unigenes 较

多, 分别有 362、362、337、232 和 231 个。在分子功能大类中, 注释为结合和催化活性的差异表达 unigenes 较多, 分别有 291 和 210 个。

表4 茅苍术二倍体和同源四倍体转录组差异表达 unigenes 的 GO 功能分类

Table 4 GO functional classification of differentially expressed unigenes in transcriptomes of diploid and autotetraploid of *Atractylodes lancea* (Thunb.) DC.

功能分类 Functional classification	差异表达 unigenes 数 Number of differentially expressed unigenes		
	上调 Up-regulated	下调 Down-regulated	总计 Total
生物过程 Biological process			
细胞过程 Cellular process	189	153	342
代谢过程 Metabolic process	170	137	307
应激反应 Response to stimulus	139	110	249
生物调节 Biological regulation	124	95	219
生物过程调节 Regulation of biological process	115	86	201
发育过程 Developmental process	107	75	182
多细胞生物过程 Multicellular organismal process	107	72	179
细胞成分组织或生物合成 Cellular component organization or biogenesis	81	72	153
多生物过程 Multi-organism process	75	59	134
生殖 Reproduction	75	46	121
生殖过程 Reproductive process	73	46	119
生物过程正调控 Positive regulation of biological process	67	47	114
定位 Localization	52	59	111
信号 Signaling	62	49	111
生物过程负调控 Negative regulation of biological process	54	40	94
生长 Growth	35	33	68
免疫系统过程 Immune system process	28	28	56
细胞增殖 Cell proliferation	13	16	29
节律过程 Rhythmic process	11	9	20
运动 Locomotion	10	9	19
行为 Behavior	1	7	8
生物附着 Biological adhesion	4	2	6
细胞杀伤 Cell killing	4	1	5
氮利用 Nitrogen utilization	2	1	3
色素沉积 Pigmentation	2	1	3
参与化学突触传递的突触前过程 Presynaptic process involved in chemical synaptic transmission	2	1	3
细胞聚集 Cell aggregation	2	0	2
解毒作用 Detoxification	2	0	2
糖类利用 Carbohydrate utilization	1	0	1
碳利用 Carbon utilization	0	1	1
细胞组分 Cellular component			
细胞 Cell	192	170	362
细胞部分 Cell part	192	170	362
细胞器 Organelle	177	160	337
细胞器部分 Organelle part	119	113	232
细胞膜 Cytomembrane	119	112	231
细胞膜部分 Cytomembrane part	68	72	140
含蛋白质复合体 Protein-containing complex	54	46	100
膜封闭管腔 Membrane-enclosed lumen	49	40	89
细胞外区域 Extracellular region	39	36	75
细胞连接 Cell junction	34	37	71

续表4 Table 4 (Continued)

功能分类 Functional classification	差异表达 unigenes 数 Number of differentially expressed unigenes		
	上调 Up-regulated	下调 Down-regulated	总计 Total
	共质体 Symplast	32	34
细胞外区域部分 Extracellular region part	19	20	39
超分子复合物 Supramolecular complex	6	1	7
突触 Synapse	3	4	7
突触部分 Synapse part	3	3	6
类核 Nucleoid	1	2	3
其他有机体 Other organism	0	2	2
其他有机体部分 Other organism part	0	2	2
病毒粒子 Virion	0	1	1
病毒粒子部分 Virion part	0	1	1
分子功能 Molecular function			
结合 Binding	154	137	291
催化活性 Catalytic activity	110	100	210
转录调控因子活性 Transcription regulator activity	27	20	47
信号传感器活性 Signal transducer activity	16	11	27
结构分子活性 Structural molecular activity	15	9	24
转运活性 Transporter activity	8	15	23
分子功能调节剂 Molecular function regulator	11	5	16
分子载体活性 Molecular carrier activity	1	2	3
抗氧化活性 Antioxidant activity	1	0	1
营养库活性 Nutrient reservoir activity	1	0	1
劫持分子功能 Hijacked molecular function	0	1	1

2.3.2 KEGG 富集结果 茅苍术二倍体和同源四倍体共有 155 个差异表达 unigenes 注释到 KEGG 通路中,并对其进行 KEGG 分类和通路富集分析,结果分别见表 5 和表 6。结果显示:这些差异表达 unigenes 被注释到代谢、生物系统、遗传信息处理、环境信息处理和细胞过程 5 个大类中,以注释到代谢中的差异表达 unigenes 最多(75),占注释到 KEGG 通路的差异表达 unigenes 总数的 48.39%。这些差异表达 unigenes 富集在 8 条代谢通路中,包括次生代谢物生物合成(ko01110)、淀粉和蔗糖代谢(ko00500)、脂肪酸延伸(ko00062)、光合作用-天线蛋白(ko00196)、NF- κ B 信号通路(ko04064)、植物激素信号转导(ko04075)、植物 MAPK 信号通路(ko04016)和核糖体(ko03010)。

2.4 高可信度差异表达 unigenes 分析

筛选出的茅苍术二倍体和同源四倍体叶片和全株的共有高可信度差异表达 unigenes 有 298 个,其中,156 个共有高可信度差异表达 unigenes 有功能注释信息,而且其中的 58 个在 KEGG 代谢途径中有注释信息,包括热激蛋白基因(*DnaK/Hsp70*)、ACC 氧化

酶基因(*ACO*)、乙烯受体基因(*ETR/ERS*)、脂肪酸脱氢酶基因(*FAD2*)、丙二烯氧化物合酶基因(*AOS*)、脂氧合酶基因(*LOX*)、长链-3-羟酰基-CoA 脱水酶基因(*PHS1/PAS2*)和 14-3-3 结构域蛋白基因(*YWHAE*)。茅苍术二倍体和同源四倍体叶片和全株的这些共有高可信度差异表达 unigenes 的表达分析见表 7。结果表明:*DnaK/Hsp70*、*ACO*、*FAD2* 和 *PHS1/PAS2* 基因的表达量在同源四倍体中上调,而

表 5 茅苍术二倍体和同源四倍体转录组差异表达 unigenes 的 KEGG 分类
Table 5 KEGG classification of differentially expressed unigenes in transcriptomes of diploid and autotetraploid of *Atractylodes lancea* (Thunb.) DC.

分类 Classification	差异表达 unigenes Differentially expressed unigenes	
	数量 Number	比例/% ¹⁾ Percentage ¹⁾
	代谢 Metabolism	75
生物系统 Organismal system	50	32.26
遗传信息处理 Genetic information processing	42	27.10
环境信息处理 Environmental information processing	24	15.48
细胞过程 Cellular process	16	10.32

¹⁾ 占注释到 KEGG 通路的差异表达 unigenes 总数的比例 Percentage of total differentially expressed unigenes annotated to KEGG pathway.

表 6 茅苍术二倍体和同源四倍体转录组差异表达 unigenes 的 KEGG 通路富集

Table 6 KEGG pathway enrichment of differentially expressed unigenes in transcriptomes of diploid and autotetraploid of *Atractylodes lancea* (Thunb.) DC.

编号 ID	KEGG 通路 KEGG pathway	差异表达 unigenes Differentially expressed unigenes	
		数量 Number	比例/% ¹⁾ Percentage ¹⁾
ko01110	次生代谢物生物合成 Biosynthesis of secondary metabolites	34	45.33
ko00500	淀粉和蔗糖代谢 Starch and sucrose metabolism	9	12.00
ko00062	脂肪酸延伸 Fatty acid elongation	4	5.33
ko00196	光合作用-天线蛋白 Photosynthesis-antenna proteins	3	4.00
ko04064	NF- κ B 信号通路 NF- κ B signaling pathway	9	37.50
ko04075	植物激素信号转导 Plant hormone signal transduction	7	29.17
ko04016	植物 MAPK 信号通路 MAPK signaling pathway-plant	6	25.00
ko03010	核糖体 Ribosome	18	36.00

¹⁾ 占注释到代谢中差异表达 unigenes 总数的比例 Percentage of total differentially expressed unigenes annotated to metabolism.

表 7 茅苍术二倍体和同源四倍体叶片和全株部分共有高可信度差异表达 unigenes 的表达分析

Table 7 Analysis on expression of some common differentially expressed unigenes with high confidence in leaves and whole plants of diploid and autotetraploid of *Atractylodes lancea* (Thunb.) DC.

基因 Gene	叶片 Leaf		全株 ¹⁾ Whole plant ¹⁾		KEGG 通路 KEGG pathway
	编号 ID	log ₂ FC ²⁾	编号 ID	log ₂ FC ²⁾	
<i>DnaK/Hsp70</i>	TRINITY_DN240443_c1_g3	6.5	CL16556.Contig1	1.4	DnaK 伴侣/热激蛋白 70 Chaperone DnaK/heat shock protein 70
<i>ACO</i>	TRINITY_DN253517_c0_g1	4.1	Unigene26829	1.5	1-氨基环丙烷-1-羧酸氧化酶 1-aminocyclopropane-1-carboxylate oxidase
<i>ETR/ERS</i>	TRINITY_DN248948_c2_g6	-7.7	CL12271.Contig1	-13.1	乙烯受体 Ethylene receptor
<i>FAD2</i>	TRINITY_DN239013_c2_g4	4.9	CL895.Contig4	1.4	Δ 12 脂肪酸脱氢酶 Δ 12 fatty acid dehydrogenase
<i>FAD2</i>	TRINITY_DN252202_c0_g1	2.4	CL7290.Contig9	1.3	Δ 12 脂肪酸脱氢酶 Δ 12 fatty acid dehydrogenase
<i>FAD2</i>	TRINITY_DN253467_c1_g1	2.3	CL5079.Contig8	1.7	Δ 12 脂肪酸脱氢酶 Δ 12 fatty acid dehydrogenase
<i>FAD2</i>	TRINITY_DN240260_c2_g1	2.5	CL9030.Contig2	1.4	Δ 12 脂肪酸脱氢酶 Δ 12 fatty acid dehydrogenase
<i>FAD2</i>	TRINITY_DN252202_c0_g3	2.5	CL7290.Contig6	1.3	Δ 12 脂肪酸脱氢酶 Δ 12 fatty acid dehydrogenase
<i>FAD2</i>	TRINITY_DN253467_c1_g6	2.4	CL5079.Contig7	1.4	Δ 12 脂肪酸脱氢酶 Δ 12 fatty acid dehydrogenase
<i>AOS</i>	TRINITY_DN249223_c1_g1	-2.7	CL8770.Contig1	-3.5	丙二烯氧化物合酶 Allene oxide synthase
<i>LOX1_5</i>	TRINITY_DN253747_c1_g3	-3.6	CL6551.Contig6	-2.9	亚油酸 13S-脂氧合酶 2-1 Linoleate 13S-lipoxygenase 2-1
<i>LOX2S</i>	TRINITY_DN253747_c1_g1	-3.2	CL6551.Contig6	-2.9	亚油酸 13S-脂氧合酶 2-1 Linoleate 13S-lipoxygenase 2-1
<i>LOX2S</i>	TRINITY_DN250560_c1_g2	-2.6	CL6551.Contig9	-2.7	亚油酸 13S-脂氧合酶 2-1 Linoleate 13S-lipoxygenase 2-1
<i>LOX2S</i>	TRINITY_DN254019_c3_g3	-5.5	CL16911.Contig1	-7.7	亚油酸 13S-脂氧合酶 2-1 Linoleate 13S-lipoxygenase 2-1
<i>PHS1/PAS2</i>	TRINITY_DN244034_c1_g3	6.1	CL11187.Contig1	2.9	长链-3-羟酰基-CoA 脱水酶 Long chain-3-hydroxyacyl-CoA dehydratase
<i>YWHAE</i>	TRINITY_DN237061_c3_g2	-2.0	CL5058.Contig2	-7.6	14-3-3 结合域蛋白 14-3-3 domain-containing protein

¹⁾ 数据来自文献[10] The datums from reference [10].

²⁾ FC: 差异倍数 Fold change.

ETR/ERS、*AOS*、*LOX* 和 *YWHAE* 基因的表达量在同源四倍体中下调。

3 讨论和结论

本研究中, 茅苍术二倍体和同源四倍体的 unigenes 在 GO、NR、String、SwissProt、KEGG 和 Pfam 数据库的比对结果表明不是所有 unigenes 都能在已知数据库中比对出功能, 这一结果与目前测序的许多植物的相关研究结果相符^[11-13]。这主要是因为多数非模式植物尚未完成全基因组测序工作, 致使许多植物的基因组信息不全, 不可能通过数据库比对得知所

有 unigenes 的功能。值得注意的是, 与 NR 数据库比对结果显示, 茅苍术与同为菊科的 *C. cardunculus* var. *scolymus* 和向日葵的同源性较高, 说明同科植物的 unigenes 序列同源性更高。

植物多倍化后的性状差异主要由基因表达变化引起, 且基因表达只发生微小改变也可能引起表型变异^[14-15]。因此, 基因差异表达是茅苍术同源四倍体植株性状变异的基础。本研究发现, 注释为细胞过程和代谢过程以及细胞、细胞部分、细胞器、细胞器部分和细胞膜的茅苍术二倍体和同源四倍体叶片的差异表达 unigenes 较多, 说明茅苍术的倍性变化会影响其细胞发育和代谢; 还有部分差异表达 unigenes 注释为

信号和信号传感器活性,说明茅苍术同源四倍体叶片在信号转导方面与二倍体也存在差异。在其他植物多倍体研究中,酸枣〔*Ziziphus jujuba* var. *spinosa* (Bunge) Hu ex H. F. Chow.〕四倍体的差异表达基因富集在糖类和氨基酸代谢、信号传递以及能量代谢方面,并且这些基因的表达量在淀粉和蔗糖代谢及植物激素传导通路中显著上调^[15],杨树(*Populus* spp.)三倍体中参与糖类代谢和细胞生长的差异表达基因的表达量也上调^[16]。本研究的KEGG富集结果表明:茅苍术二倍体和同源四倍体的差异表达 unigenes 富集在脂肪酸延伸、淀粉和蔗糖代谢及次生代谢物生物合成等通路。淀粉、蔗糖和脂肪均为植物细胞活动提供能量,并且,在植物体内,淀粉和蔗糖代谢、脂肪酸延伸及次生代谢物生物合成过程均伴随着能量转化和物质积累。分析表明:差异表达基因可参与植物体的能量转化和物质积累过程,为细胞发育提供充足的能量和原料、催化相关酶活性,这也可能是茅苍术同源四倍体叶片较二倍体大且厚的一个重要原因。

植物在适应环境胁迫过程中需要能量、物质、激素和酶等共同发挥作用。在受到环境胁迫时,植物体会产生大量的转导信号,与淀粉和蔗糖等代谢过程相关的基因的表达量明显上调,体内的淀粉和蔗糖等代谢过程加强^[17],使植物体适应逆境条件。在本研究的GO功能分类中,注释到生物过程类的细胞过程、代谢过程、应激反应、生物调节和生物过程调节等分类的差异表达 unigenes 较多,且在KEGG富集结果中,部分差异表达 unigenes 富集于植物激素信号转导、植物MAPK信号通路和NF- κ B信号通路共3个信号转导相关通路中。其中,植物MAPK信号通路是将细胞膜表面信号放大并传递到细胞内,从而引起细胞内转录及代谢水平变化的信号转导途径之一,与植物生长发育和抗逆信号转导具有密切关系^[18]。茅苍术同源四倍体叶片在能量转化、物质积累和信号转导方面都与二倍体叶片存在差异,这些差异能够影响植物体对环境变化的响应,据此推断茅苍术同源四倍体在应对逆境胁迫时可能存在更强的适应能力,以保证整个植株包括其药用部位根状茎的正常生长。后续研究应对不同倍性茅苍术药用部位根茎进行转录组和代谢组联合分析,以揭示二者在外观性状、生长发育及抗逆性等方面的分子生物学差异机制,为进一步定位茅苍术的药用成分合成关键基因及通过基因工程手段进行茅苍术育种等奠定基础。

将茅苍术二倍体和同源四倍体叶片和全株^[10]共有的差异表达 unigenes 进行比较,共筛选出298个共有高可信度差异表达 unigenes,其中,叶片高可信度差异表达 unigenes (包括 *DnaK/Hsp70*、*ACO*、*ETR/ERS*、*FAD2*、*AOS*、*LOX*、*PHS1/PAS2* 和 *YWHAE* 等基因)可能与同源四倍体出现叶变大、变厚,叶色深绿,气孔增大等表型变化和植株抗逆性相关^[19-34]。本研究中,*DnaK/Hsp70*、*ACO*、*FAD2* 和 *PHS1/PAS2* 基因的表达量在茅苍术同源四倍体叶片和全株中均上调,而 *ETR/ERS*、*AOS*、*LOX* 和 *YWHAE* 基因的表达量在茅苍术同源四倍体叶片和全株中均下调,这可能是茅苍术同源四倍体植株抗逆性强于二倍体的原因。另外,茅苍术二倍体和同源四倍体叶片和全株^[10]非共有的差异表达 unigenes 也值得研究,这些基因可能调控其植株叶片和地下部(如根状茎)的生长发育。

本研究结果显示:茅苍术同源四倍体叶片在细胞发育、代谢、能量转化、物质积累及信号转导方面与二倍体间存在差异,差异表达 unigenes 可能参与叶片发育调控。高可信度差异表达基因 *DnaK/Hsp70*、*ACO*、*FAD2* 和 *PHS1/PAS2* 的表达量在茅苍术同源四倍体叶片和全株中均上调,而 *ETR/ERS*、*AOS*、*LOX* 和 *YWHAE* 的表达量均下调,这些基因直接参与茅苍术同源四倍体叶片发育和抗逆性,可作为茅苍术倍性育种的关键候选基因。

参考文献:

- [1] SATTTLER M C, CARVALHO C R, CLARINDO W R. The polyploidy and its key role in plant breeding [J]. *Planta*, 2016, 243: 281-296.
- [2] 林玉凤. 穿心莲同源四倍体诱导、鉴别及有效成分含量测定研究[D]. 广州: 广州中医药大学中药学院, 2016: 35.
- [3] 李雅婷, 王红娟, 向增旭. 甜叶菊二倍体与同源四倍体生理特征及 AFLP 分析[J]. *核农学报*, 2015, 29(11): 2103-2109.
- [4] XIANG Z X, LIANG H H, TANG X L, et al. Physiological and genetic diversity analysis of diploid and autotetraploid *Platycodon grandiflorum* A. De Candolle [J]. *Plant Cell, Tissue and Organ Culture*, 2019, 136: 597-603.
- [5] 国家药典委员会. 中华人民共和国药典: 2020年版(一部) [M]. 北京: 中国医药科技出版社, 2020: 161.
- [6] 许梦云. 道地药材茅苍术的遗传多样性分析[D]. 镇江: 江苏大学农业工程学院, 2009: 1.
- [7] 孙玲, 王月, 陈闽, 等. 薯蓣茎生珠芽发生相关的转录组和PEBP家族基因表达特性分析[J]. *植物资源与环境学报*, 2021, 30(3): 8-19.
- [8] LI C, WANG Y, HUANG X, et al. *De novo* assembly and

- characterization of fruit transcriptome in *Litchi chinensis* Sonn and analysis of differentially regulated genes in fruit in response to shading[J]. *BMC Genomics*, 2013, 14: 552.
- [9] 王红娟. 茅苍术同源四倍体离体诱导与鉴定及其遗传变异研究[D]. 南京: 南京农业大学园艺学院, 2015: 21.
- [10] 高真, 张成才, 罗丽娜, 等. 茅苍术二倍体及其同源四倍体的生理特征与转录组差异分析[J]. *南京农业大学学报*, 2020, 43(6): 1024-1032.
- [11] 马彦军, 段慧荣, 魏佳, 等. NaCl胁迫下黑果枸杞转录组测序分析[J]. *生物技术通报*, 2020, 36(2): 100-109.
- [12] 单春苗, 王晨凯, 施圆圆, 等. 多花黄精甾体皂苷生物合成途径分析及关键酶基因研究[J]. *中国中药杂志*, 2020, 45(12): 2847-2857.
- [13] 何江, 马艺沔, 杨伟俊, 等. 基于转录组测序挖掘新塔花黄酮类物质生物合成相关基因[J]. *中国中药杂志*, 2019, 44(15): 3178-3186.
- [14] BALAO F, HERRERA J, TALAVERA S. Phenotypic consequences of polyploidy and genome size at the microevolutionary scale: a multivariate morphological approach[J]. *New Phytologist*, 2011, 192: 256-265.
- [15] 李萌, 郭焯, 刘松珊, 等. 二倍体及其同源四倍体酸枣的生理特征和转录组分析[J]. *北京林业大学学报*, 2019, 41(7): 57-67.
- [16] CHENG S, HUANG Z, LI Y, et al. Differential transcriptome analysis between *Populus* and its synthesized allotriploids driven by second-division restitution[J]. *Journal of Integrative Plant Biology*, 2015, 57(12): 1031-1045.
- [17] 汤晓丽. 海滨锦葵盐胁迫转录组分析和抗盐基因研究[D]. 烟台: 中国科学院烟台海岸带研究所, 2016: 43.
- [18] 徐良伟, 吴小祝, 贾明良, 等. MAPK级联及其在植物抗病防卫反应中的研究进展[J]. *激光生物学报*, 2019, 28(6): 488-495.
- [19] HAQ S U, KHAN A, ALI M, et al. Heat shock proteins: dynamic biomolecules to counter plant biotic and abiotic stresses[J]. *International Journal of Molecular Sciences*, 2019, 20: 5321.
- [20] ONO K, HIBINO T, KOHINATA T, et al. Overexpression of DnaK from a halotolerant cyanobacterium *Aphanathece halophytica* enhances the high-temperature tolerance of tobacco during germination and early growth[J]. *Plant Science*, 2001, 160: 455-461.
- [21] SUGINO M, HIBINO T, TANAKA Y, et al. Overexpression of DnaK from a halotolerant cyanobacterium *Aphanathece halophytica* acquires resistance to salt stress in transgenic tobacco plants[J]. *Plant Science*, 1999, 137: 81-88.
- [22] ACHARD P, CHENG H, DE GRAUWE L, et al. Integration of plant responses to environmentally activated phytohormonal signals[J]. *Science*, 2006, 311: 91-94.
- [23] 李立芹, 王西瑶, 鲁黎明. 马铃薯ACC氧化酶参与低钾胁迫的功能研究[C]//中国作物学会马铃薯专业委员会, 湖北省农业农村厅, 恩施州人民政府. 马铃薯产业与健康消费(2019). 哈尔滨: 黑龙江科学技术出版社, 2019: 2.
- [24] IWAI T, MIYASAKA A, SEO S, et al. Contribution of ethylene biosynthesis for resistance to blast fungus infection in young rice plants[J]. *Plant Physiology*, 2006, 142: 1202-1215.
- [25] SHI Y, TIAN S, HOU L, et al. Ethylene signaling negatively regulates freezing tolerance by repressing expression of *CBF* and type-A *ARR* genes in *Arabidopsis*[J]. *The Plant Cell*, 2012, 24: 2578-2595.
- [26] CAO Y R, CHEN S Y, ZHANG J S. Ethylene signaling regulates salt stress response: an overview[J]. *Plant Signaling and Behavior*, 2008, 3(10): 761-763.
- [27] WANG Y, WANG T, LI K, et al. Genetic analysis of involvement of *ETR1* in plant response to salt and osmotic stress[J]. *Plant Growth Regulation*, 2008, 54: 261-269.
- [28] 冯莹莹, 张钟秀, 董先娟, 等. 白木香丙二烯氧化物合酶基因的表达分析[J]. *药学学报*, 2017, 52(12): 1962-1969.
- [29] ZHOU G, QI J, REN N, et al. Silencing *OsHI-LOX* makes rice more susceptible to chewing herbivores, but enhances resistance to a phloem feeder[J]. *The Plant Journal*, 2009, 60: 638-648.
- [30] 常丽丽, 王丹, 彭存智, 等. 木薯14-3-3蛋白家族成员 *MeGRF3* 基因的原核表达及多克隆抗体制备[J]. *西北植物学报*, 2020, 40(3): 375-381.
- [31] LIU Q, YANG J, ZHANG S, et al. *OsGF14b* positively regulates panicle blast resistance but negatively regulates leaf blast resistance in rice[J]. *Molecular Plant-Microbe Interactions*, 2016, 29(1): 46-56.
- [32] HERNÁNDEZ M L, PADILLA M N, SICARDO M D, et al. Effect of different environmental stresses on the expression of oleate desaturase genes and fatty acid composition in olive fruit[J]. *Phytochemistry*, 2011, 72: 178-187.
- [33] KARGIOTIDOU A, DELI D, GALANOPOULOU D, et al. Low temperature and light regulate *delta 12* fatty acid desaturases (*FAD2*) at a transcriptional level in cotton (*Gossypium hirsutum*) [J]. *Journal of Experimental Botany*, 2008, 59(8): 2043-2056.
- [34] 万静雅, 涂行浩, 魏芳, 等. 超长链脂肪酸的研究进展[J]. *中国食物与营养*, 2019, 25(8): 5-11.

(责任编辑: 佟金凤)