

姜永强, 王丽慧, 孙雪梅. 基于转录组挖掘叶用莴苣中花青素生物合成代谢关键基因[J]. 江苏农业学报, 2025, 41(1): 150-160.  
doi: 10.3969/j.issn.1000-4440.2025.01.017

# 基于转录组挖掘叶用莴苣中花青素生物合成代谢关键基因

姜永强<sup>1,2</sup>, 王丽慧<sup>1,2</sup>, 孙雪梅<sup>1,2</sup>

(1. 青海大学农林科学院, 青海 西宁 810016; 2. 青海省蔬菜遗传与生理重点实验室, 青海 西宁 810016)

**摘要:** 为探究调控叶用莴苣叶片花青素合成的主效基因, 本研究对叶用莴苣大速生(绿色)和紫霞(紫色)的转录组数据进行分析。结果表明, 大速生和紫霞比较组中共检测到7 232个差异表达基因, 其中相对表达量上升基因4 214个, 相对表达量下降基因3 018个。KEGG 数据库富集分析结果表明, 富集到代谢过程中的差异表达基因最多, 花青素生物合成和次生代谢产物-其他抗生素生物合成通路中差异表达基因富集程度较高。在大速生和紫霞比较组中共鉴定到 55 个转录因子家族, 差异表达基因大多数属于 MYB、bHLH、C<sub>2</sub>C<sub>2</sub> 和 AP2/ERF 转录因子家族基因。从 MYB 转录因子基因中筛选出  $\log_2 FC$  绝对值较大的 LOC111917799, 命名为 *LsMYBI* 基因。瞬时表达试验结果验证了 *LsMYBI* 基因可以调控花青素的生物合成。本研究结果为叶用莴苣叶色的分子遗传机理及进一步的基因功能研究奠定了基础。

**关键词:** 叶用莴苣; 花青素; 转录组; 差异表达基因

**中图分类号:** S636.2 **文献标识码:** A **文章编号:** 1000-4440(2025)01-0150-11

## Identification of key genes involved in anthocyanin biosynthesis in *Lactuca sativa* L. based on transcriptome analysis

JIANG Yongqiang<sup>1,2</sup>, WANG Lihui<sup>1,2</sup>, SUN Xuemei<sup>1,2</sup>

(1. Academy of Agriculture and Forestry Sciences, Qinghai University, Xining 810016, China; 2. Qinghai Key Laboratory of Vegetable Genetics and Physiology, Xining 810016, China)

**Abstract:** In order to explore the major effective genes regulating anthocyanin synthesis in *Lactuca sativa* L., this study analyzed the transcriptome data of Dasusheng (green) and Zixia (purple). The results showed that a total of 7 232 differentially expressed genes were detected in the comparison group of Dasusheng and Zixia, including 4 214 genes with increased relative expression and 3 018 genes with decreased relative expression. The results of KEGG database enrichment analysis indicated that the most differentially expressed genes were enriched in the metabolic process, and the enrichment degrees of the differentially expressed genes in anthocyanin biosynthesis and secondary metabolite-other antibiotic biosynthesis pathways were relatively high. A total of 55 transcription factor families were identified in the comparison group of Dasusheng and Zixia. Most of the differentially expressed genes belonged to the MYB, bHLH, C<sub>2</sub>C<sub>2</sub> and AP2/ERF transcription factor family genes. LOC111917799 with a larger

absolute value of  $\log_2 FC$  was screened out from the MYB transcription factor genes and named as *LsMYBI*. The transient expression experiment verified that the *LsMYBI* gene could regulate the biosynthesis of anthocyanins. The results of this study lay a foundation for the molecular genetic mechanism of leaf color in leaf lettuce and further research

收稿日期: 2024-06-27

基金项目: 青海省科学技术厅重点研发与转化计划项目(2024-NK-106)

作者简介: 姜永强(1998-), 男, 安徽滁州人, 硕士研究生, 主要研究方向为蔬菜生理与分子育种。(E-mail) jiangyongqiang0951@163.com

通讯作者: 孙雪梅, (E-mail) 13997091543@163.com

on gene function.

**Key words:** *Lactuca sativa* L.; anthocyanins; transcriptome; differentially expressed genes

叶用莴苣(*Lactuca sativa* L.)俗称生菜,属于菊科莴苣属,为一年生或二年生草本植物,起源于地中海沿岸。叶用莴苣的种植面积逐年扩大,现已成为全球消费量最大的蔬菜之一,深受消费者喜爱<sup>[1-2]</sup>。叶用莴苣叶片富含维生素C、维生素E、叶酸、多酚和膳食纤维等营养成分<sup>[3-4]</sup>,能够刺激胃液分泌,促进肠道蠕动,增强消化吸收功能,预防便秘,因此它被誉为“减肥蔬菜”<sup>[5]</sup>。根据叶片颜色,叶用莴苣可分为绿叶叶用莴苣和紫叶叶用莴苣两大类,其中紫叶叶用莴苣富含花青素<sup>[6]</sup>,这种天然色素不仅赋予了叶片独特的紫色,还具有抗氧化功能<sup>[7-10]</sup>。

花青素是类黄酮的一种,是广泛存在于高等植物中的水溶性次生代谢物。它不仅为果实与花朵增添了鲜艳的色彩,还能吸引传粉昆虫和鸟类,促进花粉与种子的传播,从而为植物的繁殖和遗传多样性做出贡献<sup>[11]</sup>。花青素主要分布于植物细胞的液泡中,能提高植物抗逆能力,如耐寒、耐旱及耐盐能力<sup>[12]</sup>。花青素的生物合成经过苯丙烷生物合成途径<sup>[13]</sup>,涉及多个结构基因<sup>[14]</sup>。同时,转录因子MYB、bHLH和WD40通过相互作用形成MYB-bHLH-WD40(MBW)复合体,该复合体能够调控花青素生物合成途径中关键结构基因的转录,在多种植物中具有高度保守性<sup>[15]</sup>。

转录组测序(RNA-seq)技术利用高通量测序方法,对细胞或组织内的mRNA、小RNA(Small RNA)以及非编码RNA(No-coding RNA)等转录本进行全面或部分测序与分析,从而揭示基因表达的动态变化及其调控网络<sup>[16-17]</sup>。转录组测序技术能够解析基因与生物性状之间的关联性,是研究基因表达调控机制强有力的工具<sup>[18-20]</sup>。目前花青素的生物合成途径已被深入研究。Zong等<sup>[21]</sup>通过转录组数据找到调控枸杞花青素合成的MYB转录因子,该转录因子能够促进花青素的生物合成,在转基因番茄中,MYB转录因子的功能也得到了验证。Li等<sup>[22]</sup>对比了猕猴桃红肉品种与绿肉品种的转录组数据,筛选出6种与花青素生物合成途径密切相关的基因。Tian等<sup>[23]</sup>通过分析海棠叶片的转录组数据,筛选出在低温和高光照胁迫条件下诱导花青素积累的差异表达基因。第二代测序技术以其高通量和高效率的

特点,极大地推动了不同颜色植物组织中差异表达基因的筛选。本研究拟以叶用莴苣大速生和紫霞为试验材料,利用Illumina NovaSeq进行转录组测序分析,探究叶用莴苣叶片颜色形成的关键基因。

## 1 材料与方法

### 1.1 试验材料

本试验选用大速生(绿色叶用莴苣)和紫霞(紫色叶用莴苣)2个莴苣品种,种植于青海大学园艺创新基地(36°43'6"N,101°45'15"E)。在莴苣成熟期,选取长势一致的植株,取茎基部向上第4片叶,每个品种设3个生物学重复。叶片立刻用液氮冷冻,于-80℃冰箱保存。

### 1.2 试验方法

1.2.1 总RNA的提取 利用TaKaRa MiniBEST Plant RNA Extraction Kit总RNA提取试剂盒,提取大速生和紫霞叶片的总RNA。利用Nano drop浓度测定仪测定提取到的RNA浓度。并通过1.0%琼脂糖凝胶电泳检测RNA的完整性和纯度,确保提取的RNA的质量满足文库构建和高通量测序试验的要求。

1.2.2 转录组文库的构建和测序 利用含胸腺嘧啶的Oligo(dT)磁珠技术,从总RNA中选择性富集真核生物mRNA。富集的mRNA被随机打断形成一系列mRNA短片段,随后以这些mRNA短片段作为模板,合成双链cDNA。利用PCR技术对双链cDNA片段进行扩增,构建cDNA文库。新构建的cDNA文库经Qubit2.0荧光定量仪进行定量分析,并使用Agilent 2100生物分析仪检测插入片段的大小。完成这些步骤后,使用Illumina NovaSeq 6000平台进行高通量测序。转录组文库的构建和测序工作由南京集思慧远生物科技有限公司完成。

1.2.3 测序数据组装与分析 对原始测序数据进行质量过滤,去除测序过程可能引入的低质量序列和低复杂性区域。仅保留高质量的序列用于后续的数据分析。剔除低质量数据后,对剩余的高质量序列进行数据组装。利用HISAT2软件<sup>[24]</sup>将过滤后的高质量序列与参考基因组和基因注释文件(基因登录号:GCF\_002870075.4)进行比对,对比对结果进行基因注释。

1.2.4 差异表达基因分析 本研究利用 DESeq2 软件<sup>[25]</sup>对转录组数据进行定量分析,得到大速生与紫霞之间的差异表达基因。设定的筛选阈值为,  $\log_2 FC$ 的绝对值 $\geq 1$ ,假发现率( $FDR$ ) $< 0.05$ 。 $FC$ 表示差异倍数。

1.2.5 差异表达基因功能注释、分类和花青素合成代谢通路分析 为了对筛选出的差异表达基因进行全面分析,结合多个公共数据库资源,对差异表达基因的功能进行注释和分类。数据库主要包括非冗余蛋白氨基酸序列数据库(NR)、蛋白质氨基酸序列数据库(SwissProt)、真核生物蛋白相邻类的聚簇(KOG)、京都基因与基因组百科全书(KEGG)、基因本体(GO)、同源蛋白簇(COG)和 Pfam。将大速生与紫霞品种间显著差异表达的基因进行功能定位,并进一步利用 GO 数据库和 KEGG 数据库对差异表达基因进行分类和通路分析,重点筛选和收集与花青素生物合成密切相关的基因。对花青素生物合成途径中关键基因的每千个碱基的转录每百万映射读取的片段数( $FPKM$ )值进行分析,比较它们在品种中的表达水平,并计算基因表达差异倍数。

1.2.6 结构基因相对表达量测定 从 2 个品种的叶用莴苣中提取总 RNA,通过凝胶电泳技术检测 RNA 质量,以确保其完整性和纯度。提取的总 RNA 使用天根反转录试剂盒反转录为 cDNA,并将浓度稀释 50 倍。设计各基因引物,以叶用莴苣 18S rRNA 作为内参标准,各引物序列如表 1 所示。以绿色叶用莴苣大速生品种作为对照,利用  $2^{-\Delta\Delta Ct}$  法计算目标基因的相对表达量。为确保试验结果的可靠性,进行 3 次生物学重复,利用 Graphpad prims 9.5 软件绘制图表。

1.2.7 候选基因瞬时表达验证 以紫霞品种 RNA 反转录得到的 cDNA 为模板,通过 PCR 技术扩增目标基因并将其克隆。随后,利用双酶切法将目标基因插入过表达载体 pCAMBIA2300s,构建重组质粒。将重组质粒转入农杆菌感受态细胞 GV3101 中备用。将含有目的基因的农杆菌在 LB 液体培养基中培养,培养液的  $OD_{600}$  调至 0.6~0.8。选择 5 周龄的 Samsun 烟草幼苗,注射培养液。

## 2 结果与分析

### 2.1 RNA-Seq 测序数据统计分析

如表 2 所示,各品种高质量 read 的碱基数 $\geq$

5.67 Gb,质量值 $\geq 20$ 的碱基所占的百分比( $Q20$ )和质量值 $\geq 30$ 的碱基所占的百分比( $Q30$ ) $> 92\%$ ,比对到参考基因组的高质量 read 所占比例为 89.05%~96.16%。结果表明,Illumina 测序数据质量合格,可靠性较高,可用于下一步分析。

表 1 供试引物

Table 1 Primers used in this study

引物名称	引物序列(5'→3')	退火温度(°C)
<i>Ls18s</i> -qF	GTGACTGAAGAAGGGCAATG	55
<i>Ls18s</i> -qR	CACTTTCACCCGATTACC	55
<i>LsCHI</i> -qF	CACCGCTATCGGAGTTTA	54
<i>LsCHI</i> -qR	CCGTACTGTGATCCCTTG	53
<i>LsCHS</i> -qF	AAGCGAGCACAAAGACAGA	55
<i>LsCHS</i> -qR	ACTTCCACGACAACGATA	53
<i>LsDFR</i> -qF	ACTGTTCCGTGACCCTGAT	55
<i>LsDFR</i> -qR	CCTTCTATTGTTGGCTTTAT	53
<i>LsF3'5'H</i> -qF	TGGCTGAAATACTAGGGT	53
<i>LsF3'5'H</i> -qR	TGGACTATGTAAACCGAAT	53
<i>LsF3H</i> -qF	CTAAGGAATACAGCGAGGTG	55
<i>LsF3H</i> -qR	TGAGATCGGGTTGAGGAC	55
<i>PAL</i> -qF	CCCATTATCCTACATCGC	53
<i>PAL</i> -qR	TCTTTCGGCTGTAACCTCG	53

### 2.2 差异表达基因

$FPKM$  值是转录组分析中的重要参数,可揭示基因表达的丰度差异,从而帮助筛选出在不同品种间差异显著的表达基因。以  $\log_2 FC$  绝对值 $\geq 1$ 且错误发现率( $FDR$ ) $< 0.05$ 作为筛选标准,共筛选出差异表达基因 7 232 个,其中相对表达量上升基因 4 214 个,相对表达量下降基因 3 018 个。相对表达量无显著差异基因 16 295 个。对差异表达基因进行层次聚类分析,如图 1 所示,每个品种的 3 个生物学重复聚成一簇,表明重复性较好,数据的可靠性高。

### 2.3 差异基因功能注释统计

通过蛋白质氨基酸序列数据库(SwissProt)、基因本体(GO)、京都基因与基因组百科全书(KEGG)、同源蛋白簇(COG)、真核生物蛋白相邻类的聚簇(KOG)、Pfam 和非冗余蛋白质氨基酸序列数据库(NR)对差异表达基因进行分析,分别注释到 4 874、4 523、2 187、1 868、2 749、4 913 和 5 846 个蛋白质。

表 2 转录组测序数据

Table 2 Transcriptome sequencing data

编号	高质量 read 总数	高质量 read 的 碱基数 (Gb)	GC (%)	Q20 (%)	Q30 (%)	比对到参考基因组的 高质量 read 所占比例 (%)
ZX1	18 906 478	5.67	42.44	98.15	94.32	96.16
ZX2	25 229 279	7.57	42.14	97.78	93.78	95.19
ZX3	22 877 520	6.86	41.94	97.74	93.72	95.39
DSS1	22 395 078	6.72	44.18	97.32	92.98	89.07
DSS2	25 999 522	7.80	43.98	97.46	93.36	89.83
DSS3	20 892 956	6.27	44.29	97.56	93.54	89.05

ZX:紫霞;DSS:大速生。GC:G、C 碱基数占总碱基数的百分比;Q20:质量值 $\geq 20$ 的碱基所占的百分比;Q30:质量值 $\geq 30$ 的碱基所占的百分比。

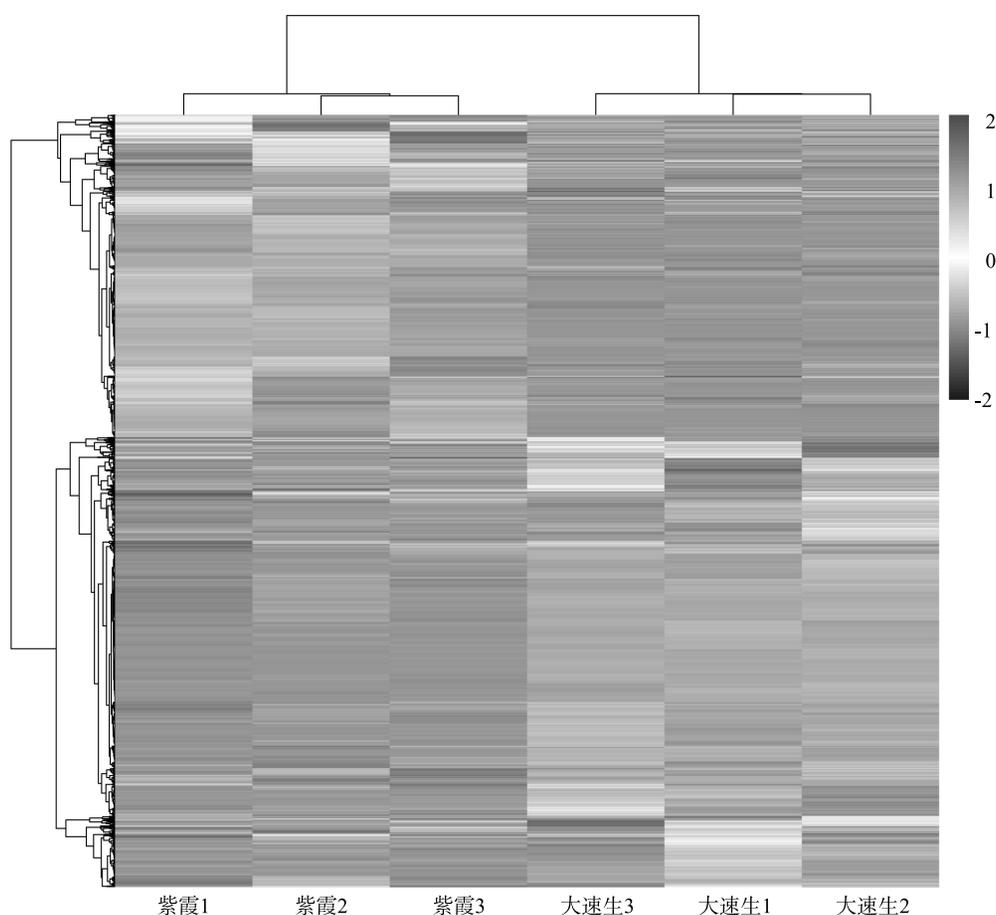


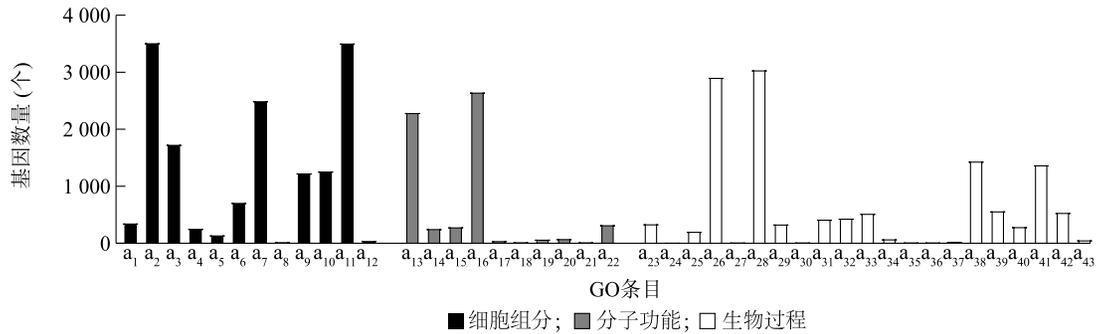
图 1 差异表达基因聚类分析结果

Fig.1 Cluster analysis results of differentially expressed genes

#### 2.4 差异基因 GO 注释

对差异表达基因进行 GO 注释分析,如图 2 所示,这些基因根据其功能和属性被分为生物过程 (Biological process, BP)、细胞组分 (Cellular component, CC)、以及分子功能 (Molecular function, MF)。在生物过程中,共注释到 12 622 个基因;在细胞组分中,共注释到 15 302 个基因;在分子功能中,共注释

到 6 031 个基因。紫霞和大速生的差异表达基因富集到 52 个 GO 条目,其中 43 个 GO 条目为差异表达基因显著富集条目。在生物过程中,3 039 个基因与细胞过程相关,2 908 个基因与代谢过程相关。在分子功能中,3 516 个基因与亚细胞结构的调控相关,3 511 个基因与细胞成分相关,2 499 个基因与细胞器功能相关。



a<sub>1</sub>: 胞外区; a<sub>2</sub>: 细胞; a<sub>3</sub>: 薄膜; a<sub>4</sub>: 细胞连接; a<sub>5</sub>: 膜包围腔; a<sub>6</sub>: 蛋白质复合物; a<sub>7</sub>: 细胞器; a<sub>8</sub>: 胞外区部分; a<sub>9</sub>: 细胞器部分; a<sub>10</sub>: 膜组件; a<sub>11</sub>: 细胞部分; a<sub>12</sub>: 超分子复合体; a<sub>13</sub>: 催化活性; a<sub>14</sub>: 结构分子活性; a<sub>15</sub>: 转运活性; a<sub>16</sub>: 结合; a<sub>17</sub>: 抗氧化活性; a<sub>18</sub>: 养分库活性; a<sub>19</sub>: 分子换能器活性; a<sub>20</sub>: 分子功能调节剂; a<sub>21</sub>: 分子载体活性; a<sub>22</sub>: 转录调控因子活性; a<sub>23</sub>: 再生产; a<sub>24</sub>: 细胞杀伤; a<sub>25</sub>: 免疫系统过程; a<sub>26</sub>: 代谢过程; a<sub>27</sub>: 细胞增殖; a<sub>28</sub>: 细胞过程; a<sub>29</sub>: 生殖过程; a<sub>30</sub>: 生物黏附; a<sub>31</sub>: 信号转导; a<sub>32</sub>: 多细胞生物过程; a<sub>33</sub>: 演变过程; a<sub>34</sub>: 生长; a<sub>35</sub>: 移动; a<sub>36</sub>: 色素积累; a<sub>37</sub>: 节律程序; a<sub>38</sub>: 刺激反应; a<sub>39</sub>: 定位; a<sub>40</sub>: 多生物过程; a<sub>41</sub>: 生物调控; a<sub>42</sub>: 细胞成分组织或生物合成; a<sub>43</sub>: 解毒。

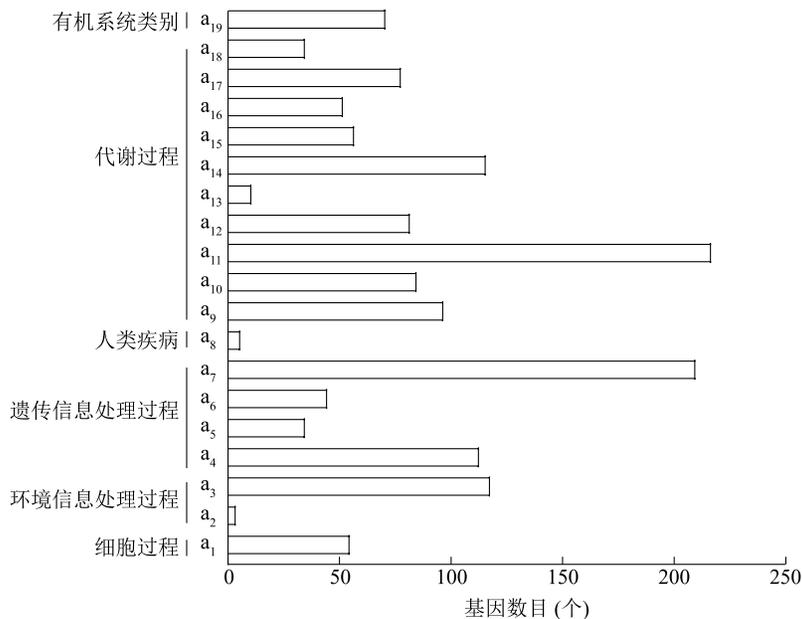
图2 差异表达基因 GO 功能分类

Fig.2 GO functional classification of differentially expressed genes

2.5 差异基因 KEGG 富集分析

如图 3 所示,KEGG 通路富集分析将差异表达基因分为 6 个类别。在细胞过程中,共注释到 55 个基因。在环境信息处理过程中,共注释到 122 个基因。在遗传信息处理过程中,共注释到 403 个基因。在代谢过程中,共注释到 830 个基因。在有机系统类别中,共注

释到 71 个基因。在人类疾病中,仅注释到 6 个基因。如表 3 所示,通过 KEGG 数据库富集分析,从筛选出的显著差异表达基因中确定其所属的代谢通路或生物学过程,最终筛选出富集程度最高的前 20 条通路。其中花青素生物合成和次生代谢产物-其他抗生素生物合成通路中差异表达基因富集程度最高。



a<sub>1</sub>: 运输和分解代谢; a<sub>2</sub>: 膜转运; a<sub>3</sub>: 信号转导; a<sub>4</sub>: 折叠、分类和退化; a<sub>5</sub>: 复制与修复; a<sub>6</sub>: 转录; a<sub>7</sub>: 翻译; a<sub>8</sub>: 内分泌与代谢疾病; a<sub>9</sub>: 氨基酸代谢; a<sub>10</sub>: 次生代谢产物生物合成; a<sub>11</sub>: 碳水化合物代谢; a<sub>12</sub>: 能量代谢; a<sub>13</sub>: 聚糖的生物合成与代谢; a<sub>14</sub>: 脂代谢; a<sub>15</sub>: 辅助因子和维生素代谢; a<sub>16</sub>: 其他氨基酸代谢; a<sub>17</sub>: 萜类化合物和聚酮类化合物代谢; a<sub>18</sub>: 核苷酸代谢; a<sub>19</sub>: 环境适应。

图3 差异表达基因 KEGG 功能分类

Fig.3 KEGG functional classification of differentially expressed genes

表 3 差异表达基因富集程度最高的前 20 条通路

Table 3 The top 20 pathways with the highest enrichment of differentially expressed genes

代谢通路编号	代谢通路名称	富集因子值	<i>q</i> 值	差异表达基因数量 (个)
ko00908	玉米素生物合成	1.546 543 408	0.811 376 618	9
ko00130	泛醌和其他萜类醌的生物合成	1.356 617 025	0.811 376 618	15
ko00350	酪氨酸代谢	1.370 354 919	0.651 639 625	21
ko00900	萜类骨架生物合成	1.481 363 418	0.464 539 149	25
ko00430	牛磺酸和亚牛磺酸代谢	2.062 057 878	0.811 376 618	4
ko00500	淀粉和蔗糖代谢	1.347 072 986	0.300 716 550	52
ko03010	核糖体	2.392 442 842	0	168
ko04075	激素信号转导	1.185 550 072	0.587 566 168	89
ko00190	氧化磷酸化	1.207 217 428	0.811 376 618	37
ko00902	单萜类生物合成	1.874 598 071	0.587 566 168	8
ko00010	糖酵解/糖异生	1.268 110 460	0.587 566 168	46
ko00052	半乳糖代谢	1.693 833 257	0.214 081 572	23
ko00062	脂肪酸延长	1.568 957 081	0.587 566 168	14
ko00073	角质、木栓质和蜡质生物合成	1.599 872 491	0.771 718 112	9
ko00020	柠檬酸循环	1.449 884 445	0.587 566 168	18
ko00906	类胡萝卜素合成	1.841 123 105	0.300 716 550	15
ko01040	不饱和脂肪酸的生物合成	1.829 244 892	0.469 095 763	11
ko00998	次生代谢产物-其他抗生素生物合成	3.436 763 130	0.771 718 112	2
ko00942	花青素生物合成	5.155 144 695	0.587 566 168	2
ko00592	$\alpha$ -亚油酸代谢	1.950 595 290	0.010 017 269	28

## 2.6 转录因子分析

转录因子在植物生长和代谢过程中具有重要的调控作用。如图 4 所示,在大速生和紫霞比较组中,共鉴定到 55 个转录因子家族,对应 442 个差异表达基因。其中差异表达基因数量较多的转录因子家族为 AP2/ERF (44 个)、MYB (39 个)、C<sub>2</sub>C<sub>2</sub> (32 个) 和 bHLH (30 个)。而 VOZ、TUB、SRS、Rcd1-like 等转录因子家族对应的差异表达基因数量最少,仅 1 个。

## 2.7 花青素合成代谢相关基因的表达差异分析

如表 4 所示,通过转录组数据分析,筛选到 19 个可能与花青素合成相关的候选基因,包括 bHLH 基因(2 个)、MYB 基因(8 个)以及关键结构基因(9 个)。这些基因在紫叶叶用莴苣中的表达量高于绿叶叶用莴苣。通过数据库筛选、比对发现,对于 bHLH 基因, LOC111892911 的功能已知,

LOC111920837 在紫叶叶用莴苣和绿叶叶用莴苣中表达差异不显著。对于 MYB 基因, LOC111893240 的 log<sub>2</sub>FC 绝对值最大,但是该基因的功能已知。LOC111917799 的 log<sub>2</sub>FC 绝对值较大,为 2.444 1,且 LOC111917799 几乎不在绿叶叶用莴苣中表达。因此最终选择 LOC111917799 作为候选基因,命名为 *LsMYB1*。

通过计算 FPKM 值,对花青素合成途径中基因表达差异进行分析。以大速生的转录组数据作为基准,紫霞中,涉及花青素合成的关键结构基因表达量显著高于大速生。如图 5 所示,紫霞 DFR 基因的相对表达量是大速生的 3 729.83 倍。紫霞 CHS 基因的相对表达量是大速生的 55.89 倍。花青素合成途径中关键基因的高表达可能是紫霞叶用莴苣花青素积累显著高于大速生叶用莴苣的主要原因。

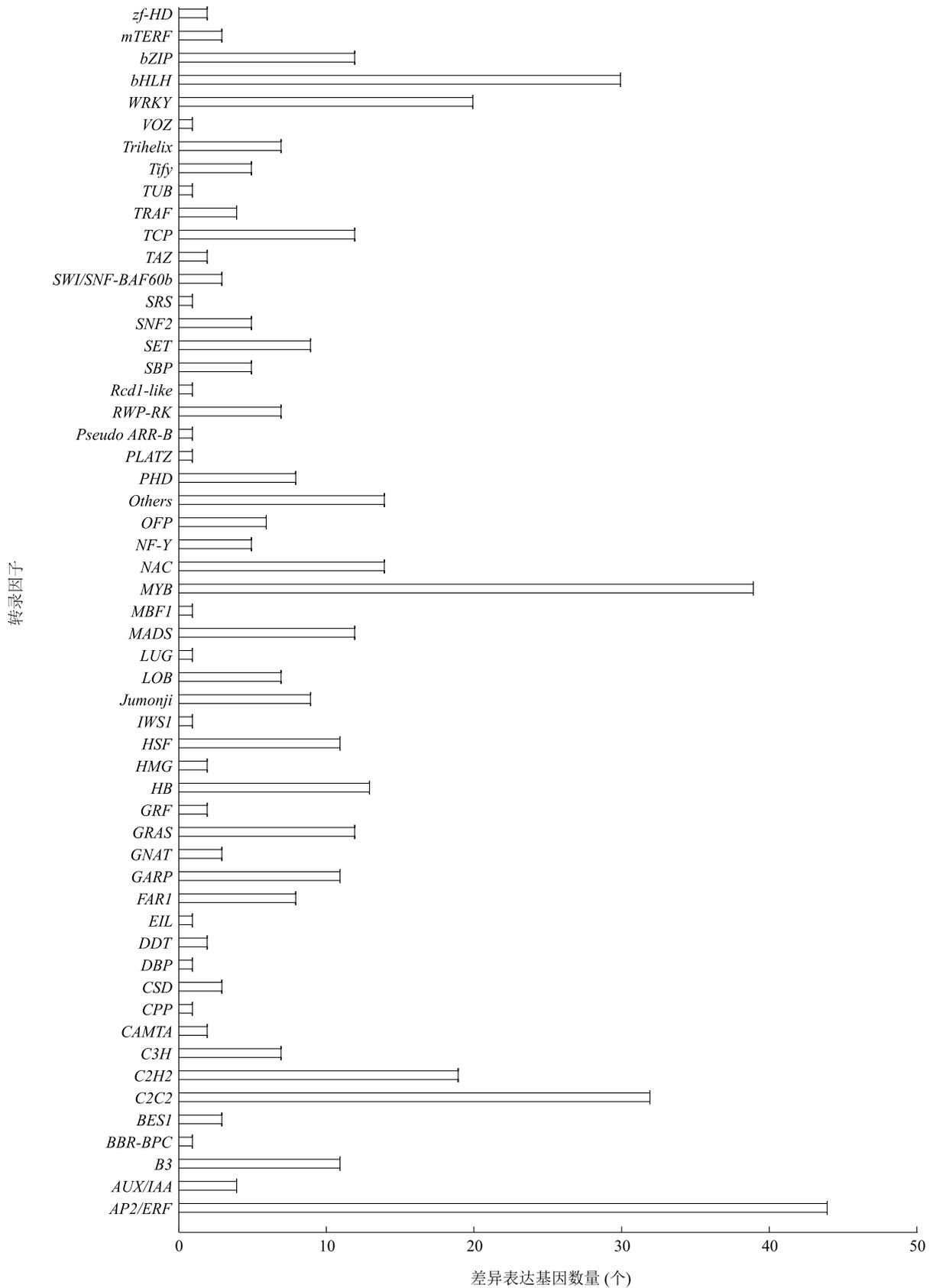


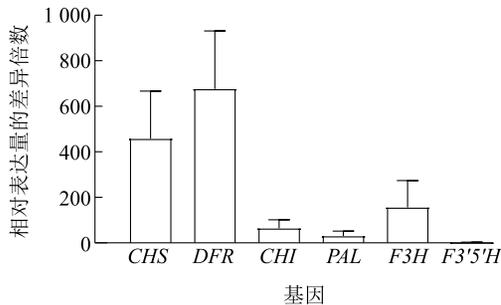
图 4 转录因子家族富集的差异表达基因数量

Fig.4 The number of differentially expressed genes enriched in the transcription factor family



## 2.8 紫霞和大速生比较组中花青素合成关键结构基因的相对表达量

依据差异基因筛选结果,以叶用莴苣的 18S rRNA 作为标准化的内参,采用实时荧光定量 RT-qPCR 测定大速生和紫霞中花青素合成相关结构基因的相对表达量。以大速生相对表达量作为对照,结果表明,紫霞中苯丙氨酸解氨酶基因(*PAL*)、查尔酮合成酶基因(*CHS*)、查尔酮异构酶基因(*CHI*)、黄烷酮 3-羟化酶基因(*F3H*)、二氢黄酮醇还原酶基因(*DFR*)和类黄酮 3'-5'-羟化酶基因(*F3'5'H*)相对表达量均高于大速生。其中,*DFR* 基因的相对表达量差异最大,表明其在花青素合成途径中具有关键作用,*F3'5'H* 基因的相对表达量差异最小,该结果与转录组数据一致(图 6)。



*CHS*:查尔酮合成酶基因;*DFR*:二氢黄酮醇还原酶基因;*CHI*:查尔酮异构酶基因;*PAL*:苯丙氨酸解氨酶基因;*F3H*:黄烷酮 3-羟化酶基因;*F3'5'H*:类黄酮 3'-5'-羟化酶基因。

图 6 紫霞和大速生比较组中花青素合成关键结构基因的相对表达量的差异倍数

Fig.6 The fold change of relative expression levels of key structural genes in the anthocyanin biosynthesis pathway in the comparison group of Dasusheng and Zixia

## 2.9 候选基因功能验证结果

将 *LsMYBI* 基因注射到 Samsun 烟草叶片的特定区域,注射点周边区域呈现紫色。这一结果表明, *LsMYBI* 基因在花青素合成中具有关键作用(图 7)。

## 3 讨论

高通量转录组测序技术能够深入探究基因表达的差异性,并揭示其背后的调控机制,从而帮助理解这些变化如何影响细胞功能<sup>[26]</sup>。在生物体面临应激反应时,差异表达基因表达水平会发生显著的变化。对这些基因进行深入研究,不仅有助于阐明其生物学功能,还能为揭示相关的分子机

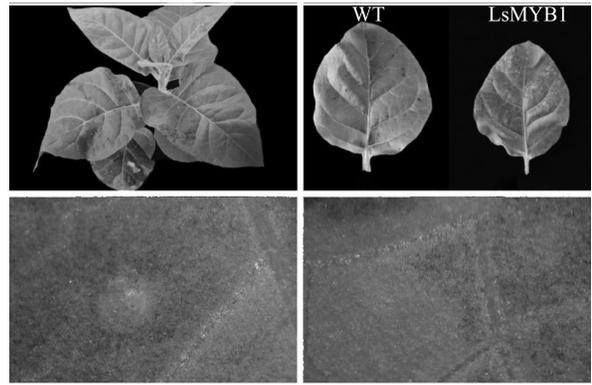


图 7 *LsMYBI* 基因在 Samsun 烟草中的瞬时表达

Fig.7 Transient expression of *LsMYBI* gene in Samsun tobacco

制提供重要的数据支持<sup>[27]</sup>。叶用莴苣种植面积较广,据 2018 年统计数据,全球叶用莴苣的收获面积超过  $1.27 \times 10^6$   $\text{hm}^2$ ,总产量达到  $2.73 \times 10^7$   $\text{t}$ <sup>[28]</sup>。紫色叶用莴苣富含花青素,花青素不仅可以增加植物自身的抗逆性,还在药用和保健领域具有重要价值<sup>[29]</sup>。

花青素是类黄酮代谢途径的重要分支产物,其生物合成受到多个转录因子的调控。转录因子的分离和功能鉴定完善了花青素生物合成的调控网络,并为花青素的应用提供了理论依据<sup>[30]</sup>。花青素的生物合成调控涉及多个关键转录因子,其中包括 MYB、bHLH 和 WD40。这些转录因子通过相互作用和协同效应,构成了一个错综复杂的调控网络<sup>[31]</sup>。MBW 复合体在花青素合成调控中起着至关重要的作用,WD40 蛋白是构成 MBW 复合体的核心部分。研究表明,所有已知的 WD40 蛋白均参与 MBW 复合体的组装。bHLH 转录因子主要通过 MYB 转录因子的相互作用,调控二氢黄酮醇还原酶(*DFR*)和花青素合成酶(*ANS*)等关键酶的表达。在 MBW 复合体中,bHLH 转录因子能够同时与 MYB 转录因子和 WD40 蛋白发生相互作用,调控花青素的生物合成。MYB 转录因子在植物不同组织,如叶片<sup>[32]</sup>、果实<sup>[33]</sup>和花朵<sup>[34]</sup>中的表达模式表现出显著特异性。果实外皮与内部的花青素积累以及花瓣的色彩变化,均受到特异性 MYB 转录因子的精确调控。此外,果实内花青素的累积不仅由遗传因素决定,还受到光照、温度等非生物因素的影响。这些环境因子通过信号转导途径激活特定的 MYB 和 bHLH 转录因子,从而参与花青素生物合成的调控,

使植物能够响应外界环境变化。深入研究非生物因子如何通过转录因子影响花青素代谢,对于揭示植物环境适应机制以及创新作物改良策略具有重要意义。

## 4 结论

本研究以大速生(绿色叶用莴苣)和紫霞(紫色叶用莴苣)为试验材料,对转录组数据进行分析。结果表明,大速生和紫霞比较组中共检测到7 232个差异表达基因,其中相对表达量上升基因4 214个,相对表达量下降基因3 018个。KEGG 数据库富集分析结果表明,富集到代谢过程中的差异表达基因最多,花青素生物合成和次生代谢产物-其他抗生素生物合成通路中差异表达基因富集程度较高。在大速生和紫霞比较组中共鉴定到55个转录因子家族,差异表达基因大多数属于 MYB、bHLH、C<sub>2</sub>C<sub>2</sub>和 AP2/ERF 转录因子家族基因。从 MYB 转录因子基因中筛选出  $\log_2 FC$  绝对值较大的 LOC111917799,命名为 *LsMYB1* 基因。瞬时表达试验结果验证了 *LsMYB1* 基因可以调控花青素的生物合成。

## 参考文献:

- [1] VIACAVA G E, ROURA S I, BERRUETA L A, et al. Characterization of phenolic compounds in green and red oak-leaf lettuce cultivars by UHPLC-DAD-ESI-QToF/MS using MSE scan mode [J]. *Journal of Mass Spectrometry*, 2017, 52(12): 873-902.
- [2] 杨 晓,魏仕伟,郭斗斗,等. 上海生菜产销现状与技术关键点分析[J]. *长江蔬菜*, 2016, 409(11): 4-8.
- [3] ZŁOTEK U, SWIECA M, JAKUBCZYK A. Effect of abiotic elicitation on main health-promoting compounds, antioxidant activity and commercial quality of butter lettuce (*Lactuca sativa* L.) [J]. *Food Chemistry*, 2014, 148: 253-260.
- [4] 杨 攀,杨诗雯,李 磊,等. 生菜研究进展综述[J]. *现代园艺*, 2020, 43(15): 34-36.
- [5] 王义国,白延波. 常见蔬菜的营养价值及生长要求[J]. *中国果菜*, 2019, 39(7): 73-76.
- [6] MEDINA-LOZANO I, BERTOLÍN J R, DÍAZ A. Nutritional value of commercial and traditional lettuce (*Lactuca sativa* L.) and wild relatives; vitamin C and anthocyanin content [J]. *Food Chemistry*, 2021, 359: 129864.
- [7] CLARK J L, ZAHRAKKA P, TAYLOR C G. Efficacy of flavonoids in the management of high blood pressure [J]. *Nutrition Reviews*, 2015, 73(12): 799-822.
- [8] LIOBIKAS J, SKEMIENE K, TRUMBECKAITE S, et al. Anthocyanins in cardioprotection: a path through mitochondria [J]. *Pharmacological Research*, 2016, 113: 808-815.
- [9] MORAIS C A, DE-ROSSO V V, ESTADELLA D, et al. Anthocyanins as inflammatory modulators and the role of the gut microbiota [J]. *The Journal of Nutritional Biochemistry*, 2016, 33: 1-7.
- [10] SEHITOGLU M H, FAROOQI A A, QURESHI M Z, et al. Anthocyanins; targeting of signaling networks in cancer cells [J]. *Asian Pacific Journal of Cancer Prevention*, 2014, 15(5): 2379-2381.
- [11] 杨 波. 马铃薯块茎花青素差异累积及光诱导累积的研究 [D]. 武汉:华中农业大学, 2019.
- [12] 韩海华,梁名志,王 丽,等. 花青素的研究进展及其在茶叶中的应用 [J]. *茶叶*, 2011, 37(4): 217-220.
- [13] NAING A H, KIM C K. Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants [J]. *Plant Molecular Biology*, 2018, 98(1): 1-18.
- [14] 高国应,伍小方,张大为,等. MBW 复合体在植物花青素合成途径中的研究进展 [J]. *生物技术通报*, 2020, 36(1): 126-134.
- [15] RAMSAY N A, GLOVER B J. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity [J]. *Trends in Plant Science*, 2005, 10(2): 63-70.
- [16] CHU Y, COREY D R. RNA sequencing: platform selection, experimental design, and data interpretation [J]. *Nucleic Acid Therapeutics*, 2012, 22(4): 271-274.
- [17] 崔 凯,吴伟伟,刁其玉. 转录组测序技术的研究和应用进展 [J]. *生物技术通报*, 2019, 35(7): 1-9.
- [18] ISLAM M S, JALALUDDIN M, GARNER J O, et al. Artificial shading and temperature influence on anthocyanin compositions in sweetpotato leaves [J]. *HortScience*, 2005, 40(1): 176-180.
- [19] KOCH K E, WU Y, XU J. Sugar and metabolic regulation of genes for sucrose metabolism: potential influence of maize sucrose synthase and soluble invertase responses on carbon partitioning and sugar sensing [J]. *Journal of Experimental Botany*, 1996, 47(S): 1179-1185.
- [20] LEYVA A, JARILLO J A, SALINAS J, et al. Low temperature induces the accumulation of phenylalanine ammonia-lyase and chalcone synthase mRNAs of *Arabidopsis thaliana* in a light-dependent manner [J]. *Plant Physiology*, 1995, 108(1): 39-46.
- [21] ZONG Y, ZHU X B, LIU Z G, et al. Functional MYB transcription factor encoding gene AN2 is associated with anthocyanin biosynthesis in *Lycium ruthenicum* Murray [J]. *BMC Plant Biology*, 2019, 19(1): 169.
- [22] LI Y K, FANG J B, QI X J, et al. Combined analysis of the fruit metabolome and transcriptome reveals candidate genes involved in flavonoid biosynthesis in *Actinidia arguta* [J]. *International Journal of Molecular Sciences*, 2018, 19(5): 1471.
- [23] TIAN J, ZHANG J, HAN Z Y, et al. McMYB12 transcription factors co-regulate proanthocyanidin and anthocyanin biosynthesis in *Malus crabapple* [J]. *Scientific Reports*, 2017, 7(1): 43715.
- [24] KIM D, PAGGI J M, PARK C, et al. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype [J]. *Nature Biotechnology*, 2019, 37(8): 907-915.
- [25] LOVE M I, HUBER W, ANDERS S. Moderated estimation of fold

- change and dispersion for RNA-seq data with DESeq2[J]. *Genome Biology*, 2014, 15(12):550.
- [26] 何平, 李林光, 王海波, 等. 遮光性套袋对桃果实转录组的影响[J]. *中国农业科学*, 2017, 50(6):1088-1097.
- [27] 黄琼林. 高良姜转录组差异表达基因分析[J]. *中药材*, 2020, 43(3):553-557.
- [28] MEDINA-LOZANO I, BERTOLÍN J R, DIAZ A. Nutritional value of commercial and traditional lettuce (*Lactuca sativa* L.) and wild relatives; vitamin C and anthocyanin content[J]. *Food Chemistry*, 2021, 359:129864.
- [29] 胡雅馨, 李京, 惠伯棣. 蓝莓果实中主要营养及花青素成分的研究[J]. *食品科学*, 2006(10):600-603.
- [30] 宋雪薇, 魏解冰, 狄少康, 等. 花青素转录因子调控机制及代谢工程研究进展[J]. *植物学报*, 2019, 54(1):133-156.
- [31] ALBERT N W, DAVIES K M, LEWIS D H, et al. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots[J]. *The Plant Cell*, 2014, 26(3):962-980.
- [32] BOREVITZ J O, XIA Y, BLOUNT J, et al. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis[J]. *The Plant Cell*, 2000, 12(12):2383-2393.
- [33] KOBAYASHI S, ISHIMARU M, HIRAOKA K, et al. Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis[J]. *Planta*, 2002, 215:924-933.
- [34] ELOMAA P, UIMARI A, MEHTO M, et al. Activation of anthocyanin biosynthesis in *Gerbera hybrida* (Asteraceae) suggests conserved protein-protein and protein-promoter interactions between the anciently diverged monocots and eudicots[J]. *Plant Physiology*, 2003, 133(4):1831-1842.

(责任编辑:成纾寒)