

紫背天葵 MBW 相关调控因子转录组测序分析

张少平^{1,2}, 赖正锋¹, 练冬梅¹, 姚运法¹, 李和平¹, 洪建基¹, 郑开斌^{1,2}, 邱珊莲^{1*}

(1. 福建省农业科学院亚热带农业研究所, 福建 漳州 363005; 2. 福建省农业科学院作物研究所, 福州 350013)

摘要: 为获取紫背天葵(*Cynura bicolor*)花青素合成相关转录调控因子 MBW 家族基因, 采用二代高通量测序技术进行全转录功能基因组测序后组装, 再通过 Pfam、SwissProt 和 Nr 数据库搜索, 共获得 138 个 MBW 相关 Unigene, 分别有 42 个 MYB、67 个 bHLH、15 个 bHLH-MYB 和 14 个 WD40, 其中目前已报道与花青素合成代谢相关的 MYB、bHLH、bHLH-MYB 和 WD40 分别为 11、33、6 和 3 个。这为进一步研究紫背天葵花青素合成的调控机理和相关基因克隆等奠定基础。

关键词: 紫背天葵; MBW; 转录组测序; 调控因子; 花青素

doi: 10.11926/jtsb.3787

Sequence Analysis of MBW Related Genes Transcriptome in *Gynura bicolor*

ZHANG Shao-ping^{1,2}, LAI Zhen-feng¹, LIAN Dong-mei¹, YAO Yun-fa¹, LI He-ping¹, HONG Jian-ji¹, ZHENG Kai-bin^{1,2}, QIU Shan-lian^{1*}

(1. Subtropical Agriculture Research Institute, Fujian Academy of Agricultural Sciences, Zhangzhou 363005, Fujian, China; 2. Crop Sciences Institute, Fujian Academy of Agricultural Sciences, Fuzhou 350013, China)

Abstract: *Gynura bicolor* is rich in anthocyanidin. In order to identify the transcription factors involved in anthocyanidin biosynthesis pathway in *G. bicolor*, the transcriptome of this plant was sequenced by Illumina HiSeq 2500 platform. By searching of 3 databases, including Pfam, SwissProt and Nr, 138 MBW-related unigenes were obtained, which containing 42 MYBs, 67 bHLHs, 15 bHLH-MYBs and 14 WD40s. Among these MBWs, 11 MYBs, 33 bHLHs, 6 bHLH-MYBs and 3 WD40s involved in anthocyanidin biosynthesis pathway were annotated in other plant species, including *Vitis vinifera*, *Dahlia pinnata*, *Gerbera*, *Catharanthus roseus*, *Prunus persica* etc. The obtaining of the 138 unigenes related to MBWs from *Gynura bicolor* provides a good basis for further research in gene cloning and the regulation of anthocyanin synthesis.

Key words: *Gynura bicolor*; MBW; Transcriptome sequencing; Regulation factor; Anthocyanidin

叶片正面绿色、背面紫色的紫背天葵(*Cynura bicolor*), 俗称紫背菜、红凤菜、观音菜及血皮菜等^[1], 为菊科(Asteraceae)三七草属植物。其嫩茎叶除含有叶绿素外还富含花青素, 可做保健蔬菜、观叶植物

及天然花青素来源^[2-4]。植物花青素兼具营养、药理以及在基因工程中改良花色等^[5-7]作用, 具有重要的研究及应用价值。花青素合成代谢途径涉及结构及调控两类基因^[8], 其中结构基因编码花青素合

收稿日期: 2017-07-05

接受日期: 2017-09-25

基金项目: 福建省公益类科研院所专项(2016R1012-4); 福建省农业科学院青年英才计划项目(YC2015-19); 福建省农业科学院青年创新团队项目(STIT2017-3-4); 福建省农业科学院生产性工程化实验室中试项目(AG2017-5)资助

This work was supported by the special Project for Public Welfare Research Institutes in Fujian (Grant No. 2016R1012-4), the Project for Young Talents of Fujian Academy of Agricultural Sciences (Grant No. YC2015-19), the Project for Innovation Team of Fujian Academy of Agricultural Sciences (Grant No. STIT2017-3-4), and the Pilot Project of Productive Engineering Laboratory of Fujian Academy of Agricultural Sciences (Grant No. AG2017-5).

作者简介: 张少平, 男, 农艺师, 主要从事特色蔬菜研究。E-mail: zspnc@163.com

* 通信作者 Corresponding author. E-mail: slqiu79@163.com

成代谢途径中所需的酶,花青素合成特有的酶主要有 DFR、ANS 和 3GT 等 3 种^[9-10]。而花青素合成代谢调节基因编码转录因子,主要通过与其结构基因的启动子结合来调控表达,从而影响花青素的时空积累。花青素调控的转录因子主要包括 MYB、bHLH 和 WD40 等 3 种^[11],其中 MYB 转录因子是一类 DNA 结合蛋白,由 MYB 结构域构成,这些结构域都是由高度保守的氨基酸连续或间隔的序列拼接后再通过折叠成 HTH 结构(螺旋-螺旋-转角-螺旋)来识别 DNA 并与之结合,达到调控基因表达的目的^[12-13]。MYB 是花青素合成代谢中涉及及最广泛的调控因子,依据 R 基序的数量,MYB 一般分为 4 类,以含有 2 个基序的 R2R3-MYB 最多^[14-15],此外还包括 R3-MYB、R1R2R3-MYB 和 R1R2R1R2-MYB 等^[16]。bHLH(碱性螺旋-环-螺旋,也作 MYC)转录因子也是由一系列保守的 bHLH 氨基酸序列构成,通过 N 末端碱性氨基酸区域结合 DNA 来调控基因表达^[17]。WD40 也是一类由一段高度保守的氨基酸序列串联重复的蛋白,这些序列都是从 N 端 GH 二肽开始到 C 端 WD 结尾,该蛋白不能直接作用 DNA,只是通过形成蛋白质骨架参与多个蛋白基因的相互作用^[18]。研究表明,花青素生物合成一般由 bHLH 和 R2R3-MYB^[19-21]或由 MYB、bHLH 和 WD40 形成 MBW 三聚复合体^[22-23]来调节和影响结构基因表达。

近年来,二代高通量转录组测序技术已广泛应用于多种生物特定状态下全转录功能基因组测序,大量新的功能基因被发现甚至被利用^[24-25]。本研究以 Illumina HiSeq 2500 技术进行紫背天葵嫩茎叶高通量测序,通过多种生物信息学方法进行大量的基因簇(Unigene)功能注释,然后再通过 Pfam、SwissProt 和 Nr 等 3 个公共数据库进行花青素相关调控基因 MBW (MYB、bHLH 和 WD40)搜索及相关信息分析,为紫背天葵或其他植物花青素合成代谢相关调控因子的研究奠定基础。

1 材料和方法

1.1 材料和转录组测序

从福建省农业科学院亚热带农业研究所野特菜资源圃选取紫背天葵健壮枝条扦插于穴盘,约 45 d 后将整盆植株小苗送至北京百迈克测序公司进行 Illumina HiSeq 2500 测序平台测序,参照张少平等^[26]

的方法,对 3 株紫背天葵嫩茎叶进行总 RNA 提取、等量组合 RNA 池、mRNA 富集及反转录成 cDNA、测序文库制备等,最后进行测序数据拼接。

1.2 功能注释及 MBW 相关数据分析

将测序拼接获取的紫背天葵 Unigene 通过非冗余蛋白数据库(Nr)、蛋白质序列数据库(SwissProt)、蛋白质家族域数据库(Pfam)、基因本体论数据库(GO)、蛋白质直系同源数据库(COG)以及东京基因与基金组百科全书(KEGG)等数据库进行注释分析。根据注释获取的结果,进一步进行 MYB、bHLH 和 WD40 等调控因子关键词搜索,根据是否与花青素合成代谢相关再进行归类,同时分别对获取的 MBW 相关调控因子进行 FPKM 及核苷酸序列长度等分析。

2 结果和分析

2.1 MYB 相关的 Unigene 信息分析

根据六大数据库注释,仅从 Pfam、SwissProt 和 Nr 等 3 个公共数据库中共获取不同 MYB 相关调控因子 42 个。从 Nr 数据库注释信息可知(表 1),编码 ID 为 c30845~c23218 中共有 11 个 MYB 调控因子与花青素合成代谢相关,这些花青素合成相关调控因子的 FPKM 值为 0~838.1,序列长度为 219~2 261 bp。其中 Nr 数据库注释到的 11 个花青素合成代谢相关调控因子中,分别来自葡萄(*Vitis vinifera*, 3 个 Unigene)、非洲菊(*Gerbera*, 3 个)、长春花(*Catharanthus roseus*, 3 个)以及向日葵(*Helianthus annuus*)和甜菜(*Beta vulgaris*)各 1 个。这 11 个 MYB 相关调控因子在 SwissProt 数据库中共注释到 9 个 Unigene,来自拟南芥(*Arabidopsis thaliana*, 8 个 Unigene)和马铃薯(*Solanum tuberosum*, 1 个);此外,以上 MYB 相关调控因子在 Pfam 数据库中只注释到 5 个 Unigene。

编码 ID 为 c6955~c20765 的共 31 个 MYB 相关调控因子中,目前还未见报道与花青素合成代谢相关,因此其功能代谢调控还有待进一步深入研究。所有这些 MYB 调控因子的 FPKM 值为 2.32~185.63,序列长度为 205~1 891 bp,主要来自拟南芥、马铃薯及芝麻(*Sesamum indicum*)等。

2.2 bHLH 相关的 Unigene 信息分析

从 Pfam、SwissProt 和 Nr 等 3 个公共数据库中,共获得 67 个 bHLH 相关调控因子。从 Nr 数据库注

表 1 MYB 相关 Unigene 信息分析

Table 1 Information of MYB-related Unigene

编码 ID	序列长度 (bp) Sequence length	FPKM	Pfam	SwissProt 数据库及匹配物种 SwissProt and matching plant	Nr 数据库及匹配物种 Nr and matching plant
c30845	275	0	-	RAX2 拟南芥 <i>Arabidopsis thaliana</i>	transcription factor 向日葵 <i>Helianthus annuus</i>
c26652	219	1.19	Myb	MYB12 拟南芥 <i>A. thaliana</i>	MYB9A 非洲菊 <i>Gerbera</i>
c35787	271	2.61	-	REVEILLE 拟南芥 <i>A. thaliana</i>	MYB 长春花 <i>Catharanthus roseus</i>
c9265	517	3.08	-	-	MYB9A 非洲菊 <i>Gerbera</i>
c13604	1 486	9.14	Myb	Myb 拟南芥 <i>A. thaliana</i>	myb 甜菜 <i>Beta vulgaris</i>
c16240	1 315	21.37	-	MYB1R1 马铃薯 <i>Solanum tuberosum</i>	MYB1R1 葡萄 <i>Vitis vinifera</i>
c23990	1 698	25.48	Myb	PHR1-LIKE 拟南芥 <i>A. thaliana</i>	PHR1-LIKE 葡萄 <i>V. vinifera</i>
c19205	1 815	39.71	-	protein 拟南芥 <i>A. thaliana</i>	MYB 长春花 <i>Catharanthus roseus</i>
c5632	484	44.93	-	-	MYB9A 非洲菊 <i>Gerbera</i>
c18261	1 249	80.86	Myb	Myb 拟南芥 <i>A. thaliana</i>	LOC100250267 葡萄 <i>Vitis vinifera</i>
c23218	2 261	838.1	Myb	REVEILLE 拟南芥 <i>A. thaliana</i>	MYB 长春花 <i>Catharanthus roseus</i>
c6955	442	2.32	Myb	MYB108 拟南芥 <i>A. thaliana</i>	
c30556	279	2.51	-	-	MYB12 棉花 <i>Gossypium hirsutum</i>
c32563	229	3.33	-	-	MYB 青蒿 <i>Artemisia annua</i>
c26963	819	3.43	-	GAMYB 水稻 <i>Oryza sativa</i>	
c891	205	3.95	-	MYB2 大麦 <i>Hordeum vulgare</i>	MYB 青蒿 <i>A. annua</i>
c7533	712	4.27	-	-	
c13290	1 274	4.96	Myb	Myb 拟南芥 <i>A. thaliana</i>	
c5298	512	5.84	-	MYB3R-1 拟南芥 <i>A. thaliana</i>	
c7239	587	5.98	-	-	MYB39 巨桉 <i>Eucalyptus grandis</i>
c8216	819	6.17	Myb	Myb 拟南芥 <i>A. thaliana</i>	
c11735	927	6.98	Myb	Myb 拟南芥 <i>A. thaliana</i>	myb 芝麻 <i>Sesamum indicum</i>
c8045	977	7.72	Myb	MYB1R1 马铃薯 <i>Solanum tuberosum</i>	MYB1R1 马铃薯 <i>Solanum tuberosum</i>
c20206	1 590	8.35	Myb	PHR1-LIKE 拟南芥 <i>A. thaliana</i>	
c12244	1 024	8.41	-	Myb 金鱼草 <i>Antirrhinum majus</i>	
c7069	804	8.62	-	MYB39 拟南芥 <i>Arabidopsis thaliana</i>	
c8214	280	9.16	-	REVEILLE 拟南芥 <i>A. thaliana</i>	MYB14 黄芩 <i>Scutellaria baicalensis</i>
c15581	609	9.87	-	MYB1R1 马铃薯 <i>Solanum tuberosum</i>	MYB1R1 芝麻 <i>Sesamum indicum</i>
c12197	748	10.10	Myb	MYB108 拟南芥 <i>Arabidopsis thaliana</i>	MYB24 芝麻 <i>S. indicum</i>
c19447	1 035	10.51	Myb	Myb 拟南芥 <i>A. thaliana</i>	
c4165	430	11.51	Myb	TRY 拟南芥 <i>A. thaliana</i>	
c8971	581	13.44	Myb	MYB1R1 拟南芥 <i>A. thaliana</i>	Myb 可可 <i>Theobroma cacao</i>
c12815	679	13.51	-	-	
c10353	584	14.04	-	-	MYB108 芝麻 <i>Sesamum indicum</i>
c23347	1 891	16.92	Myb	Myb3R 拟南芥 <i>A. thaliana</i>	PF11 马铃薯 <i>Solanum tuberosum</i>
c4487	1 338	19.36	Myb	PHR1-LIKE 拟南芥 <i>A. thaliana</i>	
c23405	697	25.69	Myb	PHR1-LIKE 拟南芥 <i>A. thaliana</i>	
c21395	1 509	42.19	Myb	PHR1-LIKE 拟南芥 <i>A. thaliana</i>	
c24000	1 551	56.94	Myb	MYB1R1 马铃薯 <i>Solanum tuberosum</i>	
c17340	1 217	119.11	Myb	MYB1R1 马铃薯 <i>S. tuberosum</i>	MYB1R1 芝麻 <i>Sesamum indicum</i>
c17367	1 215	137.47	Myb	MYB1R1 马铃薯 <i>S. tuberosum</i>	MYB1R1 烟草 <i>Nicotiana sylvestri</i>
c20765	1 280	185.63	Myb	MYB1R1 马铃薯 <i>S. tuberosum</i>	MYB1R1 马铃薯 <i>Solanum tuberosum</i>

释信息可知(表 2), 编码 ID 为 c36981~c20263 中共有 33 个 bHLH 调控因子与花青素合成代谢相关, 其 FPKM 为 1.03~244.34, 序列长度为 202~1 708 bp。这些调控因子分别来自葡萄(14 个 Unigene)、桃(*Prunus persica*, 3 个)、柑橘(*Citrus sinensis*, 2 个)、大丽花(*Dahlia pinnata*, 2 个)、莲(*Nelumbo nucifera*, 2 个)以及猴面花(*Erythranthe guttata*)、甜菜、苹果(*Malus domestica*)、番茄(*Lycopersicon esculentum*)、草莓(*Fragaria vesca*)和川桑(*Citrus clementina*)各 1

个。这些调控因子在 SwissProt 数据库中共注释到 30 个 Unigene, 来自拟南芥(28 个 Unigene)和豌豆(*Pisum sativum*, 2 个)。此外, 以上 bHLH 相关调控因子在 Pfam 数据库中共注释到了 18 个 Unigene。

编码 ID 为 c37200~c18704 中共有 34 个 bHLH 相关调控因子, 目前还未见报道与花青素合成代谢相关, 其 FPKM 值为 0.64~89.51, 序列长度为 239~1 713 bp, 主要来自拟南芥、烟草(*Nicotiana sylvestri*)、马铃薯和可可(*Theobroma cacao*)等。

表 2 bHLH 相关的 Unigene 信息分析

Table 2 Information of bHLH-related Unigenes

编码 ID	序列长度 (bp) Sequence length	FPKM	Pfam	SwissProt 及匹配物种 SwissProt and matching plant	Nr 及匹配物种 Nr and matching plant
c36948	241	1.03	bHLH	bHLH69 拟南芥 <i>Arabidopsis thaliana</i>	CISIN 柑橘 <i>Citrus sinensis</i>
c4119	465	2.18	-	bHLH 豌豆 <i>Pisum sativum</i>	bHLH 大丽花 <i>Dahlia pinnata</i>
c37054	228	2.23	-	bHLH123 拟南芥 <i>A. thaliana</i>	unnamed protein 葡萄 <i>Vitis vinifera</i>
c38537	202	2.70	-	-	bHLH111 莲 <i>Nelumbo nucifera</i>
c30369	834	3.13	bHLH	bHLH10 拟南芥 <i>A. thaliana</i>	MIMGU 猴面花 <i>Erythranthe guttata</i>
c26703	618	3.44	-	bHLH30 拟南芥 <i>A. thaliana</i>	bHLH30 葡萄 <i>Vitis vinifera</i>
c33179	278	4.20	-	-	bHLH111 葡萄 <i>V. vinifera</i>
c6673	1 028	4.28	-	bHLH78 拟南芥 <i>A. thaliana</i>	bHLH78 甜菜 <i>Beta vulgaris</i>
c8134	845	5.19	bHLH	bHLH35 拟南芥 <i>A. thaliana</i>	bHLH35 葡萄 <i>Vitis vinifera</i>
c12339	358	5.51	-	bHLH 豌豆 <i>Pisum sativum</i>	bHLH 大丽花 <i>Dahlia pinnata</i>
c9561	526	5.66	-	bHLH62 拟南芥 <i>A. thaliana</i>	bHLH62 莲 <i>Nelumbo nucifera</i>
c13180	1 224	6.21	bHLH	bHLH87 拟南芥 <i>A. thaliana</i>	bHLH87 苹果 <i>Malus domestica</i>
c16877	1 276	6.98	bHLH	bHLH91 拟南芥 <i>A. thaliana</i>	VITISV 葡萄 <i>Vitis vinifera</i>
c9056	997	7.00	bHLH	bHLH51 拟南芥 <i>A. thaliana</i>	bHLH51 葡萄 <i>V. vinifera</i>
c12434	1 221	7.05	bHLH	bHLH18 拟南芥 <i>A. thaliana</i>	PRUPE 桃 <i>Prunus persica</i>
c6713	832	7.27	bHLH	bHLH30 拟南芥 <i>A. thaliana</i>	bHLH30 苹果 <i>Malus domestica</i>
c5988	831	8.20	bHLH	bHLH35 拟南芥 <i>A. thaliana</i>	bHLH35 葡萄 <i>Vitis vinifera</i>
c18956	2 018	8.71	bHLH	bHLH49 拟南芥 <i>A. thaliana</i>	unnamed protein 葡萄 <i>V. vinifera</i>
c3791	627	12.01	-	-	bHLH27 番茄 <i>Lycopersicon esculentum</i>
c8862	1 303	14.10	bHLH	bHLH18 拟南芥 <i>A. thaliana</i>	bHLH25 草莓 <i>Fragaria vesca</i>
c18980	1 569	16.70	-	bHLH112 拟南芥 <i>A. thaliana</i>	bHLH112 葡萄 <i>Vitis vinifera</i>
c20181	1 420	17.82	bHLH	bHLH74 拟南芥 <i>A. thaliana</i>	unnamed protein 葡萄 <i>V. vinifera</i>
c11737	1 216	23.27	bHLH	bHLH130 拟南芥 <i>A. thaliana</i>	bHLH130 葡萄 <i>V. vinifera</i>
c19260	347	24.59	-	bHLH148 拟南芥 <i>A. thaliana</i>	L484 川桑 <i>Citrus clementina</i>
c24244	1 418	27.57	-	bHLH121 拟南芥 <i>A. thaliana</i>	PRUPE 桃 <i>Prunus persica</i>
c21509	1 498	39.50	bHLH	bHLH63 拟南芥 <i>A. thaliana</i>	PRUPE 桃 <i>P. persica</i>
c15867	673	42.92	-	bHLH147 拟南芥 <i>A. thaliana</i>	IBH1 葡萄 <i>Vitis vinifera</i>
c22983	1 683	48.35	bHLH	bHLH130 拟南芥 <i>A. thaliana</i>	bHLH130 葡萄 <i>V. vinifera</i>
c23869	1 656	49.02	-	bHLH144 拟南芥 <i>A. thaliana</i>	bHLH144 葡萄 <i>V. vinifera</i>
c19750	1 191	80.54	bHLH	bHLH79 拟南芥 <i>A. thaliana</i>	bHLH79 柑橘 <i>Citrus sinensis</i>
c22667	1 623	102.39	bHLH	bHLH77 拟南芥 <i>A. thaliana</i>	bHLH77 葡萄 <i>Vitis vinifera</i>
c16906	981	124.07	-	bHLH113 拟南芥 <i>A. thaliana</i>	PRUPE 桃 <i>Prunus persica</i>
c20263	1708	244.34	bHLH	bHLH130 拟南芥 <i>A. thaliana</i>	VITISV 葡萄 <i>V. vinifera</i>
c37200	342	0.64	bHLH	bHLH71 拟南芥 <i>A. thaliana</i>	bHLH71 烟草 <i>Nicotiana sylvestri</i>
c35160	239	1.04	bHLH	bHLH94 拟南芥 <i>A. thaliana</i>	bHLH96 大豆 <i>Glycine max</i>
c33248	353	1.84	-	bHLH30 拟南芥 <i>A. thaliana</i>	bHLH30 烟草 <i>Nicotiana sylvestri</i>
c35444	240	2.07	-	bHLH137 拟南芥 <i>A. thaliana</i>	bHLH137 鹰嘴豆 <i>Cicer arietinum</i>
c25167	628	2.46	-	bHLH90 拟南芥 <i>A. thaliana</i>	JCGZ 麻疯树 <i>Jatropha curcas</i>
c1361	310	2.90	-	LHW 拟南芥 <i>A. thaliana</i>	bHLH157 马铃薯 <i>Solanum tuberosum</i>
c9688	274	3.43	-	-	bHLH110 青梅 <i>Prunus mume</i>
c29575	285	4.06	bHLH	bHLH18 拟南芥 <i>A. thaliana</i>	PHAVU 菜豆 <i>Phaseolus vulgaris</i>
c6469	927	4.24	bHLH	bHLH128 拟南芥 <i>A. thaliana</i>	bHLH128 烟草 <i>Nicotiana sylvestri</i>
c26816	457	4.46	-	bHLH140 拟南芥 <i>A. thaliana</i>	bHLH140 苜蓿 <i>Medicago truncatula</i>
c8621	602	4.52	-	bHLH68 拟南芥 <i>A. thaliana</i>	bHLH68 杨树 <i>Populus euphratica</i>
c28239	619	4.69	bHLH	bHLH69 拟南芥 <i>A. thaliana</i>	JCGZ 麻疯树 <i>Jatropha curcas</i>
c7501	614	4.73	-	-	bHLH118 杨树 <i>Populus euphratica</i>
c26819	418	4.96	bHLH	bHLH14 拟南芥 <i>A. thaliana</i>	POPTR 杨树 <i>P. euphratica</i>
c5121	855	5.13	bHLH	bHLH123 拟南芥 <i>A. thaliana</i>	JCGZ 麻疯树 <i>Jatropha curcas</i>
c10203	729	5.20	-	bHLH61 拟南芥 <i>A. thaliana</i>	bHLH61 可可 <i>Theobroma cacao</i>
c5178	487	5.37	-	-	bHLH137 芝麻 <i>Sesamum indicum</i>
c12315	966	6.34	-	bHLH137 拟南芥 <i>A. thaliana</i>	bHLH 烟草 <i>Nicotiana sylvestri</i>
c25899	1 182	8.74	bHLH	bHLH25 拟南芥 <i>A. thaliana</i>	helix-loop-helix 可可 <i>Theobroma cacao</i>
c19371	1 713	9.95	bHLH	bHLH78 拟南芥 <i>A. thaliana</i>	bHLH62 芝麻 <i>Sesamum indicum</i>
c4039	1 093	11.01	bHLH	bHLH118 拟南芥 <i>A. thaliana</i>	bHLH118 烟草 <i>Nicotiana sylvestri</i>

续表 (Continued)

编码 ID	序列长度 (bp) Sequence length	FPKM	Pfam	SwissProt 及匹配物种 SwissProt and matching plant	Nr 及匹配物种 Nr and matching plant
c20569	1 692	14.30	-	bHLH68 拟南芥 <i>A. thaliana</i>	bHLH68 大豆 <i>Glycine max</i>
c19712	970	21.16	bHLH	bHLH80 拟南芥 <i>A. thaliana</i>	bHLH80 木本棉 <i>Gossypium arboreum</i>
c17759	1 311	21.98	bHLH	bHLH71 拟南芥 <i>A. thaliana</i>	bHLH71 马铃薯 <i>Solanum tuberosum</i>
c19247	1 159	24.83	bHLH	bHLH79 拟南芥 <i>A. thaliana</i>	DNA binding protein 蓖麻 <i>Ricinus communis</i>
c14606	1 068	29.80	bHLH	bHLH30 拟南芥 <i>A. thaliana</i>	bHLH30 马铃薯 <i>Solanum tuberosum</i>
c14668	983	36.96	-	-	bHLH130 芝麻 <i>Sesamum indicum</i>
c21688	2 743	38.58	-	bHLH145 拟南芥 <i>A. thaliana</i>	transcription factore 蓖麻 <i>Ricinus communis</i>
c19843	1 496	42.20	-	bHLH48 拟南芥 <i>A. thaliana</i>	unnamed protein 中粒咖啡 <i>Coffea canephora</i>
c21235	1 114	47.89	-	bHLH47 拟南芥 <i>A. thaliana</i>	bHLH47 烟草 <i>Nicotiana sylvestri</i>
c19260	347	70.89	-	bHLH147 拟南芥 <i>A. thaliana</i>	CICLE 巴西橡胶树 <i>Morus notabilis</i>
c22359	1 267	83.75	bHLH	HB11 拟南芥 <i>A. thaliana</i>	bHLH64 马铃薯 <i>Solanum tuberosum</i>
c16839	1 137	86.30	bHLH	bHLH104 拟南芥 <i>A. thaliana</i>	unnamed protein 中粒咖啡 <i>Coffea canephora</i>
c18704	1 635	89.51	-	bHLH144 拟南芥 <i>A. thaliana</i>	bHLH1 巴西橡胶树 <i>Morus notabilis</i>

2.3 MYB 和 bHLH 结合蛋白相关的 Unigene 信息

在 Pfam 公共数据库中, 共注释到 15 个 bHLH 和 R2R3-MYB 相关结合蛋白, 这些结合蛋白在 Nr 和 SwissProt 数据库中大多注释为 bHLH 或相应的基因别称。从 Nr 数据库注释信息可知(表 3), c9372、c6720、c15452、c24314、c24516 和 c23161 等 6 个 Unigene 与花青素合成代谢相关, 其 FPKM 值为 2.39~67.68, 序列长度为 360~2 664 bp。Nr 数据库注释到的 6 个花青素合成代谢相关调控因子, 涉及到的植物有大丽花(2 个 Unigene)以及莲、紫背天葵 (*Gynura bicolor*)、葡萄和甘野菊(*Chrysanthemum*

boreale)各 1 个。这 6 个调控因子在 SwissProt 数据库中共注释到 5 个 Unigene, 涉及植物有拟南芥(3 个 Unigene)和豌豆(2 个)。

编码 ID 为 c33892~c21592 中共有 9 个相关结合蛋白调控因子, 目前还未见报道与花青素合成代谢相关, 因此还有待进一步深入研究。从全转录功能基因组测序分析可知, 这些调控因子的 FPKM 值为 2.44~55.9, 序列长度为 215~2 330 bp, 涉及的植物主要有拟南芥, 此外还有梨(*Pyrus*)、马铃薯、杨树(*Populus trichocarpa*)、蓖麻(*Ricinus communis*)和烟草等。

表 3 bHLH 及 R2R3-MYB 相关 Unigene 信息分析

Table 3 Information of bHLH and R2R3-MYB-related Unigenes

编码 ID	序列长度 (bp) Sequence length	FPKM	Pfam	SwissProt 及匹配物种 SwissProt and matching plant	Nr 及匹配物种 Nr and matching plant
c9372	360	2.39	bHLH and R2R3-MYB	-	EMB1444 莲 <i>Nelumbo nucifera</i>
c6720	484	2.91	bHLH and R2R3-MYB	helix-loop-helix 豌豆 <i>Pisum sativum</i>	bHLH 大丽花 <i>Dahlia pinnata</i>
c15452	1 774	15.47	bHLH and R2R3-MYB	EGL1 GN 拟南芥 <i>Arabidopsis thaliana</i>	GbMYC1 紫背天葵 <i>Gynura bicolor</i>
c24314	2 254	39.40	bHLH and R2R3-MYB	BHLH 豌豆 <i>Pisum sativum</i>	bHLH 大丽花 <i>Dahlia pinnata</i>
c24516	2 664	41.33	bHLH and R2R3-MYB	EMB1444 拟南芥 <i>Arabidopsis thaliana</i>	LHW 葡萄 <i>Vitis vinifera</i>
c23161	2 099	67.68	bHLH and R2R3-MYB	MYC4 拟南芥 <i>A. thaliana</i>	MYC21 甘野菊 <i>Chrysanthemum boreale</i>
c33892	215	2.44	bHLH and R2R3-MYB	-	EMB1444 梨 <i>Pyrus</i>
c34341	401	2.61	bHLH and R2R3-MYB	bHLH157 拟南芥 <i>A. thaliana</i>	bHLH157 马铃薯 <i>Solanum tuberosum</i>
c26695	575	4.42	bHLH and R2R3-MYB	MYC2 拟南芥 <i>A. thaliana</i>	
c18084	2 029	6.43	bHLH and R2R3-MYB	bHLH157 拟南芥 <i>A. thaliana</i>	
c8080	1 201	9.83	bHLH and R2R3-MYB	bHLH155 拟南芥 <i>A. thaliana</i>	
c24011	2 187	20.59	bHLH and R2R3-MYB	EMB1444 拟南芥 <i>A. thaliana</i>	
c11203	2 330	21.96	bHLH and R2R3-MYB	LHW 拟南芥 <i>A. thaliana</i>	transcription factor 杨树 <i>Populus trichocarpa</i>
c24059	1 054	30.78	bHLH and R2R3-MYB	EMB1444 拟南芥 <i>A. thaliana</i>	helix-loop-helix 蓖麻 <i>Ricinus communis</i>
c21592	1 686	55.90	bHLH and R2R3-MYB	bHLH3 拟南芥 <i>A. thaliana</i>	bHLH3 烟草 <i>Nicotiana sylvestri</i>

2.4 WD40 相关的 Unigene 信息分析

在 Pfam、SwissProt 及 Nr 等 3 个公共数据库中, 共注释到 14 种 WD40 相关调控因子。从 Nr 数据库

注释信息可知(表 4), 有 c7535、c23326 和 c20388 等 3 个 Unigene 与花青素合成代谢相关, 其中 c7535 序列长度为 562 bp, FPKM 值为 6.64, 在 Nr 数据

库中来自葡萄,而在 Pfam 数据库中来自拟南芥;c23326 序列长度为 2 438 bp,FPKM 值为 113.28,在 Nr 数据库中来自柑橘,而在 Pfam 数据库中未注释到;而 c20388 序列长度为 1 672 bp,FPKM 值为 130.54,在 Nr 数据库中来自莲,而在 Pfam 数据库中来自拟南芥。

在编码 ID 为 c30814~c23391 中共有 11 个 WD40

相关调控因子,还未见报道与花青素合成代谢相关。从全转录功能基因组测序分析可知,FPKM 值为 3.09~190.22,序列长度为 296~3 872 bp,在 Nr 及 SwissProt 数据库中所有 Unigene 主要来自可可和拟南芥,此外还有胡杨(*Populus euphratica*)、烟草、马铃薯以及紫草(*Lithospermum erythrorhizon*)、衣藻(*Chlamydomonas smithii*)和苜蓿(*Medicago sativa*)等。

表 4 WD40 相关 Unigene 信息分析

Table 4 Information of WD40-related Unigenes

编码 ID	序列长度 (bp) Sequence length	FPKM	Pfam	SwissProt 及匹配物种 SwissProt and matching plant	Nr 及匹配物种 Nr and matching plant
c7531	562	6.64	WD domain	WD40 拟南芥 <i>Arabidopsis thaliana</i>	Cajal body protein 葡萄 <i>Vitis vinifera</i>
c23326	2 438	113.28	WD40	-	CICLE 柑橘 <i>Citrus sinensis</i>
c20388	1 672	130.54	WD domain	Aladin 拟南芥 <i>A. thaliana</i>	aladin isoform X2 莲 <i>Nelumbo nucifera</i>
c30814	296	3.09	WD domain	-	Transducin/WD40 可可 <i>Theobroma cacao</i>
c5312	503	3.18	WD domain	LEC14B 紫草 <i>Lithospermum erythrorhizon</i>	Transducin/WD40 可可 <i>T. cacao</i>
c28754	460	3.54	-	-	Transducin/WD40 可可 <i>T. cacao</i>
c821	488	5.77	-	-	Transducin/WD40 可可 <i>T. cacao</i>
c11884	1 521	7.27	WD domain	-	Transducin/WD40 可可 <i>T. cacao</i>
c19965	1 062	8.78	con80 domain	WD40 拟南芥 <i>A. thaliana</i>	katanin p80 WD40 胡杨 <i>Populus euphratica</i>
c22639	3 016	18.80	-	protein ligase 苜蓿 <i>Medicago sativa</i>	Transducin/WD40 可可 <i>T. cacao</i>
c23024	1 535	24.53	WD domain	WD40 拟南芥 <i>A. thaliana</i>	small nucleolar RNA 烟草 <i>Nicotiana sylvestri</i>
c23036	1 363	53.92	WD domain	Mut11 衣藻 <i>Chlamydomonas smithii</i>	Transducin/WD40 可可 <i>T. cacao</i>
c24003	3 872	79.09	WD domain	complexity 拟南芥 <i>A. thaliana</i>	Transducin/WD40 可可 <i>T. cacao</i>
c23391	2 302	190.22	WD40	-	LOC102604753 马铃薯 <i>Solanum tuberosum</i>

3 结论和讨论

植物花青素具有突出的药用和营养价值,同时在基因工程改良花色中具有重要的经济效益。因此,对花青素代谢途径及其分子调控机理的研究一直是现代分子生物学领域的热点^[27]。多年来,关于植物花青素合成代谢途径的研究大多集中在结构基因上,而有关调控基因的研究也主要是在少量模式植物中^[28],从其他高等植物中分离到的 MBW (MYB、bHLH、WD40)相关调控因子成员相对单一,因此,大多成员的特征特性及相关功能未能得到进一步分析。本研究采用二代高通量测序技术对紫背天葵全转录功能基因组进行测序,再分别通过 Pfam、SwissProt 及 Nr 等 3 个公共数据库进行 MYB、bHLH、WD40 等关键词搜索,其中 Pfam 注释到 MBW 相关基因共 83 个,包括 23 个 MYB、34 个 bHLH、15 个 bHLH-MYB 及 11 个 WD40 或相关基因别称;SwissProt 注释到 MBW 相关基因共 115 个,包括 34 个 MYB、60 个 bHLH、13 个 bHLH-MYB 和 8 个 WD40 或相关基因别称;Nr 共注释到 MBW 相关基因共 118 个,包括 26 个 MYB、67 个 bHLH、

11 个 bHLH-MYB 和 14 个 WD40 或相关基因别称。从 3 大数据库中共获得 138 个 MBW 相关 Unigene,包括 42 个 MYB、67 个 bHLH、15 个 bHLH-MYB 和 14 个 WD40,以上注释到 MBW 相关调控因子 Unigene 以来自拟南芥的最多。

在所有 138 个 MBW 相关调控因子中,目前已报道与花青素合成代谢相关的 MYB、bHLH、bHLH-MYB 和 WD40 分别为 11、33、6 和 3 个,分别来自葡萄、大丽花、非洲菊、长春花、桃、柑橘、向日葵、甜菜、猴面花等近 20 种植物。植物花青素合成代谢相关调控基因的研究已引起广泛关注,在紫背天葵中也有少量报道^[29-30],但有关 MBW 相关调控因子的系统研究仍集中在少量模式植物。目前从拟南芥中分离克隆出 MYB 家族成员 339 个,bHLH 家族成员 150 个,WD 家族成员 237 个,而从水稻中分离克隆出的 MYB 家族成员也达到 230 个^[31]。本研究从紫背天葵转录功能基因组测序中获得的 MBW 相关调控基因较之偏少,这可能是由于不同植物的本质特征决定,同时,也可能是由于 MBW 相关调控因子在参与植物合成代谢过程中受到光照、低温、干旱、激素以及病原感染等逆境胁迫的

影响, 同一植物在不同状态下 MBW 相关调控因子表达量不同甚至存在变异, 从而导致不同 MBW 相关调控因子家族基因数量差异明显。本研究获得的 138 个 MBW 相关 Unigene, 通过进一步鉴定和分析, 可了解紫背天葵花青素合成调控机理, 为进一步利用基因工程开展植物遗传改良等具有重要的理论指导意义。

参考文献

- [1] ZHANG S P, QIU S L, DENG Y, et al. Study and application of the anthocyanin from *Gynura bicolor* [J]. Chin Agric Sci Bull, 2015, 31(22): 157–162.
张少平, 邱珊莲, 邓源, 等. 紫背天葵花青素相关研究与应用 [J]. 中国农学通报, 2015, 31(22): 157–162.
- [2] MIYAZAWA M, NAKAHASHI H, USAMI A, et al. Chemical composition, aroma evaluation, and inhibitory activity towards acetylcholinesterase of essential oils from *Gynura bicolor* DC. [J]. J Nat Med, 2016, 70(2): 282–289. doi: 10.1007/s11418-015-0961-1.
- [3] TEOH W Y, TAN H P, LING S K, et al. Phytochemical investigation of *Gynura bicolor* leaves and cytotoxicity evaluation of the chemical constituents against HCT 116 cells [J]. Nat Prod Res, 2016, 30(4): 448–451. doi: 10.1080/14786419.2015.1017726.
- [4] CHEN J, MANGELINCKX S, LÜ H, et al. Profiling and elucidation of the phenolic compounds in the aerial parts of *Gynura bicolor* and *G. divaricata* collected from different Chinese origins [J]. Chem Biodiv, 2015, 12(1): 96–115. doi: 10.1002/cbdv.201400134.
- [5] da SILVA DOS PASSOS A P, MADRONA G S, MARCOLINO V A, et al. The Use of thermal analysis and photoacoustic spectroscopy in the evaluation of maltodextrin microencapsulation of anthocyanins from juçara palm fruit (*Euterpe edulis* Mart.) and their application in food [J]. Food Technol Biotechnol, 2015, 53(4): 385–396. doi: 10.17113/ftb.53.04.15.3726.
- [6] LIPPERT E, RUEMMELE P, OBERMEIER F, et al. Anthocyanins prevent colorectal cancer development in a mouse model [J]. Digestion, 2017, 95(4): 275–280. doi: 10.1159/000475524.
- [7] FERNANDES I, MARQUES C, ÉVORA A, et al. Pharmacokinetics of table and Port red wine anthocyanins: a crossover trial in healthy men [J]. Food Funct, 2017, 8(5): 2030–2037. doi: 10.1039/C7FO00329C.
- [8] ALBERT N W, DARIES K M, LEWIS D H, et al. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots [J]. Plant Cell, 2014, 26(3): 962–980. doi: 10.1105/tpc.113.122069.
- [9] JIA X Y, SHEN J, LIU H, et al. Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato [J]. Planta, 2015, 242(1): 283–293. doi: 10.1007/s00425-015-2305-5.
- [10] LI W, WANG B, WANG M, et al. Cloning and characterization of a potato *StAN11* gene involved in anthocyanin biosynthesis regulation [J]. J Integr Plant Biol, 2014, 56(4): 364–372. doi: 10.1111/jipb.12136.
- [11] XU W J, DUBOS C, LEPINIEC L. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes [J]. Trends Plant Sci, 2015, 20(3): 176–185. doi: 10.1016/j.tplants.2014.12.001.
- [12] JUN J H, LIU C G, XIAO X R, et al. The Transcriptional repressor MYB2 regulates both spatial and temporal patterns of proanthocyanidin and anthocyanin pigmentation in *Medicago truncatula* [J]. Plant Cell, 2015, 27(10): 2860–2879. doi: 10.1105/tpc.15.00476.
- [13] HATLESTAD G J, AKHAVAN N A, SUNNADENIYA R M, et al. The beet *Y* locus encodes an anthocyanin MYB-like protein that activates the betalain red pigment pathway [J]. Nat Genet, 2015, 47(1): 92–96. doi: 10.1038/ng.3163.
- [14] LI S T, ZACHGO S. TCP3 interacts with R2R3-MYB proteins, promotes flavonoid biosynthesis and negatively regulates the auxin response in *Arabidopsis thaliana* [J]. Plant J, 2013, 76(6): 901–913. doi: 10.1111/tpj.12348.
- [15] PÉREZ-DÍAZ J R, PÉREZ-DÍAZ J, MADRID-ESPINOZA J, et al. New member of the R2R3-MYB transcription factors family in grapevine suppresses the anthocyanin accumulation in the flowers of transgenic tobacco [J]. Plant Mol Biol, 2016, 90(1–2): 63–76. doi: 10.1007/s11103-015-0394-y.
- [16] WANG F Q, SUO Y F, WEI H, et al. Identification and characterization of 40 isolated *Rehmannia glutinosa* MYB family genes and their expression profiles in response to shading and continuous cropping [J]. Int J Mol Sci, 2015, 16(7): 15009–15030. doi: 10.3390/ijms160715009.
- [17] GOOSSENS J, MERTENS J, GOOSSENS A. Role and functioning of bHLH transcription factors in jasmonate signalling [J]. J Exp Bot, 2017, 68(6): 1333–1347. doi: 10.1093/jxb/erw440.
- [18] SERRAMITO-GÓMEZ I, BOADA-ROMERO E, PIMENTEL-MUIÑOS F X. Unconventional autophagy mediated by the WD40 domain of ATG16L1 is derailed by the T300A Crohn disease risk polymorphism [J]. Autophagy, 2016, 12(11): 2254–2255. doi: 10.1080/15548627.2016.1216303.
- [19] NEMIE-FEYISSA D, HEIDARI B, BLAISE M, et al. Analysis of interactions between heterologously produced bHLH and MYB proteins that regulate anthocyanin biosynthesis: Quantitative interaction kinetics by microscale thermophoresis [J]. Phytochemistry, 2015, 111: 21–26. doi: 10.1016/j.phytochem.2015.01.004.

- [20] BULGAKOV V P, VEREMEICHIK G N, GRIGORCHUK V P, et al. The *rolB* gene activates secondary metabolism in *Arabidopsis* calli via selective activation of genes encoding MYB and bHLH transcription factors [J]. *Plant Physiol Biochem*, 2016, 102: 70–79. doi: 10.1016/j.plaphy.2016.02.015.
- [21] CHEZEM W R, CLAY N K. Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs [J]. *Phytochemistry*, 2016, 131: 26–43. doi: 10.1016/j.phytochem.2016.08.006.
- [22] XIE Y, TAN H J, MA Z X, et al. DELLA proteins promote anthocyanin biosynthesis via sequestering MYBL2 and JAZ suppressors of the MYB/bHLH/WD40 Complex in *Arabidopsis thaliana* [J]. *Mol Plant*, 2016, 9(5): 711–721. doi: 10.1016/j.molp.2016.01.014.
- [23] XU W J, GRAIN D, BOBET S, et al. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in *Arabidopsis* seed [J]. *New Phytol*, 2014, 202(1): 132–144. doi: 10.1111/nph.12620.
- [24] SANGWAN R S, TRIPATHI S, SINGH J, et al. De novo sequencing and assembly of *Centella asiatica* leaf transcriptome for mapping of structural, functional and regulatory genes with special reference to secondary metabolism [J]. *Gene*, 2013, 525(1): 58–76. doi: 10.1016/j.gene.2013.04.057.
- [25] XU L F, YANG P P, YUAN S X, et al. Transcriptome analysis identifies key candidate genes mediating purple ovary coloration in Asiatic hybrid lilies [J]. *Int J Mol Sci*, 2016, 17(11): 1881. doi: 10.3390/ijms17111881.
- [26] ZHANG S P, HONG J J, QIU S L, et al. Sequencing and analysis of the transcriptome of *Gynura bicolor* [J]. *Acta Hort Sin*, 2016, 43(5): 935–946. doi: 10.16420/j.issn.0513-353X.2016-0140.
- 张少平, 洪建基, 邱珊莲, 等. 紫背天葵高通量转录组测序分析 [J]. *园艺学报*, 2016, 43(5): 935–946. doi: 10.16420/j.issn.0513-353X.2016-0140.
- [27] LI Y, GAO Z R, ZHANG C, et al. Research progress on the molecular regulation mechanism of anthocyanin biosynthesis pathway [J]. *Chin J Ecol*, 2015, 34(10): 2937–2942. doi: 10.13292/j.1000-4890.2015.0281.
- 李莹, 高振蕊, 张驰, 等. 花青素合成途径中分子调控机制的研究进展 [J]. *生态学杂志*, 2015, 34(10): 2937–2942. doi: 10.13292/j.1000-4890.2015.0281.
- [28] LI S N, WANG W Y, GAO J L, et al. MYB75 phosphorylation by MPK4 is required for light-induced anthocyanin accumulation in *Arabidopsis* [J]. *Plant Cell*, 2016, 28(11): 2866–2883. doi: 10.1105/tpc.16.00130.
- [29] SHIMIZU Y, MAEDA K, KATO M, et al. Isolation of anthocyanin-related MYB gene, *GbMYB2*, from *Gynura bicolor* leaves [J]. *Plant Biotechnol*, 2010, 27(5): 481–487.
- [30] SHIMIZU Y, MAEDA K, KATO M, et al. Co-expression of *GbMYB1* and *GbMYC1* induces anthocyanin accumulation in roots of cultured *Gynura bicolor* DC. plantlet on methyl jasmonate treatment [J]. *Plant Physiol Biochem*, 2011, 49(2): 159–167. doi: 10.1016/j.plaphy.2010.11.006.
- [31] LIU X F, LI F, YIN X R, et al. Recent advances in the transcriptional regulation of anthocyanin biosynthesis [J]. *Acta Hort Sin*, 2013, 40(11): 2295–2306. doi: 10.3969/j.issn.0513-353X.2013.11.023.
- 刘晓芬, 李芳, 殷学仁, 等. 花青苷生物合成转录调控研究进展 [J]. *园艺学报*, 2013, 40(11): 2295–2306. doi: 10.3969/j.issn.0513-353X.2013.11.023.