

阳桃 R2R3-MYB 家族成员鉴定及其在木质素合成过程中的表达

赵亚梅¹, 陈 蕾¹, 吴春梅¹, 秦 思¹, 翟俊文¹, 任 惠², 吴沙沙¹

(¹ 福建农林大学园林学院 / 兰科植物保护与利用国家林业和草原局重点实验室, 福州 350002;

² 广西壮族自治区农业科学院园艺研究所, 南宁 530007)

摘要: MYB 家族已被证实在植物生长发育、类黄酮合成、环境胁迫、木质素生物合成等多个生理方面发挥着重要的作用。为探究阳桃 MYB 家族在木质素生物合成过程中的作用, 本研究基于全基因组数据, 对阳桃 (*Averrhoa carambola* L.) R2R3-MYB 转录因子进行了全面的生物信息学分析, 并采用 RT-qPCR 技术, 对 8 个阳桃 R2R3-MYB 基因在阳桃小枝 3 个发育时期 (T1、T3、T6)、果和叶的表达情况进行了验证。结果显示, 阳桃基因组中共包含 57 个 1R-MYB、100 个 R2R3-MYB、4 个 3R-MYB, 100 个阳桃 R2R3-MYB (*AcMYBs*) 基因不均匀地分布在 11 条染色体上, 被分为 24 个亚组, 相同亚组的基因具有相似但不完全相同的基因结构和保守基序。片段重复和串联重复在阳桃 MYB 家族的扩张中发挥了重要的作用, 且重复产生的基因都经历了纯化选择。同时, 阳桃 R2R3-MYB 家族与拟南芥 (*Arabidopsis thaliana* (L.) Heynh.) 及毛果杨 (*Populus trichocarpa* Torr. & A. Gray) 都具有较强的共线性关系。RT-qPCR 结果表明, 除 *AcMYB05* 外, *AcMYB41*、*AcMYB65*、*AcMYB84*、*AcMYB87*、*AcMYB92*、*AcMYB97* 和 *AcMYB100* 都在 T3 和 T6 时期呈现出较高的表达, 在 T1 和果、叶中几乎不表达。本研究表明, 7 个 *AcMYBs* 的表达具有明显组织特异性, 其可能参与了阳桃木质素生物合成, 为进一步研究 MYB 家族对木质素生物合成的影响提供了重要的参考价值。

关键词: 阳桃; 木质素; R2R3-MYB 转录因子; 全基因组鉴定; 表达模式

Identification of the R2R3-MYB Family Members in *Averrhoa carambola* and Their Expression Involved in Lignin Biosynthesis

ZHAO Ya-mei¹, CHEN Lei¹, WU Chun-mei¹, QIN Si¹,
ZHAI Jun-wen¹, REN Hui², WU Sha-sha¹

(¹ College of Landscape Architecture, Fujian Agriculture and Forestry University/Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization, Fuzhou 350002;

² Horticulture Research Institute of Guangxi Academy of Agricultural Science, Nanning 530007)

Abstract: MYB family transcription factors have been proved to play an important role in biological process, such as plant growth and development, flavonoid synthesis, environmental stress, lignin biosynthesis. To explore if the MYB family members are involved in the lignin biosynthesis in *Averrhoa carambola* L., a genome-wide bioinformatics analysis of R2R3-MYB transcription factor has been conducted, followed by the

收稿日期: 2021-04-12 修回日期: 2021-04-22 网络出版日期: 2021-05-12

URL: <http://doi.org/10.13430/j.cnki.jpgr.20210412002>

第一作者研究方向为园林植物与应用, E-mail: yamei_zhao@126.com

通信作者: 吴沙沙, 研究方向为园林植物资源与应用研究, E-mail: shashawu1984@126.com

基金项目: 广西科技重大专项 (桂科 AA17204045-4); 广西农业科学院基本科研业务专项 (桂农科 2021YT046); 福建农林大学艺术学院、园林学院 (合署) 学科专业建设项目 (YSYL-bdpy-1)

Foundation projects: Science and Technology Major Project of Guangxi (Guike AA17204045-4), Guangxi Academy of Agricultural Sciences Special Project for Basic Scientific Research (Guinongke 2021YT046), Disciplinary Professional Construction Project of College of Art & College of Landscape Architecture (YSYL-bdpy-1)

RT-qPCR analysis to valid the expression levels of the eight R2R3-MYB genes at different stages in twigs (T1, T3, T6), fruits and leaves. Fifty-seven 1R-MYB, 100 R2R3-MYB and four 3R-MYB have been identified. Out of them, 100 R2R3-MYB (*AcMYBs*) were detected to be unevenly distributed on 11 chromosomes and they were divided into 24 subclades (A1-A24; the genes from the same subclade present similarity on gene structures and functional motifs). The family gene expansion might associated to the segmental and tandem duplications, and the duplicated genes were undergone purifying selection. The sequence homology of R2R3-MYB gene family was observed in *A. carambola* L., *Arabidopsis thaliana* (L.) Heynh. and *Populus trichocarpa* Torr. & A. Gray. RT-qPCR showed that except for *AcMYB05*, the genes *AcMYB41*, *AcMYB65*, *AcMYB84*, *AcMYB87*, *AcMYB92*, *AcMYB97* and *AcMYB100* were found with higher expression in T3 and T6 and rare transcripts in T1, fruit and leaves. Collectively, the transcriptional profiles suggested the tissue-specific expression of seven *AcMYBs*, which possibly participate in the lignin biosynthesis pathway of *A. carambola* L., therefore providing reference for further analyzing the MYB family members that interplay with the lignin biosynthesis.

Key words: *Averrhoa carambola* L.; lignin; R2R3-MYB transcription factors; genome-wide identification; expression pattern analysis

木质素作为一种复杂的有机聚合物,是植物次生细胞壁 (SCW, secondary cell wall) 中重要的组成部分,它不仅能增强植物的机械强度,提高细胞壁的刚性、疏水性,还能促进矿物质在维管束的运输^[1]。木质素的生物合成除了受各种酶的影响外,还受到一个复杂的转录因子网络 (如 MYB、NAC、bHLH 等) 调控,其中 MYB 转录因子是植物 SCW 形成或木质素生物合成过程中最重要的转录因子^[2]。

MYB 蛋白具有高度保守的 DNA 结合结构域,通常由 1~4 个保守的 R 结构域组成 (R0、R1、R2 和 R3)^[3],每个保守的结构域都具有 3 个 α 螺旋结构,第 2 个和第 3 个 α 螺旋通过 3 个均匀分布的色氨酸形成了 HTH (Helix-Turn-Helix) 螺旋结构^[4],且每个重复中的第 3 个螺旋是与 DNA 直接接触的 DNA 识别螺旋^[5]。根据保守结构域 R 的数量,MYB 蛋白可以分为 4 种类型: 1R-MYB、R2R3-MYB、3R-MYB 和 4R-MYB,其中 R2R3-MYB 转录因子是植物中最主要的类型^[3]。

R2R3-MYB 转录因子在植物生长发育的过程中具有重要的作用,如调节生物和非生物胁迫,影响木质素、花色苷的合成等^[4]。Chen 等^[6]研究表明,棉花 (*Gossypium* spp.) 转录因子 GbMYB05 能赋予棉花和转基因烟草 (*Nicotiana tabacum* L.) 不同程度的耐寒性。水稻 (*Oryza sativa* L.) 中过量表达 OsMYB91 和 OsMYB3R-2 能提升其对盐胁迫的耐受性、耐寒性,增加细胞有丝分裂指数^[7-8]。对拟南芥 (*Arabidopsis thaliana* (L.) Heynh.) 的研究表明,至少有 17 个 R2R3-MYB 参与调控植物 SCW 的形

成^[9-10],其中,AtMYB58、AtMYB63 和 AtMYB85 是木质素生物合成特异性的转录因子^[11],AtMYB46 和 AtMYB83 属于同一亚组,它们不仅是木质素合成途径的调节因子,而且能有效地激活 SCW 形成的整个过程^[12-13]。

作为植物中最丰富的 MYB 类型,拟南芥^[3]、大豆 (*Glycine max* (L.) Merr.)^[14]、毛果杨 (*Populus trichocarpa* Torr. & A. Gray)^[15] 等多个物种已经在全基因组的水平上对 R2R3-MYB 转录因子进行了分析。阳桃 (*Averrhoa carambola* L.) 是热带亚热带重要的观赏和经济树种,但由于枝条脆弱,不耐台风,限制了其应用和栽培范围。基于木质素含量与抗逆性具有重要的相关性,通过阳桃 R2R3-MYB 的鉴定及其对木质素生物合成影响的研究,旨在为选育抗逆性强的阳桃品种提供理论支撑。目前,阳桃两个基因组已公布^[16-17],使得在全基因组水平上分析 R2R3-MYB 成为可能。本研究基于基因组数据对阳桃 R2R3-MYB (*AcMYBs*) 进行了生物信息学分析,通过转录组数据及实时荧光定量 PCR (RT-qPCR) 技术研究 R2R3-MYB 的表达模式,分析 MYB 基因与阳桃木质素生物合成的关系,为进一步研究 MYB 基因在阳桃木质素生物合成过程中的功能提供参考。

1 材料与方法

1.1 阳桃 MYB 成员的全基因组预测

首先在 BIG 数据库 (<https://bigd.big.ac.cn/gwh>) 中下载阳桃的全基因组数据 (GWHABKE00000000),

同时在 Pfam (<http://pfam.xfam.org/>) 数据库下载 MYB 结构域 (PF00249) 作为搜索模型, 利用 HMM 3.0 软件中的 hmm search 筛选阳桃中含有该结构域的蛋白质序列。然后, 从 TAIR 数据库 (<https://www.arabidopsis.org/>) 下载拟南芥的 MYB 蛋白质序列并构建本地数据库, 在阳桃的蛋白质数据中执行 BLASTP 搜索。将以上两种方法所得到的蛋白质序列合并, 去除重复、没有完整编码框和测序不完全的蛋白质序列。再利用 SMART (<http://smart.embl.de/>)、Pfam、NCBI-CDD (<https://www.ncbi.nlm.nih.gov/cdd>) 数据库对候选蛋白质进行验证, 从网站 ExpASy (<https://web.expasy.org/protparam/>) 获得 *AcMYBs* 的等电点和分子量。

1.2 保守结构域分析和系统发育分析

利用 Clustal^[18] 对最终确认的 *AcMYBs* 进行多序列比对, 在线软件 WebLogo3 (<http://weblogo.berkeley.edu/logo.cgi>) 生成 R2R3 结构域的 logos 图片。为探究 *AcMYBs* 之间的系统发育关系及功能, 从 TAIR 数据库中下载拟南芥 R2R3-MYB 蛋白质序列作为参考, 采用 MEGA 7.0 中的邻接法 (Neighbor-joining) 进行系统发育树的构建, Bootstrap 值设置为 1000 次重复^[19]。

1.3 基因结构、蛋白基序及表达分析

使用在线网站 GSDS (<http://gsds.cbi.pku.edu.cn/>) 分析 *AcMYBs* 的基因结构, 通过 MEME (<http://meme-suite.org/tools/meme>) 对 *AcMYBs* 蛋白序列保守 motif 进行分析, 并使用 TBtools^[20] 将得到的 motif 结果进行进一步绘制。

于 2020 年 11 月采集广西农业科学院苗圃种植的大果甜阳桃一号作为实验样品, T1、T3、T6 3 个时期分别混合取样第 1 片和第 2 片叶子之间的小枝、第 3 片和第 4 片叶子之间的小枝、第 6 片和第 7 片叶子之间的小枝进行转录组测序, 每个样本 3 个生物学重复, 获取 *AcMYBs* 在不同时期的 FPKM 值, 使用 Tbttool 进行热图的绘制。

1.4 蛋白互作网络预测及染色体定位、共线性分析

利用在线网站 STRING (<https://string-db.org/>), 以拟南芥作为参考, 最低互动评分设置为 0.700 以探究 *AcMYBs* 蛋白互作网络。从阳桃基因组 GFF 文件获取染色体定位信息, 得到每个基因在染色体上的分布情况。以拟南芥染色体分析结果作为参照^[21], 鉴定 *AcMYBs* 中的串联重复基因, 并分析其选择压力。采用 MCScanX 进行共线性分析, 根据 Wang 等^[22] 的描述分析 *AcMYBs* 的重复模式。

1.5 RNA 提取与实时荧光定量 PCR

从鉴定出可能参与阳桃木质素生物合成的 *AcMYBs* 中挑选 8 个基因 (*AcMYB05*、*AcMYB41*、*AcMYB65*、*AcMYB84*、*AcMYB87*、*AcMYB92*、*AcMYB97* 和 *AcMYB100*), 通过 RT-qPCR 定量分析它们在 T1、T3、T6 3 个时期及果、叶中的表达, 每个样本 3 个生物学重复。

采用天根多糖多酚植物总 RNA 提取试剂盒提取 T1、T3、T6、果和叶的总 RNA, 使用 TIANGEN RNA prep pure Plant kit 试剂盒反转录合成 cDNA。*MYB* 家族成员的 RT-qPCR 引物采用 Premier 5.0 设计 (表 1)。

表 1 荧光定量 PCR 引物

Table 1 Primers used for fluorescent quantitative PCR

基因名称 Gene name	正向引物 (5'-3') Forward primer	反向引物 (5'-3') Reverse primer
<i>AcMYB05</i>	AAGGAGGAAGAGGACAGCA	CTTATCGGGAGCAATCAATC
<i>AcMYB41</i>	CCCTGGAAGGACTGATAATG	GTAGTTGGAAGTGGCTTGTGT
<i>AcMYB65</i>	CGCGACTAACGAAGAAGAA	TGAAACAACCAAGACCAAAA
<i>AcMYB84</i>	GCAACCAACAGAAGGTGAA	CAACATCCATAGCCGTGAG
<i>AcMYB87</i>	TGGCAATAGGTGGTCAAAG	TATCAGTCCTTCTGGCAAC
<i>AcMYB92</i>	ACAACAACGACAATCACTTCA	CTCCATCAAACCTTCAAAATC
<i>AcMYB97</i>	TCAGGGAAGAGTTGTAGGTTG	TTCGATGTGAAGCCAAAAG
<i>AcMYB100</i>	TGCTGTCAAGAATCACTACCA	TCAGTCCTAAAGAAACTCCCA
<i>TUA</i>	CTTGGTGTTTAGTGCCGTTG	CCTGAGGAGAAGGATAGATGGT

以 *TUA* 作为内参基因, RT-qPCR 采用 TB Green *Premix Ex Taq II* (Tli RNaseH Plus) 试剂盒 (TaKaRa) 在 Applied Biosystems 7500 Real-Time PCR 仪中进行, 每个反应设置 3 次重复。扩增程序为: 95 °C 预变性 30 s; 95 °C 变性 5 s, 60 °C 退火 34 s, 40 个循环; 溶解曲线: 95 °C 反应 15 s, 60 °C 反应 1 min, 95 °C 反应 15 s。最后根据 $2^{-\Delta\Delta Ct}$ 法计算目的基因的相对表达量。

2 结果与分析

2.1 阳桃 MYB 基因鉴定

经筛选和多重验证后从阳桃全基因组中共鉴定

出 57 条 1R-MYB 序列, 100 条 R2R3-MYB 序列, 4 条 3R-MYB 序列, 没有发现 4R-MYB。R2R3-MYB 转录因子是当前研究最多的 MYB 蛋白, 具有丰富多样的功能特点, 该研究将 100 个 R2R3-MYB 基因分别命名为 *AcMYB1~AcMYB100*。

利用 ExPASy 进行理化分析显示, *AcMYB* (本研究中指 *AcMYB1~AcMYB100*) 氨基酸数目为 172~939 aa, 平均数量为 316 aa, 等电点范围为 4.71~9.95, 蛋白分子量范围为 19690.60~105910.14 Da (表 2)。

表 2 *AcMYBs* 信息分析

Table 2 Information analysis of *AcMYBs*

基因名称 Gene name	基因 ID Gene ID	登录号 Accession	染色体定位 (bp) Chromosome location	氨基酸数目 (aa) Amino acid number	等电点 PI	分子量 (Da) MW
<i>AcMYB01</i>	Yangtao2000027	MW831432	Chr.4: 18194099~18195533	328	5.19	36642.02
<i>AcMYB02</i>	Yangtao2000134	MW831433	Chr.4: 18194099~18195533	267	8.47	29609.57
<i>AcMYB03</i>	Yangtao2000154	MW831434	Chr.4: 15786264~15788167	419	6.57	47893.63
<i>AcMYB04</i>	Yangtao2000203	MW831435	Chr.4: 14889922~14891253	373	6.01	42118.87
<i>AcMYB05</i>	Yangtao2000558	MW831436	Chr.4: 10603216~10605774	345	5.10	38687.91
<i>AcMYB06</i>	Yangtao2000643	MW831437	Chr.4: 9730851~9732298	275	6.09	31204.41
<i>AcMYB07</i>	Yangtao2000645	MW831438	Chr.4: 9661718~9663167	275	7.58	31242.45
<i>AcMYB08</i>	Yangtao2000646	MW831439	Chr.4: 9651039~9652354	344	7.19	38579.44
<i>AcMYB09</i>	Yangtao2000915	MW831440	Chr.4: 7522666~7524804	554	6.49	60200.66
<i>AcMYB10</i>	Yangtao2001492	MW831441	Chr.4: 7522666~7524804	344	5.83	38338.79
<i>AcMYB11</i>	Yangtao2001770	MW831442	Chr.4: 1299163~1299873	172	4.71	19690.60
<i>AcMYB12</i>	Yangtao2002343	MW831472	Chr.6: 28645824~28646801	239	5.18	28134.30
<i>AcMYB13</i>	Yangtao2002561	MW831473	Chr.6: 27081351~27082427	359	9.44	41145.09
<i>AcMYB14</i>	Yangtao2002669	MW831474	Chr.6: 26281097~26283049	388	5.23	41859.80
<i>AcMYB15</i>	Yangtao2003413	MW831475	Chr.6: 19014330~19017609	381	7.68	41732.98
<i>AcMYB16</i>	Yangtao2003593	MW831476	Chr.6: 16310888~16313079	337	5.69	37453.45
<i>AcMYB17</i>	Yangtao2004186	MW831461	Chr.7: 3739736~3740903	264	5.97	29488.04
<i>AcMYB18</i>	Yangtao2004304	MW831462	Chr.7: 4503266~4505016	303	5.47	34767.10
<i>AcMYB19</i>	Yangtao2004305	MW831463	Chr.7: 4526219~4527415	318	6.60	36419.63
<i>AcMYB20</i>	Yangtao2004306	MW831464	Chr.7: 4562829~4564150	316	6.46	35881.00
<i>AcMYB21</i>	Yangtao2004448	MW831465	Chr.7: 5683671~5684485	194	9.26	22330.36
<i>AcMYB22</i>	Yangtao2004452	MW831466	Chr.7: 5714786~5715692	224	9.95	25598.63
<i>AcMYB23</i>	Yangtao2004454	MW831467	Chr.7: 5734184~5735251	274	7.15	30848.62
<i>AcMYB24</i>	Yangtao2004455	MW831468	Chr.7: 5765008~5766256	299	6.56	34122.07
<i>AcMYB25</i>	Yangtao2004555	MW831469	Chr.7: 6651686~6653521	286	6.95	31647.53

表 2 (续)

基因名称 Gene name	基因 ID Gene ID	登录号 Accession	染色体定位 (bp) Chromosome location	氨基酸数目 (aa) Amino acid number	等电点 PI	分子量 (Da) MW
<i>AcMYB26</i>	Yangtao2005084	MW831470	Chr.7: 11994103~11995656	392	6.01	43198.48
<i>AcMYB27</i>	Yangtao2005184	MW831422	Chr.10: 11772754~11774158	312	7.00	34848.12
<i>AcMYB28</i>	Yangtao2005196	MW831423	Chr.10: 11929101~11930553	259	5.01	30096.92
<i>AcMYB29</i>	Yangtao2005197	MW831424	Chr.10: 11937385~11938828	270	6.25	30816.33
<i>AcMYB30</i>	Yangtao2005198	MW831425	Chr.10: 11972591~11974280	246	5.31	28213.38
<i>AcMYB31</i>	Yangtao2005199	MW831426	Chr.10: 11980548~11982291	250	5.92	28636.80
<i>AcMYB32</i>	Yangtao2005308	MW831427	Chr.10: 12945010~12947347	561	5.25	60165.70
<i>AcMYB33</i>	Yangtao2005533	MW831428	Chr.10: 15071461~15072668	311	6.16	35221.63
<i>AcMYB34</i>	Yangtao2005726	MW831429	Chr.10: 16841670~16843452	318	5.33	35989.90
<i>AcMYB35</i>	Yangtao2006484	MW831430	Chr.10: 22343000~22344443	293	8.69	33482.94
<i>AcMYB36</i>	Yangtao2006487	MW831431	Chr.10: 22359665~22360508	211	9.15	24680.67
<i>AcMYB37</i>	Yangtao2006648	MW831448	Chr.1: 14840440~14842233	360	9.20	41190.08
<i>AcMYB38</i>	Yangtao2006759	MW831449	Chr.1: 13572351~13573680	187	5.82	20975.65
<i>AcMYB39</i>	Yangtao2006780	MW831450	Chr.1: 13291266~13295148	255	8.89	29395.07
<i>AcMYB40</i>	Yangtao2006781	MW831451	Chr.1: 13214161~13217884	251	9.41	28605.73
<i>AcMYB41</i>	Yangtao2006812	MW831452	Chr.1: 12728094~12729938	319	5.35	36083.58
<i>AcMYB42</i>	Yangtao2007013	MW831453	Chr.1: 10616421~10617450	268	5.81	30915.69
<i>AcMYB43</i>	Yangtao2007112	MW831454	Chr.1: 9850697~9853666	359	5.37	40873.42
<i>AcMYB44</i>	Yangtao2007119	MW831455	Chr.1: 9767427~9770189	192	7.72	21927.72
<i>AcMYB45</i>	Yangtao2007134	MW831456	Chr.1: 9643710~9644735	221	8.84	24782.61
<i>AcMYB46</i>	Yangtao2007313	MW831457	Chr.1: 8181631~8182976	341	5.81	37940.83
<i>AcMYB47</i>	Yangtao2007336	MW831458	Chr.1: 7983211~7985297	365	9.68	41210.73
<i>AcMYB48</i>	Yangtao2007344	MW831459	Chr.1: 7874135~7875965	398	6.17	43460.89
<i>AcMYB49</i>	Yangtao2007780	MW831460	Chr.1: 5037308~5039969	267	6.53	30268.24
<i>AcMYB50</i>	Yangtao2008295	MW831491	Chr.3: 30202870~30203640	257	5.48	29235.85
<i>AcMYB51</i>	Yangtao2008763	MW831492	Chr.3: 33904432~33905384	284	6.80	32073.91
<i>AcMYB52</i>	Yangtao2009064	MW831407	Chr.11: 8167258~8168408	255	5.42	28903.60
<i>AcMYB53</i>	Yangtao2009174	MW831408	Chr.11: 8993396~8994521	286	8.88	33101.37
<i>AcMYB54</i>	Yangtao2009754	MW831409	Chr.11: 13642579~13643944	361	5.67	40702.20
<i>AcMYB55</i>	Yangtao2009862	MW831410	Chr.11: 14510192~14511166	325	7.67	35340.31
<i>AcMYB56</i>	Yangtao2009886	MW831411	Chr.11: 14673637~14674380	248	8.84	27942.66
<i>AcMYB57</i>	Yangtao2010565	MW831416	Chr.8: 10128209~10129924	317	5.58	35721.04
<i>AcMYB58</i>	Yangtao2010785	MW831417	Chr.8: 11698048~11700350	315	6.02	34569.88
<i>AcMYB59</i>	Yangtao2010995	MW831418	Chr.8: 13184812~13186093	299	8.89	33729.78
<i>AcMYB60</i>	Yangtao2011126	MW831419	Chr.8: 14204167~14205764	412	6.38	45654.35
<i>AcMYB61</i>	Yangtao2011187	MW831420	Chr.8: 14686843~14687988	293	9.07	33047.06
<i>AcMYB62</i>	Yangtao2011269	MW831421	Chr.8: 15445570~15448470	279	8.39	30419.21

表 2 (续)

基因名称 Gene name	基因 ID Gene ID	登录号 Accession	染色体定位 (bp) Chromosome location	氨基酸数目 (aa) Amino acid number	等电点 PI	分子量 (Da) MW
<i>AcMYB63</i>	Yangtao2011669	MW831493	Chr.2: 14185440~14187084	355	7.59	39690.01
<i>AcMYB64</i>	Yangtao2012098	MW831482	Chr.2: 5841707~5843024	299	8.89	33442.19
<i>AcMYB65</i>	Yangtao2012155	MW831483	Chr.2: 6261204~6263042	314	4.73	34967.73
<i>AcMYB66</i>	Yangtao2012287	MW831484	Chr.2: 7168740~7170378	278	8.43	30861.64
<i>AcMYB67</i>	Yangtao2012590	MW831485	Chr.2: 9246490~9247624	264	8.26	29286.00
<i>AcMYB68</i>	Yangtao2012705	MW831486	Chr.2: 10194120~10195540	347	7.17	38704.94
<i>AcMYB69</i>	Yangtao2012710	MW831487	Chr.2: 10247035~10249849	425	9.01	46914.41
<i>AcMYB70</i>	Yangtao2012777	MW831488	Chr.2: 10756528~10758402	255	5.24	28698.57
<i>AcMYB71</i>	Yangtao2012904	MW831489	Chr.2: 11634485~11635537	233	7.61	26911.41
<i>AcMYB72</i>	Yangtao2013140	MW831443	Chr.9: 19492313~19494069	363	8.04	39519.61
<i>AcMYB73</i>	Yangtao2013156	MW831444	Chr.9: 19663050~19664775	433	6.64	47201.40
<i>AcMYB74</i>	Yangtao2013209	MW831445	Chr.9: 20118755~20123200	494	6.13	55707.76
<i>AcMYB75</i>	Yangtao2013374	MW831446	Chr.9: 21423026~21424428	239	9.25	26948.01
<i>AcMYB76</i>	Yangtao2013536	MW831447	Chr.9: 22716584~22717608	242	5.57	27584.75
<i>AcMYB77</i>	Yangtao2014410	MW831477	Chr.2: 31117835~31121394	428	5.77	46402.24
<i>AcMYB78</i>	Yangtao2014499	MW831478	Chr.2: 30216142~30217278	231	7.04	26292.64
<i>AcMYB79</i>	Yangtao2014722	MW831479	Chr.5: 849342~850935	347	9.47	39327.28
<i>AcMYB80</i>	Yangtao2015089	MW831480	Chr.5: 3827839~3829092	355	5.99	39650.92
<i>AcMYB81</i>	Yangtao2016765	MW831494	Chr.2: 3685493~3686368	260	8.76	28770.59
<i>AcMYB82</i>	Yangtao2016946	MW831412	Chr.7: 26170519~26172519	232	8.73	26093.48
<i>AcMYB83</i>	Yangtao2016948	MW831413	Chr.7: 26156013~26160772	939	5.34	105910.14
<i>AcMYB84</i>	Yangtao2017119	MW831414	Chr.7: 24442540~24444401	356	5.92	39851.10
<i>AcMYB85</i>	Yangtao2017717	MW831501	Chr.6: 1423362~1424231	290	6.13	31615.16
<i>AcMYB86</i>	Yangtao2017927	MW831496	Chr.9: 8862336~8863127	197	8.95	22888.96
<i>AcMYB87</i>	Yangtao2018016	MW831497	Chr.9: 7531400~7533059	295	5.59	32881.60
<i>AcMYB88</i>	Yangtao2018377	MW831499	Chr.1: 2386501~2387531	197	8.80	22167.92
<i>AcMYB89</i>	Yangtao2018378	MW831500	Chr.1: 2375404~2376777	230	9.47	25931.51
<i>AcMYB90</i>	Yangtao2018800	MW831490	Chr.1: 16453359~16456655	266	6.60	29853.38
<i>AcMYB91</i>	Yangtao2019560	MW831471	Chr.1: 25599124~25602595	594	5.01	64437.57
<i>AcMYB92</i>	Yangtao2019821	MW831415	Chr.1: 32464397~32466446	376	6.05	42155.86
<i>AcMYB93</i>	Yangtao2020261	MW831503	Chr.8: 6078172~6086039	400	9.62	41921.29
<i>AcMYB94</i>	Yangtao2021157	MW831481	Chr.8: 8820684~8821913	328	6.12	36348.78
<i>AcMYB95</i>	Yangtao2021313	MW831505	Chr.11: 3116870~3120317	393	7.62	44686.97
<i>AcMYB96</i>	Yangtao2022101	MW831495	Chr.11: 1178677~1180524	266	7.00	30507.94
<i>AcMYB97</i>	Yangtao2022283	MW831506	Chr.4: 25505326~25507057	245	7.68	28561.81
<i>AcMYB98</i>	Yangtao2022998	MW831502	Chr.5: 23401912~23403036	230	6.98	26603.13
<i>AcMYB99</i>	Yangtao2023034	MW831504	Chr.8: 21113937~21115479	292	8.42	32927.80
<i>AcMYB100</i>	Yangtao2023976	MW831498	Chr.8: 24516237~24517298	310	9.30	34984.09

2.2 *AcMYBs* 保守结构域分析

logos图显示, *AcMYBs* 在每个重复的结构域中都具有高度保守的色氨酸残基(W), 其分别位于阳桃 R2 结构域的第 9、29 和 49 位, R3 结构域的第 27 和 46 位。在 R3 结构域中, 第 1 个色氨酸残基位于第 8 位, 现被苯丙氨酸所取代, 第 2、3 的 3 个色氨酸

残基高度保守。除高度保守的色氨酸残基外, R2 结构域中的 E¹³、D¹⁴、L¹⁷、G²⁵、L⁵³ 以及 R3 结构域中的 G²⁴、A³¹、R³⁷ 也存在高度保守的情况(图 1)。此外, 还有一部分相对保守的氨基酸残基, 即 R2 结构域的 G³⁷、L³⁸、R⁴⁰、K⁴³、C⁴⁵、R⁴⁶、L⁴⁷、R⁴⁸、N⁵¹, R3 结构域中的 P¹、E¹²、I³⁰、G³⁶、T³⁸、N⁴⁰、K⁴³、N⁴⁴。

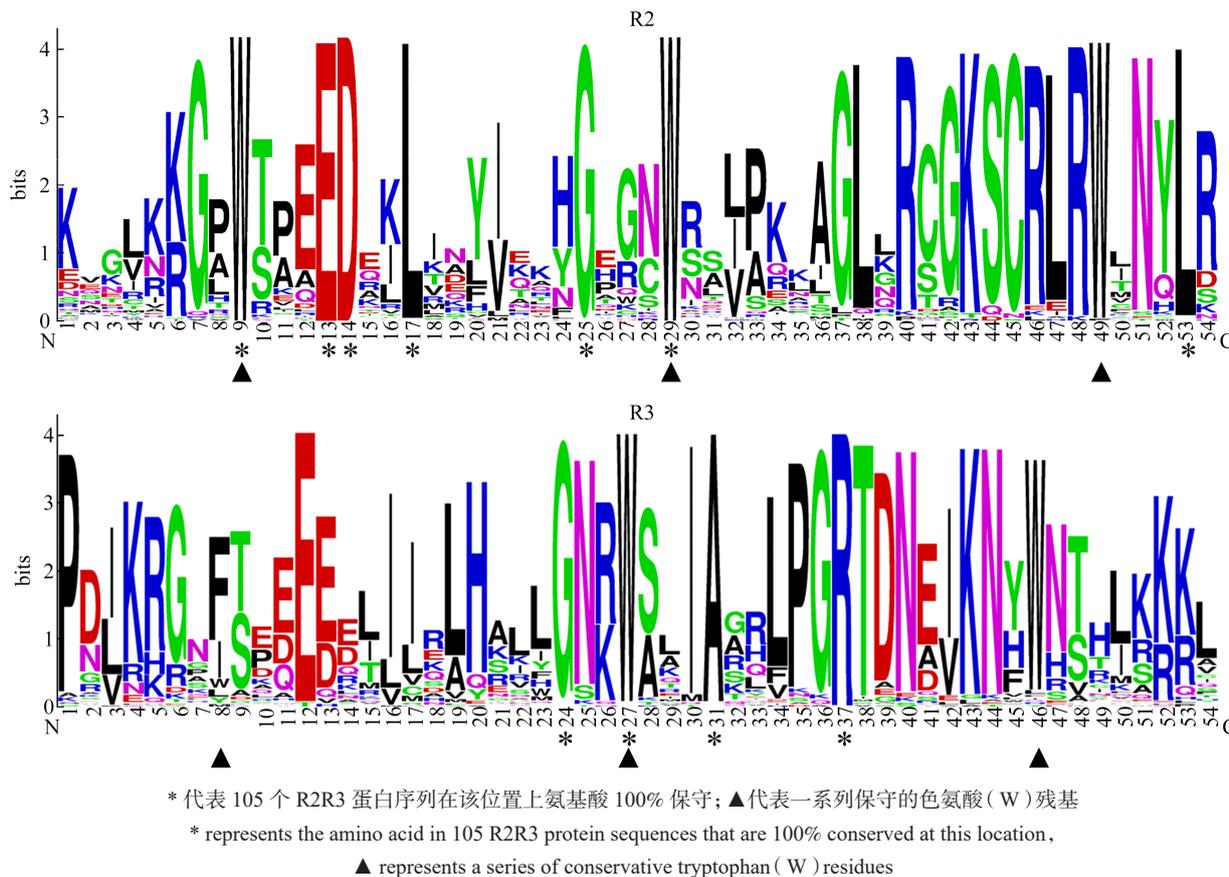


图 1 *AcMYBs* 的 R2R3 保守结构域序列标签

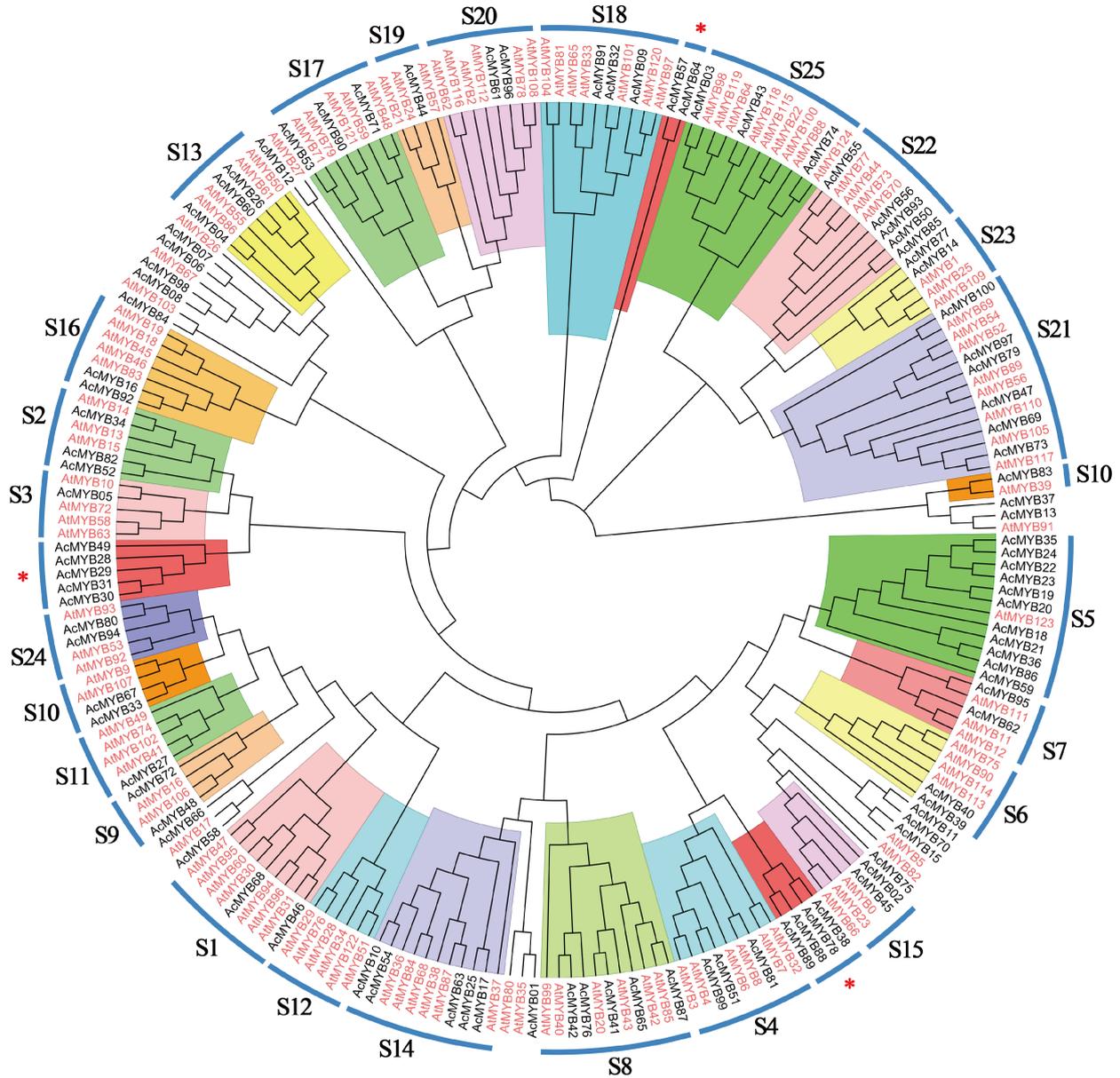
Fig.1 The conserved domain logos of the R2R3 domain in *AcMYBs*

2.3 *AcMYBs* 系统进化分析

为进一步研究 *AcMYBs* 的功能, 将 100 个 *AcMYBs* 与 125 个拟南芥 R2R3-MYB 基因进行系统发育树的构建。根据 Dubos 等^[4]提出的分类方法, 拟南芥 R2R3-MYB 被分为 25 个亚组(S1~S25)。从图 2 可以看出, 大部分 *AcMYBs* 都与拟南芥聚集, 但少数 *AcMYBs* 和 *AtMYBs* 没有被分到 25 个亚组内。S12 亚组不包含阳桃 MYB 基因, 且 11 个阳桃 MYB 产生了新的阳桃特有的系统发育分支。此外, S1、S6、S18 亚组中 *AcMYBs* 数量与拟南芥相比明显较少。

R2R3-MYB 蛋白与植物生长的多个方面都有不同程度的联系, 如黄酮物质的合成、生物和非生

物的胁迫、花色苷物质的合成等^[23]。在拟南芥 25 个亚组的 R2R3-MYB 基因中, 每个亚组都有其相应的功能作用, 其中 4 个亚组(S3、S8、S13 和 S21)及 *AtMYB46*、*AtMYB83* 都被验证参与木质素生物合成^[4]。由于同源基因通常具有相似的功能, 通过拟南芥可以推测阳桃中对应的直系同源基因的功能。因此与拟南芥 S3、S8、S13 和 S21 同一亚组的 *AcMYBs* (*AcMYB04*、*AcMYB05*、*AcMYB26*、*AcMYB41*、*AcMYB42*、*AcMYB47*、*AcMYB60*、*AcMYB65*、*AcMYB69*、*AcMYB73*、*AcMYB76*、*AcMYB79*、*AcMYB87*、*AcMYB97*、*AcMYB100*) 以及 *AtMYB46*、*AtMYB83* 的同源基因 *AcMYB16*、*AcMYB92* 都可能参与木质素生物合成。



黑色: 阳桃, 红色: 拟南芥; 不同的色块代表不同的亚组; *: 阳桃特有的亚组
 Black: *A. carambola* L., Red: *A. thaliana* (L.) Heynh., each subclade was labelled with different colour code,
 *: *Averrhoa*-specific subclade

图 2 R2R3-MYB 蛋白系统进化分析和亚组分类

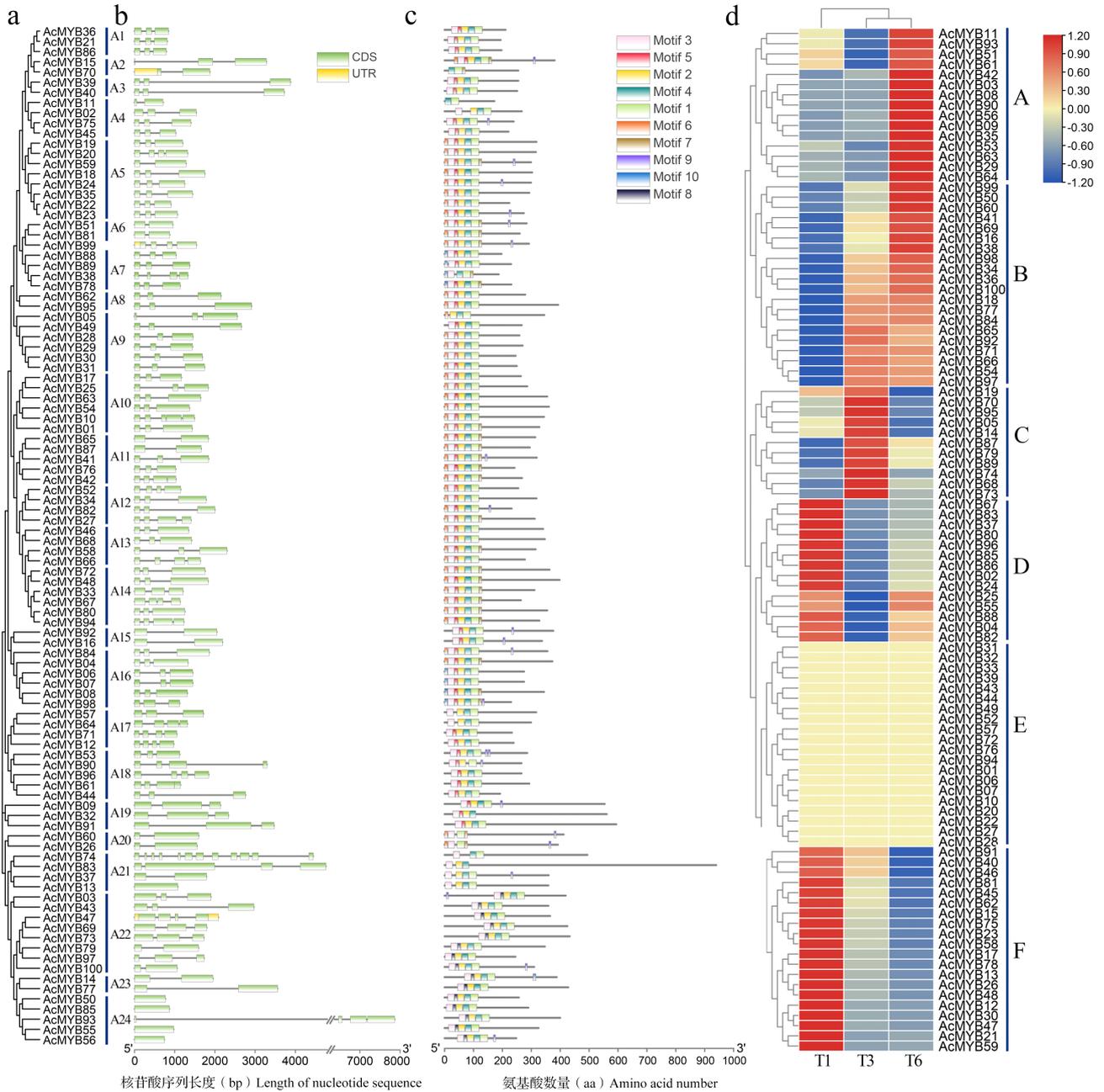
Fig.2 Phylogenetic relationships and subclade division of R2R3-MYB gene family

2.4 *AcMYBs* 基因结构及蛋白基序分析

结合基因结构信息和系统进化树, 100 个 *AcMYBs* 被分为 24 个亚组 (A1~A24, 图 3a)。大部分 *AcMYBs* 都包含 1~11 个不等的内含子 (图 3b), 只有 5 个 *AcMYBs* (*AcMYB13*、*AcMYB50*、*AcMYB55*、*AcMYB56* 和 *AcMYB85*) 没有内含子。大多数序列由 2 个内含子和 3 个外显子组成, 另外 13 个 *AcMYBs* 由 2 个外显子和 1 个内含子组成, 而 *AcMYB74* 具有 12 个外显子和 11 个内含子。*AcMYBs* 基因家族的系统发育分析显示, 同一亚组的序列具有相似的外显

子和内含子结构, 如 A9 和 A16 亚组都具有 2 个内含子和 3 外显子。

使用 MEME 工具识别 100 条 *AcMYBs* 蛋白序列的保守基序 (图 3c)。结果显示, *AcMYBs* 蛋白保守基序的数量为 2~8, 大多数 *AcMYBs* 蛋白都具有 Motif 1、Motif 2、Motif 3、Motif 4 和 Motif 5。与基因结构一样, 进化树中同一亚组的成员显示出相似的保守基序, A7 亚组的成员都包含 Motif 10, A24 亚组都包含 Motif 8。



a: *AcMYBs* 系统进化树; b: *AcMYBs* 基因结构图; c: *AcMYBs* 蛋白保守基序分布图; d: *AcMYBs* 表达分析图。

A1~A24 表示 *AcMYBs* 分组; A~F 表示 *AcMYBs* 6 种不同的表达模式

a: Phylogenetic tree of *AcMYBs*, b: the gene structure *AcMYBs*, c: conserved motif distribution of *AcMYBs*,

d: expression analyses of *AcMYBs*. A1-A24 represent the grouping of *AcMYBs*,

A-F represent six different expression patterns of *AcMYBs*

图 3 *AcMYBs* 系统发育树、基因结构、蛋白基序和表达分析

Fig.3 The phylogenetic tree, gene structure, motif and expression analyses of *AcMYBs*

2.5 *AcMYBs* 表达模式分析

为探究 *AcMYBs* 的表达模式,利用 3 个发育时期的小枝转录组数据对 100 个 *AcMYBs* 进行了表达分析(图 3d)。结果显示,100 个基因被分成 6 组不同的表达模式,聚集在同一组中的基因表达模式相同。A 组中 15 个基因都在 T6 时期呈现出最高的

表达量;B 组的基因在 T1 时期几乎不表达,但大部分基因都伴随着阳桃小枝的发育,表达量逐渐增加,少数基因在 T3 时期的表达量高于 T6;C 组基因在 T3 时期表达量最高,但在 T1 和 T6 时期表达量较弱;D 组基因在 T1 时期的表达量最高,其余 2 个时期几乎不表达;E 组中的基因在 3 个时期都不表达;

剩下 20 个基因都在 F 组, 和 D 组一样, F 组的基因在 T1 时期表达量最高, 但伴随着小枝的发育表达量逐渐降低。

2.6 AcMYBs 蛋白互作网络预测

以南芥蛋白数据库作为参考, 对在 T1、T3 和 T6 时期具有差异表达的 AcMYB41 (AtMYB20)、AcMYB84 (AtMYB80)、AcMYB100 (AtMYB69) 蛋

白互作关系进行分析(图 4)。结果显示 AcMYB41 (AtMYB20) 与 WRKY2、NAC066、KNAT7、NST1、NAC073 等蛋白发生互作, AcMYB84 (AtMYB80) 与 MS1、AMS、NAC010、UND 等蛋白发生互作, AcMYB100 (AtMYB69) 与 MYB56、HB18、NAC012、AT3G27220 等蛋白发生互作, 故推测这些基因在阳桃生长发育过程中可能具有功能的相关性。

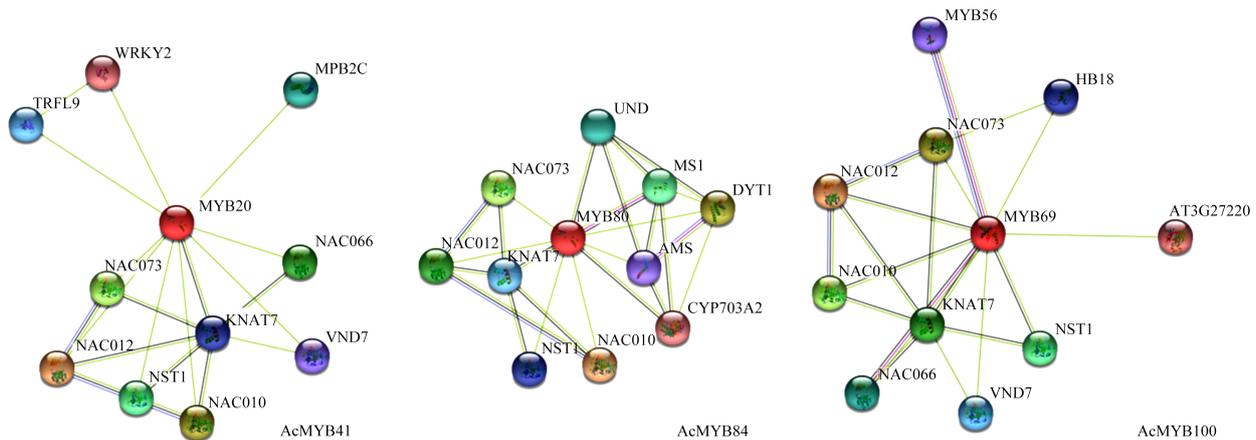


图 4 AcMYB 家族成员蛋白互作网络

Fig.4 Interaction network of AcMYB family member proteins

2.7 AcMYBs 染色体定位分析及共线性分析

为确定 AcMYBs 的遗传差异和基因复制关系, 根据基因位置信息, 得到 AcMYBs 在染色体上的分布情况(图 5)。100 个 AcMYBs 不均匀地分布在 11 条染色体上面, AcMYBs 数量较多的染色体为 1 号染色体和 7 号染色体, 分别有 18 个和 13 个基因, 3 号染色体上的成员数最少, 只有 2 个基因。

通过对阳桃全基因组的分析, AcMYBs 共包含 9 对片段复制事件, 包括染色体之间的 8 次复制事件以及同一染色体内的 1 次复制事件 (AcMYB17 和 AcMYB25)。同时, 多个 AcMYBs 是通过串联复制产生的, 7 号染色体上有最多的串联复制事件, 且这些串联复制事件发生在 7 号染色体上的同一区域。串联复制的基因 AcMYB07 和 AcMYB08、AcMYB39 和 AcMYB40、AcMYB88 和 AcMYB89 之间的基因结构模式和基序几乎完全相同, 并且它们对应的序列也非常相似。

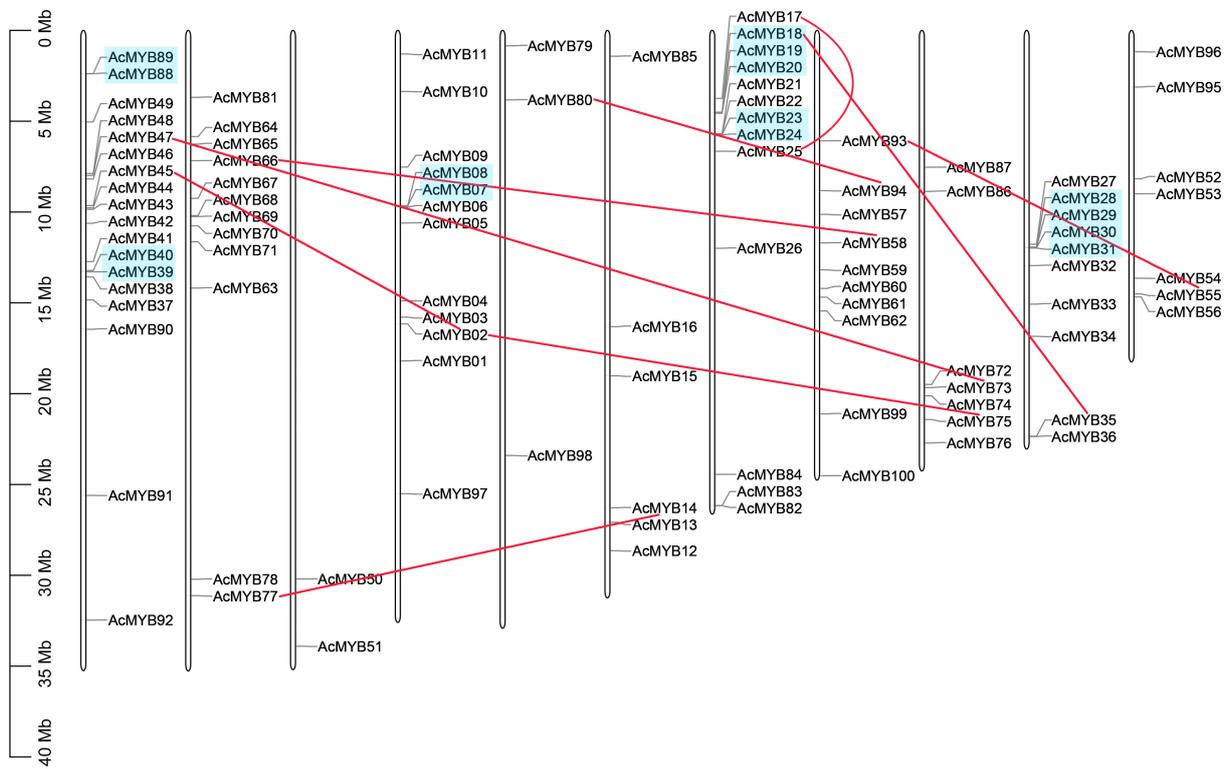
AcMYBs 重复基因的 Ka/Ks 结果表明, 串联重复的 Ka/Ks 在 0.158625744~0.529598234 之间, 平均值为 0.31350837975, 片段重复的 Ka/Ks 在 0.159839867~0.366825458 之间, 平均值为 0.253675666 (表 3)。值得注意的是 AcMYB02-AcMYB45、AcMYB02-AcMYB75 和 AcMYB39-AcMYB40 没有 Ks 值 (NaN),

即其核苷酸序列差异性不大。

为进一步研究 AcMYBs 的进化机制, 结合拟南芥和毛果杨进行了共线性分析(图 6)。结果表明, 阳桃与毛果杨及拟南芥之间都具有较强的共线性关系, 大部分的基因都呈现出一对一的关系, 但 1 号染色体上的 AcMYB91 和 AcMYB92、5 号染色体上的 AcMYB79、6 号染色体上的 AcMYB85 等部分基因只和毛果杨发生共线性关系, 且他们对应多条毛果杨序列。

2.8 RT-qPCR 验证 AcMYBs 与木质素合成的关系

结合系统发育树和表达分析, 选择 8 个与拟南芥木质素合成相关的同源基因 AcMYBs, 通过 RT-qPCR 的方法验证其在 T1、T3、T6、果和叶中的表达(图 7)。结果显示, 除 AcMYB05 外, 7 个 AcMYBs 都存在相似的表达模式。AcMYB65 在 T3 中表达量最高, T6 较 T3 有所下降, 但 T6 的表达还是明显高于 T1, 在果和叶中表达量最低, 而 AcMYB05 在 T1 及果中呈现出较高的表达, T3 和 T6 时期表达较弱。此外, AcMYB65、AcMYB84、AcMYB87、AcMYB97 和 AcMYB100 呈现出明显的组织特异性, 在枝中呈现出较高的表达量, 但在果和叶中几乎不表达。值得注意的是, AcMYB100 在 T3 中的表达约为 T6 时期的 7 倍, AcMYB92 在 T6 时期的表达水平相比于 T3



红线: 片段复制, 蓝框: 串联重复, 下同; 从左至右依次为 Chr.1~Chr.11
 Red lines: segmental duplicated, BLue boxes: tandemly duplicated, the same as below,
 from left to right are Chr.1-Chr.11

图 5 *AcMYBs* 染色体定位
 Fig.5 Chromosomal location information of *AcMYBs*

表 3 *AcMYBs* 基因对的复制模型

Table 3 Duplication models for *AcMYBs* gene pairs

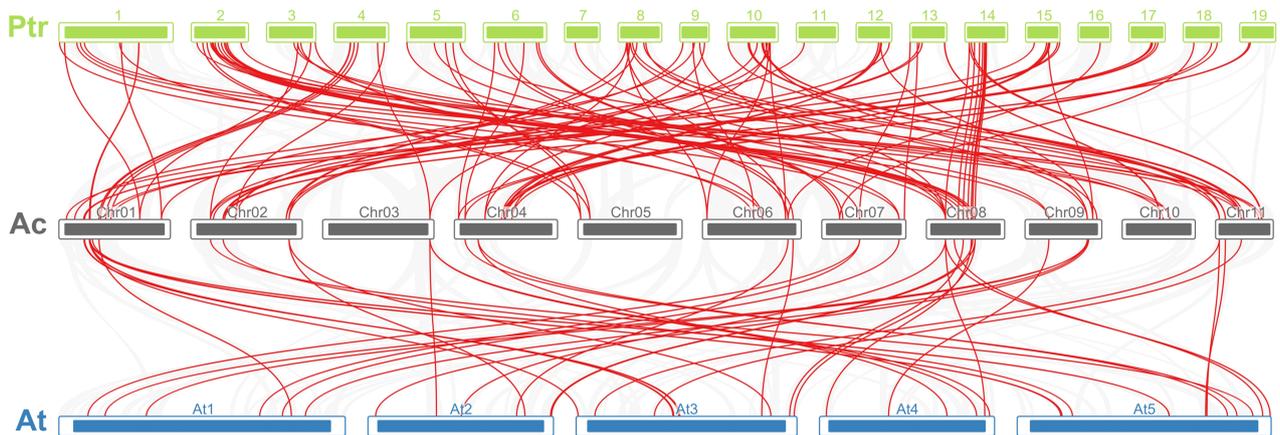
基因 1 Gene 1	基因 2 Gene 2	非同义替换率 Ka	同义替换率 Ks	非同义替换率 / 同义替换率 Ka/Ks	复制方式 Duplication
<i>AcMYB02</i>	<i>AcMYB45</i>	0.416547562	NaN	NaN	S
<i>AcMYB02</i>	<i>AcMYB75</i>	0.457660888	NaN	NaN	S
<i>AcMYB07</i>	<i>AcMYB08</i>	0.263486474	1.661057455	0.158625744	T
<i>AcMYB17</i>	<i>AcMYB25</i>	0.431433117	1.398763215	0.308438993	S
<i>AcMYB18</i>	<i>AcMYB19</i>	0.509306965	1.807123842	0.281832907	T
<i>AcMYB18</i>	<i>AcMYB35</i>	0.514092091	2.463635808	0.208672113	S
<i>AcMYB19</i>	<i>AcMYB20</i>	0.181496523	0.426366218	0.425682232	T
<i>AcMYB23</i>	<i>AcMYB24</i>	0.43481587	1.963398363	0.22146085	T
<i>AcMYB28</i>	<i>AcMYB29</i>	0.543456721	1.544204616	0.351933102	T
<i>AcMYB29</i>	<i>AcMYB30</i>	0.243746262	0.699756836	0.348329948	T
<i>AcMYB30</i>	<i>AcMYB31</i>	0.072503586	0.136902997	0.529598234	T
<i>AcMYB39</i>	<i>AcMYB40</i>	0.486235577	NaN	NaN	T
<i>AcMYB47</i>	<i>AcMYB73</i>	0.46900875	1.278561069	0.366825458	S

表 3(续)

基因 1 Gene 1	基因 2 Gene 2	非同义替换率 Ka	同义替换率 Ks	非同义替换率 / 同义替换率 Ka/Ks	复制方式 Duplication
<i>AcMYB66</i>	<i>AcMYB58</i>	0.370458326	1.342171612	0.276014127	S
<i>AcMYB77</i>	<i>AcMYB14</i>	0.387722005	1.483127575	0.261421884	S
<i>AcMYB80</i>	<i>AcMYB94</i>	0.29208288	1.501578531	0.194517219	S
<i>AcMYB88</i>	<i>AcMYB89</i>	0.295393346	1.549774994	0.190604021	T
<i>AcMYB93</i>	<i>AcMYB55</i>	0.342360176	2.141894779	0.159839867	S

S: 片段重复; T: 串联复制; NaN: 无数据; 表中基因按基因名顺序排列

S: segmental duplicated, T: tandem duplication, NaN: No number, the genes in the table are arranged in order of gene name

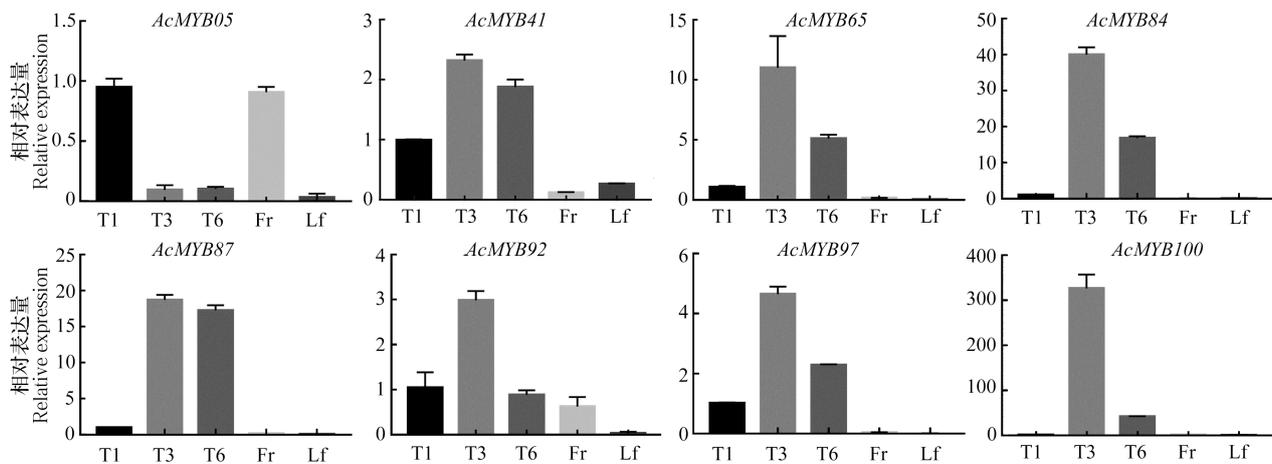


Ptr: 毛果杨; Ac: 阳桃; At: 拟南芥; 灰色背景线代表全基因组中的共线性区块; 红色高亮线条代表阳桃 R2R3-MYB 基因在 3 个物种间的共线性关系; 数字代表物种中染色体名称

Ptr: *Populus trichocarpa* Torr. & A. Gray, Ac: *Averrhoa carambola* L., At: *Arabidopsis thaliana* (L.) Heynh, the gray background line representing the collinear block in the whole genome, the red highlighted line representing the collinear relationship between the *AcMYBs* in the three species, the number represents the name of the chromosome in the species

图 6 阳桃、拟南芥和毛果杨中 R2R3-MYB 基因的共线性分析图

Fig.6 Synteny analyses of R2R3-MYB genes between *Arabidopsis thaliana* (L.), *Averrhoa carambola* L. and *Populus trichocarpa* Torr. & A. Gray



T1、T3、T6: 转录组取样的 3 个时期; Fr: 果; Lf: 叶

T1, T3, T6: sample collection for transcriptome, Fr: fruit, Lf: leaf

图 7 8 个 *AcMYBs* 在不同组织的表达

Fig.7 Expression of eight *AcMYBs* in different tissues

下调约3倍,相比之下 *AcMYB41* 和 *AcMYB87* 在 T3 和 T6 时期的表达水平没有显著的变化。RT-qPCR 在小枝 3 个时期中的表达与转录组数据一致,验证了转录组测序的可靠性。

3 讨论

MYB 基因家族作为植物中最大的转录因子家族之一,已经被证明在植物次生代谢、细胞分化、生物和非生物胁迫、花青苷的合成等方面具有重要的作用^[1]。在猕猴桃 (*Actinidia deliciosa* (A. Chev.) C. F. Liang & A. R. Ferguson) 中, *AdMYB7* 通过对代谢途径基因的转录激活,以调节类胡萝卜素和叶绿素色素的积累^[24];烟草 *PbrMYB5* 在增强低温胁迫的耐受性方面发挥着积极作用^[25];拟南芥中 *AtMYB60* 和 *AtMYB96* 是 ABA 信号传导、气孔运动、干旱胁迫和抗病性所必需的基因^[26]。本研究利用阳桃全基因组数据,共鉴定出 57 个 1R-MYB、100 个 R2R3-MYB 和 4 个 3R-MYB。 *AcMYBs* 的数量低于大豆 (244 个)^[14]、亚麻 (*Linum usitatissimum* L., 167 个)^[27]、毛果杨 (192 个)^[15]、水稻 (102 个)^[28],但高于毛梗李 (*Prunus salicina* Lindl., 96 个)^[29]、菠萝 (*Ananas comosus* L., 94 个)^[30]、小麦 (*Triticum aestivum* L., 22 个)^[31]。表明不同物种中的 R2R3-MYB 基因数量存在较大的差异性,但各植物中 R2R3-MYB 转录因子都是 MYB 基因家族中最主要的类型。R2R3-MYB 具有非常保守的结构,其 N 末端为保守的 DNA 结合域,高度可变的 C 末端为激活或抑制结构域^[4]。与大部分植物一样, *AcMYBs* 也同样都包含 R2R3 结构域所具有代表性的色氨酸 (W) 残基,这些保守的色氨酸残基是植物 MYB 基因家族的标志,在序列特异性 DNA 结合中具有关键性的作用。阳桃 R2 结构域中 3 个色氨酸残基都高度保守,但 R3 结构域中第 1 个色氨酸 (W) 残基被 F 所替换,色氨酸残基的取代能识别新的靶基因或者导致针对靶基因的 DNA 连接活性丧失^[5]。

阳桃与拟南芥的系统进化树显示,大部分 *AcMYBs* 都与拟南芥处于同一分支,表明阳桃和拟南芥 R2R3-MYB 成员具有相近的进化起源,但一些分支中缺少拟南芥或阳桃 MYB 成员,证明这两个物种的 R2R3-MYB 基因可能存在多样性。阳桃自身的系统发育树分析可以将 100 个 *AcMYBs* 分为 24 个亚组,同一亚组中的基因通常表现出相同的内含子、外显子模式,包括内含子的位置和数量,这些

结果表明阳桃内含子模式不是随机的,而是高度保守的。与前人的研究一致,大多数 *AcMYBs* 都不超过 2 个内含子,但 *AcMYB74* 和 *AtMYB124* 却拥有 2 个物种中最为复杂的外显子和内含子结构 (12 个外显子, 11 个内含子),经证明 *AcMYB74* 是 *AtMYB124* 的同源基因,这表明同源基因不仅具有相似的内含子结构,且内含子在进化过程中具有高度保守性^[14]。已有研究表明,剪切调控序列的密度随着内含子长度的增加 (< 1.5 kb) 而增加,内含子长度的增加 (> 1.5 kb) 与剪切位点强度的增加呈正相关^[32],因此, *AcMYBs* 中各种内含子的模式可能反映其表达或剪切调控的进化保守性。此外,进化保守基因具有更长的内含子长度,并且进化保守水平和真核基因内含子长度之间呈现正相关^[33],在阳桃中, *AcMYB39*、*AcMYB40*、*AcMYB77* 和 *AcMYB93* 较其他基因具有更长的内含子长度,表明其在进化过程中可能更保守,相较于其他基因,他们可能参与更重要的生物学功能。保守基序分析显示,相同亚组中的 *AcMYBs* 成员通常具有相似的基序组成,并且在不同亚组中观察到显著差异,这表明特定亚组中的蛋白质结构是保守的。总的来说,同一亚组中相似的基序组成、基因结构有力地支持了亚组分类的可靠性。

由于植物幼嫩的组织木质素含量低,故将木质化高的组织与幼嫩组织进行比较是鉴定木质素生物合成相关基因的理想方法^[27]。因此,本研究选取阳桃幼嫩的小枝 T1、木质化程度高的小枝 T6 及中间小枝 T3 作为转录组的样品,以鉴定木质素生物合成相关的 MYB 基因在阳桃木质化过程中的表达情况。表达量热图显示,一些基因在小枝的不同时期呈现出表达特异性, *AcMYB05*、*AcMYB79* 和 *AcMYB89* 在 T3 时期表达最高,在 T1 和 T6 时期几乎不表达; *AcMYB41*、*AcMYB65*、*AcMYB92*、*AcMYB97* 和 *AcMYB100* 在 3 个时期呈现出较为明显的表达差异,在 T1 时期几乎不表达,伴随着阳桃枝的发育表达量逐渐增加,由于阳桃 T6 和 T3 时期样品的木质化程度远高于 T1 时期,故猜测这些基因可能参与阳桃小枝的生长发育和木质素生物合成过程。

基因复制事件一直以来都发生在植物进化的整个过程中,片段重复和串联重复是基因家族扩展的主要原因,在单个簇中一些密切相关的基因在给定的染色体上彼此靠近,不超过一个插入的注释基因即串联重复,而不同染色体上或同一条

染色体不相邻的基因复制, 被称为片段重复事件^[21, 34]。Du 等^[14]在大豆中共检测到 92 对旁系同源基因和 13 个串联重复的基因, 这些基因占大豆 R2R3-MYB (244) 数量的 80.7%; 毛果杨的 19 条染色体上共发现 80 对片段重复产生的基因, 其中 30 对基因仍然处于 13 个重复区块保守的位置, 表明他们可能是全基因组复制的结果^[35]。在本研究中, 100 个 *AcMYBs* 中共鉴定出 31 个复制产生的基因, 其中 16 个涉及片段重复, 14 个涉及串联重复, 而 *AcMYB18* 既涉及片段重复又涉及串联重复, 表明在阳桃基因组中片段重复和串联重复对 *AcMYB* 转录因子的扩展具有重要的影响。异义 (Ka) 和同义 (Ks) 的替代率是评价重复事件正向选择压力的基础, Ka/Ks 值为 1 表示中性选择, 而 Ka/Ks < 1 表示纯化选择, Ka/Ks > 1 表示正向选择^[36]。对 *AcMYBs* 重复基因进行选择压力分析显示所有重复基因的 Ka/Ks 值都小于 1, 说明这些基因都是在纯化选择的作用下进化的, 但 *AcMYB02-AcMYB45*、*AcMYB02-AcMYB75* 和 *AcMYB39-AcMYB40* 没有 Ks 值, 说明基因复制可能导致了核酸水平的突变, 但氨基酸序列保持不变^[37]。种间共线性分析显示阳桃与毛果杨及拟南芥之间都存在大量的共线性基因, 表明阳桃和毛果杨及拟南芥之间都存在相似的起源, 共线性的结果为研究 MYB 家系在不同物种间的进化提供了新的理论依据。

大量研究表明, MYB 转录因子是调控木质素生物合成的关键因子之一。Omer 等^[38]将银合欢 (*Leucaena leucocephala* (Lam.) de Wit) 中的 *LlMYB1* 基因导入烟草中进行过表达分析, 结果烟草中的木质素含量明显降低; 巨桉 (*Eucalyptus grandis* W. Mill ex Maiden) 中 *AtMYB46* 的同源基因 *EgMYB2* 能与位于木质素合成酶启动子上的 AC 原件结合, 进而调控木质素的合成^[39]; Zhao 等^[9]在水稻中敲除 *OsMYB46*, 导致其细胞壁成分 (木质素、纤维素) 含量降低, 且 SCW 生物合成基因表达下调。拟南芥 *AtMYB85* 及 S21 亚组被证明与木质素的合成具有重要的联系^[40], *AtMYB85* 同源基因 *ZmMYB167* 的表达明显影响玉米 (*Zea mays* L.) 的生长、组织木质化及木质素单体 S/G 的比值^[41]。在阳桃中, RT-qPCR 结果显示 *AtMYB85* 的同源基因 *AcMYB41*、*AcMYB65*、*AcMYB87* 以及 S21 亚组中的 *AcMYB97*、*AcMYB100* 在小枝中呈现出明显的表达差异, 其在木质化程度高的 T3 和 T6 时期的表达远

高于 T1, 但在果和叶中几乎不表达, 由于基因的表达模式与其功能作用具有密切的相关性^[42], 表明其可能是阳桃木质素生物合成过程中的特异性基因。*AtMYB103* 的同源基因 *AcMYB84* 及 *AtMYB83*、*AtMYB46* 的同源基因 *AcMYB92* 在 T3 中具有最高的表达量, *AcMYB92* 在 T3 中的表达量约为 T6 的 3 倍。*AtMYB46* 能激活拟南芥 SCW, 调节木质素的沉积^[13, 43], 且具有功能保守性, 水稻 *OsMYB46*、玉米 *ZmMYB46*、柳枝稷 (*Panicum virgatum* L.) *PvMYB46* 都被证明具有相似的功能^[44], 故推测 *AcMYB92* 可能也具有激活阳桃 SCW, 促进木质素合成的功能。*AtMYB63* 和 *AtMYB58* 是木质素生物合成的活化剂^[11], 与其他的基因不同, 在阳桃中 *AcMYB05* 在 T1 和果中的表达量相对较高, 在 T3 和 T6 时期反而表达较弱, 推测 *AcMYB05* 发生了亚功能化和新功能化。

参考文献

- [1] Schuetz M, Benske A, Smith R A, Watanabe Y, Tobimatsu Y, Ralph J, Demura T, Ellis B, Samuels A L. Laccases direct lignification in the discrete secondary cell wall domains of protoxylem. *Plant Physiology*, 2014, 166 (2): 798-807
- [2] Hua C W, Goué N, Saidi M N, Legay S, Sivadon P, Goffner D, Grima-Pettenati J. Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. *Plant Science*, 2013, 4: 189
- [3] Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology*, 2001, 4 (5): 447-456
- [4] Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. *Trends in Plant Science*, 2010, 15 (10): 573-581
- [5] Ogata K, Kanei-Ishii C, Sasaki M, Hatanaka H, Nagadoi A, Enari M, Nakamura H, Nishimura Y, Ishii S, Sarai A. The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and *trans*-activation. *Nature Structural & Molecular Biology*, 1996, 3 (2): 178-187
- [6] Chen T Z, Li W J, Hu X H, Guo J R, Liu A M, Zhang B L. A cotton MYB transcription factor, *GbMYB5*, is positively involved in plant adaptive response to drought stress. *Plant and Cell Physiology*, 2015, 56 (5): 917-929
- [7] Zhu N, Cheng S F, Liu X Y, Du H, Dai M Q, Zhou D X, Yang W J, Zhao Y. The R2R3-type MYB gene *OsMYB91* has a function in coordinating plant growth and salt stress tolerance in rice. *Plant Science*, 2015, 236: 146-156
- [8] Ma Q, Dai X Y, Xu Y Y, Guo J, Liu Y J, Chen N, Xiao J, Zhang D J, Xu Z L, Zhang X S, Chong K. Enhanced tolerance to chilling stress in *OsMYB3R-2* transgenic rice is mediated by alteration in cell cycle and ectopic expression of stress genes. *Plant Physiology*, 2009, 150 (1): 244-256
- [9] Zhao K, Bartley L E. Comparative genomic analysis of the R2R3 MYB secondary cell wall regulators of *Arabidopsis*,

- poplar, rice, maize, and switchgrass. *BMC Plant Biology*, 2014, 14(1): 135
- [10] Zhong R, Ye Z H. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant and Cell Physiology*, 2015, 56(2): 195-214
- [11] Zhou J, Lee C, Zhong R, Ye Z H. *MYB58* and *MYB63* are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *The Plant Cell*, 2009, 21(1): 248-266
- [12] McCarthy R L, Zhong R, Ye Z H. *MYB83* is a direct target of *SND1* and acts re-dundantly with *MYB46* in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant and Cell Physiology*, 2009, 50(11): 1950-1964
- [13] Ko J H, Jeon H W, Kim W C, Kim J Y, Han K H. The *MYB46/ MYB83*-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Annals of Botany*, 2014, 114(6): 1099-1107
- [14] Du H, Yang S S, Liang Z, Feng B R, Liu L, Huang Y B, Tang Y X. Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biology*, 2012, 12(1): 106
- [15] Wilkins O, Nahal H, Foong J, Provart N J, Campbell M M. Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant Physiology*, 2009, 149(2): 981-993
- [16] Wu S S, Sun W, Xu Z C, Zhai J W, Li X P, Li C R, Zhang D Y, Wu X Q, Shen L M, Chen J H, Ren H, Dai X Y, Dai Z W, Zhao Y M, Chen L, Cao M X, Xie X Y, Liu X D, Peng D H, Dong J W, Hsiao Y Y, Chen S L, Tsai W C, Lan S R, Liu Z J. The genome sequence of star fruit (*Averrhoa carambola*). *Horticulture Research*, 2020, 7(1): 95
- [17] Fan Y N, Sahu S K, Yang T, Mu W X, Wei J P, Cheng L, Yang J L, Mu R C, Liu J, Zhao J M, Zhao Y X, Xu X, Liu X, Liu H. Dissecting the genome of star fruit (*Averrhoa carambola* L.). *Horticulture Research*, 2020, 7(94): 1-10
- [18] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 1994, 22(22): 4673-4680
- [19] Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics*, 2008, 9(4): 299-306
- [20] Chen C J, Chen H, Zhang Y, Thomas H, Frank M H, He Y H, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, 2020, 13(8): 1194-1202
- [21] Cannon S B, Mitra A, Baumgarten A, Young N D, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biology*, 2004, 4: 10
- [22] Wang Y, Tang H, deBarry J D, Tan X, Li J, Wang X, Tae-ho L, Jin H, Marler B, Guo H. MCSanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, 2012, 40(7): e49
- [23] Li X L, Xue C, Li J M, Qiao X, Li L T, Yu L A, Huang Y H, Wu J. Genome-wide identification, evolution and functional divergence of MYB transcription factors in Chinese white pear (*Pyrus bretschneideri*). *Plant and Cell Physiology*, 2016, 57(4): 824-847
- [24] Ampomah-Dwamena C, Thrimawithana A H, Dejnopratt S, Lewis D, Espley R V. A kiwifruit (*Actinidia deliciosa*) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation. *New Phytologist*, 2019, 221(1): 309-325
- [25] Xing C H, Liu Y, Zhao L Y, Zhang S L, Huang X S. A novel MYB transcription factor regulates ascorbic acid synthesis and affects cold tolerance. *Plant Cell and Environment*, 2019, 42(3): 832-845
- [26] Seo P J, Park C M. *MYB96*-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in *Arabidopsis*. *New Phytologist*, 2010, 186(2): 471-483
- [27] Tombuloglu H. Genome-wide identification and expression analysis of R2R3, 3R- and 4R-MYB transcription factors during lignin biosynthesis in flax (*Linum usitatissimum*). *Genomics*, 2019, 112(1): 782-795
- [28] Chen Y H, Yang X Y, He K, Liu M H, Li J G, Gao Z F, Lin Z Q, Zhang Y F, Wang X X, Qiu X M, Shen Y P, Zhang L, Deng X H, Luo J C, Deng X W, Chen Z L, Gu H Y, Qu L J. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology*, 2006, 60(1): 107-124
- [29] Liu C Y, Hao J J, Qiu M Q, Pan J J, He Y H. Genome-wide identification and expression analysis of the MYB transcription factor in Japanese plum (*Prunus salicina*). *Genomics*, 2020, 112(6): 4875-4886
- [30] Liu C Y, Xie T, Chen C J, Luan A P, Long J M, Li C H, Ding Y Q, He Y H. Genome-wide organization and expression profiling of the R2R3-MYB transcription factor family in pineapple (*Ananas comosus*). *BMC Genomics*, 2017, 18(1): 503
- [31] Zhang L C, Zhao G Y, Jia J Z, Liu X, Kong X Y. Molecular characterization of 60 isolated wheat MYB genes and analysis of their expression during abiotic stress. *Journal of Experimental Botany*, 2012, 63(1): 203-214
- [32] Dewey C N, Rogozin I B, Koonin E V. Compensatory relationship between splice sites and exonic splicing signals depending on the length of vertebrate introns. *BMC Genomics*, 2006, 7(1): 311
- [33] Gorlova O Y, Fedorov A, Logothetis C, Amos C I, Gorlov I. Genes with a large intronic burden show greater evolutionary conservation on the protein level. *BMC Evolutionary Biology*, 2014, 14(1): 50
- [34] Kent W J, Baertsch R, Hinrichs A, Miller W, Haussler D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 2003, 100(20): 11484-11489
- [35] Tuskan G A, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 2006, 313(5793): 1596-1604
- [36] Yang Z H. PAML 4: phylogenetic analysis by maximum

- likelihood. *Molecular Biology and Evolution*, 2007, 24(8): 1586-1591
- [37] Wang J, Liu Y, Tang B Q, Dai X Z, Xie L L, Liu F, Zou X X. Genome-wide identification and capsaicinoid biosynthesis-related expression analysis of the R2R3-MYB gene family in *Capsicum annuum* L. *Frontiers in Genetics*, 2020, 11: 1678
- [38] Omer S, Kumar S, Khan B M. Over-expression of a subgroup 4 R2R3 type MYB transcription factor gene from *Leucaena leucocephala* reduces lignin content in transgenic tobacco. *Plant Cell Reports*, 2013, 32(1): 161-171
- [39] Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Grima-Pettenati J. *EgMYB2*, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant Journal*, 2005, 43(4): 553-567
- [40] Zhong R, Lee C, Zhou J, McCarthy R L, Ye Z H. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *The Plant Cell*, 2008, 20(10): 2763-2782
- [41] Bhatia R, Dalton S, Roberts L A, Moron-Garcia O M, Iacono R, Kosik O, Gallagher J A, Bosch M. Modified expression of *ZmMYB167* in *Brachypodium distachyon* and *Zea mays* leads to increased cell wall lignin and phenolic content. *Scientific Reports*, 2019, 9(1): 8800
- [42] 刘婷婷, 李晓曼, 张晓辉, 王海平, 阳文龙, 宋江萍, 贾会霞, 刘同金, 李锡香. 萝卜全基因组中 SPL 基因家族成员的鉴定与分析. *植物遗传资源学报*, 2021, 22(4): 1145-1156
- Liu T T, Li X M, Zhang X H, Wang H P, Yang W L, Song J P, Jia H X, Liu T J, Li X X. Genome-wide identification and analysis of SPL gene family in radish. *Journal of Plant Genetic Resources*, 2021, 22(4): 1145-1156
- [43] Zhong R, Richardson E A, Ye Z H. The *MYB46* transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. *The Plant Cell*, 2007, 19(9): 2776-2792
- [44] Zhong R, Yuan Y, Spiekerman J J, Guley J T, Egbosiuba J C, Ye Z H. Functional characterization of NAC and MYB transcription factors involved in regulation of biomass production in switchgrass (*Panicum virgatum*). *PLoS ONE*, 2015, 10(8): e0134611