

藏红花柱头小RNA高通量测序及生物信息学分析

卢晓慧^{1,2}, 朱志明³, 岳俊阳^{2,4*}

¹嘉兴职业技术学院信息技术分院, 浙江嘉兴314000

²安徽农业大学园艺学院, 合肥230036

³嘉兴天禾藏红花专业合作社, 浙江嘉兴314000

⁴合肥工业大学食品科学与工程学院, 合肥230009

摘要: 通过构建藏红花开花前2 d与开花日两个不同时期柱头的小RNA文库并进行高通量测序, 分别获得3 711 809和2 303 043条序列。借助生物信息学的方法, 共鉴定出164条保守的miRNA。其中, 随开花过程下调表达和上调表达的miRNA分别是71和46条。靶基因预测及功能富集分析显示它们可能参与了硫辛酸生物合成、细胞程序性凋亡、花粉细胞的分化、脱落酸介导的信号途径等生物过程。本文筛选的藏红花成熟过程中的miRNA为今后脱辅基类胡萝卜素代谢调控网络的研究提供了可能的线索。

关键词: 藏红花; 小RNA; miRNA; 高通量测序; 生物信息学

藏红花(*Crocus sativus*)是一种鸢尾科番红花属的多年生草本植物, 又称番红花、西红花。藏红花原产于希腊、土耳其、西班牙等地中海沿岸国家, 并在明朝时期从印度经西藏传入我国, 因此得名。1965年, 国内开始逐步引种人工栽培藏红花, 至今已在江苏、浙江、上海、山东等地广泛种植。据《中华人民共和国药典》记载, 藏红花是一种名贵的中药材, 具有活血化瘀、凉血解毒、舒经活络、消肿止痛、解郁安神等功效, 可用于预防和治疗脑血栓、脉管炎、血亏体虚、产后瘀阻、惊悸癫狂等疾病(国家药典委员会2010)。藏红花的干燥柱头作为主要药用部分, 其中含有丰富的藏红花素(crocin)、藏红花酸(crocetin)、藏红花醛(safranal)和藏红花甙(picrocrocic)等活性成分, 能全面增强人体免疫力, 有效抑制肿瘤细胞生长(李伟平等2011)。

藏红花素、藏红花酸、藏红花醛和藏红花甙同属于脱辅基类胡萝卜素物质, 经由类胡萝卜素氧化生成(Baba等2015a)。目前该生物合成途径主要发现于番红花属的植物中, 现已分离鉴定出CsCCD、CsZCD、CsULT1等多个功能基因(Bouvier等2003)。另外, 通过高通量测序和生物信息学分析相结合的方法, 脱辅基类胡萝卜素生物合成途径中的一些重要功能基因及其相关的调控因子也得到初步鉴定。Baba等(2015b)通过转录组测序在藏红花柱头中发现许多脱辅基类胡萝卜素代谢相关的功能基因具有高水平的表达, 在开花之时表达最强。同样地, Jain等(2016)也借助转录组测序

的方法构建了脱辅基类胡萝卜素生物合成的调控网络。

微小RNA (microRNA, 简为miRNA)是一类长18~30 nt的内源性单链非编码小RNA (small RNA, 简为smRNA)分子, 广泛存在于动植物中(Bartel 2004)。研究表明, miRNA具有高度保守性, 在植物个体发育和环境适应过程中发挥着重要的调控作用(Axtell 2013)。miRNA通过碱基互补配对的方式识别靶基因, 并特异性切割靶基因的mRNA或抑制mRNA的翻译从而来调节基因的表达水平, 表现为一种转录后水平的负调控方式(Brodersen等2008)。近年来, 围绕miRNA的鉴定、靶基因的识别、靶基因的功能分析及调控网络的构建已经开展了较多的研究。相比较传统的直接分离法, 高通量测序可以有效分离鉴定一些低丰度表达或时空特异性表达的miRNA (Robin等2016)。再结合生物信息学的方法, 不仅可以快速鉴定保守的miRNA, 还可能预测新的miRNA, 同时还能够分析miRNA的表达特征(Normand和Yanai 2013)。因此, 本文通过构建藏红花开花前2 d与开花日两个不同时期柱头的小RNA文库, 利用Illumina平台进行高通量测序, 而后借助生物信息学的分析方法鉴定了藏红花中保守的miRNA, 筛选出两个时期差异表达的miRNA, 并基于已公开发表的藏红花转

收稿 2018-01-08 修订 2018-04-26

资助 嘉兴科技公益性技术项目(2017AY33032)。

* 通讯作者(aaran.yue@gmail.com)。

录组的数据, 预测了差异表达miRNA的靶基因, 以期发现对藏红花柱头发育成熟及其中类胡萝卜素物质代谢合成具有重要调控功能的miRNA。

1 材料与方法

1.1 材料

本文用来进行小RNA测序的藏红花(*Crocus sativus* L.)为人工栽培的秋花类品种, 采自浙江嘉兴天禾藏红花种植基地。参考Rubio等(2008)文献, 藏红花的香气在开花之日急剧增大, 因此本文选择藏红花开花前2 d与开花日两个时期的柱头组织样品, 分别取名为2DBA (2 days before anthesis)和DOA (day of anthesis)。样品采摘后, 迅速用液氮速冻并保存于-80°C冰箱待用。

1.2 方法

1.2.1 藏红花柱头小RNA文库构建与测序

两个不同时期的柱头组织样品2DBA和DOA分别在液氮中研磨至粉末状, 加入TRIzol试剂(Invitrogen, 美国)混合均匀, 然后提取总RNA, 并使用安捷伦2100生物分析仪(Agilent Technologies, 美国)测定所提取RNA的含量和质量, 只保留28S和18S比值接近2.0的样品。接下来, 使用15%变性的聚丙烯酰胺凝胶电泳分离上述样品, 回收纯化18~30 nt的序列部分。最后, 在T₄ RNA连接酶(TaKaRa, 大连)的作用下分别在5'端和3'端加上接头序列, 并经RT-PCR扩展后进行Illumina高通量测序。

1.2.2 测序序列的基本分析

对Illumina高通量测序所获得的原始序列(raw reads)进行如下的基本分析: (1)去除5'端接头序列(GTTCAGAGTTCTACAGTCCGACGATC)和3'端接头序列(TCGTATGCCGCTTCTGCTTG); (2)通过ASCII值进行质量控制, 只保留质量值(Q=ASCII值-64)大于9的高质量序列; (3)去除小于18 nt和大于30 nt的序列, 得到纯净序列(clean reads); (4)统计18~30 nt之间每一个长度下的序列数量(tags)和序列读数(reads); (5)设定阈值E为0.1, 使用本地BLASTN程序把所有获得的纯净序列逐步与藏红花cDNA序列、miRBase数据库(版本21) (Kozomara和Griffiths-Jones 2014)、Rfam数据库(版本12.1) (Nawrocki等2014)进行比对, 以确定编码区的短片段、潜在的miRNA、rRNA、tRNA、snRNA、snoRNA等各类小RNA序列。

1.2.3 保守miRNA的鉴定

因为与藏红花cDNA匹配的序列很可能是基因编码区的短片段, 所以先把这部分序列排除在外, 而把剩下的序列继续与miRBase数据库中所有搜集到的植物前体miRNA序列和成熟miRNA序列进行比对(E值为0.1)。对于匹配上的序列, 只保留不多于2个错配碱基的部分。另外, 为了减小高通量测序过程中可能引入的错误, 只把读数在3及其以上的小RNA序列鉴定为藏红花保守的miRNA。

1.2.4 差异表达miRNA的筛选

使用R语言中Bioconductor的DESeq包(Anders和Huber 2010)来对2DBA和DOA两个时期柱头中保守miRNA的表达量进行标准化和显著性检验($P < 0.05$)。将符合条件的miRNA定义为差异表达miRNA。

1.2.5 miRNA靶基因的预测

通过psRNATarget在线工具(Dai和Zhao 2011)将筛选到的差异表达miRNA与藏红花cDNA (Jain等2016)进行序列比对分析, 可以预测miRNA作用的靶基因。在比对结果的筛选过程中, 采用了以下较为严格的标准: (1) miRNA序列中2~8 nt位置的碱基称为种子区域(seed region), 需要保证该区段严格的匹配; (2)最大的期望值设定为3; (3)互补得分值设定为20; (4)在结合位点处进行解链的最大能量值为25; (5)通过在9~11 nt位置是否严格匹配来判断miRNA调控靶基因的方式。

1.2.6 靶基因功能富集分析

基因本体论(gene ontology, GO)是一个在生物信息学领域中广泛使用的注释方法, 可以实现对基因序列的直观语言描述, 主要涵盖三个方面: 生物过程(biological process, BP)、细胞组分(cellular component, CC)和分子功能(molecular function, MF)。本文借助Blast2GO工具(Conesa等2005)对miRNA作用的靶基因进行了生物过程、细胞组分和分子功能三个方面的综合GO注释。同时, 使用超几何分布检验($P < 0.05$)对每一个注释的GO条目(GO term)进行富集分析, 公式如下:

$$f(k; n, m, N) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}$$

$$P(\text{值}) = 1 - f(k; n, m, N)$$

其中, N 代表藏红花中所有基因获得注释的GO条目总数, m 代表某个特定GO条目在所有基因

中的出现数, n 代表所有靶基因获得注释的GO条目数, k 代表特定GO条目在靶基因中的出现数。

2 实验结果

2.1 藏红花柱头小RNA的种类组成

基于Illumina高通量测序, 分别在藏红花2DBA和DOA柱头中各获得15 000 000条原始序列读数

(raw reads)。通过去除5'端和3'端接头、污染序列及低质量的序列, 最后共保留14 294 654和14 534 515条长度在18~30 nt之间的纯净序列读数(clean reads), 见表1。

对测序获得的数据根据长度进行分析可知, 无论是2DBA还是DOA柱头, 其中都以21 nt长度的序列占有的读数比例最高(图1-A)。同时, 这些读

表1 藏红花柱头小RNA测序数据的基本分析

Table 1 Statistics of sequencing reads in the stigmas of *C. sativus*

读数类型	2DBA柱头		DOA柱头	
	读数	比例/%	读数	比例/%
总读数	15 000 000	100	15 000 000	100
低质量读数	27 536	0.184	30 035	0.2
高质量读数	14 972 464	99.816	14 969 965	99.8
无3'接头	95 157	0.634	90 816	0.605
无插入片段	98 193	0.655	94 051	0.627
无5'接头	33 410	0.223	22 823	0.152
小于18 nt的片段数	345 947	2.306	188 028	1.254
大于30 nt的片段数	102 815	0.685	37 873	0.252
多聚A	2 288	0.015	1 859	0.012
纯净读数	14 294 654	95.298	14 534 515	96.897

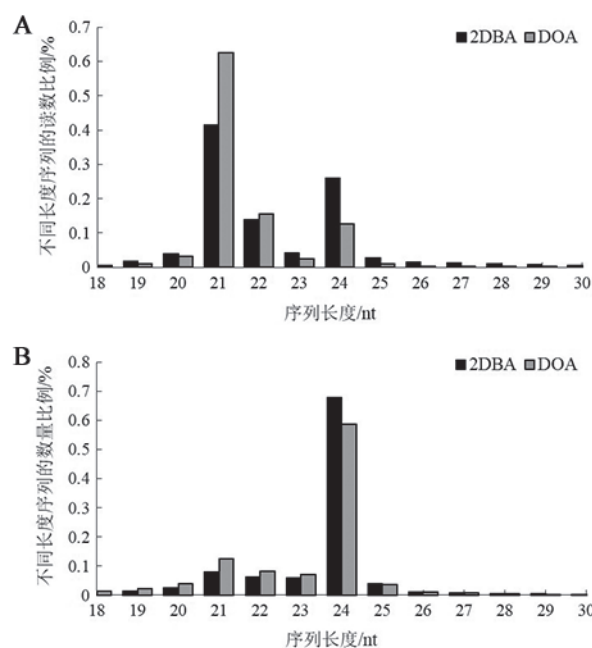


图1 藏红花2DBA和DOA柱头中不同长度序列小RNA的读数与数量

Fig.1 Distribution of small RNAs with different sequence length according to their total reads and unique tags in the stigmas of *C. sativus*

数共分布在3 711 809和2 303 043条序列中(tags), 但其中却以24 nt长度的序列数量居多(图1-B)。接下来, 把这些序列与藏红花cDNA序列、miRBase数据库、Rfam数据库进行比对, 分别发现在2DBA和DOA柱头中各有2 037 143和1 219 065条序列匹配上cDNA, 以及346 222和225 007条序列匹配上Rfam数据库(表2)。

2.2 藏红花柱头保守miRNA的鉴定及家族分类

通过小RNA高通量测序的方法可以有效发现并鉴定藏红花中保守的miRNA。经过上述逐步分析并进行严格的筛选, 分别在藏红花2DBA和DOA柱头中鉴定出132和112条保守的miRNA。其中, 同时存在于两个时期的保守miRNA有80条, 其余的52和32条miRNA分别为各时期所特有(图2-A)。

基于Rfam数据库中已构建的miRNA家族分类特征, 可以分别把2DBA和DOA柱头中保守的miRNA各归类到44和40个miRNA家族中。其中, 31个家族是两者所共有(图2-B)。而2DBA柱头中MIR894家族成员最多, 为10个; DOA柱头中以MIR828家族成员为最多, 共有11个(图2-C)。

表2 藏红花柱头小RNA的种类组成

Table 2 Categorization of unique tags in the stigmas of *C. sativus*

序列分类	2DBA柱头		DOA柱头	
	数量	比例/%	数量	比例/%
总序列数	3 711 809	100	2 303 043	100
cDNA片段	2 037 143	54.883	1 219 065	52.933
rRNA	113 236	3.051	65 537	2.846
tRNA	26 610	0.717	17 256	0.749
snoRNA	38 336	1.033	27 181	1.18
snRNA	24 076	0.649	18 233	0.792
其他比对Rfam的小RNA	143 964	3.879	96 800	4.203
保守的miRNA	52	0.001	42	0.002
未分类的小RNA	1 328 392	35.788	858 929	37.295

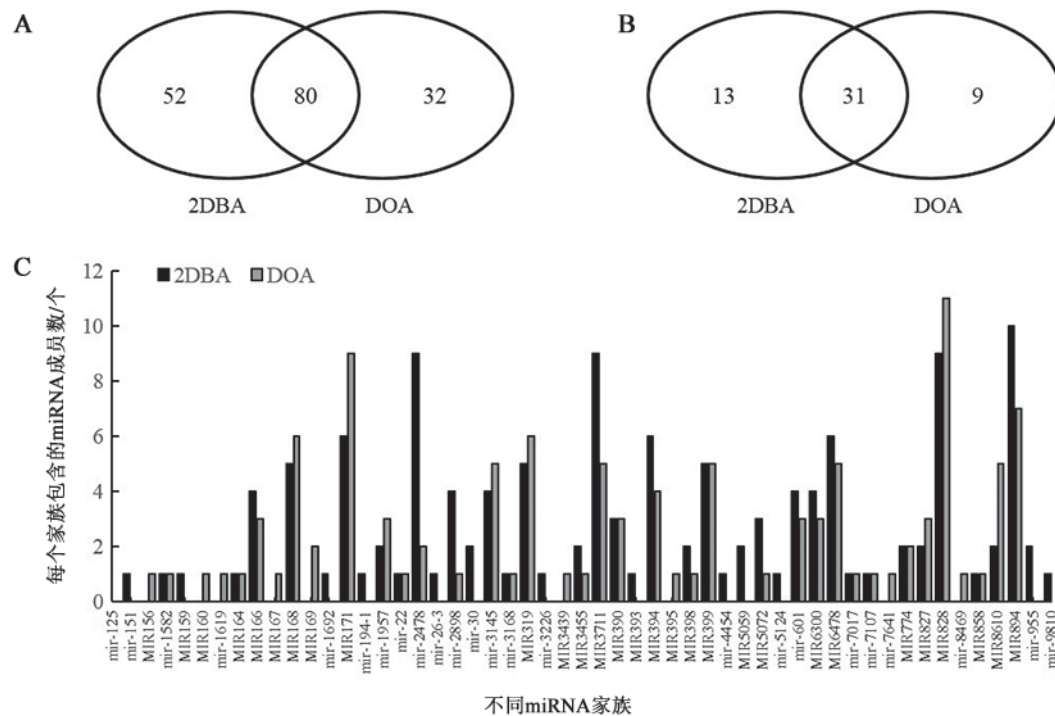


图2 藏红花2DBA和DOA柱头中共同具有的保守miRNA数量及其家族分布
Fig.2 Distribution of conserved miRNA reads and their family members that shared between 2DBA and DOA stigmas of *C. sativus*

A: 保守的miRNA数量; B: miRNA家族数量; C: 每个miRNA家族所包含的成员数。

2.3 藏红花柱头保守miRNA的表达水平及其在两个时期内的表达模式

小RNA高通量测序不仅能够鉴定物种中保守的miRNA, 而且可以获得不同miRNA的表达水平。通过统计已鉴定到的164条保守miRNA的表达水平及其家族分布发现(图3): 在2DBA时期, MIR168家族具有最高的表达水平, 为6 181个读数, 接下来

依次为MIR894 (3 675个读数)、MIR171 (2 226个读数)和MIR828 (916个读数); 而在DOA时期, MIR168家族仍然具有最高的表达水平, 为4 348个读数, 此外读数大于1 000的家族还有MIR828 (2 224个读数)和MIR171 (1 386个读数)。

比较2DBA和DOA两个时期内相同家族中miRNA的表达模式, 有29个家族的总miRNA表达

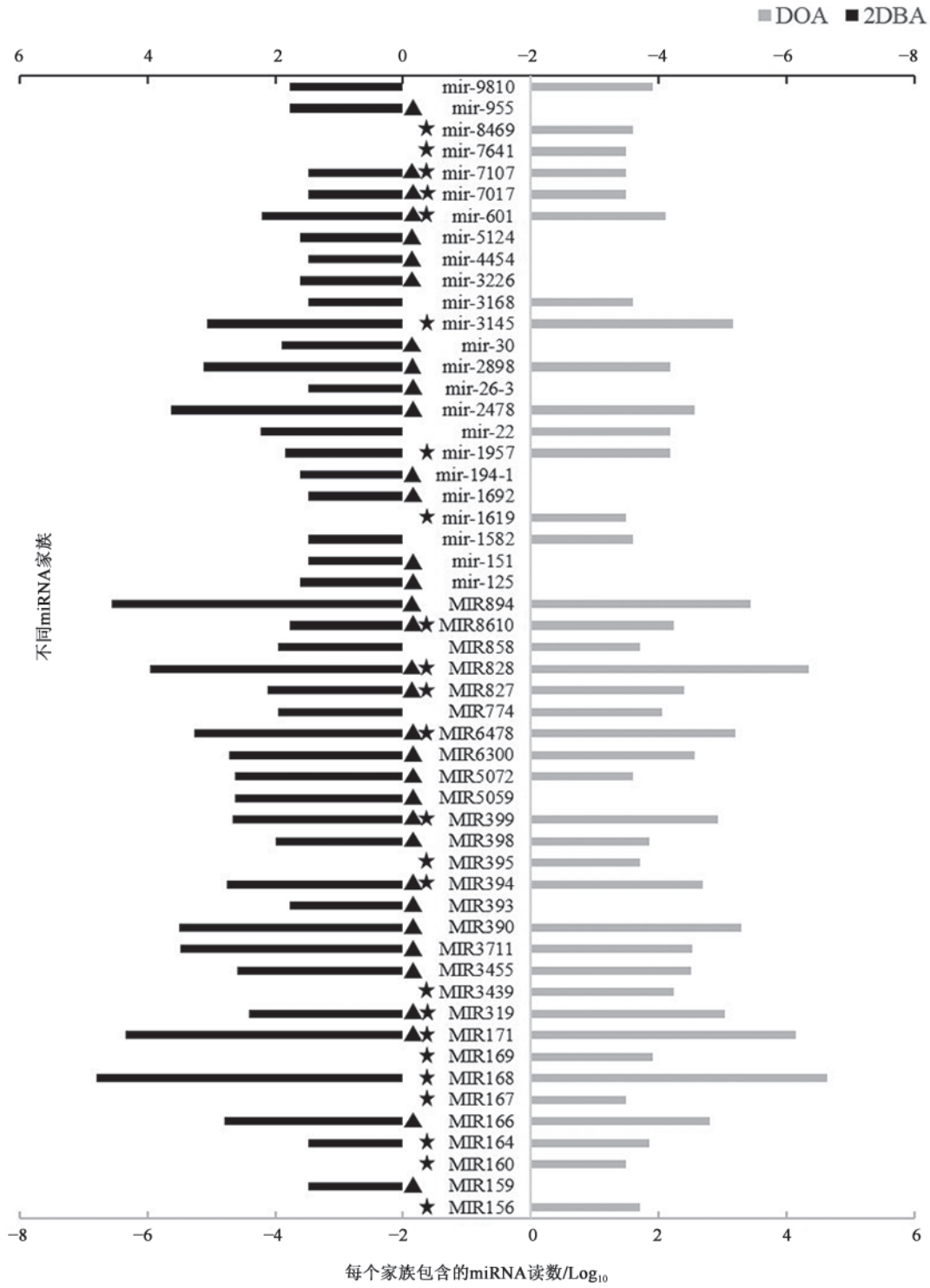


图3 藏红花2DBA和DOA柱头中保守miRNA的表达水平

Fig.3 Expression levels of conserved miRNAs in the 2DBA and DOA stigmas of *C. sativus*

每个家族中miRNA的总读数值经过对数处理。三角形表示该家族中含有下调表达的miRNA,五角星表示该家族中含有上调表达的miRNA。

水平降低, 22个家族的总miRNA表达水平升高, 另外有2个家族(mir-7017和mir-7107)的总miRNA表达水平不变。其中, 表达水平的差异倍数在2倍以上的共有32个miRNA家族, 包含18个下调表达的家族和14个上调表达的家族。

2.4 差异表达miRNA的筛选和靶基因的预测及功能分析

在2DBA和DOA柱头中, 分别有52和32条特异表达的miRNA (图2-A), 直接被认为是差异表达的miRNA。另外, 对于80条同时在两个时期都有表达的miRNA, 借助R语言的DESeq包筛选出33条差异表达的miRNA ($P < 0.05$)。其中, 在2DBA柱头中表达水平高的有19条, 在DOA柱头中表达水平高的有14条。因此, 本文在藏红花开花阶段的柱头

成熟过程中共筛选到71条下调表达和46条上调表达的miRNA。家族分类显示下调表达和上调表达的miRNA分别来自34 (带三角形标记)和24 (带五角星标记)个不同的家族(图3)。

miRNA主要通过切割靶基因的转录本或抑制其翻译来实现功能, 因此找到miRNA作用的靶基因是理解其生物学功能的基础。借助psRNATarget在线工具, 能够在71条下调表达的miRNA中, 预测到52个靶基因; 在上调表达的46条miRNA中, 预测到49个靶基因。随后, 通过GO注释, 发现这些靶基因的功能主要集中在生物调控(biological regulation, GO:0065007)、细胞过程(cellular process, GO:0009987)、代谢过程(metabolic process, GO:0008152)、色素(pigmentation, GO:0043473)、应激反应(response to stimulus, GO:005-

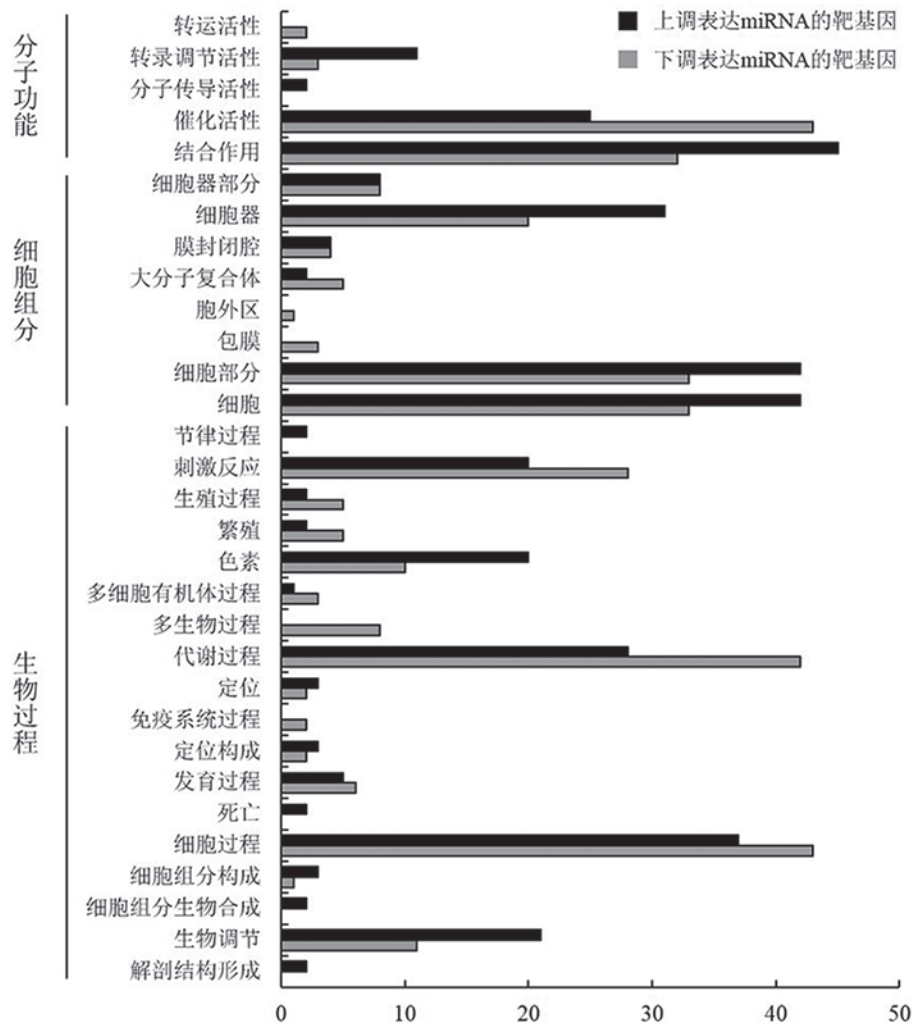


图4 藏红花柱头差异表达miRNA的靶基因的GO注释

Fig.4 GO annotation of the predicted target genes of differentially expressed miRNAs in the stigmas of *C. sativus*

0896)、细胞(cell, GO:0005623)、细胞器(organelle, GO:0043226)、分子结合(binding, GO:0005488)、酶催化(catalytic activity, GO:0003824)等(图4)。

为了查看靶基因的具体功能, 本文进一步通过超几何分布检验找到显著富集的GO条目。如表3所示, 在下调表达miRNA的靶基因中最显著富集的GO条目有GO:0009105 (lipoic acid biosynthetic process, 硫辛酸生物合成过程)和GO:0017140 (lipoic acid synthase activity, 硫辛酸合成活性), 其 P 值都是 $1.99E-5$ 。此外, 在下调表达miRNA的靶基因中显

著富集的GO条目还有GO:0009107 (lipoate biosynthetic process, 共轭硫辛酸生物合成过程)、GO:0016591 (DNA-directed RNA polymerase II, holoenzyme, DNA定向RNA聚合酶II全酶)、GO:0016992 (lipoate synthase activity, 共轭硫辛酸合成活性)等。而在上调表达miRNA的靶基因中最显著富集的GO条目有GO:0043068 (positive regulation of programmed cell death, 正向调控程序性细胞死亡)和GO:0030785 [(ribulose-bisphosphate carboxylase)-lysine N-methyltransferase activity, (二磷酸核酮糖羧化酶)-赖氨酸N-甲基转移酶活性] 其 P 值都是 $1.39E-5$ 。此外, 在上调表达miRNA的靶基因中显著富集的GO条目还有GO:0045893 (positive regulation of transcription, DNA-dependent, DNA依赖的正向转录调控)、GO:0008026 (ATP-dependent helicase activity, 依赖ATP的解旋酶活性)等(表4)。

表3 随藏红花柱头成熟过程表达下调miRNA的靶基因中最显著富集的10个GO条目

Table 3 The top ten GO terms enriched in the predicted target genes of down-regulated miRNAs during the ripening of stigmas in *C. sativus*

GO分类及条目	GO描述	P 值
生物过程		
GO:0009105	硫辛酸生物合成过程	1.99E-05
GO:0009107	共轭硫辛酸生物合成过程	2.37E-04
GO:0015074	DNA整合作用	8.47E-03
细胞组分		
GO:0016591	DNA定向RNA聚合酶II全酶	2.99E-05
GO:0005665	DNA定向RNA聚合酶II核心酶	8.47E-04
分子功能		
GO:0017140	硫辛酸合成活性	1.99E-05
GO:0016992	共轭硫辛酸合成活性	4.18E-05
GO:0032549	核糖核甙酸结合作用	1.58E-03
GO:0015391	阳离子同向转运活性	4.25E-03
GO:0051539	4Fe-4S结合作用	5.32E-03

表4 随藏红花柱头成熟过程表达上调miRNA的靶基因中最显著富集的10个GO条目

Table 4 The top ten GO terms enriched in the predicted target genes of up-regulated miRNAs during the ripening of stigmas in *C. sativus*

GO分类及条目	GO描述	P 值
生物过程		
GO:0043068	正向调控程序性细胞死亡	1.39E-05
GO:0045893	DNA依赖的正向转录调控	3.50E-04
GO:0048235	花粉精细胞的分化	1.50E-03
GO:0010044	铝离子反应	3.51E-03
GO:0010017	红光与红外光系统途径	3.51E-03
GO:0009789	正向调控脱落酸介导的信号途径	4.96E-03
GO:0009266	对温度刺激的反应	5.69E-03
分子功能		
GO:0030785	(二磷酸核酮糖羧化酶)-赖氨酸N-甲基转移酶活性	1.39E-05
GO:0008026	依赖ATP的解旋酶活性	4.64E-04
GO:0004386	解旋酶活性	1.65E-03

3 讨论

本文通过高通量测序的方法研究了藏红花开花阶段柱头成熟过程中两个时期的小RNA表达情况, 分别在2DBA和DOA柱头中各获得14 294 654和14 534 515条序列的纯净读数, 分布在3 711 809和2 303 043条序列中。其中, 24 nt长度的小RNA序列最多, 这与拟南芥(Lu等2005)、水稻(Morin等2008)、番茄(Wu等2016)、蓝莓(Yue等2017)等物种中报道的结果一样, 说明它们都是主要经由DCL2加工而成(Bouche等2006)。

通过同源比对,本文鉴定了藏红花中保守的miRNA,两个时期的miRNA序列分别可以归类到44和40个miRNA家族中。其中一些家族呈现出高水平的表达量,如MIR168、MIR171、MIR828、MIR894等(图3),暗示它们可能参与到藏红花开花时期柱头的成熟过程。已有报道表明MIR168是植物中特有的miRNA家族,可以影响植株的生长速率与开花时间,并且参与了植物光周期的调控过程中(韩璐和栾雨时2014)。MIR171也是植物中高度保守的miRNA家族,其作用的靶基因GRAS主要参与到植物的生长发育、信号转导、应激反应等过程。这些研究结果表明MIR168和MIR171很可能在藏红花柱头的成熟过程中也发挥着重要的调控功能(王艳芳等2015)。另外, MIR828一般认为是花青素合成的负调控因子,因此高水平表达的MIR828会抑制花青素的合成(谢焯等2013),从而很可能通过改变细胞中次级代谢途径的流向来促进藏红花素等类胡萝卜素物质的生物合成。而关于MIR894家族,目前研究的还比较少,其在藏红花柱头发育过程中的功能还需要进一步阐明。

根据Rubio等(2008)的报道,藏红花在开花之前香气很淡,而到开花之日香气急剧增大;与此同时,主要香气成分 β -紫罗兰酮、二氢- β -紫罗兰酮等也在开花之日达到最大的浓度。因此,本文比较了2DBA和DOA两个时期柱头中miRNA的表达水平,共筛选出117条差异表达的miRNA。靶基因预测得到众多的氧化还原酶、转移酶和转录因子等,如Radical-SAM酶家族、12-氧-植物二烯酸还原酶、磷酸核糖基转移酶、植物富含亮氨酸重复序列型类受体蛋白激酶、ARP蛋白复合体、BEL1-like蛋白家族、SPX蛋白家族、MYB转录因子、WRKY转录因子。其中,没有发现编码直接参与类胡萝卜素物质生物合成结构蛋白的靶基因,这或许是因为miRNA主要是通过与转录因子的相互作用来实现其调控功能(D'Ario等2017)。而本文鉴定到的转录因子中,MYB和WRKY都已被报道广泛参与了植物次生代谢产物生物合成途径的调控(Chezem和Clay 2016; 李岢和周春江2014)。同时, Jain等(2016)通过对藏红花5组转录组数据的分析也推测MYB和WRKY极有可能调控类胡萝卜素的生物合成,但其具体的调控功能和途径仍需要进一步的实验验证。

基于靶基因的GO功能注释,得以富集到一些重要的生物学过程,如硫辛酸生物合成的过程、细胞程序性凋亡、花粉细胞的分化、脱落酸介导的信号途径等(表3和4)。硫辛酸作为一种存在于线粒体上的辅酶,是一些酶行使催化功能所必需的分子(Raddatz和Bisswanger 1997)。硫辛酸主要在氧化脱羧反应中发挥关键作用,催化酰基的产生和转移。本文中获得的富集的硫辛酸生物合成过程来自于下调表达miRNA作用的靶基因(Radical-SAM酶家族),说明硫辛酸的合成会随着藏红花柱头的成熟而逐步加强,暗示其可能也是脱辅基类胡萝卜素的氧化脱羧反应。而来自于上调表达miRNA调控的细胞程序性凋亡、花粉细胞的分化、脱落酸介导的信号途径会受到抑制,这将会阻止细胞的分化和凋亡并推迟开花的时间从而有利于藏红花柱头的成熟。

藏红花作为昂贵的香料、染料和中药材,阐明其合成和积累藏红花素、藏红花酸、藏红花醛和藏红花甙等活性成分的生物途径具有重要意义。本文筛选到的差异表达miRNA及其功能靶基因可以为今后全面解析脱辅基类胡萝卜素生物合成调控网络的研究提供基础数据。

参考文献(References)

- Anders S, Huber W (2010). Differential expression analysis for sequence count data. *Genome Biol*, 11: R106
- Axtell MJ (2013). Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol*, 64: 137–159
- Baba SA, Malik AH, Wani ZA, et al (2015a). Phytochemical analysis and antioxidant activity of different tissue types of *Crocus sativus* and oxidative stress alleviating potential of saffron extract in plants, bacteria, and yeast. *S Afr J Bot*, 99: 80–87
- Baba SA, Mohiuddin T, Basu S, et al (2015b). Comprehensive transcriptome analysis of *Crocus sativus* for discovery and expression of genes involved in apocarotenoid biosynthesis. *BMC Genomics*, 16: 698
- Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116: 281–297
- Bouche N, Laussergues D, Gascioli V, et al (2006). An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J*, 25: 3347–3356
- Bouvier F, Suire C, Mutterer J, et al (2003). Oxidative remodeling of chromoplast carotenoids: Identification of the ca-

- rotenoid dioxygenase *CsCCD* and *CsZCD* genes involved in *Crocus* secondary metabolite biogenesis. *Plant Cell*, 15: 47–62
- Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, et al (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science*, 320: 1185–1190
- Chezem WR, Clay NK (2016). Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry*, 131: 26–43
- Conesa A, Götz S, García-Gómez JM, et al (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21: 3674–3676
- D'Ario M, Griffiths-Jones S, Kim M (2017). Small RNAs: Big impact on plant development. *Trends Plant Sci*, 22: 1056–1068
- Dai X, Zhao PX (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res*, 39: W155–W159
- Han L, Luan YS (2014). Progress in the research of plant miR168. *Genomics Appl Biol*, 2: 438–444 (in Chinese with English abstract) [韩璐, 栾雨时(2014). 植物miR168的研究进展. *基因组学与应用生物学*, 2: 438–444]
- Jain M, Srivastava PL, Verma M, et al (2016). *De novo* transcriptome assembly and comprehensive expression profiling in *Crocus sativus* to gain insights into apocarotenoid biosynthesis. *Sci Rep*, 6: 22456
- Kozomara A, Griffiths-Jones S (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*, 42: D68–D73
- Li K, Zhou CJ (2014). Research progress in WRKY transcription factors in plants. *Plant Physiol J*, 50: 1329–1335 (in Chinese with English abstract) [李崧, 周春江(2014). 植物WRKY转录因子的研究进展. *植物生理学报*, 50: 1329–1335]
- Li WP, Zhang Y, Ding ZS (2011). Research progress of saffron. *J Beijing Union Univ*, 25: 55–58 (in Chinese with English abstract) [李伟平, 张云, 丁志山(2011). 西红花的研究进展. *北京联合大学学报*, 25: 55–58]
- Lu C, Tej SS, Luo S, et al (2005). Elucidation of the small RNA component of the transcriptome. *Science*, 309: 1567–1569
- Morin RD, Aksay G, Dolgosheina E, et al (2008). Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res*, 18: 571–584
- Nawrocki EP, Burge SW, Bateman A, et al (2014). Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res*, 43: D130–D137
- Normand R, Yanai I (2013). An introduction to high-throughput sequencing experiments: design and bioinformatics analysis. *Methods Mol Biol*, 1038: 1–26
- Raddatz G, Bisswanger H (1997). Receptor site and stereospecificity of dihydrolipoamide dehydrogenase for *R*- and *S*-lipoamide: A molecular modeling study. *J Biotechnol*, 58: 89–100
- Robin JD, Ludlow AT, LaRanger R, et al (2016). Comparison of DNA quantification methods for next generation sequencing. *Sci Rep*, 6: 24067
- Rubio A, Rambla JL, Santaella M, et al (2008). Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in beta-ionone release. *J Biol Chem*, 283: 24816–24825
- The National Pharmacopoeia Commission (2010). People's Republic of China Pharmacopoeia (Part One). Beijing: China Medical Science and Technology Press, 120–121 [国家药典委员会(2010). 中华人民共和国药典(一部). 北京: 中国医药科技出版社, 120–121]
- Wang YF, Zhou RL, Zhao YH (2015). Phylogenetic analysis of miR-171 gene family and target gene prediction. *Life Sci Res*, 19: 479–483 (in Chinese with English abstract) [王艳芳, 周瑞莲, 赵彦宏(2015). miR-171基因家族进化分析及靶基因预测. *生命科学研究*, 19: 479–483]
- Wu P, Wu Y, Liu CC, et al (2016). Identification of arbuscular mycorrhiza (AM)-responsive microRNAs in tomato. *Front Plant Sci*, 7: 429
- Xie Y, Sun Y, Li DN, et al (2013). Anthocyanin synthesis induced by sucrose in *Arabidopsis* microRNA828. *Plant Physiol J*, 49: 188–194 (in Chinese with English abstract) [谢焯, 孙毅, 李淡宁等(2013). 拟南芥microRNA828负调控蔗糖诱导的花青素合成. *植物生理学报*, 49: 188–194]
- Yue JY, Lu XH, Zhang H, et al (2017). Identification of conserved and novel microRNAs in blueberry. *Front Plant Sci*, 8: 1155

Bioinformatics analysis for high-throughput sequencing of small RNAs in the stigmas of saffron (*Crocus sativus*)

LU Xiao-Hui^{1,2}, ZHU Zhi-Ming³, YUE Jun-Yang^{2,4,*}

¹College of Information Technology, Jiaxing Vocational Technical College, Jiaxing, Zhejiang 314000, China

²College of Horticulture, Anhui Agricultural University, Hefei 230036, China

³Jiaxing Tianhe Saffron Specialized Cooperative, Jiaxing, Zhejiang 314000, China

⁴School of Food Science and Engineering, Hefei University of Technology, Hefei 230009, China

Abstract: In the present study, we employed the high-throughput sequencing technology to profile miRNAs in saffron stigmas from two different stages (2 days before anthesis and day of anthesis), with a total of 3 711 809 and 2 303 043 small RNA tags obtained, respectively. Bioinformatics analysis identified 164 conserved miRNAs, among which 71 miRNAs were down-regulated and 46 miRNAs were up-regulated. Furthermore, the potential target genes of these conserved miRNAs were predicted and subjected to gene ontology (GO) annotation. Enrichment analysis of the GO-represented biological processes revealed that these target genes were potentially involved in lipoic acid biosynthetic process, positive regulation of programmed cell death, pollen sperm cell differentiation, and positive regulation of abscisic acid mediated signaling. This study provides a useful resource for further elucidation of the regulatory roles of miRNAs that participate in apocarotenoid biosynthesis during saffron stigma development.

Key words: saffron; small RNA; miRNA; high-throughput sequencing; bioinformatics

Received 2018-01-08 Accepted 2018-04-26

This work was supported by the Jiaxing Science and Technology Public Welfare Technology Project (2017AY33032).

*Corresponding author (aaran.yue@gmail.com).