

## 研究论文

# 红花(*Carthamus tinctorius* L.)不同开花时期的转录组测序及黄酮合成相关基因的验证

刘秀明<sup>1,2</sup> 张宇<sup>2</sup> 姚娜<sup>1</sup> 崔喜燕<sup>1</sup> 王立勇<sup>1</sup> 董园园<sup>1</sup> 王南<sup>1</sup> 李晓薇<sup>1,2</sup>  
李海燕<sup>1,2\*</sup> 李校堃<sup>1\*</sup>

<sup>1</sup>吉林农业大学生物反应器与药物开发教育部工程研究中心, 长春 130118;

<sup>2</sup>吉林农业大学生命科学学院, 长春 130118)

**摘要** 应用454测序技术对红花(*Carthamus tinctorius* L.)初花期和盛花期两个样本进行转录组测序, 分别获得了577 664和562 930条Clean reads片段, 其平均长度分别为427 bp和436 bp。差异表达分析结果发现, 在初花期获得40 139个基因, 盛花期获得39 768个基因, 在两个时期同时表达的基因有28 316个。COG/KOG分析中, 获得了大量与代谢相关的功能注释信息, 注释到比例最高的为功能预测, 占24.681%。KEGG Pathway富集分析中, 找到红花黄酮化合物合成途径中的关键酶基因, 利用Real-time PCR方法验证了三个基因的正确性。表达量分析表明, 查尔酮合酶在两个红花品种中各个开花时期的相对表达量最高。

**关键词** 红花; 转录组测序; 黄酮合成

## Analysis of Genes Involved in Safflower Flavonoids Biosynthesis Based on Safflower Transcriptome in Different Flowering

Liu Xiuming<sup>1,2</sup>, Zhang Yu<sup>2</sup>, Yao Na<sup>1</sup>, Cui Xiyan, Wang Liyong<sup>1</sup>, Dong Yuanyuan<sup>1</sup>, Wang Nan<sup>1</sup>, Li Xiaowei<sup>1,2</sup>,  
Li Haiyan<sup>1,2\*</sup>, Li Xiaokun<sup>1,2\*</sup>

<sup>1</sup>Ministry of Education Engineering Research Center of Bioreactor and Pharmaceutical Development, Jilin Agricultural University, Changchun 130118, China; <sup>2</sup>College of Life Sciences, Jilin Agricultural University, Changchun 130118, China)

**Abstract** 454 sequencing technology was applied to sequencing transcriptome of safflower samples of early flowering and full-bloom stage, respectively. We obtained 577 664 and 562 930 Clean reads fragments, with average length 427 bp and 436 bp, respectively. 40 139 genes were expressed in early flowering in transcriptome. 39 768 genes were expressed in full-bloom flowering through differential expression analysis. 28 316 genes were expressed simultaneously in these two samples. Massive annotation information related to metabolic was obtained through COG/KOG analysis. Among them, 24.681% was highest in annotating as basic functionality. Safflower flavonoids metabolic pathway was found and verified by Real-time PCR in KEGG Pathway analysis. Quantitative PCR analysis showed that the CHS had the highest expression level in the flowering period of two safflower var.

**Keywords** safflower; transcriptome sequencing; flavonoids biosynthesis

收稿日期: 2015-07-03 接受日期: 2015-08-19

国家高技术研究发展计划(863计划)(批准号: 2011AA100606)和国家自然科学基金(批准号: 31101172、31201237)资助的课题

\*通讯作者。Tel: 0431-84532885, E-mail: hyl99@163.com; Tel: 0431-84533348, E-mail: xiaokunli@163.net

Received: July 3, 2015 Accepted: August 19, 2015

This work was supported by the National High Technology Research and Development Program of China (863 Program) (Grant No.2011AA100606) and the National Natural Science Foundation of China (Grant No.31101172, 31201237)

\*Corresponding authors. Tel: +86-431-84532885, E-mail: hyl99@163.com; Tel: +86-431-84533348, E-mail: xiaokunli@163.net

网络出版时间: 2015-09-17 11:20:15 URL: <http://www.cnki.net/kcms/detail/31.2035.Q.20150917.1120.002.html>

红花(*Carthamus tinctorius* L.)属于菊科双子叶植物,是我国重要的传统中草药,既可以作为食用和药用,也可以用于染料和油料等。红花的化学成分较多,研究表明,红花中含有黄酮类、脂肪酸类、烷基二醇类、多糖类成分、多种酚类化合物及木质素等<sup>[1-2]</sup>。黄酮类化合物是红花中最主要的化学成分之一,目前已经从红花中分离出多种黄酮化合物<sup>[3-4]</sup>。红花的主要功效是活血化瘀,因此,被广泛地应用于活血通经、抗氧化、抗癌等方面<sup>[5]</sup>。

转录组(transcriptome)即转录后所有mRNA的总称,是研究基因的结构和功能的基础,是发掘新的功能基因的重要手段<sup>[6]</sup>。转录组学作为一种新的研究方法,利用全部基因的表达调控、蛋白质功能等信息来解决生物学问题,将基因组学研究带入了一个高速发展的时代<sup>[7]</sup>。转录组学不仅可以研究转录组样本中每个基因的表达水平的变化,也可以预测转录组的定位和注释及每个基因在基因组中的功能结构<sup>[8-10]</sup>。454测序技术是基于焦磷酸测序法的新型超高通量测序技术,其优点是快速、准确率高、产生的片段数量多,适合于一些未知基因组序列的物种<sup>[11]</sup>。目前,454测序已经应用于橄榄<sup>[12]</sup>、金银花<sup>[13]</sup>、足叶草<sup>[14]</sup>、豌豆<sup>[15]</sup>等物种中,但在红花中的研究还未见报道。本研究首次对红花花瓣转录组进行454测序,获得了大量数据,并挖掘了红花黄酮合成途径中的关键酶基因,以期对红花黄酮化合物代谢调控的深入研究奠定基础。

## 1 材料与方法

### 1.1 材料

测序材料的红花品种采用新疆塔城红花(*Carthamus tinctorius* L.),实验材料取自红花的初花期和盛花期花瓣。红花种植于吉林农业大学生物反应器与药物开发教育部工程研究中心实验基地,常规管理,分别于开花的初花期和盛花期采集花瓣,迅速放于液氮中,用锡箔纸包好后于-80℃保存备用。

### 1.2 RNA的提取及cDNA文库的合成

RNA提取采取Trizol-plus方法(TaKaRa公司),采用PolyATtract<sup>®</sup> mRNA纯化系统(Promega公司)完成mRNA的分离工作。将分离后的mRNA按照Reasy RNA Cleaning Kit(Qiagen公司)进行浓缩,试验过程中严格按照操作说明书进行,得到的mRNA进行Agilent 2100生物芯片分析系统分析。采用Reasy RNA

Cleaning Kit对获得的Fragment产物进行回收。以回收的mRNA为模板,用逆转录酶(MMLV)和随机引物进行反转录,合成cDNA的第一条链。先取14 μL RNA,加入2 μL N6 Radom Primer(400 μmol/L),68℃水浴10 min后加入2 μL DTT、2 μL dNTP、2 μL MMLV RTase、6 μL 5×First Strand Buffer和2 μL H<sub>2</sub>O,总体积为30 μL,42℃水浴30 min。反应结束后,向上述反应液中加入5 μL 10× DNA Polymerase I Buffer、1 μL DNA Polymerase I(10 U)、1 μL RNase H(6 U)、13 μL H<sub>2</sub>O,总体积为50 μL,16℃温浴2 h。反应完毕后,采用MinElute DNA Cleaning Kit回收产物,用于测序。测序工作委托北京奥莱博生物科技有限公司进行转录组测序,即高通量454测序,454测序采用Roche GS-FLX测序仪配合GS FLX Titanium试剂盒完成。

### 1.3 Unigenes序列的获得及分析

经base calling软件将测序得到的原始数据转化为序列数据,获得的数据称为raw reads,结果以sff文件格式存储起来,分析前需要将其转换成fasta格式的文件,存储为reads序列,提取的工具采用Roche公司的sffinfo。采用In-house developed program将addaptor序列去接头,并采用tagdust(<http://genome.gsc.riken.jp>)<sup>[16]</sup>、Seqclean(<http://compbio.dfci.harvard.edu>)<sup>[17]</sup>等将原始序列中获得的低质量、短序列(<50 bp)、载体及镶嵌序列等污染序列去除,最后得到的序列成为clean reads,后续的分析都将基于Clean reads数据。去污染后序列使用Mira<sup>[18]</sup>、CAP3(<http://seq.cs.iastate.edu>)等序列拼接软件进行拼接,从而获得s序列。

### 1.4 差异表达及相关性分析

Unigenes在样本中的reads被用来表示Unigenes在该样本的表达丰度,通过RPM和RPKM<sup>[19]</sup>值进行标准化处理,RPKM归一后的表达数据作为后续的分析对象。将归一化后的倍数绝对值≥2的Unigenes作为差异表达的Unigenes,而倍数绝对值≥2并且FDR修正值<0.001的Unigenes为显著差异表达的Unigenes。

### 1.5 COG/KOG功能注释

COG/KOG被认为是对获得的基因产物进行直系同源分类的数据库,每个COG/KOG蛋白都被假定来自祖先蛋白,该数据库是基于细菌、藻类及真核生物具有完整基因组的编码蛋白和系统进化关系进

表1 荧光定量PCR引物

Table 1 The primers for Real-time PCR

基因 Gene	引物(5'→3') Primers (5'→3')
18S rRNA	F: GAG AAA CGG CTA CCA CAT CCA A R: TCG TTT GAG CCC GGT ATT GTT A
Unigene_rep_c37404	F: CTC CGG TTC TCT CTG TAA T R: ACC GCA CGT AGC CTG GCG GT
Unigene_rep_c21005	F: GGC CAC TAT CCT ACC ATC GA R: ATT GCT CGG ATT GAT GCC GGA G
Unigene_rep_c32019	F: CTT GGC GGT GTT TGA TAA AGA R: CCA GAT GCA GAC AAT TGG T
Unigene_c16845	F: AGA TCT AGA TCT GGT ATC G R: ATG GAG CAA CAA GCT CAA TT
Unigene_rep_c20979	F: GTT CTT ATT TGT TTC TAA G R: GCA ACA TGT TCC AAC GGT TGC T
Unigene_c16567	F: GTA GTT CAT CAA ACA ACT C R: TTG GAT AGC ATG AGC AAG
Unigene_s48497	F: GTA TGC GTG TCT GTG TTT CA R: ATT CAG CAA CAG TGA CAA
Unigene_c27184	F: GTA CAT CAA GCC TCT ATC G R: AGA GTC AAG GTG GTT ATG
Unigene_rep_c21224	F: GTG CTG CTG GTT ATG GTG GTG A R: AGC TTT AGT CCA TGC CTG T

行构建的, 然后将Unigene和COG/KOG数据库进行比对, 预测Unigenes可能的功能, 并对其做功能分类统计, COG/KOG注释采用blastx。

### 1.6 KEGG Pathway功能注释

KEGG Pathway(Kyoto Encyclopedia of Genes and Genomes Pathway)注释采用blastx与GenMAPP

2.1(<http://www.genmapp.org/>)<sup>[20]</sup>。KEGG Pathway是Pathway的显著性富集分析, 用来分析基因产物在细胞中的代谢途径及生物学功能的数据库。KEGG Pathway可将基因从生物生化的功能上进行划分, 从而了解基因参与哪些细胞的活动, 进一步研究基因在生物学中的复杂行为及意义, 根据KEGG注释信息能进一步得到Unigenes的显著性富集的Pathway注释。

### 1.7 红花转录组测序数据的验证

为了验证红花花瓣转录组测序结果的准确性, 从转录组中获得的Unigenes中随机选取9个Unigenes (unigene\_rep\_c32019、unigene\_c16845、unigene\_c16567、unigene\_s48497、unigene\_c27184、unigene\_rep\_c20979、unigene\_rep\_c21005、unigene\_rep\_c21224和unigene\_rep\_c37404)设计特异性引物(表1), 进行荧光定量PCR分析。将分析结果与测序中的RPKM进行比较, 确定转录组测序的准确性。荧光定量PCR反应体系如下: Primix EX Taq 10 μL, Primer F (10 μmol/L) 0.4 μL, Primer R(10 μmol/L) 0.4 μL, Dye II(50×) 0.4 μL, cDNA模板2 μL, H<sub>2</sub>O 6.8 μL, 总体积20 μL。采用两步法PCR扩增程序。第一步: 预变性, 95 °C 30 s; 第二步: 95 °C 30 s, 62 °C 30 s, 30个循环。反应结束后确认Real-time PCR的扩增曲线和熔解曲线以及各个样本的不同Ct值, 计算出各个基因的相

表2 红花转录组中获得的黄酮代谢关键酶基因的候选序列

Table 2 Sequence in flavonoids metabolism of safflower from transcriptome

关键酶基因 Key enzyme	代码 Code	454测序中挑选Unigenes Unigenes in 454 sequencing	差异表达 Differentially expression
Chalcone synthase	2.3.1.74	Unigene_c16567 (2)	Up-regulated
		Unigene_rep_c22629	
		Unigene_rep_c21289	Down-regulated
		Unigene_c9855 (-2)	
		Unigene_rep_c33061 (-2)	
		Unigene_rep_c32553 (-2)	
		Unigene_rep_c30381 (-2)	
		Unigene_c13467	
		Unigene_rep_c31854	
		Unigene_c27184	
Chalcone isomerase	5.5.1.6	Unigene_c27184	
Anthocyanidin synthase	1.14.11.19	Unigene_rep_c32431	Up-regulated
		Unigene_c5812	
		Unigene_rep_c20891	
		Unigene_c20381 (2)	
		Unigene_s42103 (2)	
		Unigene_c2096	
		Unigene_rep_c32424	
		Unigene_c9873	
		Unigene_rep_c35725	Down-regulated
		Unigene_s48497	
Unigene_rep_c33425			
Unigene_c20561			

对表达量。相对表达量的计算方法采取目的基因和18S rRNA内参基因之间Ct值之间的差值计算, 计算公式为 $2^{-\Delta\Delta C_t}$ [21]。

### 1.8 红花黄酮化合物合成相关基因的挖掘

在红花的转录组测序结果中获得的Unigenes, 找到红花黄酮化合物合成途径, 通过对转录组测序结果的分析, 对KEGG Pathway富集及差异表达分析, 将找到的到红花黄酮化合物生物合成代谢途径, 根据代谢通路中获得的注释信息, 找到一些在红花黄酮化合物合成途径中的一些关键酶基因的候选Unigenes(表2), 并设计特异性引物进行Real-time PCR验证。

### 1.9 红花黄酮化合物合成相关基因的表达分析

以红花的两个品种“吉红油”和“吉红二号”为试材, 分别对花蕾期、初花期、盛花期及衰落期进行相对定量分析。对测序正确的中间片段, 以18S rRNA作为内参, 设计特异性引物(表3), 进行荧光定量PCR分析。

## 2 结果

### 2.1 样品准备及RNA提取

为了获得更全面的数据信息, 本研究选择了开花后第1 d作为初花期, 开花后第5 d作为盛花期(图1), 于这两个时期取花瓣样品用于测序。提取的总RNA首先用Nanodrop 2000分光光度计进行RNA浓度检测, 其浓度为 $1\ 124\ \mu\text{g}/\mu\text{L}$ ,  $D_{260/280}$ 的比值为1.96,  $D_{260/230}$ 的比值为1.92, 说明提取的RNA样品纯度较高。经琼脂糖凝胶电泳发现, 电泳结果中有明显的28S和18S条带出现, 无明显拖尾现象(图2), 说明所获得RNA样品完整性较好, 可用于下一步实验。

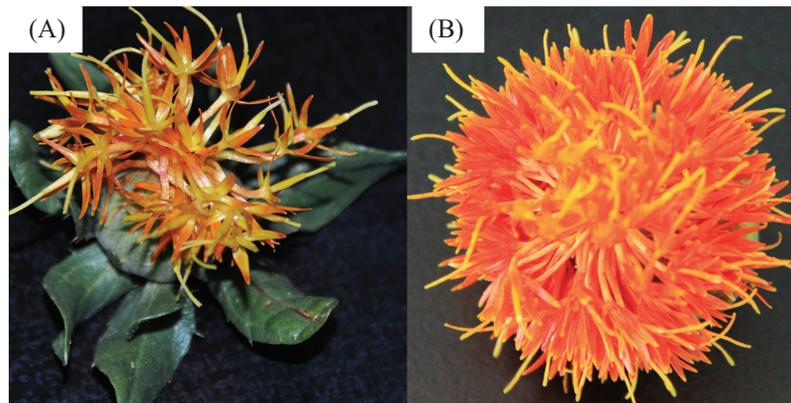
### 2.2 转录组数据产量及拼接结果

初花期和盛花期分别获得583 440和567 884条序列, 去除低质量、短序列(<50 bp)、载体和镶嵌序列等污染序列, 最终在初花期和盛花期分别获得用于分析的577 664和562 930条Clean reads片段, 其平均长度分别为427 bp和436 bp(表4)。使用MIRA拼

表3 荧光定量PCR引物

Table 3 The primers for Real-time PCR

基因	代码	引物(5'→3')
Gene	Code	Primers (5'→3')
18S rRNA		GAG AAA CGG CTA CCA CAT CCA A TCG TTT GAG CCC GGT ATT GTT A
Chalcone synthase	EC 2.3.1.74	TTG GAC CAG GTC GAG GAA AAA AAG ACA CAA GCG CTC GAC ATG
Chalcone isomerase	EC 5.5.1.6	AAG ATA CTT GGC AAT GGT TGC G CGG TAT GCA ACA TGC CGA A
Anthocyanidin synthase	EC 1.14.11.19	CAG CCC AAT CAA GCT TTT GC TGC CGG TGA AGG AGA AGA AA



A: 初花期; B: 盛花期。

A: the early period; B: the full period.

图1 不同开花时期的红花花瓣

Fig.1 Safflower flower phenotypes of different flowering period

表4 数据输出与拼接结果

数据统计	初花期	盛花期
Statistics of data	Early	Full
Low quality	50	67
Short reads (<50 bp)	796	928
Contamination sequences	2 785	2 694
Vector sequences	2 145	1 265
Clean reads	577 664	562 930
Average length	427 bp	436 bp

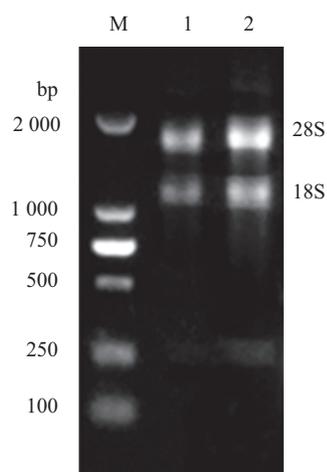


图2 红花花瓣总RNA的电泳图

Fig.2 Total RNA from the flowers of safflower

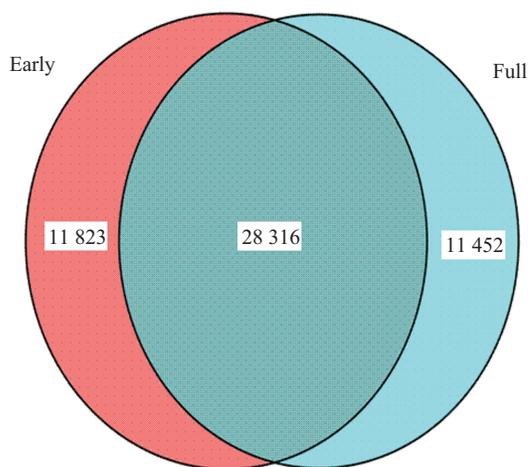


图3 差异表达基因统计

Fig.3 Gene expression statistics

接软件对去处理后的Clean reads序列进行拼接, 获得Unigenes序列。总计获得51 591个Unigenes, 其中最长的Unigene是5 109 bp, 平均长度为679 bp。

### 2.3 差异表达及相关性分析

将Unigenes在样本的reads通过RPM和RPKM值

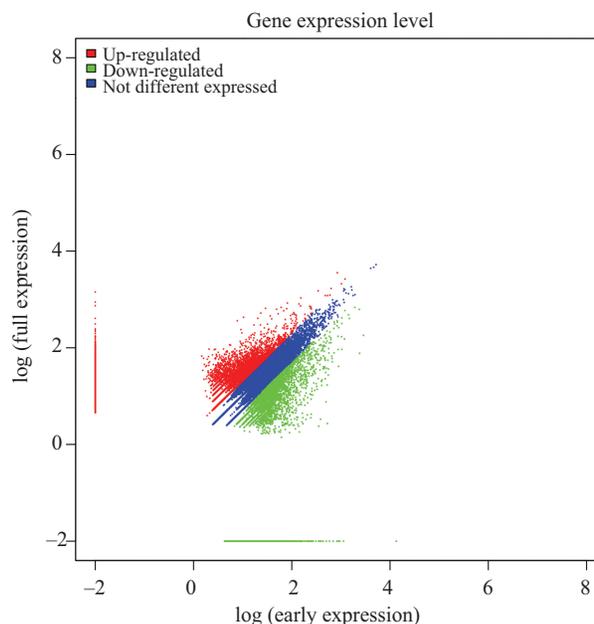


图4 样品差异表达及相关性分析

Fig.4 The bitmap of correlation and different expressed genes in two groups

进行归一化处理, 其中倍数绝对值 $\geq 2$ 的Unigenes为差异表达基因, 倍数绝对值 $\geq 2$ 且FDR修正值 $< 0.001$ 为显著差异表达基因。根据这个处理标准, 在初花期获得40 139个基因, 盛花期获得39 768个基因, 在两个时期同时表达的基因有28 316个基因(图3)。测序结果中, 两个样本间的差异表达分析见图4, 其中, 红色(上调)和绿色(下调)的点表示显著表达差异的基因(指差异倍数 $> 2$ 或样本特异性表达), 蓝色的点表示差异不显著的表达基因。测序结果表明, 差异表达的基因共有34 464个, 其中上调基因为18 043个, 下调基因为16 421个。

### 2.4 COG/KOG功能注释

为了比较样本间基因的功能注释信息, 将差异表达基因初花期和盛花期的Unigenes进行COG/KOG比对(图5), 结果表明, 两个样本中注释到各个功能组中的基因数量基本一致, 没有显著差异。

### 2.5 KEGG Pathway分析

对所有的Unigenes进行KEGG Pathway富集分析, 找到红花黄酮化合物代谢途径(图6), 图中带框的表明是在红花转录组中注释到的信息, 红色表示上调基因, 绿色表示下调基因。通过对红花转录物组的综合比对和序列分析, 在合成途径中我们找到了一些在红花转录组中存在序列的黄酮合成的关键酶基因(蓝色标示), 包括查尔酮合酶(chalcone synthase, CHS; EC

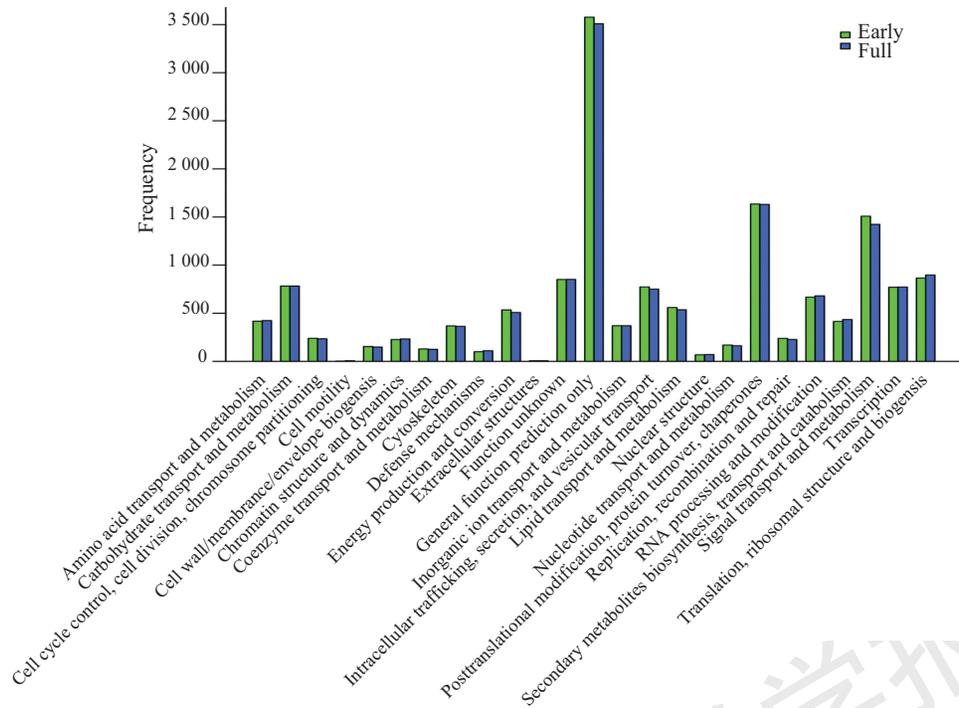
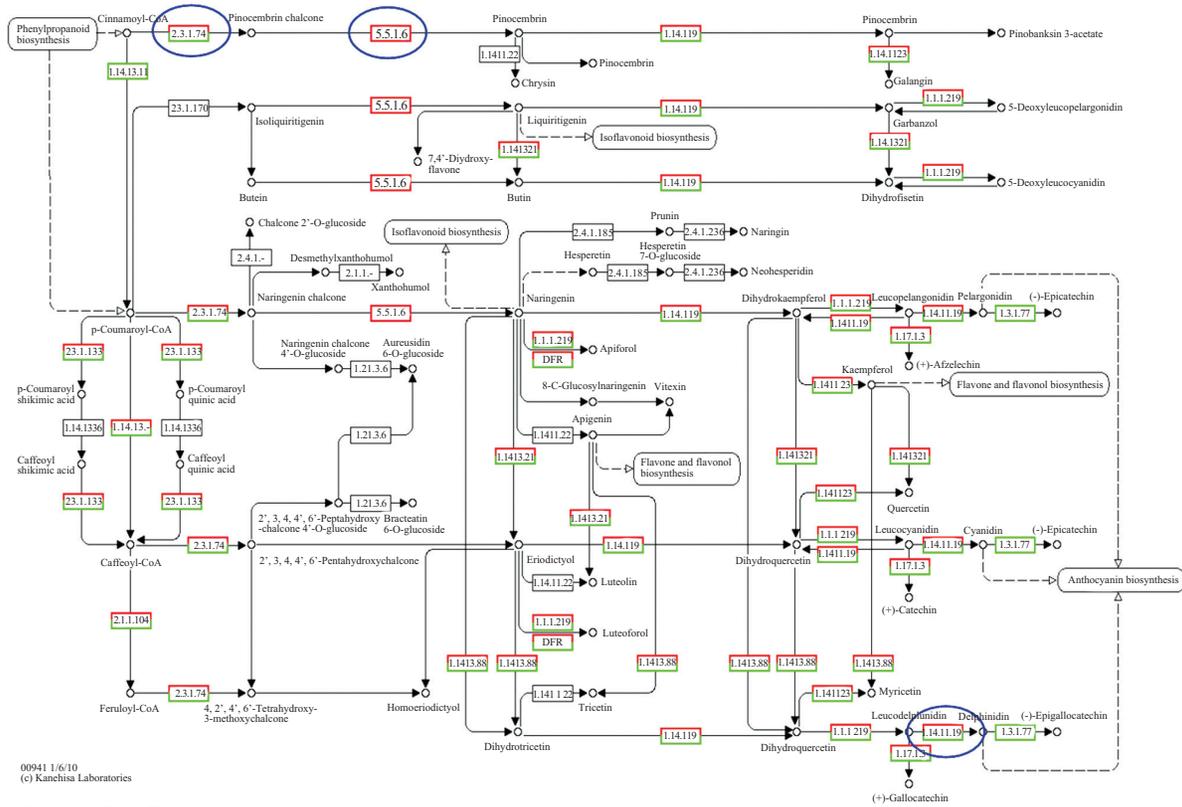


图5 样本间的COG/KOG功能分类  
Fig.5 COG/KOG distribution of two samples

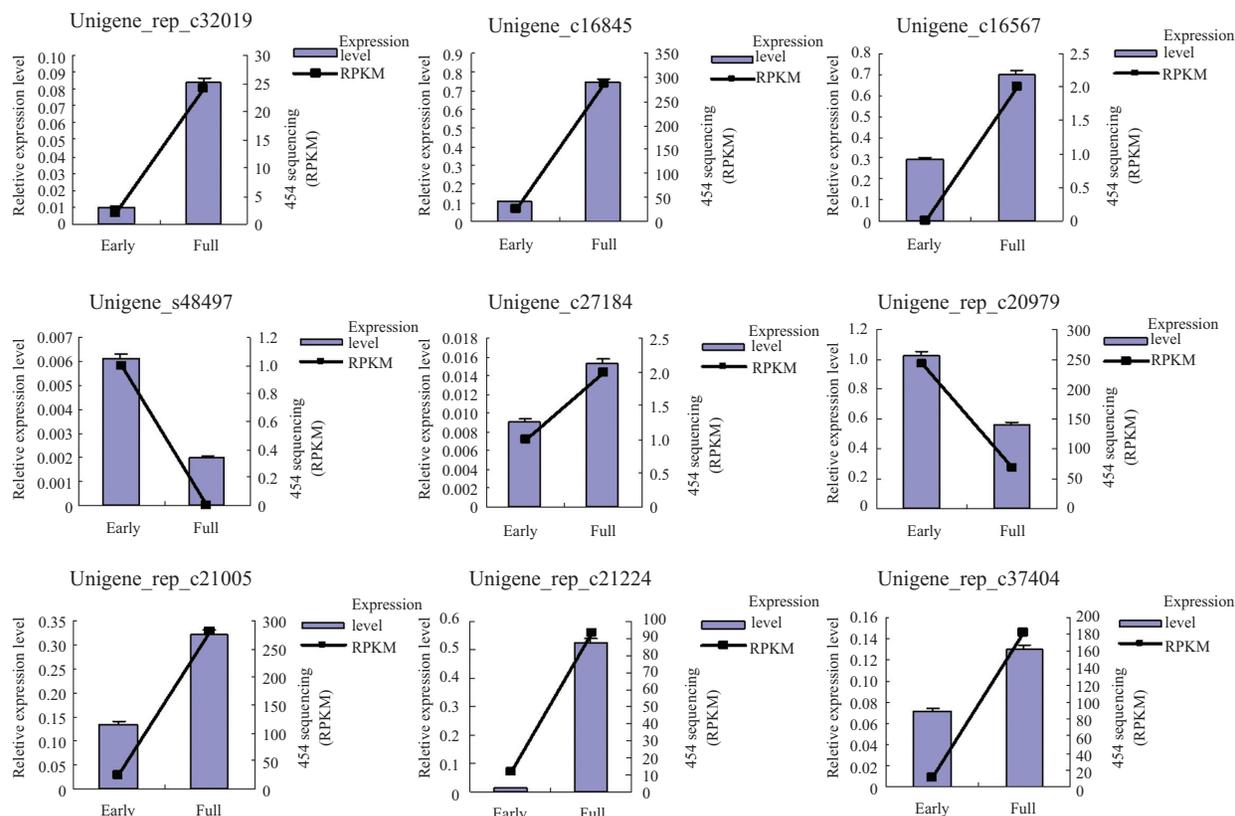
FLAVONOID BIOSYNTHESIS



蓝色: 挑选的关键酶基因。  
Blue: the gene of key enzyme.

图6 红花转录组测序中获得的黄酮代谢途径

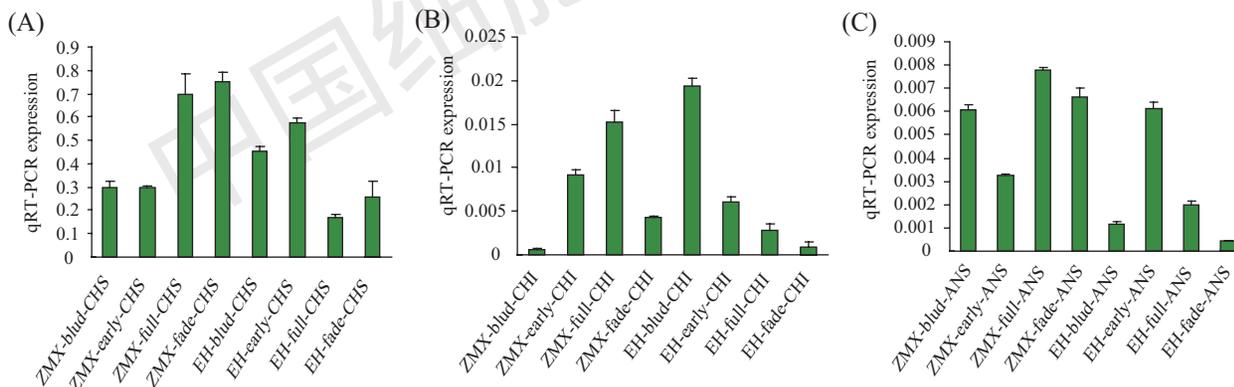
Fig.6 Representation of the metabolic pathways for flavonoid biosynthesis in safflower transcriptome



18S基因的相对表达量为1。Early: 初花期; Full: 盛花期。  
The relative expression of 18S is 1. Early: the early period; Full: the full period.

图7 候选unigenes的Real-time PCR验证

Fig.7 Real-time PCR validation of the selected unigenes



A: *CHS*基因的相对表达量; B: *CHI*基因的相对表达量; C: *ANS*基因的相对表达量。ZXM: 吉红油; EH: 吉红二号; blud: 花蕾期; early: 初花期; full: 盛花期; fade: 衰落期。

A: relative expression of *CHS*; B: relative expression of *CHI*; C: relative expression of *ANS*. ZXM: Zimeixi; EH: Jihongerhao; blud: the bud period; early: the early period; full: the full period; fade: the fade period.

图8 红花黄酮合成途径关键酶基因在不同花期的相对表达量

Fig.8 Relative expression of the key genes from the flavonoid synthesis in safflower

2.3.1.74)、查尔酮异构酶(chalcone isomerase, CHI; EC 5.5.1.6)和花青素合成酶(anthocyanidin synthase, ANS; EC 1.14.11.19)。

## 2.6 红花转录组测序数据的验证

根据本实验获得的相对荧光定量条件对转录

组测序中挑选的9个基因分别在初花期和盛花期进行检测, 由图7中可以看出, 候选的9个基因在初花期及盛花期的表达量均有差别, 除unigene\_s48497和unigene\_rep\_c20979在初花期的相对表达量高于盛花期以外, 其他的候选基因在盛花期的表达量均高

于初花期,这与红花转录组测序中得到的表达丰度一致。

## 2.7 红花黄酮化合物合成相关基因的验证

对挑选的三个关键酶基因中的22个Unigenes进行RT-PCR验证的结果表明,查尔酮合酶(unigene\_c16567)、查尔酮异构酶(unigene\_c27184)和花青素合成酶(unigene\_s48497)中的三个候选基因获得了特异性片段,片段大小分别为197 bp、209 bp、221 bp,将其连接到pEASY-T1载体上,将含有三个基因的载体菌液送去上海生工生物工程股份有限公司测序,测序结果与转录组中获得的序列完全吻合,说明这三个候选基因确实存在于红花中。

## 2.8 红花黄酮化合物合成相关基因的表达分析

由图8中各个基因的相对表达量看出, *CHS*、*CHI*、*ANS*三个基因在不同品种不同开花时期表现出明显的差异。查尔酮合酶在两个红花品种中的各个开花时期的的相对表达量最高,即使在其表达量最低的“吉红二号”盛花期,也要显著高于其他两个基因的表达量。查尔酮异构酶表达量最高的出现在“吉红二号”花蕾期,其次为“吉红油”的初花期和盛花期,其他时期的表达量均较低。花青素合成酶基因在“吉红油”中各个时期都表现相对较高的表达量,而在“吉红二号”中,初花期的表达量高明显高于花蕾期、盛花期和衰落期,衰落期的表达量最低。

## 3 讨论

454测序技术主要的优势在于覆盖深度广,读长相对较长,特别适合一些未知基因组序列的物种<sup>[22]</sup>。本研究基于红花黄酮化合物代谢途径的重要性,对红花不同开花时期的花瓣进行转录组测序来挖掘与红花黄酮合成相关的基因,这些基因的表达可能会与红花黄酮的合成相关。本研究采用454测序技术对红花的初花期和盛花期花瓣进行测序,在两个样本中分别获得了577 664和 562 930条Clean reads,平均长度分别为427 bp和436 bp。获得的数据数量与进行454测序的金银花<sup>[23]</sup>和足叶草<sup>[24]</sup>中得到的数量相当,但多于蚕豆<sup>[25]</sup>中获得的转录本。红花中所有的测序结果都将为红花的进一步研究提供大量理论数据。

查尔酮合酶(*CHS*)基因是黄酮化合物合成途径的第一个关键酶基因,它参与黄酮合成、色素合成、抗性等多种生理生化反应。不同的植物、不同的组织部位及不同时期*CHS*的表达量不同。红花*CHS*基

因在不同品种、不同花期的表达量差异较大,但表达量都相对较高,分别在“吉红油”姊妹系的衰落期和“吉红二号”的初花期表达量最高, *CHS*在红花的不同开花时期的相对表达量均高于*CHI*和*ANS*,说明红花*CHS*基因在红花黄酮合成中具有非常重要的作用。本研究中, *CHI*在“吉红二号”中的表达量呈一定的规律性,花蕾期表达量最高,随着花开放的过程, *CHI*的表达量逐渐下降,这与蔡雪玲等<sup>[26]</sup>研究*CHI*基因在白花水仙中的表达趋势是一致的。但*CHI*基因在“吉红油”中盛花期表达量最高,而花蕾期表达量最低,说明红花*CHI*基因也可能参与红花花色变化和色素的合成。花青素合成酶(*ANS*)作为黄酮化合物合成通路末端的关键酶,在花色形成及果实着色中发挥重要作用<sup>[27]</sup>。红花的*ANS*在“吉红二号”中的花蕾期表达量较低,在初花期表达量迅速上升,到盛花期又下降。红花的初花期为浅黄色,到盛花期为橙色,说明在红花花发育过程中, *ANS*可能参与花色的变化,导致表达量升高又下降,这与鹤望兰<sup>[28]</sup>花青素合成酶基因的表达规律一致。而*ANS*在“吉红油”各个开花时期的表达量均较高。以上说明, *ANS*的表达在不同品种不同组织中具有一定的差异。

本研究对红花不同开花时期进行转录组测序,获得了大量的转录组数据,并挖掘到红花黄酮化合物合成途径中的关键酶基因,为红花的资源开发研究及黄酮化合物代谢调控研究奠定了基础。

## 参考文献 (References)

- 1 Kim EO, Lee KT, Choi SW. Chemical comparison of germinated and ungerminated-safflower (*Carthamus tinctorius*) seeds. *J Korean Soc Food Sci Nutr* 2008; 37: 1162-7.
- 2 Kim EO, Oh JH, Lee SK. Antioxidant properties and quantification of phenolic compounds from safflower (*Carthamus tinctorius* L.) seeds. *Food Sci* 2007; 16: 71-7.
- 3 范莉,赵海誉,濮润,韩健,王宝荣,果德安. 红花的黄酮类化学成分研究. *中国药学杂志*(Fan Li, Zhao Haiyu, Pu Run, Han Jian, Wang Baorong, Guo Dean. Study on Flavonoids in the Flower of *Carthamus tinctorius* L. *Chin Pharm J* 2011; 46(5): 333-7.
- 4 Yaginuma S, Shiraishi T, Igarashi K. Developmental transition of the flavonoid contents in safflower leaves during stress-loaded cultivation. *Biosci Biotechnol Biochem* 2003; 67(8): 1691-8.
- 5 Zhou X, Tang L, Xu Y, Zhou G, Wang Z. Towards a better understanding of medicinal uses of *Carthamus tinctorius* L. in traditional Chinese medicine: A phytochemical and pharmacological review. *J Ethnopharmacol* 2014; 1(151): 27-43.
- 6 周华,张新,刘腾云,余发新. 高通量转录组测序的数据分析与基因发掘. *江西科学*(Zhou Hua, Zhang Xin, Liu Tengyun, Yu Faxin. Data processing and gene discovery of high-throughput

- transcriptome sequencing. *Jiangxi Science* 2012 ; 30(5): 607-11.
- 7 薛建江, 邱景富. 病原菌感染宿主的转录组学研究进展. *河北北方学院学报(医学版)*(Xue Jianjiang, Qiu Jingfu. *Journal of Hebei North University, Medical Edition*) 2007; 24(5): 63-6.
- 8 Costa V, Angelini C, de Feis I, Ciccodicola A. Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol* 2010; 2010: 853916.
- 9 Ruan Y, Le Ber P, Ng HH, Liu ET. Interrogating the transcriptome. *Trends Biotechnol* 2004; 22(1): 23-30.
- 10 Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* 2009; 10(1): 57-63.
- 11 Wang ZW, Jiang C, Wen Q, Wang N, Tao YY, Xu LA. Deep sequencing of the *Camellia chekiangoleosa* transcriptome revealed candidate genes for anthocyanin in biosynthesis. *Gene* 2014; 538(1): 1-7.
- 12 Alagna F, Agostino ND, Torchia L, Gomez-Jimenez MC. Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *BMC Genomics* 2009; 10: 399.
- 13 He L, Xu X, Li Y, Li C, Zhu Y, Yan H, *et al.* Transcriptome analysis of buds and leaves using 454 pyrosequencing to discover genes associated with the biosynthesis of active ingredients in *Lonicera japonica* Thunb. *PLoS One* 2013; 8(4): e62922.
- 14 Bhattacharyya D, Sinha R, Hazra S, Chattopadhyay S. *De novo* transcriptome analysis using 454 pyrosequencing of the Himalayan Mayapple, *Podophyllum hexandrum*. *BMC Genomics* 2013; 14: 748.
- 15 Duarte J, Rivière N, Baranger A, Aubert G, Burstin J, Cornet L, *et al.* Transcriptome sequencing for high throughput SNP development and genetic mapping in Pea. *MC Genomics* 2014; 15(1): 126.
- 16 Lassmann T, Hayashizaki Y, Daub CO. Tag Dust—a program to eliminate artifacts from next generation sequencing data. *Bioinformatics* 2009; 25(21): 2839-40.
- 17 Chen YA, Lin CC, Wang CD, Wu HB, Hwang PI. An optimized procedure greatly improves EST vector contamination removal. *BMC Genomics* 2007; 8: 416.
- 18 Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WE, Wetter T, *et al.* Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 2004; 14(6): 1147-59.
- 19 Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 2008; 5(7): 621-8.
- 20 Salomonis N, Hanspers K, Zambon AC, Vranizan K, Lawlor SC, Dahlquist KD. GenMAPP 2: New features and resources for pathway analysis. *BMC Bioinformatics* 2007; 8: 217.
- 21 Li H, Dong Y, Yang J, Liu X, Wang Y, Yao N. *De novo* transcriptome of safflower and the identification of putative genes for oleosin and the biosynthesis of flavonoids. *PLoS One* 2012; 7(2): e30987.
- 22 Zhang J, Franks R G, Liu X, Kang M, Keebler JE, Schaff JE. *De novo* sequencing, characterization, and comparison of inflorescence transcriptomes of *Cornus canadensis* and *C. florida* (Cornaceae). *PLoS One* 2013; 8(12): e82674.
- 23 He L, Xu X, Li Y, Li C, Zhu Y, Yan H. Transcriptome analysis of buds and leaves using 454 pyrosequencing to discover genes associated with the biosynthesis of active ingredients in *Lonicera japonica* Thunb. *PLoS One* 2013; 8(4): e62922.
- 24 Bhattacharyya D, Sinha R, Hazra S, Datta R, Chattopadhyay S. *De novo* transcriptome analysis using 454 pyrosequencing of the Himalayan Mayapple, *Podophyllum hexandrum*. *BMC Genomics* 2013; 14: 748.
- 25 Suresh S, Park JH, Cho GT, Datta R, Chattopadhyay S. Development and molecular characterization of 55 novel polymorphic cDNA-SSR markers in faba bean (*Vicia faba* L.) using 454 pyrosequencing. *Molecules* 2013; 18(2): 1844-56.
- 26 蔡雪玲, 陈晓静, 叶一江, 何玮毅, 申艳红. 中国水仙查尔酮异构酶基因的克隆与表达分析. *热带作物学报*(Cai Xueling, Chen Xiaojing, Ye Yijiang, He Weiyi, Shen Yanhong. *Cloning and expression analysis of chalcone isomerase gene from *Narcissus tazetta* var. *chinensis*. Chinese Journal of Tropical Crops*) 2011; 32(12): 2287-92.
- 27 Saito K, Kobayashi M, Gong Z, Tanaka Y, Yamazaki M. Direct evidence for anthocyanidin synthase as a 2-oxoglutarate-dependent oxygenase: Molecular cloning and functional expression of cDNA from a red form of *Perilla frutescens*. *Plant J* 1999; 17(2): 181-9.
- 28 樊荣辉, 黄敏玲, 吴建设, 钟淮钦. 鹤望兰花青素合成酶基因SrANS的克隆及表达分析. *中国细胞生物学学报*(Fan Ronghui, Huang Minling, Wu Jianshe, Zhong Huaiqin. *Cloning and expression of anthocyanidin synthase in *Strelitzia reginae* banks. Chinese Journal of Cell Biology*) 2013; 35(11): 1620-5.