

酸橙黄酮和生物碱类物质生物合成的转录组分析

潘媛^{1,2,3}, 谭均^{1,2,3},
赵晓^{1,2,3}, 吕卉^{1,2,3}, 杨仕春^{4*}, 陈大霞^{1,2,3*}¹重庆市中药研究院; ²中国中医科学院中药资源中心重庆分中心;³重庆市中药良种选育与评价工程技术研究中心; ⁴重庆市江津区多种经营技术推广中心, 重庆 400065

摘要:本研究采用高通量转录组技术平台 Illumina HiSeq 4 000 构建不同发育时期酸橙果实转录组数据库, 挖掘酸橙生物碱与黄酮类成分生物合成关键酶基因, 在分子水平上进一步研究其药用成分的生物合成。经 Trinity 组装获得 133 155 条 Unigenes, 平均长度 1 892 bp。通过 KEGG 代谢通路分析, 发现酸橙转录组中有 2 640 条 Unigenes 参与 13 条次生代谢产物生物合成相关的代谢通路。进一步分析参与酸橙生物碱和黄酮类生物合成相关基因, 筛选获得 9 条参与生物碱和黄酮类生物合成的关键酶基因。本研究首次从两类成分入手探讨酸橙有效成分生物合成动态转录变化, 丰富了酸橙的遗传信息, 为进一步解析酸橙药用成分生物合成的调控机制提供科学依据。

关键词:酸橙; 黄酮类; 生物碱类; 转录组

中图分类号: R932

文献标识码: A

文章编号: 1001-6880(2024) Suppl-0001-12

DOI: 10.16333/j.1001-6880.2024.S.001

Transcriptome analysis of biosynthesis of flavonoid
and alkaloid from *Citrus aurantium* L.PAN Yuan^{1,2,3}, TAN Jun^{1,2,3},
ZHAO Xiao^{1,2,3}, LYU Hui^{1,2,3}, YANG Shi-chun^{4*}, CHEN Da-xia^{1,2,3*}¹Chongqing Academy of Chinese Material Medica; ²Chongqing Sub-center of National Resource,
Center for Chinese Materia Medica, China Academy of Chinese Medical Science;³Chongqing Engineering Research Center for Fine Variety Breeding Techniques of Chinese Materia Medica;⁴Jiangjin District of Chongqing City Variety of Management Technology Promotion Center, Chongqing 400065, China

Abstract: Illumina HiSeq 4000 technology platform was used to construct the transcriptome database of *Citrus aurantium* L. at different developmental stages, and the key enzyme genes for the biosynthesis of alkaloids and flavonoids were mined to further study the biosynthesis of the medicinal components at the molecular level. A total of 133 155 Unigenes with an average length of 1 892 bp were obtained by Trinity assembly. According to KEGG metabolic pathway analysis, 2 640 Unigenes were found to be involved in 13 metabolic pathways related to the biosynthesis of secondary metabolites. The genes involved in the biosynthesis of alkaloids and flavonoids were further analyzed, and nine key enzyme genes involved in the biosynthesis of alkaloids and flavonoids were screened. In this study, the dynamic transcriptional changes of active components biosynthesis were studied from two components for the first time, which enriched the genetic information and provided scientific basis for further analysis of the regulatory mechanism of biosynthesis of medicinal components of *C. aurantium*.

Key words: *Citrus aurantium* L.; flavonoid; alkaloid; transcriptome酸橙 (*Citrus aurantium* L.) 为芸香科柑橘属植物, 主要栽培于湖北、江西、浙江、四川、重庆和湖南等省。枳实 (*Aurantii Fructus Immaturus*) 和枳壳 (*Aurantii Fructus*) 分别为芸香科柑橘属植物酸橙及其栽培变种的干燥幼果与未成熟果实^[1]。枳实 5~6 月收集自落果实, 枳壳 7 月果皮尚绿时采收。因采收期不同而分开入药^[2]。研究表明, 枳实、枳壳有效化学成分基本一致, 但质量分数不同^[3], 枳实

收稿日期: 2023-07-04

接受日期: 2023-11-15

基金项目: 重庆市江津区现代农业产业技术体系首席专家项目和英才计划 ([2022]23 号); 重庆市英才计划“创新领军人才” (CQYC20200301170)

* 通信作者 E-mail: 17837@163.com; young369037@163.com

和枳壳中主要化学成分含有黄酮类、挥发油以及生物碱。枳实具有破气消积、化痰散痞的功效,特征成分是生物碱辛弗林;枳壳具有理气宽中、行滞消胀的功效,特征成分是黄酮类成分,以二氢黄酮为主,并大多以糖苷形式存在,如柚皮苷、新橙皮苷等^[4]。

近年来,黄酮类化合物在医药、农业等领域的应用不断拓展,市场需求不断扩大,其可持续性生产制造引起了广泛关注。黄酮类化合物在传统提取和化学合成方面都一定的问题,因此,需要寻找更优的生产途径以大量、稳定、绿色可持续地供应活性黄酮化合物^[5]。黄酮类化合物合成前体主要是经聚酮途径生成的丙二酸单酰辅酶和经莽草酸-苯丙素途径生成的4-香豆酰辅酶A^[6],再由不同酶催化进入下游代谢,合成不同类型的黄酮类物质。生物碱辛弗林应用价值极高,但在药用植物体内的含量极低,寻找便捷高效获得辛弗林的新途径是现今需要解决的关键问题。生物碱的生物合成前体也是莽草酸途径产生的色氨酸、酪氨酸和苯丙氨酸,再由不同类型的酶催化合成下游各种类型的生物碱,研究证明,合成辛弗林的有效前体是酪胺^[7,8]。目前酸橙的研究主要集中在种质资源收集、质量评价、药效成分提取等方面,其药效成分生物合成研究主要集中

在柚皮苷和新橙皮苷^[9-13],生物碱类成分辛弗林的研究较少。因此,本研究利用RNA-Seq技术构建酸橙转录组数据库,应用公共数据库对转录组结果进行分类和功能注释,借助KEGG富集分析挖掘生物碱、黄酮类物质生物合成相关关键酶基因,为开展酸橙生物碱和黄酮类成分生物合成调控机制研究提供数据支撑。

1 材料与方法

1.1 材料

酸橙样品于2022年6~7月采自重庆市江津区枳壳种植基地,选取长势良好的健康果实,根据酸橙果实的不同发育时期,分别采集五年生小果(small fruit, SF)(果径 3.23 ± 0.28 cm)、中果(middle fruit, MF)(果径 5.16 ± 0.49 cm)和大果(big fruit, BF)(果径 7.50 ± 0.34 cm)各3份(每两株混合成1份)(见图1),分组为SF1~SF3, MF1~MF3和BF1~BF3,用去离子水清洗,切成1 cm × 1 cm × 1 cm的小块,用锡纸包好后经液氮速冻并保存于实验室超低温冰箱中,用于转录组测序和有效成分含量测定。酸橙样品均由重庆市中药研究院陈大霞研究员鉴定,为芸香科柑橘属植物酸橙*Citrus aurantium* L.的果实。

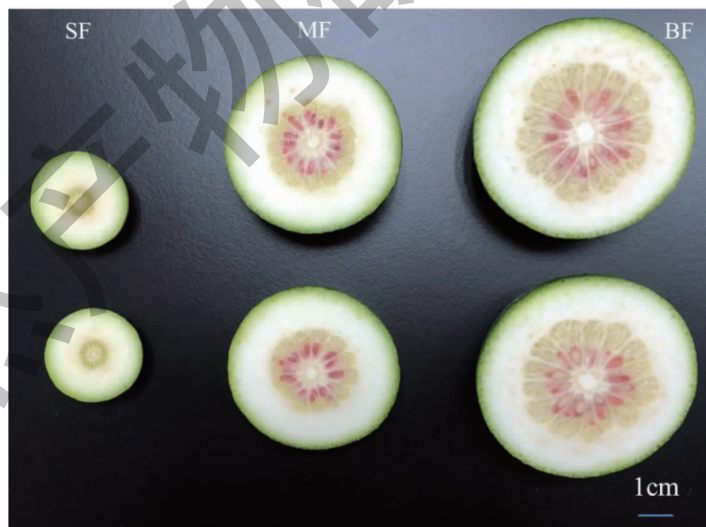


图1 不同发育时期酸橙剖面图

Fig. 1 Profile of *Citrus aurantium* L. at different developmental stages

1.2 仪器、试剂、数据库与软件

S1000 Thermal Cycler PCR仪(美国BIO-RAD); Gel Doc XR凝胶成像系统(美国BIO-RAD); ND-2000超微量分光光度计(美国Thermo Scientific); Illumina HiSeq TM 4 000高通量测序平台(美国Illu-

mina); DYY-6C型电泳仪(中国北京六一); DYCP-31F琼脂糖水平电泳槽(中国北京六一); Bio-RAD CFX 96荧光定量PCR仪(美国BIO-RAD); BS223S型万分之一电子天平(德国Sartorius); Angilent 1200高效液相色谱仪(美国Agilent); Milli-Q Advantage

A10 超纯水仪(美国 Millipore)。

TRIzol 试剂(美国 Invitrogen); 柚皮苷对照品(批号 MUST-14042110, 纯度 $\geq 98\%$)、新橙皮苷对照品(批号 MUST-14042110, 纯度 $\geq 98\%$)、辛弗林对照品(批号 MUST-14042110, 纯度 $\geq 98\%$)、乙腈(色谱纯)、甲醇(色谱纯)(国药集团化学试剂有限公司)。

Nr(非冗余蛋白序列数据库)、Swiss-Prot(高质量非冗余蛋白序列数据库)、KEGG(京都基因与基因组百科全书)和 KOG(真核生物蛋白相邻类的聚簇)等数据库。转录本拼接采用 Trinity(v2.4.0, min_kmer_cov:3, 其他参数为默认参数); DESeq2(1.6.3, padj < 0.05 & |log₂(fold change)| > 1)等软件。

1.3 方法

1.3.1 不同发育时期酸橙中生物碱和黄酮类物质

按照《中华人民共和国药典》(以下简称《中国药典》)2020 年版项下方法对酸橙样品分别进行制备和含量测定,测定酸橙中两种主要黄酮类(柚皮苷和新橙皮苷)和生物碱(辛弗林)。测定辛弗林流动相为甲醇-磷酸二氢钾(50:50),进样体积 10 μ L;检测波长为 275 nm;测定柚皮苷和新橙皮苷流动相为乙腈-水(20:80),进样体积 10 μ L;检测波长为 283 nm。

1.3.2 RNA 的提取与测序

采用 Trizol 法提取各样品总 RNA,并将提取的 RNA 进行等量混合,使用分光光度计检测 RNA 质量,以保证符合建库质量要求。总 RNA 检测合格后,将带有 Oligo(dT)的磁珠对 mRNA 分离富集,使 mRNA 片段化,分为 100~400 bp 的短片段,以 mRNA 片段为模板,用逆转录试剂盒和随机引物合成第一链 cDNA,再合成并纯化双链 DNA。纯化的 DNA 需进行末端修复及加 A 尾,利用 AMPure XP beads 选择片段大小,然后进行 PCR 扩增并纯化 PCR 产物,制备 cDNA 文库,并借助 Illumina HiSeqTM 4 000 高通量测序平台进行转录组测序。

1.3.3 测序数据质控与组装

一般将上述测序过程下机所获得的序列称作原始读序(raw data),获得原始读序后首先要进行数据质量控制,去除里面低质量的序列,从而获得干净序列。本研究采用 Trinity 软件对干净序列进行拼接,并通过序列之间的重叠信息组装得到重叠群,之后组装得到转录本。

1.3.4 Unigeness 的功能注释和分类

本研究利用 Nr、Swiss-Prot、KEGG 和 KOG 等数

据库进行相似性比对、GO 分类和注释、KEGG 代谢通路分析,进而获得所有 Unigeness 的功能注释信息。

1.3.5 基因表达水平分析

将 Trinity 软件拼接得到的转录组作为参考序列,采用了 RSEM 软件将每个时期酸橙样品的干净读序往参考序列上做比对。对比对结果进行统计,进一步得到了每个样品比对到每个基因上的读取计数(read count)数目,并对其进行每千碱基转录物片段/百万个映射片段(fragments per kilobase of transcript per million fragments mapped, FPKM)值转换,进而分析基因的表达水平。

1.3.6 差异表达基因分析

采用 DESeq2 分析差异表达基因,筛选阈值 padj < 0.05 且 |log₂fold change| > 1,用 FPKM 值表示基因的表达量。根据每个基因在不同样品的表达量进行功能注释、代谢通路富集和差异表达分析。

1.3.7 实时荧光定量 PCR 验证

采用 TRIzol 法从 100 mg 样品中提取总 RNA 进行分析。使用 M-MLV 逆转录酶将 2 μ g 总 RNA 合成第一链 cDNA。使用 Bio-RAD CFX 96 荧光定量 PCR 仪,反应条件为两步法:95 $^{\circ}$ C、3 min;95 $^{\circ}$ C、5 s,60 $^{\circ}$ C、30 s,40 个循环,以 β -actin 为内参,使用 2^{- $\Delta\Delta$ Ct}方法计算基因相对表达量。为确保实验的可靠性,对样品进行三个生物重复和三次技术重复。引物序列如表 1 所示。

表 1 引物序列
Table 1 Primer sequence

基因名称 Gene name	引物 Primer
<i>PAL</i>	F:5'-CTCAGAATCCGCCAGGGAAG-3' R:5'-AGCTGTCGGTGCCCTTATTC-3'
<i>4CL</i>	F:5'-CGCGTAAAATCGCTTCTGGG-3' R:5'-AAGGAGAGCACAAACTCGGG-3'
<i>CHS</i>	F:5'-CCGCGTTCTAGTCGTATGCT-3' R:5'-GCATCAGAAGGTCCACGGAA-3'
<i>CHI</i>	F:5'-AATTTTACAGCACAGCCGC-3' R:5'-TCCTTGCTATGGACTGCACC-3'
<i>C12RT1</i>	F:5'-TGGGTTCGCCAGGCTAAAAT-3' R:5'-CCCGAAAACCATCCCCTCAA-3'
<i>DAHPS</i>	F:5'-CGTAGGGAAGGGCAAATCGT-3' R:5'-AGGGTGCCTTGATGGTGT-3'
<i>TYDC</i>	F:5'-ATGCAGCATACGCAGGAAGT-3' R:5'-TCCGCACCTTCAACTCCATC-3'
<i>CAGT</i>	F:5'-TTTGGAAAGTGGTGAACGCT-3' R:5'-CTTGCTTTGCTCCAAACCCC-3'
<i>β-actin</i>	F:5'-CCAAGCAGCATGAAGATCAA-3' R:5'-ATCTGCTGGAAGGTGCTGA-3'

2 结果与分析

2.1 不同发育时期酸橙有效成分含量

根据《中国药典》2020 方法对酸橙果实进行辛弗林、柚皮苷和新橙皮苷含量测定。由表 2 可知,辛弗林的含量在三个时期的果实中均呈现显著的差异;柚皮苷含量在小果期显著高于其他两个时期;新橙皮苷含量在大果期显著低于其他两个时期。上述

表 2 不同发育时期果实有效成分含量($\bar{x} \pm s, n=3$)

Table 2 The content of fruit effective components at different developmental stages($\bar{x} \pm s, n=3$)

样品 Sample	含量 Content (%)		
	辛弗林 Synephrine	柚皮苷 Naringin	新橙皮苷 Neohesperidin
SF	0.62 ± 0.04 ^a	9.31 ± 0.48 ^a	10.74 ± 0.56 ^a
MF	0.44 ± 0.04 ^b	7.56 ± 0.63 ^b	9.98 ± 0.26 ^a
BF	0.27 ± 0.05 ^c	6.85 ± 0.45 ^b	7.58 ± 0.59 ^b

表 3 转录组测序数据产出质量情况

Table 3 Production quality of transcriptome sequencing data

样品 Sample	原始读序 Raw reads	干净读序 Clean reads	干净读序统计量 Clean bases	错误率 Error rate (%)	Q20 (%)	Q30 (%)	GC (%)
SF_1	49 233 786	48 644 052	7.3G	0.03	97.26	92.36	44.49
SF_2	57 548 084	56 821 344	8.52G	0.03	97.12	92.01	44.39
SF_3	64 013 758	63 050 132	9.46G	0.03	96.78	91.29	44.26
MF_1	56 480 710	55 523 914	8.33G	0.03	97.03	91.87	44.5
MF_2	52 021 526	51 243 542	7.69G	0.03	96.94	91.66	44.36
MF_3	56 879 870	56 019 466	8.4G	0.03	97.26	92.35	44.21
BF_1	53 332 942	52 445 664	7.87G	0.03	97.36	92.57	44.39
BF_2	51 559 296	50 235 014	7.54G	0.03	95.92	89.93	44.3
BF_3	54 257 030	53 304 208	8G	0.03	97.16	92.14	44.21

采用软件 Trinity 对 clean data 进行 De novo 组装,共获得 148 069 个转录本(Transcripts),平均长度 1 730 bp, N50 为 2 647 bp。并对得到的 Transcripts 序列进一步组装获得 133 155 条 Unigeness,平均长度 1 892bp, N50 为 2 682 bp。长度在 200 ~ 500 bp 的 Unigeness 序列有 19 295 条,占总体的

结果说明不同发育时期的酸橙果实中次生代谢产物有明显的差异,样品可以用于后续次生代谢产物生物合成关键酶基因的筛选。

2.2 不同发育时期酸橙转录组测序结果与功能注释

利用 Illumina HiSeq 4000 测序平台对不同发育时期的酸橙果实进行转录组测序,通过过滤去除低质量片段,各样品测序数据如表 3 所示。每个时期的样本重复 3 次,共生成 487 287 336 个干净读序(clean data),包含有 73.11 Gb 的 clean data。Q20(质量值 ≥ 20 的碱基所占百分比)分布在 95.92% ~ 97.36%, Q30(质量值 ≥ 30 的碱基所占百分比)分布在 89.93% ~ 92.57%, GC 碱基量为 44.21% ~ 44.5% (见表 3)。从测序产出数据质量评估结果可见,本次测序数据饱和度较高,为转录本的拼接提供了良好的数据支撑。

14.49%;有 24 575 条 Unigeness 的长度聚集在 500 ~ 1 000 bp, 占总 Unigeness 数的 18.46%;有 38 415 条 Unigeness 的长度聚集在 1 000 ~ 2 000 bp, 占总 Unigeness 数的 28.85%;有 50 870 条 Unigeness 的长度超过 2 000 bp, 占总 Unigeness 数的 38.20% (见表 4)。

表 4 转录组拼接长度频数分布情况

Table 4 Distribution of splicing length and frequency of lime transcriptome

类型 Type	长度 Length (200 ~ 500 bp)	长度 Length (500 ~ 1 000 bp)	长度 Length (1 000 ~ 2 000 bp)	长度 Length (>2 000 bp)	合计 Total
转录本 Transcripts	33 833	24 900	38 460	50 876	148 069
基因 Genes	19 295	24 575	38 415	50 870	133 155

拼接得到的 133 155 条 Unigenes 与 NR、NT、Pfam、KOG、Swiss-prot、GO 和 KO 等公共数据库进行比对,注释结果见表 5,分别注释了 115 526 条(86.76%)、121 869 条(91.52%)、45 916 条(34.48%)、85 199 条(63.98%)、76 974 条(57.81%)、76 974 条

(57.81%)、33 236(24.96%)条 Unigenes。在 7 大数据库中全部注释成功的 Unigenes 有 19 939 条,占总 Unigenes 数的 14.97%;至少有一个数据库注释成功的 Unigenes 有 124 180 条,占总 Unigenes 数的 93.25%。

表 5 转录组基因注释情况
Table 5 Gene annotation of lime transcriptome

数据库类别 Database class	基因数目 Number of gene	注释百分率 Percentage of annotation (%)
Annotated in NR	115 526	86.76
Annotated in NT	121 869	91.52
Annotated in KO	45 916	34.48
Annotated in SwissProt	85 199	63.98
Annotated in PFAM	76 974	57.8
Annotated in GO	76 974	57.8
Annotated in KOG	33 236	24.96
Annotated in all Databases	19 939	14.97
Annotated in at least one Database	124 180	93.25
Total Unigenes	133 155	100

通过 KEGG 富集分析,共获得 2 640 条 Unigenes 参与到 13 条与酸橙次生代谢产物生物合成相关的代谢通路中(表 6)。其中,有 937 条 Unigenes 参与到苯丙素、苯丙氨酸、酪氨酸和色氨酸的生物合成通路中,有 566 条参与到黄酮类成分的生物合成通路中,有 322 条参与到生物碱生物合成通路中,还

有部分 Unigenes 参与到萜类和花青素类成分的生物合成通路中。可见,本研究构建的酸橙转录组数据库,有助于发掘酸橙药效成分生物合成过程中的关键酶基因,以及为后续开展新化合物的分离及阐明药效物质提供数据基础。

表 6 药效物质相关的次生代谢产物生物合成途径
Table 6 Biosynthetic pathway of secondary metabolites involved in effective substances

合成通路 Synthesis pathway	通路编号 Pathway ID	基因数 Gene number	药用成分 Medicinal component
苯丙氨酸、酪氨酸和色氨酸的生物合成 Phenylalanine, tyrosine and tryptophan biosynthesis	ko00400	323	苯丙素类 Phenylpropanoids
花青素生物合成 Anthocyanin biosynthesis	ko00942	1	花青素类 Anthocyanins
黄酮和黄酮醇的生物合成 Flavone and flavonol biosynthesis	ko00944	39	黄酮类 Flavonoid
类黄酮生物合成 Flavonoid biosynthesis	ko00941	154	黄酮类 Flavonoid
异黄酮生物合成 Isoflavonoid biosynthesis	ko00943	2	黄酮类 Flavonoid
异喹啉生物碱的生物合成 Isoquinoline alkaloid biosynthesis	ko00950	142	生物碱类 Alkaloids
苯丙类生物合成 Phenylpropanoid biosynthesis	ko00940	614	苯丙素类 Phenylpropanoids
哌啶和吡啶生物碱的生物合成 Tropane, piperidine and pyridine alkaloid biosynthesis	ko00960	180	生物碱类 Alkaloid

续表 6(Continued Tab. 6)

合成通路 Synthesis pathway	通路编号 Pathway ID	基因数 Gene number	药用成分 Medicinal component
类胡萝卜素生物合成 Carotenoid biosynthesis	ko00906	371	黄酮类 Flavonoid
二萜生物合成 Diterpenoid biosynthesis	ko00904	112	萜类 Terpenoids
单萜生物合成 Monoterpenoid biosynthesis	ko00902	92	萜类 Terpenoids
倍半萜类和三萜类生物合成 Sesquiterpenoid and triterpenoid biosynthesis	ko00909	112	萜类 Terpenoids
萜类骨架生物合成 Terpenoid backbone biosynthesis	ko00900	498	萜类 Terpenoids

2.3 基因表达水平分析

一般用 FPKM 值检测基因表达水平,其优势在于同时考虑了测序深度和基因长度对片段 (fragments) 计数的影响。从不同生育时期酸橙样品

Unigeness 的密度图和盒形图 (见图 2) 可以看出,不同生育期样品基因总体表达水平在分布度和离散度上存在一定的差异,说明不同生育时期酸橙果实的基因表达存在差异。

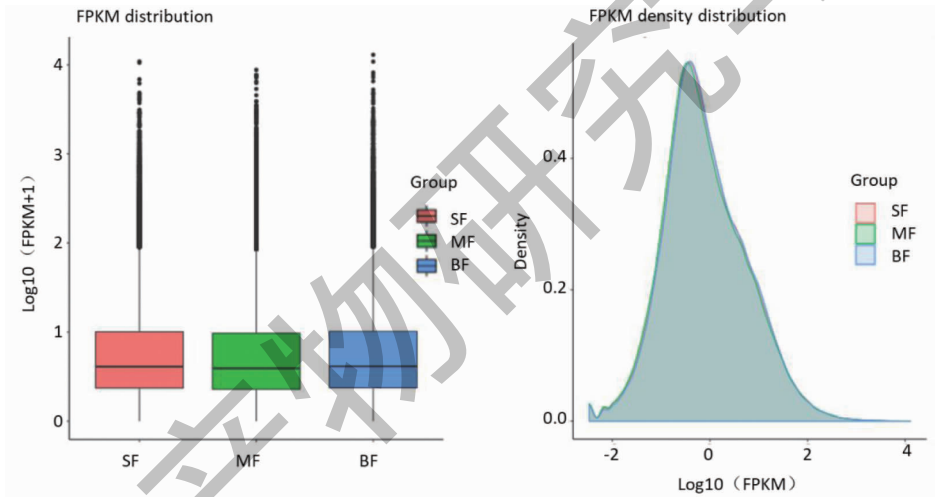


图 2 转录组基因表达水平密度分布图与盒形图

Fig. 2 Distribution map and box diagram of gene expression level density of transcriptome

2.4 差异表达基因的筛选

本研究采用 DESeq 对 3 个时期的酸橙果实进行两两比较,对筛选获得的差异表达基因 (DEGs, differentially expressed genes) 进行分析。由差异基因聚类热图 (见图 3A) 可知,酸橙小果 (SF) 的差异基因单独聚为一类,中果 (MF) 和大果 (BF) 聚为一类,3 个时期的样品存在明显的差异。从基因差异表达火山图 (图 3B ~ 3D) 可知,酸橙大果与小果相比,差异表达基因总数为 346 条,其中上调表达 139 条;中果与小果相比,差异基因总数为 1 040 条,其中上调表达 565 条;大果与中果相比,差异表达基因总数为 182 条,其中上调表达 99 条。

2.5 差异表达基因 KEGG 富集分析

差异表达基因 (DEGs) 在 KEGG 的富集结果显示,BF 与 MF 的 DEGs 被注释到 39 条代谢通路中,包括类胡萝卜素生物合成 (Ko00906)、基础转录因子 (Ko03022)、RNA 降解 (Ko03018) 和苯丙氨酸、酪氨酸和色氨酸的生物合成 (Ko00400) 等;BF 与 SF 的 DEGs 被注释到 95 条代谢通路中,包括光合作用 (Ko00195)、倍半萜类和三萜类生物合成 (Ko00900)、乙醛酸盐和二羧酸盐代谢 (Ko00630) 和苯丙素的生物合成 (Ko0094) 等;MF 与 SF 的 DEGs 被注释到 64 条 Pathway 中,包括亚油酸代谢 (Ko00591)、谷胱甘肽代谢 (Ko00480)、苯丙素的生

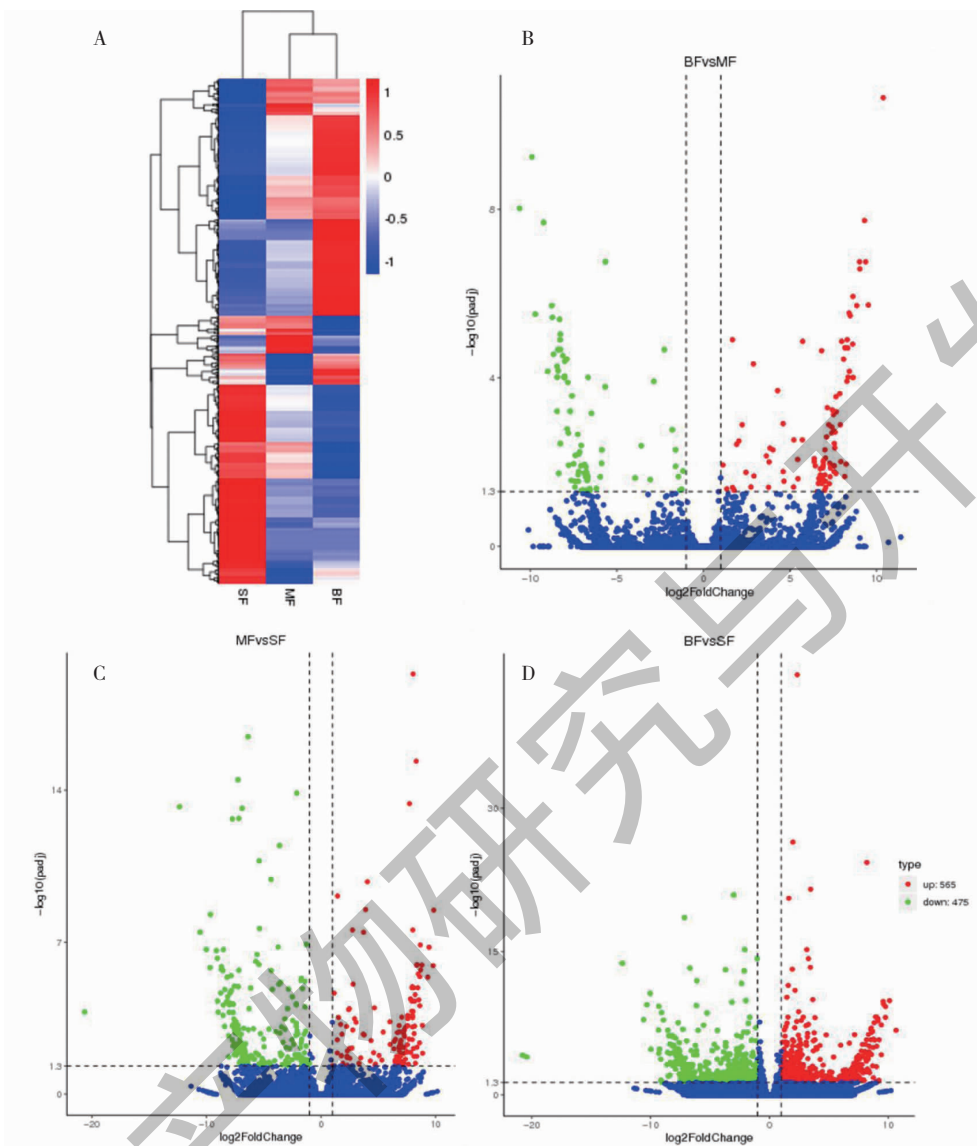


图3 转录组差异表达基因热图与火山图

Fig. 3 Heat map and volcano map of differentially expressed genes in lime transcriptome

物合成 (Ko00940) 和氨酰生物合成 (Ko00970) 等 (见图4)。其中苯丙氨酸、苯丙素、黄酮类和生物碱类成分的生物合成均富集到差异基因,此类差异基因将为后续开展酸橙生物碱和黄酮类物质生物合成关键酶基因的挖掘提供数据支撑。

2.6 参与生物碱和黄酮类生物合成差异表达基因筛选

酸橙果实中药效成分主要有黄酮类、生物碱以及挥发油,其中生物碱主要是辛弗林,黄酮类成分以二氢黄酮为主,并大多以糖苷形式存在,如柚皮苷、新橙皮苷等^[4]。酸橙果实生物碱类和黄酮类物质的生物合成的前体都经过苯丙氨酸、酪氨酸和色氨

酸的生物合成途径,由糖酵解/糖异生途径产生的磷酸烯醇式丙酮酸 (phosphoenolpyruvate, PEP) 和磷酸戊糖途径的赤藓糖-4-磷酸 (erythrose 4-phosphate, E4P) 是合成芳香族氨基酸的初级前体,PEP 和 E4P 在 3-脱氧-阿拉伯酸糖-7-磷酸合酶 (DAHP synthase, DAHPS) 的催化下缩合产生 3-脱氧-D-阿拉伯庚酮糖酸-7-磷酸,进入合成芳香族氨基酸的公共合成途径莽草酸途径。其在莽草酸激酶 (shikimate kinase, SK)、分支酸合酶 (chorismate synthase, aroC) 等酶的连续催化下合成分支酸,又在分支酸变位酶 (chorismate mutase, CM) 等酶的作用下生成酪氨酸和苯丙氨酸。生成的酪氨酸在酪氨酸脱羧酶 (tyrosine

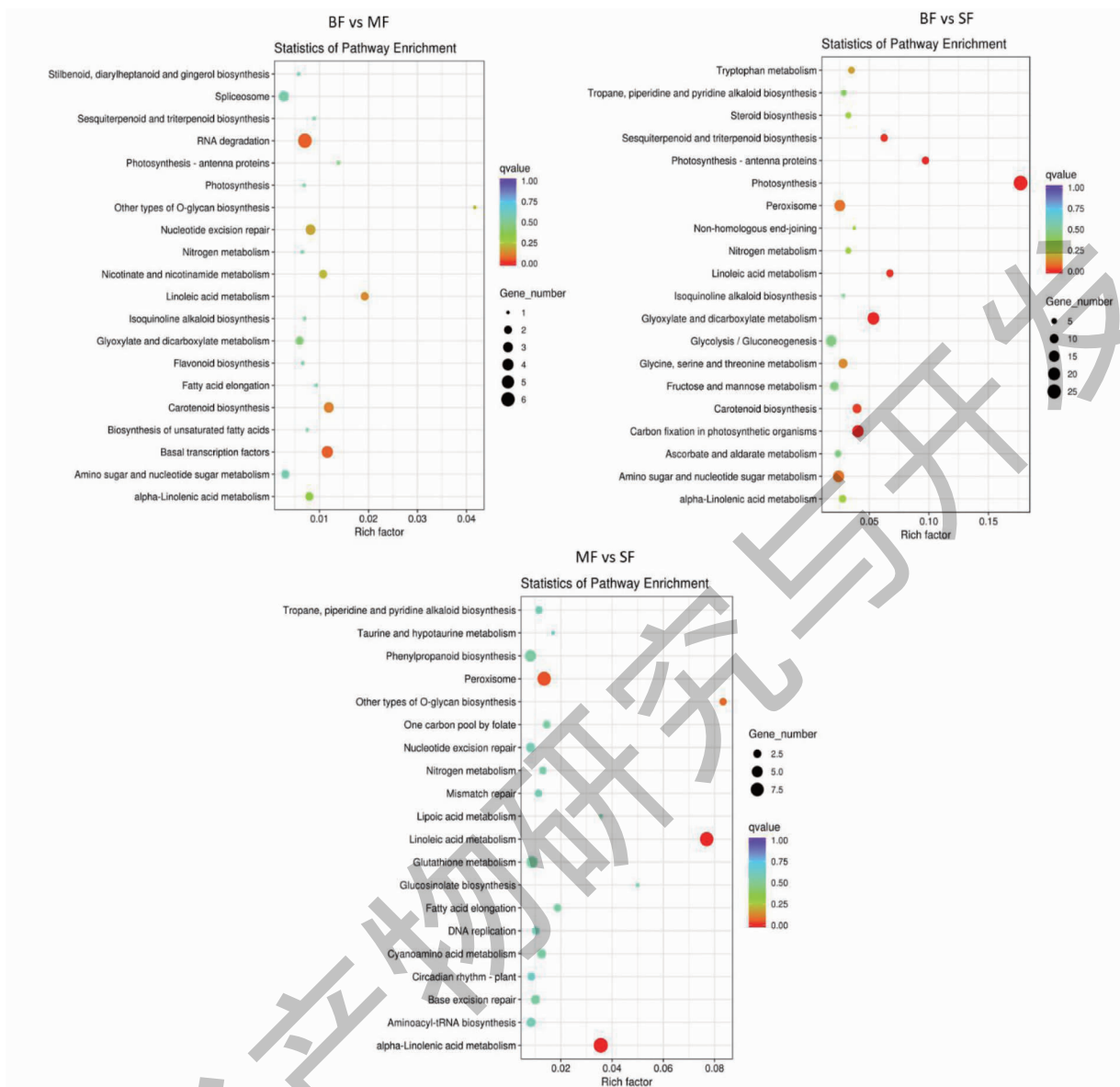


图 4 转录组差异表达基因 KEGG 富集分析

Fig. 4 KEGG enrichment analysis of transcriptome differentially expressed genes

decarboxylase, TYDC) 的作用下使酪氨酸脱羧产生 CO_2 和酪胺,又在松柏醇葡萄糖基转移酶(pine cypress alcohol glucosyltransferase, CAGT)的作用下生产辛弗林。另一条分支则是苯丙氨酸由苯丙氨酸酶(phenylalanine enzyme, PAL)催化转化为反式肉桂酸,经过肉桂酸-4-羟化酶(cinnamate-4-hydroxylase, C4H)、对香豆酸-辅酶 A 连接酶(4-coumarate: CoA ligase, 4CL)、查尔酮合成酶(chalcone synthase, CHS)、查尔酮异构酶(chalcone isomerase, CHI)、黄酮 7-O-葡萄糖苷 2'-O- β -L-鼠李糖基转移酶(flavanone 7-O-glucoside 2''-O- β -L-rhamnosyltransferase, C12RT1)等多步

反应,催化产生二羟黄酮柚皮苷和新橙皮苷(见图 5)。根据“2.4”获得的差异基因,在苯丙氨酸、酪氨酸和色氨酸生物合成途径中筛选到了差异基因 DAHPS、SK,在黄酮合成途径中筛选到差异基因 PAL、4CL、CHS、CHI、C12RT1,在生物碱合成途径中筛选到差异基因 TYDC 和 CAGT。由图 6 可知,酸橙不同发育时期 DAHPS 基因的表达呈现逐渐升高的趋势;CHI、TYDC 基因的表达呈现先升高后降低的趋势;SK、CAGT、PAL、CHS 基因的表达呈现逐渐降低的趋势;4CL、C12RT1 基因的表达则呈现先降低后升高的趋势。

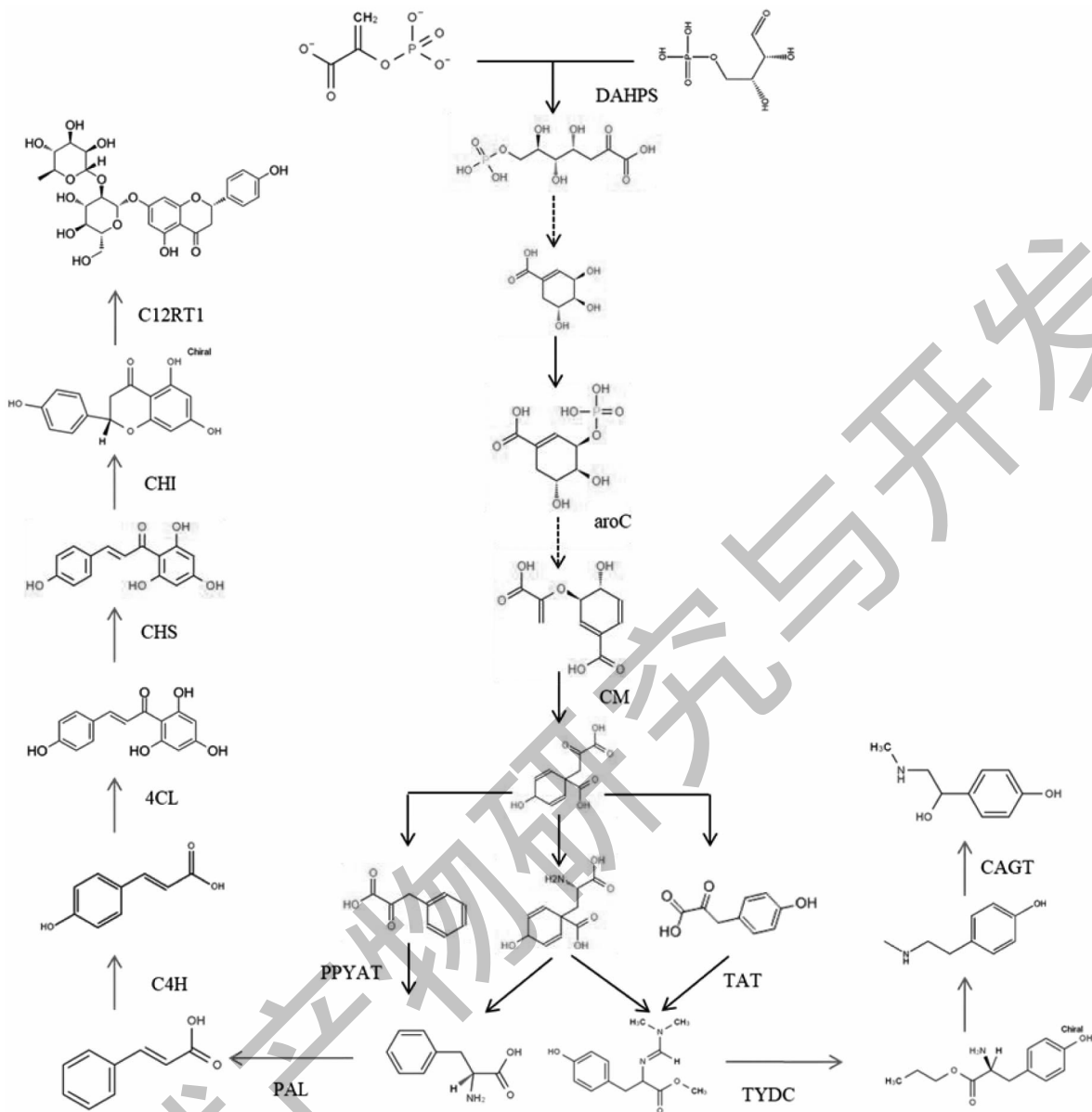


图5 生物碱与黄酮类生物合成途径及关键基因

Fig. 5 Biosynthetic pathway of alkaloids and flavonoids and key genes

2.7 参与生物碱和黄酮类生物合成差异表达基因的表达分析

为了验证转录组结果的可靠性,以 β -actin为内参基因,BF为对照,对从KEGG代谢通路中筛选出的8条与酸橙果实生物碱和黄酮类成分生物合成候选关键酶基因进行qRT-PCR检测,对获得的CT值进行相对表达量分析。由图7所示,除DAHPS和CHI基因在大果时期外,其余6个差异表达基因的相对表达量(Relative transcript level,RTL)与转录组测序结果每千个碱基的转录每百万映射读取的fragments (fragments per kilobase of exon model per

million mapped fragments, FPKM)变化趋势相类似,可见,本次测序获得了可靠的转录组数据,可以用于后续基因挖掘。

3 讨论与结论

酸橙果实作为中药枳实和枳壳的来源,是我国重要的药用植物,其主要药用成分含有黄酮类、挥发油以及生物碱。辛弗林、柚皮苷和新橙皮苷是酸橙次生代谢产物中重要的组成部分,辛弗林的药理活性已被证实可促进脂质代谢^[14],柚皮苷和新橙皮苷等具有抗氧化、抗炎、抗肿瘤、治疗心血管疾病等功效^[15,16]。近年来,随着酸橙药用成分的开发与解

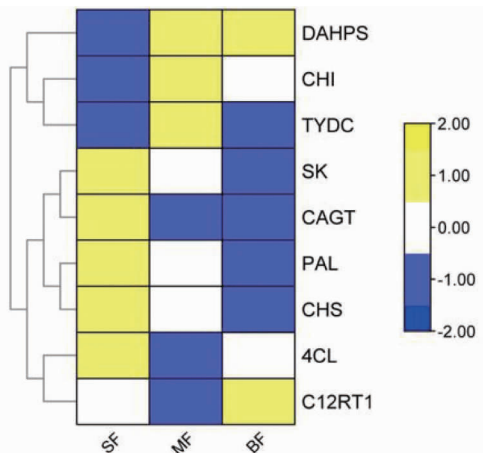


图6 生物碱与黄酮类生物合成途径关键基因表达热图

Fig. 6 Heat map of key gene expression in the biosynthetic pathway of alkaloids and flavonoids

注:不同列代表不同的样品,不同的行代表不同的基因。颜色代表了基因在样品中的表达量 FPKM 均一化后的 Score 值。Note: Different columns represent different samples, and different rows represent different genes. The color represents the score value after FPKM normalization of the gene expression in the samples.

析,越来越多的研究集中于其黄酮类和生物碱类成分生物合成方面的研究,开展酸橙次生代谢产物生物合成途径及其相关基因转录与调控方面的研究,将有助于提高酸橙药用成分含量,为其遗传改良提供基因资源。本研究对酸橙的转录组测序数据进行了统计和组装,获得 133 155 条 Unigenes,平均长度 1 892bp,将组装得到的所有 Unigenes 与七大数据库比对注释,为进一步发现并鉴定酸橙功能基因提供了大量的数据信息。

药用植物次生代谢过程是一个极其复杂的生物过程,从本研究获得的基因注释结果来看,有 13 个生物合成代谢途径 2 640 条 Unigenes 涉及到次生代谢产物的生物合成,其中包括黄酮类、萜类、生物碱等物质。酸橙中药用成分生物碱和黄酮类成分的合成前体来源于苯丙氨酸、酪氨酸和色氨酸生物合成途径,磷酸烯醇丙酮酸和赤藓糖-4-磷酸在 DAHPS 的催化下,生成 3-脱氧-D-7-磷酸阿拉伯庚酮糖酸,DAHPS 基因在不同发育时期的酸橙果实中呈逐渐升高的表达趋势,可见随着果实的发育,该基因的上调表达有助于下游物质的积累;下游 SK 基因则呈现出显著下调表达的趋势。生成苯丙氨酸和酪氨酸后,次生代谢产物朝两个方向继续合成,一条途径以苯丙氨酸为前体,在 PAL、C4H、4CL、CHS、CHI、C12RT1 的连续催化下逐步生成柚皮苷和新橙皮苷等二氢黄酮类成分,而另一条分支则在 TYDC 和 CAGT 等酶的催化下合成辛弗林。在此生物合成过

程中,本研究筛选到 9 条参与酸橙生物碱和黄酮类成分生物合成的关键酶基因,这些关键酶基因在不同时期的表达趋势存在明显的差异。结合前期测定的不同发育时期酸橙果实中辛弗林、柚皮苷和新橙皮苷的百分含量,可见仅有 PAL 和 CHS 基因的表达趋势与上述三种成分的含量变化趋势一致,其余关键基因的表达趋势与含量之间存在一定的差异。为何不同发育时期酸橙果实中关键酶基因的表达趋势与实际合成的目标成分含量趋势存在差异,还需要我们借助代谢组学、基因工程等手段继续发掘代谢流的前进方向和关键酶基因的转录与调控,然而本研究的开展将为后续的研究工作提供一手的数据资源。

随着对黄酮类物质日益增长的市场需求,生产大规模的柑橘属果实需要大量的耕地、化肥和农药,投入高,消耗高,污染高且产出低,使得此类产品的大量生产受到限制。发展绿色环保的方法是亟待解决的关键技术问题。化学合成的方法涉及化学试剂的使用以及严苛的反应条件,容易产生溶剂残留,制得的黄酮类物质用于食品、保健品、化妆品等产品中容易引起安全问题^[17]。目前的研究主要集中在利用微生物合成黄酮类物质。Hwang 等^[18]2003 年在大肠杆菌中组装了 3 个合成黄烷酮的特异性基因,并分别以苯丙氨酸和酪氨酸为底物成功合成了松属素和柚皮素。Watts 等^[19]将来自荚膜梭菌的编码酪氨酸解氨酶基因 TAL 和来自拟南芥的 4CL、CHS 导

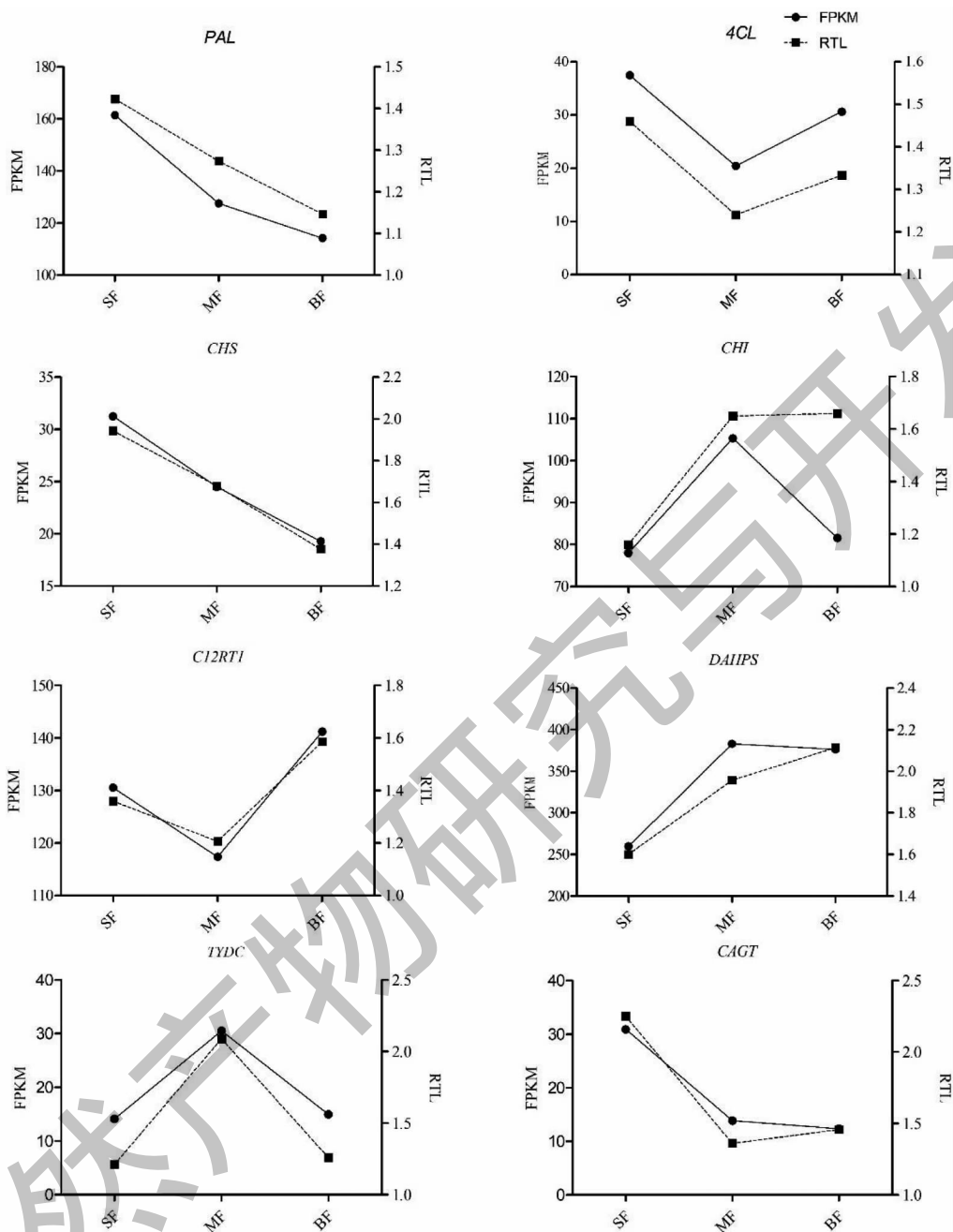


图7 qRT-PCR 和 RNA-Seq 结果比较

Fig. 7 Comparison of qRT-PCR and RNA-Seq results

入大肠杆菌中,将葡萄糖作为底物成功合成了 20.8 mg/L 柚皮素。Koopman 等^[20]以葡萄糖为底物,将来自拟南芥的 *4CL*、*CHS1*、*CHI1*、*C4H*、*PAL1*、*CPR*,来自荚膜梭菌的 *TAL1* 构建到酿酒酵母中,并在 2L 发酵罐中扩大培养,成功合成了 112.9mg/L 柚皮素。Park 等^[21]将来自拟南芥的 *AtCHS* 和 *AtCHI* 导入委内瑞拉链霉菌中,并以对香豆酸为底物合成了 35.6 mg/L 柚皮素。Lv 等^[22]在筛选得到的产油酵母中

构建柚皮素合成途径,以葡萄糖为底物成功合成了 252.4 mg/L 柚皮素。即便如此,微生物合成黄酮类物质的产量水平还较低,还需要继续挖掘合成途径中的关键酶以及调控因子以提高合成量。

药用植物高通量转录组测序通过不同发育时期、不同培育条件下不同转录组测序数据的比较分析,挖掘药用植物新的功能基因和代谢通路,为药用植物通过克隆关键酶基因实现代谢工程生产药效成

分以及种质创新等提供基础^[23]。通过酸橙果实转录组的组装分析以及生物碱、黄酮等合成相关基因的挖掘,为研究酸橙生物碱与黄酮类物质生物合成途径及调控机制、克隆关键基因及功能分析奠定了基础,也为枳实和枳壳药材品质的形成提供理论依据。

参考文献

- Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China: Vol I (中华人民共和国药典:第一部)[M]. Beijing: China Medical Science Press, 2020: 257-258.
- Yuan D, Wang Q, Wang Q, et al. Study on quality evaluation method of Fructus Aurantii and its extract[J]. Chin Pharm J (中国药学杂志), 2005, 40: 1178-1181.
- Jia Q, Bai Y, Ma Y, et al. Chemical constituents of Fructus Aurantii and Fructus Aurantii Immaturus by HPLC-ESI-MS [J]. Chin Tradit Herb Drugs (中草药), 2005, 11: 169-172.
- He YJ, Liu DB, Tang Q, et al. Research progress on chemical constituents of Fructus Aurantii and Fructus Aurantii in *Citrus aurantium* L. [J]. J Chin Med Mater (中药材), 2017, 40: 1488-1494.
- Lin CC, Chen DW, Dai JG, et al. Advances of synthetic biology of flavonoids[J]. Acta Pharm Sin (药学报), 2022, 57: 1322-1335.
- Li ZB, Li CF, Li J, et al. Molecular cloning and functional characterization of two divergent 4-coumarate: coenzyme A ligases from Kudzu (*Pueraria lobata*) [J]. Biol Pharm Bull, 2014, 37: 113-122.
- Lv XG. Screening of *Citrus aurantium* L. germplasm resources with high alkaloid content and mining of candidate genes[D]. Hanzhong: Shaanxi University of Technology (陕西理工大学), 2022.
- Yao YF, Zhao GR. Advances on metabolic engineering of *L-tyrosine* [J]. Food Ferment Ind (食品与发酵工业), 2013, 39: 132-137.
- Gao S, Zhou HR, Zhou JW, et al. Promoter-library-based pathway optimization for efficient (2S)-naringenin production from *p*-coumaric acid in *Saccharomyces cerevisiae* [J]. J Agric Food Chem, 2020, 68: 6884-6891.
- Lyu XM, Zhao GL, Ng KR, et al. Metabolic engineering of *Saccharomyces cerevisiae* for de novo production of kaempferol [J]. J Agric Food Chem, 2019, 67: 5596-5606.
- Ganesan V, Li ZH, Wang XN, et al. Heterologous biosynthesis of natural product naringenin by co-culture engineering [J]. Synth Syst Biotechnol, 2017, 2: 236-242.
- Zhang W, Liu H, Li X, et al. Production of naringenin from dxylose with co-culture of *E. coli* and *S. cerevisiae* [J]. Eng Life Sci, 2017, 17: 1021-1029.
- Lv YK, Marsafari M, Koffas M, et al. Optimizing oleaginous yeast cell factories for flavonoids and hydroxylated flavonoids biosynthesis [J]. ACS Synth Biol, 2019, 8: 2514-2523.
- Yan Han, Ding ZE, Dong M, et al. Effects of synephrine from citrus peel on weight losing of nutritional pinelosis rats [J]. J Food Sci Biotechnol (食品与生物技术学报), 2017, 36: 890-895.
- Parka K, Parkb HS, Kimb MK, et al. Flavonoids identified from Korean *Citrus aurantium* L. inhibit non-small cell lung cancer growth *in vivo* and *in vitro* [J]. J Funct Foods, 2014, 7: 287-297.
- Nagappan A, Park HS, Hong GE, et al. Anti-cancer and anti-inflammatory properties of Korean Citrus fruits (*Citrus aurantium* L.) [J]. J Korean Med Sci, 2014, 2: 73-78.
- Shan Y, Liu J, Wang Z, et al. Research progress on biosynthesis of citrus flavonoids [J]. Chin J Food Sci (中国食品学报), 2019, 19: 1-13.
- Hwang E I, Kaneko M, Ohnishi Y, et al. Production of plant-specific flavanones by *Escherichia coli* containing an artificial gene cluster [J]. Appl Environ Microbiol, 2003, 69: 2699-2706.
- Watts KT, Lee PC, Schmidi C. Exploring recombinant flavonoid biosynthesis in metabolically engineered *Escherichia coli* [J]. ChemBiochem, 2004, 5: 500-507.
- Koopman F, Beekwilder J, Crimi B, et al. De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae* [J]. Microb Cell Fact, 2012, 11: 155.
- Park SR, Ahn MS, Han AR, et al. Enhanced flavonoid production in *Streptomyces venezuelae* via metabolic engineering [J]. J Microbiol Biotechnol, 2011, 21: 1143-1146.
- Lv YK, Koffas M, Zhou J, et al. Optimizing oleaginous yeast cell factories for flavonoids and hydroxylated flavonoids biosynthesis [J]. ACS Synth Biol, 2019, 8: 2514-2523.
- Zhao ZY, Wang SY, Guo FG, et al. RNA-sequencing and its application in medicinal plants [J]. Genomics Appl Biol (基因组学与应用生物学), 2017, 36: 6-10.