

黄秋葵花和果荚转录组测序及 类黄酮代谢差异表达分析

姚运法, 张少平, 练冬梅, 赖正锋, 黄慧明, 洪建基*

(福建省农业科学院 亚热带农业研究所, 福建漳州 363005)

摘要: 采用 Illumina Hi-seq 2500 转录组高通量测序, 构建黄秋葵花和果荚转录组文库, 并利用测序评估、差异基因功能注释等生物信息学方法进行分析。研究结果表明: (1) 分别获得黄秋葵花和果荚有效数据 7.12 Gb 和 8.14 Gb, 碱基百分比(Q30)均达到 91.0% 以上。(2) 获得差异表达基因(DEGs) 1 336 个, 其中上调基因 319 个, 下调基因 1 017 个。(3) 获得功能注释的基因有 1 131 个, GO 将 455 注释基因归成 41 个功能小类, 主要涉及代谢过程、催化活性、单一生物过程和细胞过程等过程; KOG 注释了 472 个 DEGs, 涉及 23 个功能分类, 其中与次生代谢直接相关过程 O 和 Q 类别获得 111 个注释结果; KEGG 将 372 个 DEGs 注释到 80 条代谢通路中, 获得 F3H、F3'5'H、DFR、ANR、ANS、GT、LAR 共 10 个关键差异基因, 其中 F3H、DFR 在黄秋葵花中表现上调效应, F3'5'H、ANR、LAR 在黄秋葵果荚中表现显著上调效应, ANS、GT 则分别在花和果荚中均有上调或下调效应。(4) 实时定量 PCR 分析显示, 其中 7 个差异表达基因得到的相对表达量与转录组表达谱分析趋势完全一致。(5) 类黄酮代谢途径分析表明, 黄秋葵花通过 F3H、DFR、ANS、GT 途径将 NAR 催化为生成花青素苷; 果荚则将 NAR 通过 F3'5'H 将催化为 DHM, 后在 FLS 催化下生成黄酮醇类物质等; 部分 NAR 在 F3H、DFR 催化下, 生成无色飞燕草素苷元, 再分别在 ANS、LAR 作用下, 进入花青素苷元和原花青素合成途径。该研究结果丰富了黄秋葵转录组信息, 为黄秋葵类黄酮物质纯化和功能利用提供参考依据。

关键词: 黄秋葵; 转录组; 功能注释; 差异表达基因; 黄酮代谢

中图分类号: Q786; Q789; S641.9 **文献标志码:** A

Transcriptome Sequencing and Differential Expression Analysis of Flavonoid Metabolism in Flowers and Fruits of Okra

YAO Yunfa, ZHANG Shaoping, LIAN Dongmei, LAI Zhengfeng, HUANG Huiming, HONG Jianji*

(Subtropical Agriculture Research Institute, Fujian Academy of Agricultural Sciences, Zhangzhou, Fujian 363005, China)

Abstract: The cDNA of okra (flowers and fruits) were sequenced based on Illumina Hi-seq 2500 and were analyzed by using the bioinformatics methods subsequently, such as sequencing assess and gene function annotation. The result of research shows that: (1) 7.12 Gb and 8.14 Gb Clean Data were obtained respectively from the flowers and fruits and the base ratios with quality values higher than 30 in reads (Q30) were more than 91.0% from both. (2) Compared the transcriptome of the flowers and fruits, 1 131 differentially expressed genes (DEGs) were found, including 319 up-regulated genes and 1 017 down-regulated genes. (3) Annotation analysis indicated that 1 131 genes were annotated. With GO function annotation

收稿日期: 2018-08-09; **修改稿收到日期:** 2018-11-05

基金项目: 福建省公益类项目(2016R1012-1); 福建省农业科学院芳香植物创新科技团队(STIT2017-2-11)

作者简介: 姚运法(1982-), 男, 硕士, 助理研究员, 主要从事黄秋葵遗传育种与功能成分利用研究。E-mail: 124128532@qq.com

* 通信作者: 洪建基, 研究员, 主要从事黄秋葵遗传育种与综合利用研究。E-mail: 76227508@qq.com

classification, 455 functional annotations of these DEGs were divided into 41 function categories, in which many function categories were mainly involved, such as metabolic process, catalytic activity, single-organism process, cellular process. Through KOG analysis, there were 472 functional annotations of DEGs, involving 23 functional classifications. KEGG analysis showed that 372 DEGs were annotated to 80 metabolic pathways, and obtained 10 key DEGs, F3H, F3'5'H, DFR, ANR, ANS, GT, LAR. F3H and DFR showed up-regulated effect in okra flowers; F3'5'H, ANR and LAR showed down-regulated effect in okra fruits; ANS and GT showed up/down-regulated effects in flowers and fruits. (4) The relative expression levels of 7 DEGs was in line with that of the transcriptional group by qRT-PCR. (5) Through the flavonoid metabolic pathway, the research showed that anthocyanin were catalyzed by F3H, DFR, ANS(c83240), GT (c82265/c82004) from NAR in okra flowers; The formation of DHM was initiated by F3'5'H, then produced flavonols by FLS; Partly NAR was catalyzed by F3H and DFR to dephinidin, then ultimately went into anthocyanins and proanthocyanidins synthesis pathways. This study enriched transcriptome information of okra and provided reference for the purification and functional utilization of flavonoids in okra.

Key words: okra; transcriptome; function annotation; DEGs; flavonoid metabolism

黄秋葵,学名咖啡黄葵 [*Abelmoschus esculentus* (Linn.) Moench], 为锦葵科 (Malvaceae) 秋葵属 (*Abelmoschus*) 一年或多年生草本植物。原产于非洲,自 20 世纪 90 年代初引入中国,现国内各地均有栽培。黄秋葵花和果荚作为重要开发价值的器官,富含蛋白质、果胶多糖、总黄酮等营养成分^[1-4],其花主要开发黄秋葵花茶,经济附加值高^[5];果荚具有降血压、血脂和提高机体抗疲劳等功效,在黄秋葵保健品开发方面具有巨大潜力^[6]。当前,国内对黄秋葵研究主要从栽培技术、营养成分提取、功效分析和产品加工^[7-10]等方面,但其基础理论研究方面还很薄弱,例如花、果荚中类黄酮物质等主要功能成分及相关代谢途径等均属研究空白。

类黄酮(flavonoids)属于植物重要次生代谢产物之一,是指 2 个苯环(A-与 B-)通过 3 个碳原子连接形成具有 C6-C3-C6 基础结构的一类化合物^[11],由于其纯净状态呈现黄色,故称黄酮。研究表明,类黄酮物质对人体具有抗癌、抗氧化、抗动脉硬化等功能^[12-13],据其结构的差异将类黄酮主要分为黄烷酮(flavanols)、异黄酮(isoflavones)、黄酮(flavones)、黄酮醇(flavonols)、二氢黄酮(2H-flavanones)与花色素(anthocyanidins)等 6 大类^[14]。

转录组测序技术是连接基因组与代谢组的重要纽带,本团队前期对紫色黄秋葵叶片^[15]、果荚^[16]转录组测序和分析,初步探讨了黄秋葵次生代谢物质合成的遗传基础。本研究利用 Illumina Hi-seq 2500 高通量测序技术,根据黄秋葵花、果荚的转录组数据及其类黄酮代谢途径关键基因分析,探讨黄秋葵花、果荚类黄酮合成机理和关键差异表达基因,为后续黄秋葵关键基因克隆和功能验证、遗传改良

和加工利用等提供研究基础。

1 材料和方法

1.1 试验材料

试验材料为黄秋葵品种‘闽秋葵 3 号’,种植于福建省农业科学院亚热带农业研究所试验农场,2017 年 6 月 10 日种植,8 月 15 日开始采集样品(花/果荚)。采集同一株黄秋葵植株花朵(上午 9:00)和果荚(开花后 8 d),花朵只保留完整花瓣,果荚去除果柄部位。花和果荚均取 3 重复,后立即液氮速冻,置于-70 °C 超低温冰箱备用。

1.2 试验方法

1.2.1 高通量测序及数据组装 提取黄秋葵花、果荚总 RNA,分别采用 Nanodrop、Qubit2.0、Agilent 2100 技术检测 RNA 样品的纯度、浓度和完整性等,构建 cDNA 文库,再分别使用 Qubit2.0、Agilent 2100 和 Q-PCR 对文库的质量进行检测。合格后,用 Illumina HiSeqTM2500 进行测序。测序读长为 PE125。获得原生数据,进行数据过滤,去除 Reads 中测序引物、接头等人工序列,用 Trinity 对有效数据(Clean Data)进行组装。

将测序 Reads 构建 K-mer 库,去除错误的 K-mer;将高频率的 K-mer 作为种子向两端进行扩展,不断循环直到耗光 K-mer 库;再对 Contig 进行聚类,得到片段集合(Component);对每个 Component 中的 Contig 构建 De Bruijn 图;De Bruijn 图进行简化;解开 De Bruijn 图,获得转录本序列。

1.2.2 基因 FPKM 值估计与差异表达分析 采用 Bowtie^[17]将测序得到的 Reads 与单基因数据(Uni-gene)库进行比对,根据比对结果,结合 RSEM^[18]进

行表达量水平估计。利用 FPKM 值 (fragments Per Kilobase of transcript per Million mapped reads)^[19] 表示对应 Unigene 的表达丰度。使用 EB-Seq 进行差异表达分析^[20]。采用 Benjamini-Hochberg 方法对原假设试验得到的显著性 P 值进行校正,校正后 P 值,即伪发现率 (false discovery rate) 小于 0.01 且差异倍数 (fold change, FC) ≥ 2 作为筛选标准^[21]。其中,FC 表示两样品(组)间 FPKM 的比值。

1.2.3 差异基因的功能注释与代谢途径富集分析

使用 Blast 软件将 Unigene 序列与 NR、Swiss-Pro、GO、COG、KOG、KEGG 数据库比对,再用 KOBAS2.0 得到 Unigene 在 KEGG 中的 Orthology,预测氨基酸序列,使用 HMMER 软件与 Pfam 数据库分析,获得 Unigene 的注释信息。系统分析基因产物代谢途径及功能,并将差异表达基因 (DEGs) 对比到 KEGG 数据库,得到 DEGs 的代谢途径。

1.2.4 类黄酮代谢关键差异基因分析 利用注释信息检索法,对黄秋葵花、果荚类黄酮(黄烷酮、黄酮、异黄酮、黄酮醇、黄烷酮和花青素)关键词进行数据库检索,分析类黄酮物质关键基因在 KEGG 数据

库功能注释及代谢路径。

1.2.5 类黄酮代谢关键差异基因验证 取 1 μg 黄秋葵花或果荚的总 RNA,利用反转录试剂盒反转录成 cDNA,设计合成引物(表 1),采用 qRT-PCR 检测黄秋葵花、果荚组织部位与类黄酮代谢相关差异表达基因,设置 3 个重复;统计 8 个基因(含 1 个内参)在待测样品中的 Ct 值,计算相对表达量。

2 结果与分析

2.1 数据组装及分析

经高通量测序和质量控制,共获得 15.26 Gb 有效数据,其中黄秋葵花为 7.12 Gb,果荚为 8.14 Gb,碱基百分比均达到 91.0% 以上(表 2)。数据质量好,适宜后续分析。对组装结果进行统计(表 3),转录本序列组装出 154 721 条转录本序列,平均长度为 937.5 bp,N50 长 1 308 bp;由单基因序列(Unigene)组装出 65 436 条 Unigene,平均长度为 782.32 bp,N50 长 1 215 bp。

2.2 Unigene 功能注释

经与 NR、Swiss-Prot、KEGG、COG、KOG、GO 和 Pfam 数据库的比对,对 Unigene 功能注释进行

表 1 实时荧光定量 PCR 引物

Table 1 Primers for real-time quantitative PCR

基因 Gene	名称 Name	简称 Abbreviation	引物 Primer (5'→3')	
			正向 Forward	反向 Reverse
c84471	内参基因 Intrinsic genes	/	GACCATGCCGAATAAGTTGG	ATGCTCTCCTCTGCCCTTCTG
c55984	3-黄烷酮羟化酶 Flavanone 3-hydroxylase	F3H	CGACCAGATCGAGATCCTAAG	CCAAGAATACTGGCCATGAC
c82420	类黄酮 3',5'-羟化酶 Flavonoid 3',5'-hydroxylase	F3'5'H	ACGAGTCCTTCGGCCTTGC	AACAACGAGCATGGATCTGG
c76561	二氢黄酮醇 4-还原酶 Dihydroflavonol-4-reductase	DFR	CATTGACGCAAGACTCGACG	CATGCCTAGAATCACCAACC
c78393	花青素还原酶 Anthocyanidin reductase	ANR	CATGTGGAAGATGTATGCCG	CTTAGCGAGTTCAGGTACAC
c44799	花青素 3-O-葡萄糖基转移酶 Anthocyanidin 3-O-glucosyltransferase	GT	ATCGGAGGTATATGGAGTGG	CCTTCATGGAGTTAGAATCTC
c89149	无色花色素还原酶 Leucoanthocyanidin reductase	LAR	CAGTGAGCAAGCCGGTGAC	AATGCATGACAAGTACTACAAG
c83240	花青素合酶 Anthocyanidin synthase	ANS	GGAATGTCGGCATGTATGTC	GTGCGAGTTCAGTCTCTGTC

表 2 有效数据评估统计

Table 2 Statistics of clean data

样品 Sample	编号 Number	读数 Read	碱基 Base	GC 含量 GC content/%	碱基百分比 Q30/%
花 Flowers	T01	24 166 557	7 118 160 132	45.01	94.67
果荚 Fruits	T02	27,209,021	8,143,851,148	44.12	93.47

统计(表 4),在 65 436 条单基因序列中,共获得 39 245 条 Unigene 的注释结果,占单基因序列总数 59.97%。其中与 COG 数据库比对,获得 9 366 条同源序列,占单基因序列总数 14.31%;与 KOG 数据库比对,获得 20 535 条同源序列,占单基因序列总数 31.38%;与 GO 数据库比对,获得 15 931 条同源序列,占单基因序列总数 24.35%;与 KEGG 数据库比对,获得 14 110 条同源序列,占单基因序列总数 21.56%;与 Pfam 数据库比对,获得 22 580 条同源序列,占单基因序列总数 34.51%;与 Swissprot 数据库比对,获得 24 494 条同源序列,占单基因序列总数 37.43%;与 NR 数据库比对,获得 38 905 条同源序列,占单基因序列总数 59.46%。

2.3 DEGs 的筛选与功能注释

2.3.1 DEGs 筛选

通过 DEGs 的筛选,获得黄秋葵与果菜 DEGs 1 336 个,其中表达量上调的基因 319 个,下调基因 1 017 个;将 DEGs 单基因序列分别注释到 COG、GO、KEGG、KOG、Pfam、Swiss-

Prot、NR7 大数据库,共有 1 131 个基因获得功能注释,其中 GO 数据库 455 个,COG 数据库 281 个,KOG 数据库 472 个,KEGG 数据库 372 个,Pfam 数据库 844 个,Swiss-Prot 数据库 807 个,NR 数据库 1 123 个,其中 NR 数据库注释比高最高,达 99.29%。

2.3.2 GO 功能注释

GO 数据库分为 3 大类,分别为 B 生物学过程(biological process),C 细胞组成(cellular component)和 M 分子功能(molecular function),分别用来描述基因编码产物所参加的生物过程、所具有的分子功能和所处的细胞环境等^[22]。对黄秋葵、果菜进行 GO 分类统计显示,455 个 DEGs 被归到 41 个功能小类(表 5)。生物学过程中 DEGs“代谢过程”、“单一生物过程”和“细胞过程”3 个功能小类占比最高;细胞组分过程中 DEGs 在“细胞组分”、“细胞”和“细胞器”3 个功能小类占比最高;分子功能过程中 DEGs 在“催化活性”和“结合活性”2 个功能小类占比最高。

2.3.3 KOG 功能注释

将注释到 KOG 数据库的

表 3 组装结果统计分析
Table 3 Statistics of assembled results

长度区间 Length range/bp	转录本序列数量 Transcript number	总数占比 The total ratio/%	单基因序列数量 Unigene number	总数占比 The total ratio/%
200~300	22 735	14.69	16 481	25.19
300~500	31 864	20.59	17 400	26.59
500~1 000	46 183	29.85	15 604	23.85
1 000~2 000	40 518	26.19	10 994	16.80
2 000+	13 421	8.67	4 957	7.58
总数 Total number	154 721	/	65 436	/
总长度 Total length	145 051 189	/	51 192 070	/
N50 长度 N50 length	1 308	/	1 215	/
平均长度 Mean length	937.50	/	782.32	/

表 4 单基因序列注释统计表
Table 4 Unigene functions annotated

生物信息数据库 Bioinformation database	注释单基因数 Unigene annotated	单基因长度 Unigene length (300≤length<1 000)	单基因长度 Unigene length (length>1 000)	百分比 Percentage/%
COG	9 366	3 056	5 463	14.31
KOG	20 535	8 887	9 036	31.38
GO	15 931	6 800	7 065	24.35
KEGG	14 110	6 342	5 634	21.56
Pfam	22 580	8 456	12 199	34.51
Swiss-Prot	24 494	10 784	10 651	37.43
NR	38 905	18 399	14 927	59.46
总注释单基因序列数 All unigenes annotated	39 245	18 574	14 941	59.97

表 5 差异表达基因 GO 功能注释

Table 5 DEGs GO annotation

	GO 层级 GO classify	单基因序列数量 Number of unigene
生物学过程 Biological process	生物调节 Biological regulation	113
	细胞成分或生物合成 Cellular component organization or biogenesis	66
	细胞过程 Cellular process	223
	发育过程 Developmental process	67
	生长 Growth	22
	免疫系统过程 Immune system process	7
	定位 Localization	67
	运动 Locomotion	1
	代谢过程 Metabolic process	287
	多生物过程 Multi-organism process	22
	多细胞生物过程 Multicellular organismal process	39
	繁殖 Reproduction	6
	繁殖过程 Reproductive process	31
	刺激应答 Response to stimulus	126
	节律过程 Rhythmic process	1
	信号传导 Signaling	30
单一生物过程 Single-organism process	225	
细胞组分 Cellular component	细胞 Cell	172
	细胞连接 Cell junction	10
	细胞成分 Cell part	173
	胞外基质 Extracellular matrix	1
	胞外区 Extracellular region	36
	胞外区域部分 Extracellular region part	1
	复杂大分子 Macromolecular complex	21
	膜结构 Membrane	93
	膜结构部分 Membrane part	44
	细胞器 Organelle	112
细胞器部分 Organelle part	26	
分子功能 Molecular function	抗氧化活性 Antioxidant activity	5
	结合活性 Binding	196
	催化活性 Catalytic activity	254
	电子载体活性 Electron carrier activity	7
	酶调节活性 Enzyme regulator activity	6
	鸟嘌呤核苷酸交换因子活性 Guanyl-nucleotide exchange factor activity	1
	分子转导活性 Molecular transducer activity	5
	核酸结合转录因子活性 Nucleic acid binding transcription factor activity	22
	营养存储活性 Nutrient reservoir activity	4
	蛋白结合转录因子活性 Protein binding transcription factor activity	2
	受体活性 Receptor activity	3
	结构分子活性 Structural molecule activity	3
	转运活性 Transporter activity	28

472 个 DEGs 进行直系同源分类,获得 23 个功能分类,其中功能类别为 R(一般功能预测),获得 96 个注释结果,占比 20.33%;O(次生代谢产物生物合成、转运和代谢),获得 67 个注释结果,占比 14.20%;T(信号转导机制)获得 52 个注释结果,占比 11.01%;Q(次生代谢产物的合成、转运和代谢)获得 44 个注释结果,占比 9.32%。另外在 G(碳水化合物转运与代谢)、K(转录)也分别获得 44 和 39 个注释结果,分别占比 9.32%和 8.26%(表 6)。

2.3.4 KEGG 功能注释 将 DEGs 通过 KEGG 数据库比对(表 7),有 372 条基因得到注释,分别富集在 80 条代谢通路,包括植物激素信号转导(39 个)、淀粉和蔗糖代谢(33 个)、戊糖、葡萄糖醛酸转换(25 个)、氨基酸的生物合成(16 个)、类苯基丙烷生物合

成(14 个)和碳代谢(14 个)等 22 条代谢通路。本研究着重选择类苯基丙烷生物合成代谢通路,寻找类黄酮代谢差异发生的关键基因。

2.4 类黄酮代谢关键差异基因分析

据 KEGG 类黄酮代谢途径和表 8 数据分析,黄秋葵类黄酮合成代谢途径中,查尔酮合酶(chalcone synthase, CHS)和查尔酮-黄烷酮异构酶(chalcone isomerase, CHI)基因在花、果菜中均无差异表达;3-黄烷酮羟化酶(flavanone 3-hydroxylase, F3H)、二氢黄酮醇 4-还原酶(Dihydroflavonol-4-reductase, DFR)在黄秋葵花中具有显著表达优势;类黄酮 3',5'羧化酶(flavanone 3',5'-hydroxylase, F3'5'H)、花青素还原酶(anthocyanidin reductase, ANR)、无色花色苷还原酶基因(leucoanthocyanidin

表 6 差异表达基因 KOG 功能注释

Table 6 DEGs KOG annotation

COG	功能分类 Function classification	KOG 单基因序列数量 KOG numbers of unigene
A	RNA 的加工和修饰 RNA processing and modification	8
B	染色体结构和动态变化 Chromatin structure and dynamics	1
C	能源的产生和转化 Energy production and conversion	21
D	细胞周期调控,细胞分裂,染色体分配 Cell cycle control, cell division, chromosome partitioning	14
E	氨基酸转运与代谢 Amino acid transport and metabolism	29
F	核苷酸转运与代谢 Nucleotide transport and metabolism	3
G	碳水化合物转运与代谢 Carbohydrate transport and metabolism	44
H	辅酶转运与代谢 Coenzyme transport and metabolism	2
I	脂质转运与代谢 Lipid transport and metabolism	27
J	翻译、核糖体结构和生物合成 Translation, ribosomal structure and biogenesis	6
K	转录 Transcription	39
L	复制,重组和修复 Replication, recombination and repair	8
M	细胞壁/细胞膜的生物发生 Cell wall/membrane/envelope biogenesis	3
N	细胞运动 Cell motility	0
O	次生代谢产物的合成、转运和代谢 Posttranslational modification, protein turnover, chaperones	67
P	无机离子转运与代谢 Inorganic ion transport and metabolism	18
Q	次生代谢产物的合成、转运和代谢 Secondary metabolites biosynthesis, transport and catabolism	44
R	一般功能预测 General function prediction only	96
S	未知功能 Function unknown	16
T	信号转导机制 Signal transduction mechanisms	52
U	胞质运输、分泌和囊泡运输 Intracellular trafficking, secretion, and vesicular transport	7
V	防御机制 Defense mechanisms	7
Z	细胞骨架 Cytoskeleton	22

表 7 差异表达基因 KEGG 功能注释

Table 7 Differential expression genes KEGG function annotation

编号 Number	KEGG 功能注释 KEGG function annotation	单基因数列数量 Number of unigene
ko00020	三羧酸循环 Citrate cycle (TCA cycle)	6
ko00040	戊糖、葡萄糖醛酸转换 Pentose and glucuronate interconversions	25
ko00052	半乳糖代谢 Galactose metabolism	5
ko00260	甘氨酸, 丝氨酸和三氨酸代谢 Glycine, serine and threonine metabolism	7
ko00270	半胱氨酸和蛋氨酸代谢 Cysteine and methionine metabolism	5
ko00330	精氨酸和脯氨酸代谢 Arginine and proline metabolism	8
ko00360	苯丙氨酸代谢 Phenylalanine metabolism	9
ko00460	氰基氨基酸代谢 Cyanoamino acid metabolism	5
ko00500	淀粉和蔗糖代谢 Starch and sucrose metabolism	33
ko00520	氨基糖和核苷酸糖代谢 Amino sugar and nucleotide sugar metabolism	13
ko00564	甘油磷脂代谢 Glycerophospholipid metabolism	5
ko00592	α 亚麻酸代谢 α -Linolenic acid metabolism	6
ko00630	二羧酸代谢 Glyoxylate and dicarboxylate metabolism	9
ko00906	类胡萝卜素生物合成 Carotenoid biosynthesis	7
ko00940	类苯基丙烷生物合成 Phenylpropanoid biosynthesis	14
ko01200	碳代谢 Carbon metabolism	14
ko01230	氨基酸的生物合成 Biosynthesis of amino acids	16
ko03030	核苷酸修复 DNA replication	5
ko04075	植物激素信号转导 Plant hormone signal transduction	39
ko04146	过氧化物酶体 Peroxisome	5
ko04626	植物病原菌互作 Plant-pathogen interaction	7
ko04712	昼夜节律—植物 Circadian rhythm-plant	7

表 8 黄秋葵花、果荚类黄酮代谢关键基因差异表达情况

Table 8 Expression of the key genes in flowers and fruits of okra

酶 Enzyme	简称 Abbreviation	花 FPKM	果荚 FPKM	基因表达 Genes expression	基因 ID Gene ID
EC:1.14.11.23	F3H	5.20	0	Down	c55984
EC:1.14.13.88	F3'5'H	0.12	957.20	Up	c82420
EC:1.1.1.219	DFR	944.09	0.28	Down	c76561
EC:1.3.1.77	ANR	0.06	953.46	Up	c78393
EC:2.4.1.111	GT	0	45.37	Up	c44799
EC:2.4.1.111	GT	162.25	3.15	Down	c82004
EC:2.4.1.111	GT	407.13	0.16	Down	c82265
EC:1.17.1.3	LAR	1.46	433.46	Up	c89149
EC:1.14.11.9	ANS	133.65	0.74	Down	c83240
EC:1.14.11.9	ANS	0.12	260.27	Up	c94125

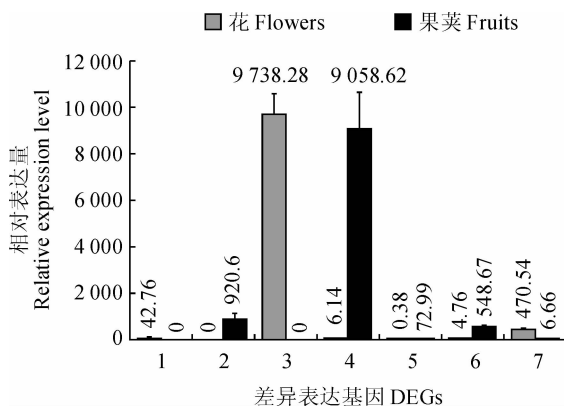
reductase, LAR)则在黄秋葵果荚中具有显著表达优势;而花青素苷合成酶(anthocyanidin, ANS)、葡萄糖基转移酶(glycosyl transferases, GT)基因则

分别在花、果荚中具有显著表达。黄秋葵花、果荚中类黄酮代谢表现为:P-香豆素-CoA(P-coumaroyl-CoA)和3分子丙二醛-CoA(malnoyl-CoA),在

CHS 催化下,生成查尔酮(chalcone),查尔酮又在 CHI 作用下形成柚皮素(naringenin, NAR),此过程花与果荚均无显著差异。黄秋葵花中 NAR 在 F3H 催化下生产二氢山奈酚(dihydokaempferol, DHK),DHK 在 DFR 作用下,生成无色天竺葵苷元,后在 ANS(c83240)作用下生成花青素苷元,花青素苷元分别在 GT(c44799/c82004)等转移酶的作用下,生成稳定的花青素苷;黄秋葵果荚则在 F3'5'H 作用下将 NAR 生成二氢杨梅素(dihydromyricetin, DHM),后在 FLS 催化下,进入黄酮醇代谢途径,部分 DHM 在 DFR、ANS(c94125)、GT(c44799)作用下生成飞燕草素苷元(delphinidin),飞燕草素苷元在 ANR 作用下,进入原花青素代谢途径,无色飞燕草素苷元在 LAR 催化下也进入原花青素代谢途径。另外,GT、ANS 在黄秋葵花和果荚中分别有 3 个和 2 个差异表达拷贝,且表达量在不同组织中具有显著互补性。

2.5 类黄酮代谢关键差异基因验证分析

将表 10 中部分 DEGs 进行荧光实时定量(RT-PCR),其中 GT 和 ANS 基因随机选取多拷贝中 1 个,共 7 个。差异表达基因中 c82420、c78393 和 c44799 3 个基因表达量上调;c55984、c76561、c89149 和 c83240 4 个基因表达量下调,以黄秋葵 c84471 基因为内参基因,进行 qRT-PCR 验证。由图 1 可知,7 个基因 qRT-PCR 分析得到的相对表达量与转录组表达谱分析趋势完全一致,但表达的变化大小存在一定差异,说明基因表达谱的分析结果基本可靠,其中 DFR 和 ANR 基因分别特异性在黄秋葵花、果荚中表达,且表达量差异极大。



1. c55984;2. c82420;3. c76561;4. c78393;5. c44799;
6. c89149;7. c94125

图 1 差异表达基因 qRT-PCR 相对表达
Fig. 1 The relative expression levels
of DEGs by qRT-PCR

3 讨论

通过 Illumina HiSeqTM2500 测序技术构建黄秋葵花、果荚转录组数据库,花和果荚分别获得 7.12 Gb 和 8.14 Gb 有效数据,碱基百分比(Q30)均达到 91.0% 以上。对 Unigene 进行功能注释,在 65 436 条单基因序列中,共获得 39 245 条 Unigene 的注释结果,占单基因序列总数 59.97%。通过 DEGs 分析,获得差异基因 1 336 个,其中上调基因 319 个,下调基因 1 017 个。获得功能注释基因有 1 131 个,GO 数据库将 455 个 DEGs 归到 41 个功能小类,代谢过程、催化活性、单一生物过程和细胞过程等 4 个功能小类占比最高;KOG 数据库将 472 个 DEGs 进行直系同源分类,获得 23 个功能分类,其中与次生代谢直接相关过程 O 和 Q 类别获得 111 个注释结果,合计占比 23.52%;通过 KEGG 数据库比对,将 372 个 DEGs 注释到 80 条代谢通路上,其中富集在类苯基丙烷生物合成途径有 14 个 DEGs,再通过类黄酮关键基因分析,共获得 10 个关键差异基因。

植物类黄酮合成代谢是目前研究最清楚的此生代谢路径之一^[23],该通路上存在 2 个重要基因群,即上游基因群(early biosynthetic genes, EBGs)与下游基因群(late biosynthetic genes, LBGs)。EBGs 主要包括 CHS、CHI、F3H、F3'H、F3'5'H 等基因^[24-25];LBGs 主要包括 DFR、FLS、ANS、ANR、GT、酰基转移酶(Acyl transferase, AT)和甲基转移酶(Aethyl transferase, MT)等^[26]。本研究发现,前体合成阶段关键酶(CHS、CHI)均无显著差异表达,这与 CHS 和 CHI 基因编码区和结构都十分保守,在不同科植物间、不同组织部位上均具有较高的保守性相一致^[27]。黄秋葵花、果荚类黄酮代谢路径中发现 F3H、F3'5'H、DFR、ANR、ANS、LAR 和 GT 存在差异表达,其中 F3H、DFR 在黄秋葵花中表现上调效应,F3'5'H、ANR、LAR 在黄秋葵果荚中表现显著上调效应,ANS、GT 则分别在花和果荚中均有上调或下调效应。NAR 作为类黄酮合成的关键分支点,黄秋葵花通过 F3H、DFR、ANS(c83240)、GT(c82265/c82004)途径,生成紫红色花色素-3-葡萄糖苷(pelagonidin-3-glucoside)^[28],这与黄秋葵花喉深红色表现可能有关。黄秋葵果荚则通过 F3'5'H、FLS 催化下,将 NAR 生成杨梅素(Myricetin)^[29]等;部分 NAR 在 F3H、DFR、ANS、LAR 作用下,进入花青素苷元(表培儿茶素)和原花青素(没食子儿茶素)合成途径。初步推断,黄秋葵花中

富含花青素苷(葵素-3-葡萄糖苷)成分,果荚富含黄酮醇(杨梅素等)、花青素苷元(表焙儿茶素)原花青素(没食子儿茶素等)等类黄酮物质。另外 ANS 和 GT 基因在黄秋葵花、果荚中均存在多拷贝和底物特异性现象^[30],推断 GT(c44799)、GT(c82004)在黄秋葵花中特异性的催化无色天竺葵素苷元,GT(c44799)在黄秋葵果荚中特异性的催化无色飞燕草苷元;ANS 分别在花和果荚里高表达,ANS(c83240)可能与黄秋葵花瓣颜色(黄色)有关,ANS(c94125)可能与果荚颜色(浓绿色)有关,具体 GT、

ANS 特异性底物种类、产物及最终呈色关系还需要进一步功能验证。

利用高通量转录组测序技术,深入挖掘黄秋葵不同组织部位 DEGs,从宏观上理清黄秋葵花、果荚差异基因及调控机理等。近年来,黄秋葵营养成分研究在国内越来越受到重视,特别是黄酮提取^[31]和功能分析^[32]逐渐成为研究热点,备受食品、医药等行业青睐。本研究丰富了黄秋葵花、果荚转录组水平生物数据信息,为类黄酮等关键基因克隆和功能验证等提供遗传基础。

参考文献:

- [1] 刘江伟,于蒙蒙,范双杰,等. 黄秋葵花的成分分析[J]. 食品科技, 2016, **41**(7): 280-283.
LIU J W, YU M M, FAN S J, *et al.* Analysis of nutritional components in okra flowers[J]. *Food Science and Technology*, 2016, **41**(7): 280-283.
- [2] 宋思圆,苏平,王丽娟,等. 黄秋葵花果胶多糖的超声提取工艺优化及其体外抗氧化活性[J]. 食品科学, 2007, **28**(2): 283-289.
SONG S Y, SU P, WANG L J, *et al.* Optimization of ultrasonic-assisted extraction by response surface method and antioxidant activities of pectic polysaccharide from okra flowers[J]. *Food Science*, 2007, **28**(2): 283-289.
- [3] 盛玮,滕井通,薛建平,等. 响应面优化超声-微波协同提取黄秋葵花中总黄酮工艺的研究[J]. 食品工业科技, 2015, (8): 290-293, 314.
SHENG W, TENG J T, XUE J P, *et al.* Optimization of ultrasonic-microwave synergistic extraction of total flavonoids from okra by response surface methodology[J]. *Science and Technology of Food Industry*, 2015, (8): 290-293, 314.
- [4] 黄阿根,陈学好,高云中,等. 黄秋葵的成分测定与分析[J]. 食品科学, 2007, **28**(10): 451-455.
HUANG A G, CHEN X H, GAO Y Z, *et al.* Determination and analysis of ingredient in okra[J]. *Food Science*, 2007, **28**(10): 451-455.
- [5] 李鹏翔. 一种黄秋葵茶饮料及其生产: 中国, 102028068[P]. 2010-12-02.
- [6] 郑云云,郑黄楠,周红玲,等. 黄秋葵果荚中多糖的提取及其累积规律[J]. 福建农业学报, 2016, **31**(1): 27-30.
ZHENG Y Y, ZHENG H N, ZHOU H L, *et al.* Polysaccharide accumulation in and extraction from okra pods [J]. *Fujian Journal of Agricultural Sciences*, 2016, **31**(1): 27-30.
- [7] 张少平,邱珊莲,张帅,等. 黄秋葵种质资源及相关品种选育研究进展[J]. 农学学报, 2017, **7**(6): 49-55.
ZHANG S P, QIU S L, ZHANG S, *et al.* Research progress of breeding and germplasm resources of okra [J]. *Journal of Agriculture*, 2017, **7**(6): 49-55.
- [8] 高玲,刘迪发,徐丽. 黄秋葵研究进展与前景[J]. 热带农业科学, 2014, (11): 22-29.
GAO L, LIU D F, XU L. Research progress and prospects of okra[J]. *Chinese Journal of Tropical Agriculture*, 2014, (11): 22-29.
- [9] 李颖,曾俊美,郑启翔,等. 黄秋葵功能成分提取技术的研究进展[J]. 现代食品, 2015, (19): 72-75.
LI Y, ZENG J M, ZHENG Q X, *et al.* Advances in the extraction of functional components of the okra [J]. *Modern Food*, 2015, (19): 72-75.
- [10] 王维婷,郭淑,刘超,等. 黄秋葵生物活性物质及加工研究进展[J]. 山东农业科学, 2015, **48**(11): 134-136, 140.
WANG W T, GUO X, LIU C, *et al.* Research progress of bioactive substances and processing of okra [J]. *Shandong Agricultural Sciences*, 2015, **48**(11): 134-136, 140.
- [11] 邹丽秋,王彩霞,匡雪君,等. 黄酮类化合物合成途径及合成生物学研究进展[J]. 中国中药杂志, 2016, **41**(22): 4 124-4 128.
ZOU L Q, WANG C X, KUANG X J, *et al.* Advance in flavonoids biosynthetic pathway and synthetic biology[J]. *China Journal of Chinese Materia Medica*, 2016, **41**(22): 4 124-4 128.
- [12] 何佳珂,于洋,陈西敬,等. 黄酮类化合物的药物代谢研究进展[J]. 中国中药杂志, 2010, (21): 2 789-2 794.
HE J K, YU Y, CHEN X J, *et al.* Research progress on drug metabolism of flavanoids[J]. *China Journal of Chinese Materia Medica*, 2010, (21): 2 789-2 794.
- [13] 延玺,刘会青,邹永青,等. 黄酮类化合物生理活性及合成研究进展[J]. 有机化学, 2008, **28**(9): 1 534-1 544.
YAN X, LIU H Q, ZOU Y Q, *et al.* Physiological activities and research advance in synthesis of flavonoids[J]. *Chinese Journal of Organic Chemistry*, 2008, **28**(9): 1 534-1 544.
- [14] SHIRLEY B W. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology and biotechnology [J]. *Plant Physiol.*, 2001, **126**(2): 485-493.
- [15] 张少平,邱珊莲,郑云云,等. 紫色黄秋葵转录组功能基因测序及分析[J]. 核农学报, 2017, **31**(4): 643-653.

- ZHANG S P, QIU S L, ZHENG Y Y, *et al.* Analysis the anthocyanin gene of *Gynura bicolor* based on sequencing of transcriptome [J]. *Acta Agriculturae Nucleatae Sinica*, 2017, **31**(4): 643-653.
- [16] 李和平,姚运法,练冬梅,等. 黄秋葵果实转录组测序及分析[J]. 生物技术通报, 2018, **34**(3): 121-127.
LI H P, YAO Y F, LIAN D M, *et al.* Sequencing and analysis of the transcription of the fruits in okra [J]. *Biotechnology Bulletin*, 2018, **34**(3): 121-127.
- [17] LANGMEAD B, TRAPNELL C, POP M. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome[J]. *Genome Biology*, 2009, **10**(3): R25.
- [18] LI B, COLIN ND. RSEM: Accurate transcript quantification from RNA Seq data with or without a reference genome [J]. *BMC Bioinformatics*, 2011, **12**: 323.
- [19] TRAPNELLI C, WILLIANMS B A, PERTEA G. Transcript assembly and quantification by RNA Seq reveals unannotated transcripts and isoform switching during cell differentiation[J]. *Nature Biotechnology*, 2010, **28**(5): 511-515.
- [20] 王 炜,郑 伟,徐晓丹,等. 基于转录组测序的滇山茶花叶呈色机理分析[J]. 西北植物学报, 2017, **37**(9): 1 720-1 727.
WANG W, ZHENG W, XU X D, *et al.* Coloring mechanism of mosaic leaves in *Camellia reticulata* based on sequencing of transcriptome [J]. *Acta Botanica Boreali-Occidentalia Sinica*, 2017, **37**(9): 1 720-1 727.
- [21] 陈 菁,郑 伟,王谈笑,等. 钩苞大丁草高通量转录组测序及差异表达分析[J]. 西北植物学报, 2017, **37**(3): 470-477.
CHEN Q, ZHENG W, WANG T X, *et al.* Transcriptome analysis of *Gerbera delabayi* based on high-throughput sequencing technology and differential expression analysis[J]. *Acta Botanica Boreali-Occidentalia Sinica*, 2017, **37**(3): 470-477.
- [22] 马 婧,成铁龙,孙灿岳,等. 草麻黄高通量转录组分析及黄酮类代谢途径相关基因的鉴定[J]. 浙江农业学报, 2016, **28**(4): 609-617.
MA J, CHENG T L, SUN C Y, *et al.* Characterization of transcriptome reveals pathway of flavonoids in *Ephedra sinica* Stapf [J]. *Acta Agriculturae Zhejiangensis*, 2016, **28**(4): 609-617.
- [23] FORKMANN G. Flavonoids as flower pigments: the formation of the natural spectrum and its extension by genetic engineering [J]. *Plant Breeding*, 1991, **106**(1): 1-26.
- [24] KUBASEK W L, SHIRLEY B W, MCKILLOP A. Regulation of flavonoid biosynthetic genes in germinating *Arabidopsis* seedlings [J]. *Plant Cell*, 1992, **4**: 1 229-1 236.
- [25] PELLETIER M K, SHIRLEY B W. Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings. Coordinate regulation with chalcone synthase and chalcone isomerase [J]. *Plant Physiology*, 1996, **111**: 339-345.
- [26] 戴思兰,洪 艳. 基于花青素苷合成和呈色机理的观赏植物花色改良分子育种[J]. 中国农业科学, 2016, **49**(3): 529-542.
DAI S L, HONG Y. Molecular breeding for flower colors modification on ornamental plants based on the mechanism of anthocyanins biosynthesis and coloration [J]. *Scientia Agricultura Sinica*, 2016, **49**(3): 529-542.
- [27] 葛翠莲,黄春辉,徐小彪. 果实花青素生物合成研究进展[J]. 园艺学报, 2012, **39**(9): 1 655-1 664.
GE C L, HUANG C H, XU X B. Research on anthocyanins biosynthesis in fruit[J]. *Acta Horticulturae Sinica*, 2012, **39**(9): 1 655-1 664.
- [28] 牛生洋,姜建福,樊秀彩,等. 葡萄小白玫瑰及其突变品种小红玫瑰果皮颜色变异机理分析[J]. 园艺学报, 2017, **44**(2): 245-254.
NIU S Y, JIANG J F, FAN X C, *et al.* Analysis on molecular basis of the color mutation in muscat rouge and Muscat Blanc Grapes [J]. *Acta Horticulturae Sinica*, 2017, **44**(2): 245-254.
- [29] 曹运琳,邢梦云,徐昌杰,等. 植物黄酮醇生物合成及其调控研究进展[J]. 园艺学报, 2018, **45**(1): 177-192.
CAO Y L, XING M Y, XU C J, *et al.* Biosynthesis of flavonol and its regulation in plants[J]. *Acta Horticulturae Sinica*, 2018, **45**(1): 177-192.
- [30] 陈嘉景,彭昭欣,石梅艳,等. 柑橘中类黄酮的组成与代谢研究进展[J]. 园艺学报, 2016, **43**(2): 384-400.
CHEN J J, PENG Z X, SHI M Y, *et al.* Advances in on flavonoid composition and metabolism in citrus[J]. *Acta Horticulturae Sinica*, 2016, **43**(2): 384-400.
- [31] 李加兴,陈 选,邓佳琴,等. 黄秋葵黄酮的提取工艺和体外抗氧化活性研究[J]. 食品科学, 2014, **35**(10): 121-125.
LI J X, CHEN X, DENG J Q, *et al.* Extraction and antioxidant activity *in vitro* of okra flavonoids [J]. *Food Science*, 2014, **35**(10): 121-125.
- [32] 苏 平,宋思圆,吴秋敏,等. 黄秋葵黄酮的酵母微胶囊制备及其对油脂的抗氧化能力[J]. 中国食品学报, 2016, **16**(4): 153-158.
SU P, SONG S Y, WU Q M, *et al.* Preparation of the flavonoid microcapsules of okra flowers and the antioxidant capacity in oil [J]. *Journal of Chinese Institute of Food Science and Technology*, 2016, **16**(4): 153-158.

(编辑:宋亚珍)