

• 药材与资源 •

七叶树 R2R3-MYB 转录因子基因家族鉴定及表达分析

解雪华¹, 秦 晨^{3#}, 管崧淳⁴, 孟祥霄¹, 曹 雪¹, 陈伟强¹, 王飞飞², 万会花^{1*}, 孙 伟^{1*}

1. 中国中医科学院中药研究所 道地药材品质保障与资源持续利用全国重点实验室, 北京 100700

2. 云南特色植物提取实验室有限公司, 云南 昆明 650106

3. 黑龙江中医药大学药学院, 黑龙江 哈尔滨 150040

4. 北京林业大学园林学院, 北京 100083

摘要: 目的 系统解析七叶树 *Aesculus chinensis* R2R3-MYB 转录因子家族成员信息, 明确其不同组织器官中的表达模式, 为该类基因在七叶树中生物学功能的研究提供参考。方法 利用生物信息学方法对七叶树的 R2R3-MYB 转录因子家族成员进行鉴定, 并对其进化关系、基因结构、理化性质、染色体定位和基因表达模式进行分析。结果 在七叶树全基因组中共鉴定出 129 条 R2R3-MYB 基因, 分布在 20 条染色体上, 进化分析结果显示 R2R3-MYB 家族基因被聚类到 34 个亚组中, 其中 S15 亚组分布的七叶树 R2R3-MYB 成员最多为 9 个; 共线性基因对在进化树的分支距离更近, 且遗传距离较近的基因具有类似的内含子-外显子结构; 在七叶树 R2R3-MYB 基因家族成员中共鉴定出 53 种不同类型的顺式作用元件; 对七叶树不同组织部位的转录组数据分析, 结合系统进化分析及七叶皂苷合成关键酶转录因子结合位点预测, 推测 *AcMYB043*、*AcMYB059*、*AcMYB022*、*AcMYB012*、*AcMYB075*、*AcMYB111* 和 *AcMYB125* 基因可能参与七叶皂苷生物合成的调控。结论 通过对七叶树全基因组 129 个七叶树 R2R3-MYB 基因进行生物信息学分析, 筛选出 7 个可能参与七叶皂苷生物合成调控的候选基因, 为后续解析七叶皂苷转录调控机制研究奠定分子基础, 对七叶皂苷的合成生物学研究和高产七叶皂苷七叶树的新品种培育具有重要指导意义。

关键词: 转录因子; 七叶树; 进化分析; 表达模式; 七叶皂苷

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2025)20-7512-14

DOI: 10.7501/j.issn.0253-2670.2025.20.021

Identification and expression analysis of R2R3-MYB transcription factor gene family in *Aesculus chinensis*

XIE Xuehua¹, QIN Chen³, GUAN Songchun⁴, MENG Xiangxiao¹, CAO Xue¹, CHEN Weiqiang¹, WANG Feifei², WAN Huihua¹, SUN Wei¹

1. State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-di Herbs, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

2. Yunnan Characteristic Plant Extraction Laboratory, Kunming 650106, China

3. College of Pharmacy, Heilongjiang University of Chinese Medicine, Harbin 150040, China

4. School of Landscape Architecture, Beijing Forestry University, Beijing 100083, China

Abstract: Objective To systematically analyze the members of the R2R3-MYB transcription factor family in Qiyeshu (*Aesculus chinensis*) and their expression patterns in different tissues, Which provides a reference for the study of the biological function of R2R3-MYB transcription factors in *A. chinensis*. **Methods** In this study, bioinformatics methods were used to identify the R2R3-MYB transcription factor family members of *A. chinensis*, and their evolutionary relationships, gene structure, physicochemical

收稿日期: 2025-06-13

基金项目: 国家自然科学基金项目 (82304673); 中国中医科学院科技创新工程项目 (CI2023E002); 云南特色植物提取实验室开放研究项目基金资助 (YKKF2024020)

作者简介: 解雪华, 博士, 研究方向为药用植物生物工程。E-mail: xie513812@163.com

#共同第一作者: 秦 晨, 硕士, 研究方向为药用植物生物工程。E-mail: qc18193735932@163.com

*通信作者: 孙 伟, 研究员, 研究方向为中药资源学。E-mail: wsun@icmm.ac.cn

万会花, 副研究员, 研究方向为中药资源与天然产物化学。E-mail: hhwan@icmm.ac.cn

properties, chromosomal localization and gene expression patterns were analyzed. **Results** The results showed that a total of 129 R2R3-MYB genes were identified in the whole genome of *A. chinensis*, distributed on 20 chromosomes. The results of evolutionary analysis showed that the R2R3-MYB family genes of *A. chinensis* were clustered into 34 subgroups, and the maximum number of R2R3-MYB in the S15 subgroup was nine. The collinearity gene pairs were closer on the branches in the phylogenetic tree and the genes with closer genetic distance have similar intron-exon structures. The promoter region of each member contains numerous cis-acting elements. A total of 53 different types of cis-acting elements were identified in the promoter regions of these R2R3-MYB family members. Combined with phylogenetic analysis and prediction of the transcription factors binding sites of key enzymes for escin synthesis, it is speculated that *AcMYB043*, *AcMYB059*, *AcMYB012*, *AcMYB022*, *AcMYB075*, *AcMYB111* and *AcMYB125* genes are likely to be involved in the regulation of escin biosynthesis. **Conclusion** Through bioinformatics analysis of 129 R2R3-MYB genes on the whole genome of horse chestnut, a total of seven candidate genes were screened. This study lays a molecular foundation for subsequent analysis of the transcriptional regulatory mechanism of escin, and is of great significance for the synthetic biology research of *A. chinensis* and the cultivation of new varieties with high-yield aescin.

Key words: transcription factor; *Aesculus chinensis* Bunge; evolutionary analysis; expression mode; aescin

七叶树 *Aesculus chinensis* Bunge 是七叶树科七叶树属落叶乔木，是一种重要的药用兼观赏用树种。七叶树属植物全世界有 30 余种，中国有 10 余种^[1]，其中七叶树、浙江七叶树、天师栗、欧洲七叶树等全国各地均有栽培^[2]。七叶树干燥成熟的种子被称为娑罗子，是《中国药典》2025 年版收录的常用中药，其主要药效成分七叶皂苷属于三萜皂苷类化合物^[3]，具有极高的药用价值，如抗水肿、抗肿瘤、对脑及脊髓缺血损伤的保护作用、对胃肠道的保护作用等，临床应用研究主要集中于治疗脑出血、创伤肢体肿胀等方面^[4-5]。市场上已有多种以七叶皂苷为活性成分的制剂应用于临床治疗，如“Reparil”（德国）、七叶皂苷素钠片（中国）等。七叶皂苷主要存在于种子娑罗子中，结构复杂，尚不能人工合成，原材料来源单一导致提取量逐年下降。七叶树种群数量较小，基因多样性较高，生长周期长，生态条件对七叶树的生长发育及有效成分的积累和成分组成影响显著^[6]，这些因素均导致七叶树种子质量均一性较差，影响其入药效果；为了扩大七叶皂苷原材料来源，推动七叶树植物资源的综合开发和应用，迫切需要解析七叶皂苷生物合成途径及调控机制。

MYB 转录因子是植物基因组中最大的转录因子家族之一，具有高度保守的 MYB 结构域^[7]，广泛参与细胞分化、次生代谢调控、环境胁迫以及病虫害的侵袭等过程^[8-12]。根据 MYB 转录因子包含的 R 结构的数量，可将其分为 4 种类型：1R-MYB（MYB 相关和 R3-MYB）、R2R3-MYB、3R-MYB（R1R2R3-MYB）和 4R-MYB。1R-MYB 主要参与植物的形态发生、次生代谢、生物钟控制及花和果实的发育等

生理过程^[13-14]，3R-MYB 在细胞周期控制、抗病等方面起到重要的调控作用^[15-16]，4R-MYB 是数量最少的一类 MYB 转录因子，植物中 4R-MYB 蛋白质功能的研究相对较少^[17]。R2R3-MYB 蛋白是存在最为广泛的一类 MYB 转录因子，许多植物编码超过 100 个 R2R3-MYB 转录因子，并在生长发育过程中表现出不同的功能^[18-19]。Liu 等^[20]通过对七叶树红色叶和绿色叶的转录组差异基因分析及瞬时表达发 *AcMYB113* 基因可通过激活七叶树叶片花青素生物合成途径促进花青素的合成。研究发现桦树 *Betula platyphylla* Sukaczew 中 *BpMYB21* 和 *BpMYB61* 参与了三萜途径的关键基因的表达，证实植物中 MYB 具备很大的调控三萜类化合物生产的潜力^[21]。因此，R2R3-MYB 转录因子的鉴定对七叶树植物生长发育及功效成分合成调控机制的解析至关重要，目前七叶树 R2R3-MYB 基因的鉴定和功能研究相对匮乏。

七叶树中已报道的七叶皂苷类化合物有 120 余种^[22-23]，然而有关七叶皂苷生物合成调控机制的研究还未见报道。植物三萜骨架在细胞色素 P450 单加氧酶（P450）的催化下进行位点特异性氧化，形成不同的三萜苷元^[24]，已有的研究发现调控植物三萜生物合成的转录因子多为 bHLH、AP2/ERF、bZIP 和 WRKY 4 种类型，MYB 调控三萜皂苷生物合成的报道相对较少，本实验室前期基于基因簇及共线性分析鉴定了七叶树七叶皂苷合成相关基因簇，并验证基因簇中 *AcOSC6*、*AcCYP716A278*、*AcCYP716A275* 及 *AcCSL1* 等基因参与了七叶皂苷的生物合成^[25]。进一步对七叶皂苷合成相关关键酶的转录因子结合位点进行预测分析，发现 MYB

转录因子结合位点在各个关键酶基因中数量显著高于其他转录因子结合位点,推测 MYB 转录因子可能在七叶皂苷的生物合成过程中发挥着重要的调控作用,MYB 转录因子的鉴定为揭示七叶皂苷生物合成的调控机制提供新的线索。本研究通过利用七叶树的高质量基因组,对七叶树 *R2R3-MYB* 基因家族成员进行了全基因组鉴定,并对 *R2R3-MYB* 基因的染色体定位、基因结构、系统发育关系、顺式元件等进行了详细分析。旨在为更深入研究七叶树 *R2R3-MYB* 转录因子的功能提供理论依据,为进一步探究七叶皂苷合成的调控机制奠定基础。

1 仪器与材料

本研究植物材料为六十五年生中华七叶树 *A. chinensis* Bunge,栽培于中国医学科学院药用植物研究所药用植物园内,经中国医学科学院药用植物研究所姚辉研究员和中国中医科学院中药研究所孙伟研究员鉴定,中国医学科学院药用植物研究所保留凭证。

2 方法

2.1 七叶树 *R2R3-MYB* 基因家族成员的鉴定

七叶树基因组数据和氨基酸序列来自 2023 年发布的七叶树全基因组数据;从 NCBI 数据库(<https://www.ncbi.nlm.nih.gov>)下载文冠果 *Xanthoceras sorbifolium* 全基因组序列和氨基酸序列信息及注释文件;从 TAIR 数据库(<https://www.arabidopsis.org/>)下载拟南芥 *Arabidopsis thaliana* L. 基因组序列和氨基酸序列信息及注释文件。使用 UniProt 在线数据库下载 *R2R3MYB-family* 保守结构域序列比对文件(PF00249),利用 HMMER 软件构建隐马尔可夫模型,在七叶树蛋白序列中对 *R2R3-MYB* 基因家族序列进行搜索;同时以拟南芥 *R2R3MYB* 蛋白作为种子序列,使用 BLAST 程序在七叶树基因组数据库中进行比对,合并获得序列并删除重复,初步获得候选 *R2R3-MYB* 转录因子蛋白序列。通过 CD-Search 工具(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)和 Pfam(<http://pfam.xfam.org>)数据库进一步筛选。获得了七叶树 *R2R3-MYB* 基因家族成员。并根据 *R2R3-MYB* 在染色体上的顺序位置进行定位,命名方式为 AcMYB 序号。

2.2 七叶树 *R2R3-MYB* 基因家族理化性质分析

使用 ExPASy 网站(<http://web.expasy.org/>)

protparam/)在线网站预测七叶树 *R2R3-MYB* 转录因子的理化性质,包括编码序列长度、相对分子量、理论等电点、不稳定指数、亲水性。使用 WOLF PSORT(<https://www.genscript.com/wolf-psort.html>)在线工具预测七叶树 *R2R3-MYB* 转录因子的亚细胞定位。

2.3 七叶树 *R2R3-MYB* 基因家族的染色体定位

从七叶树基因组 gff 文件中提取 *R2R3-MYB* 基因家族成员的染色体位置信息,利用 TBtools 软件可视化七叶树 *R2R3-MYB* 基因的染色体图谱,根据基因在染色体上的位置对七叶树 *R2R3-MYB* 基因进行命名。

2.4 七叶树 *R2R3-MYB* 基因家族的系统发育分析

利用 MEGA-X 软件对七叶树和拟南芥的 *R2R3-MYB* 氨基酸序列进行比对,并根据多序列比对结果进行系统发育分析;采用邻接法(neighbor-joining, NJ)构建系统发育树,bootstrap 重复次数为 1 000 次,使用 FigTree 软件对进化树进行美化。

2.5 七叶树 *R2R3-MYB* 基因结构和保守基序分析

进一步对七叶树 *R2R3-MYB* 基因的结构特征进行分析,利用 NCBI 在线工具 CD-search 预测 NAC 蛋白保守结构域;利用 MEME 在线工具(<https://meme-suite.org/meme/tools/meme>)对七叶树 *R2R3-MYB* 蛋白序列进行保守基序分析,motif 数量参数设置为 5,其他为默认参数^[26];利用 TBtools 软件从七叶树基因组 gff 文件提取七叶树 *R2R3-MYB* 基因家族成员的外显子和内含子等结构信息,并进行可视化处理。

2.6 七叶树 *R2R3-MYB* 基因家族共线性分析

使用 McScanX 软件对七叶树、文冠果和拟南芥 3 个物种的基因组进行共线性分析,解析 *R2R3-MYB* 家族成员在 33 个物种中的分化关系及 *R2R3-MYB* 家族成员的内部进化情况,利用 TBtools 工具进行可视化处理,并计算非同义替换率(K_a)、同义替换率(K_s)的比值 K_a/K_s 。

2.7 七叶树 *R2R3-MYB* 基因家族顺式作用元件分析

利用 perl 脚本根据七叶树基因组 gff 注释文件,从七叶树 *R2R3-MYB* 基因中提取上游 2 000 bp 序列,通过 PlantCARE 网站(<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)对启动子区域进行顺式作用元件分析,了解其潜在作用和机制。

2.8 七叶树 *R2R3-MYB* 基因家族的表达模式分析

七叶树不同组织转录组测序数据为本实验室所测（国家基因组科学数据中心登录号：CRA009093），原始转录组数据经 Trimmomatic-0.36 过滤低质量数据，利用 STAR_2.6.0a 软件比对到基因组。转录本丰度基于 HTSeq 软件计算，并将原始计数转化为 TPM（Transcripts Per Million）值^[27]，提取 *R2R3-MYB* 基因家族成员在七叶树叶片、花、果皮、种子、枝条的表达量。利用 Tbtools 软件绘制七叶树 *R2R3-MYB* 基因家族成员在不同组织部位的表达模式热图。

表 1 *R2R3-MYB* 基因 RT-qPCR 引物序列
Table 1 RT-qPCR primers sequences for *R2R3-MYB* genes

基因	基因 ID	正向引物（5'-3'）	反向引物（5'-3'）
<i>AcActin7</i>	AC6A172T37.1	CAAAGGCCAACAGAGAGAAGA	TGTACGACCACTGGCATAAAG
<i>AcMYB096</i>	AC12A47T159.1	GGAAATAGATGGTCGGCGATAG	TGGCTTTGCTTGATCCTCTT
<i>AcMYB095</i>	AC12C31T16.1	GAACTACTGGAACACCCATCTT	GTGGGTTTCTTCCTCCTTGT
<i>AcMYB121</i>	AC18C19T19.1	GCGGAATACAGGGAACAGATAG	TCTGCTCCAGAACTTGTAAG
<i>AcMYB022</i>	AC2P272T130.1	CAGAGCAGTTGATACGAGTGAG	CGTGTGAGAGTGATGGAGAAAG
<i>AcMYB129</i>	AC20A136T89.1	TTCGACAGAAATGGAGGAGAAA	CGGGAGCTTGTCGGATTAT
<i>AcMYB036</i>	AC3A151T63.1	CCTGAGGCCAGACCTTAAAC	CTGGCAATCTCCAGCTATCA
<i>AcMYB043</i>	AC4A203T102.1	TGGGAGAGTGCTAGACTTGA	GTTCCATATGCGGAGGAAGTAG
<i>AcMYB012</i>	AC1A295T186.1	GAGTAGTGTTGCCCTCTGATAAC	GTGCGTTCAAATGCCTCAAATA
<i>AcMYB061</i>	AC7A2T154.1	CTGGGAAGAGTTGCAGATTGA	ATCGAGAGTGGAGATCGAGAA
<i>AcMYB066</i>	AC7A207T92.1	CTCAAGATGCAGAGGTGAGAAA	TCGCCGTGGTTGTTGATATAG
<i>AcMYB079</i>	AC9A224T178.1	CTGCTGCTGCTTCATCTACTAC	CTCCATCTCCATCTCCATCTCT
<i>AcMYB012</i>	AC1A295T186.1	GGAAGCCAGAGGAGGATCTAATA	CCTCTCATCAACCCTGATTTCTC
<i>AcMYB079</i>	AC9A224T178.1	GGCAACAGCAGATGGTCTAA	TCTTCTTGATGTGGGTGTTCC

3 结果与分析

3.1 七叶树 *R2R3-MYB* 家族基因的鉴定，理化性质分析以及染色体定位分析

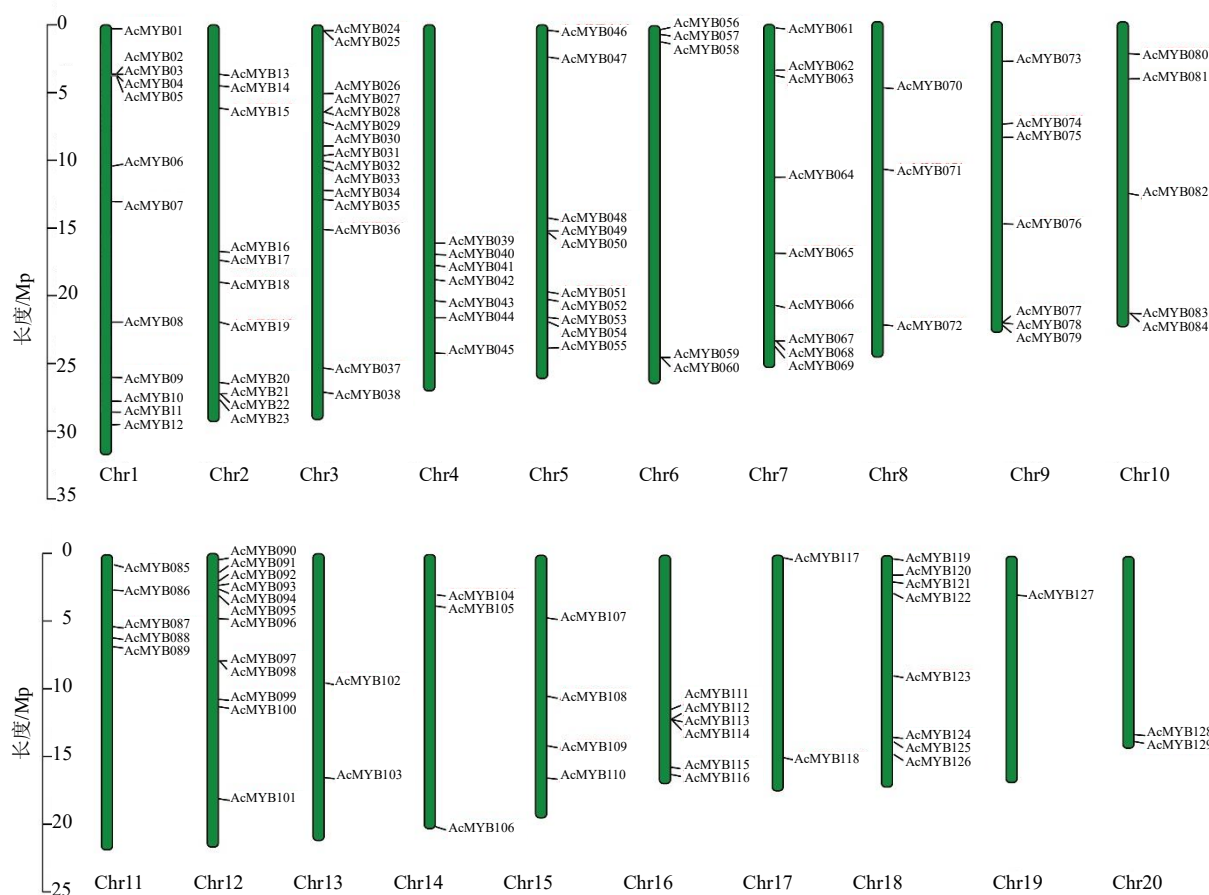
从七叶树基因组中共鉴定出 129 个 *R2R3-MYB* 基因家族成员，根据 *R2R3-MYB* 基因在染色体上的位置，命名为 *AcMYB001*~*AcMYB129*。分析表明七叶树 *R2R3-MYB* 家族成员的蛋白相对分子质量在 15 387.62~61 839.75，*AcMYB064* 编码蛋白氨基酸数目最多（561 个氨基酸），*AcMYB122* 编码蛋白氨基酸数目最少（129 个氨基酸）；理论等电点（PI）在 4.84~10.20，不稳定系数在 33.49~81.19，除 *AcMYB023*、*AcMYB054*、*AcMYB120* 属于稳定蛋白外，其余蛋白均为不稳定蛋白，推测可能和植物抗逆境能力相关；所有成员的亲水性系数均小于 0，在 -1.067~-0.437，均为亲水性蛋白。七叶树 *R2R3-MYB* 基因家族成员的亚细胞定位预测发现，129 个

2.9 实时荧光定量 PCR（real-time fluorescence quantitative PCR，RT-qPCR）检测

使用 FastPure Universal Plant Total RNA Isolation Kit 试剂盒（RC411-01，诺唯赞生物科技股份有限公司，南京）提取总 RNA；使用 Rever Tra Ace qPCR RT Mix with gDNA Remover（Toyobo 公司，日本）合成 cDNA；使用 Primer Premier 5.0 软件设计 RT-qPCR 引物（表 1），以七叶树 *Actin* 基因为内参；使用 ChamQ SYBR qPCR Master Mix（诺唯赞生物科技股份有限公司，南京）进行 RT-qPCR 反应，反应程序：95℃预变性 30s；95℃变性 5s，57℃退火 15s，72℃延伸 30s，40 个循环。每个 PCR 反应重复 3 次。采用 $2^{-\Delta\Delta C_t}$ 法对数据进行定量分析。

成员均被预测定位在细胞核中。

进一步对 129 个 *R2R3-MYB* 基因进行染色体分布研究，结果如图 1 所示。129 个 *R2R3-MYB* 基因不均等的分布在七叶树 20 条染色体上，3 号染色体（Chr3）上 *R2R3-MYB* 基因家族成员最多，共 15 个；19 号染色体（Chr19）上 *R2R3-MYB* 基因家族成员最少，仅 1 个；13 号、17 号和 20 号染色体上 *R2R3-MYB* 基因家族成员均为 2 个；其余染色体上至少鉴定到 3 个 *R2R3-MYB* 基因。由于七叶树的染色体数量较多，每条染色体分布的 *R2R3-MYB* 基因数量相对较低，大多数染色体的中心区域 *R2R3-MYB* 基因家族成员分布较少，部分 *R2R3-MYB* 家族基因在同一染色体上的相同区域集中分布，如 *AcMYB090*、*AcMYB091*、*AcMYB092*、*AcMYB093*、*AcMYB094* 和 *AcMYB095* 基因在 12 号染色体（Chr12）顶端紧密分布。



Chr1~Chr20 表示七叶树的 20 条染色体，左边的标尺代表染色体的物理长度。

Chr1—Chr20 represent the 20 chromosomes of *A. chinensis*, the scale on the left indicated the physical length of chromosomes in megabases.

图 1 染色体上七叶树 *R2R3-MYB* 基因的位置信息

Fig. 1 Position information of *R2R3-MYB* genes on chromosome of *A. chinensis*

3.2 七叶树 *R2R3-MYB* 基因家族的系统发育分析和共线性分析

为了更好的理解七叶树 *R2R3-MYB* 基因家族的进化关系，通过 MEGA-X 中 Clustal W 对 129 个七叶树 *R2R3-MYB* 蛋白序列和 125 个拟南芥 *R2R3-MYB* 蛋白序列进行多重比对，采用 NJ 构建系统进化树，根据文献中拟南芥 *R2R3-MYB* 蛋白的分类方法及其进化树结果，将七叶树 *R2R3-MYB* 家族成员进行分类，254 个 *R2R3-MYB* 被分成了 34 个亚组（图 2）。其中七叶树与拟南芥共同聚类到 30 个亚家族分支中，2 个拟南芥亚族（S10、S25）单独聚类成不同亚家族分支（无七叶树 *MYB* 基因）；有 2 个分支的七叶树 *MYB* 分别单独聚类（A1、A3），S21 亚群所含成员最多，含有 6 个七叶树 *R2R3-MYB* 蛋白和 8 个拟南芥 *R2R3-MYB* 蛋白；S15 亚群聚类到的七叶树 *R2R3-MYB* 蛋白最多，包含 9 个七叶树

R2R3-MYB 蛋白；E2 亚群所含成员最少，只含 2 个七叶树 *R2R3-MYB* 蛋白。

不同的颜色表示不同的分组，蓝色圆圈符号代表拟南芥的 *R2R3-MYB* 蛋白质序列，红色星星符号代表七叶树的 *R2R3-MYB* 蛋白质序列。

共线性分析结果（图 3）显示，七叶树 *R2R3-MYB* 基因在不同染色体上具有较高的共线相关性，所有染色体上均检测到共线对。112 个七叶树基因组中的 *R2R3-MYB* 基因组成了 89 个全基因组复制基因对，表明这些基因来源于 1 次或多次复制事件，基因组复制事件可能是七叶树 *R2R3-MYB* 基因家族成员进化的主要原因。通过计算 *R2R3-MYB* 共线性基因对的 K_a/K_s 值，分析 *R2R3-MYB* 共线性基因的进化方向，共得到 82 个共线性基因对的 K_a/K_s 值，共线性基因对的 K_a/K_s 介于 0.02~0.61，说明这些 *R2R3-MYB* 转录因子家族成员在进化过程中经历

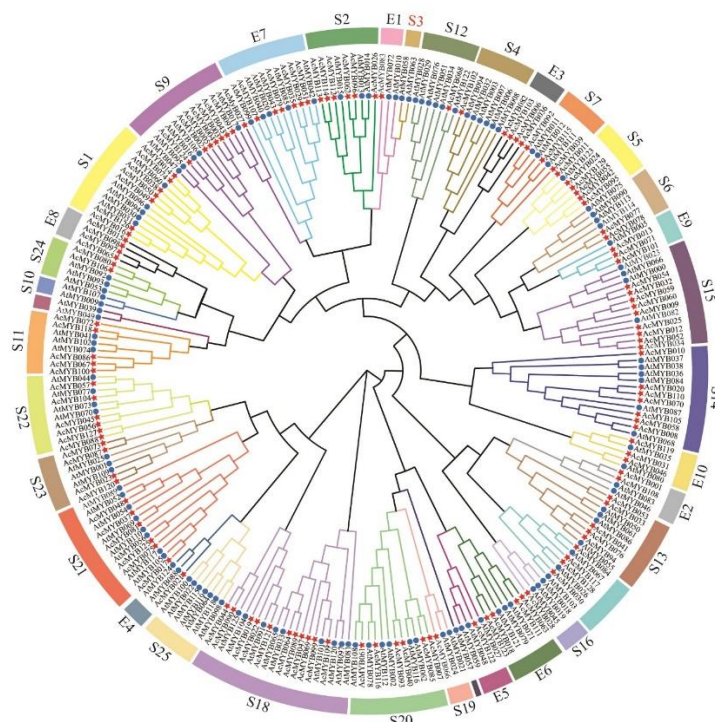
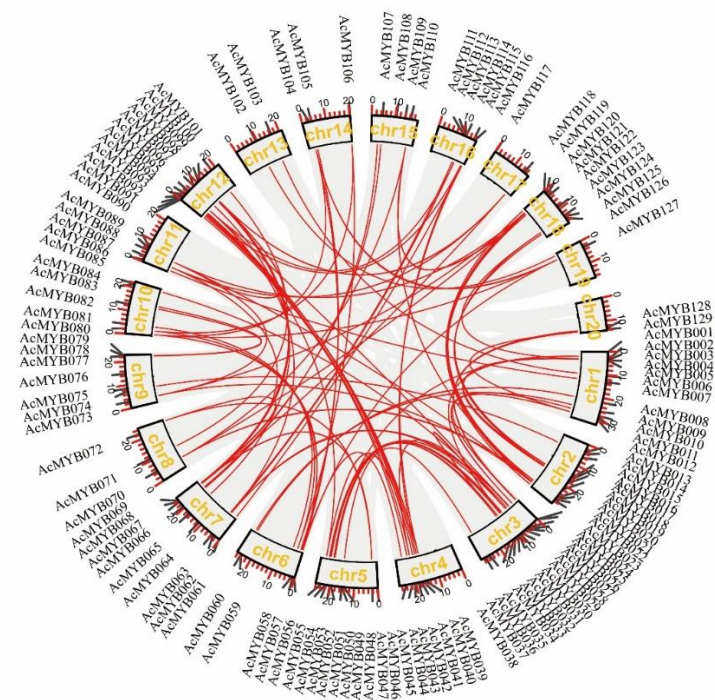


图 2 七叶树与拟南芥 R2R3-MYB 家族成员的系统进化分析

Fig. 2 Phylogenetic analysis of R2R3-MYB family members in *A. chinensis* and *A. thaliana*



灰色线条表示七叶树基因组中所有的复制基因对，红色线条表示七叶树 R2R3-MYB 基因间形成的复制基因对。

The gray lines represent all replication gene pairs in the whole genome of *A. chinensis*, the red lines represent the replication gene pairs formed between the R2R3-MYB genes of *A. chinensis*.

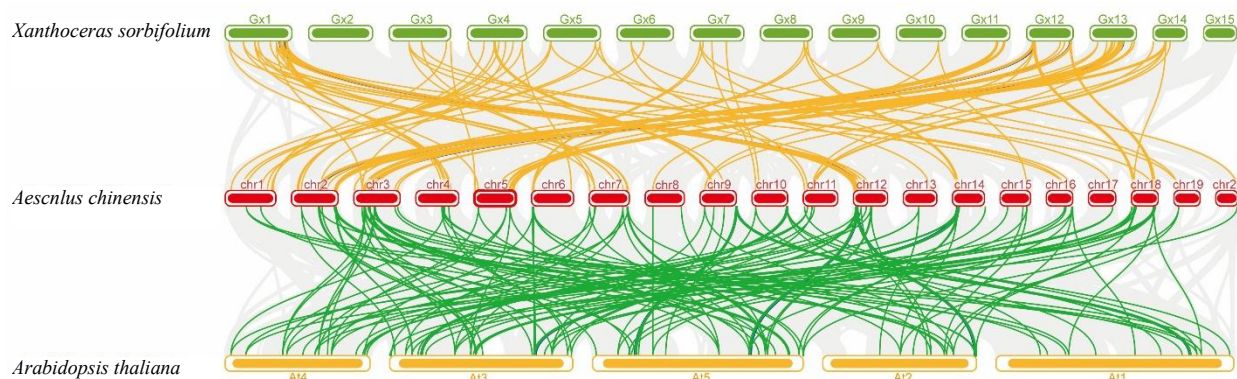
图 3 七叶树 R2R3-MYB 基因全基因组复制分析

Fig. 3 Genome-wide duplication analysis of R2R3-MYB genes in *A. chinensis*

了纯化选择。

Sun 等^[25]对 13 种无患子目植物构建了高置信度系统发育树, 分析发现文冠果和七叶树亲缘关系密切, 因此, 本研究将七叶树的 R2R3-MYB 家族成员与文冠果和拟南芥的 R2R3-MYB 家族成员进行

共线性分析(图 4)。结果显示, 七叶树基因组的 77 个 R2R3-MYB 家族成员与拟南芥的 R2R3-MYB 成员组成 136 对共线性基因对, 七叶树基因组的 81 个 R2R3-MYB 家族成员与文冠果的 R2R3-MYB 成员组成 100 对共线性基因对。



Gx1~Gx15 代表文冠果的 15 条染色体, Chr1~Chr20 代表七叶树的 20 条染色体, At1~At5 代表拟南芥的 5 条染色体; 灰色线条代表 3 个物种所有的共线基因对, 黄色线条表示文冠果和七叶树的共线性 R2R3-MYB 基因, 绿色线条表示拟南芥和七叶树共线性 R2R3-MYB 基因。Gx1—Gx15 represent 15 chromosomes of *X. sorbifolium*, Chr1—Chr20 represent 20 chromosomes of *A. chinensis*, and At1—At5 represent 5 chromosomes of *A. thaliana*. The gray lines represent all collinear gene pairs of the three species, the yellow lines represent the collinearity R2R3-MYB genes of *X. sorbifolium* and *A. chinensis*, and the green lines represent the collinearity R2R3-MYB genes of *A. thaliana* and *A. chinensis*.

图 4 七叶树、文冠果和拟南芥 R2R3-MYB 基因的共线性分析

Fig. 4 Synteny analysis of R2R3-MYB genes among *A. chinensis*, *X. sorbifolium* and *A. thaliana*

3.3 七叶树 R2R3-MYB 家族成员保守结构域和基因结构分析

为了进一步了解七叶树 R2R3-MYB 蛋白的结构特征, 利用 MEME 在线网站对 129 条 AcR2R3-MYB 转录因子家族成员进行保守元件分析, 共预测获到 20 个 motif, 多数成员均包含 motif3、motif7、motif1、motif4、motif2, 表明这 5 个 Motifs 是七叶树 R2R3-MYB 家族中较为保守的基序, 除 AcMYB17、AcMYB75、AcMYB125 外均依次排列在蛋白质的 N 端。同时发现一些 motif 为个别类群所特有, 如图 5 所示, 基序 16、18 和 19 仅存在于 S23 组, 而基序 9 仅存在于 S9 组, 相同亚家族 R2R3-MYB 成员的保守基序组成和排列具有相似性。

基因结构分析发现, 该家族成员的外显子数量在 1~4 (图 5), 遗传距离较近的基因具有类似的内含子-外显子结构, 表明基因功能可能存在保守性, 不同分支成员的结构差异预示了进化过程中为适应环境而产生的功能多样性。多序列比对结果表明, R2R3-MYB 蛋白序列较为保守, 存在明显的 R2-R3 结构域, 进一步确定了所鉴定的蛋白序列为 R2R3-MYB 基因家族成员。

3.4 七叶树 R2R3-MYB 基因顺式作用元件分析

对 129 个七叶树 R2R3-MYB 基因起始密码子上游 2 000 bp 的启动子区域序列进行分析, 共鉴定出 53 种不同类型的顺式作用元件 (图 6), 分别归属于光响应元件、激素响应元件和逆境胁迫响应元件 3 种类型, 数量最多的为光响应元件, 具有 1 404 个, 所有启动子均含有大量光响应元件; 激素响应元件中, 茉莉酸甲酯响应元件共有 328 个, 脱落酸响应元件共有 226 个; 逆境响应元件中, 厌氧诱导元件共有 274 个, 干旱响应元件有 104, 低温响应元件有 82 个。还有一些出现频率较低, 仅在部分七叶树 R2R3-MYB 基因中出现的顺式元件, 如生长素反应调控元件、细胞周期调控元件、胚乳特异性表达元件和种子特异性元件等。每个 R2R3-MYB 基因的启动子含有不同数量和种类的顺式作用元件, 表明它们通过不同的信号通路参与生物和非生物胁迫反应, 在植物的各种生物学过程中发挥作用。

3.5 七叶树 R2R3-MYB 基因家族的表达模式分析

基因表达模式通常与基因的功能有关, 为了明确七叶树 R2R3-MYB 基因在七叶树不同组织部位中的表达模式, 利用转录组数据 (图 7) 对七叶树 R2R3-MYB 基因在花瓣、果皮、种子、叶和枝条 5 个不同

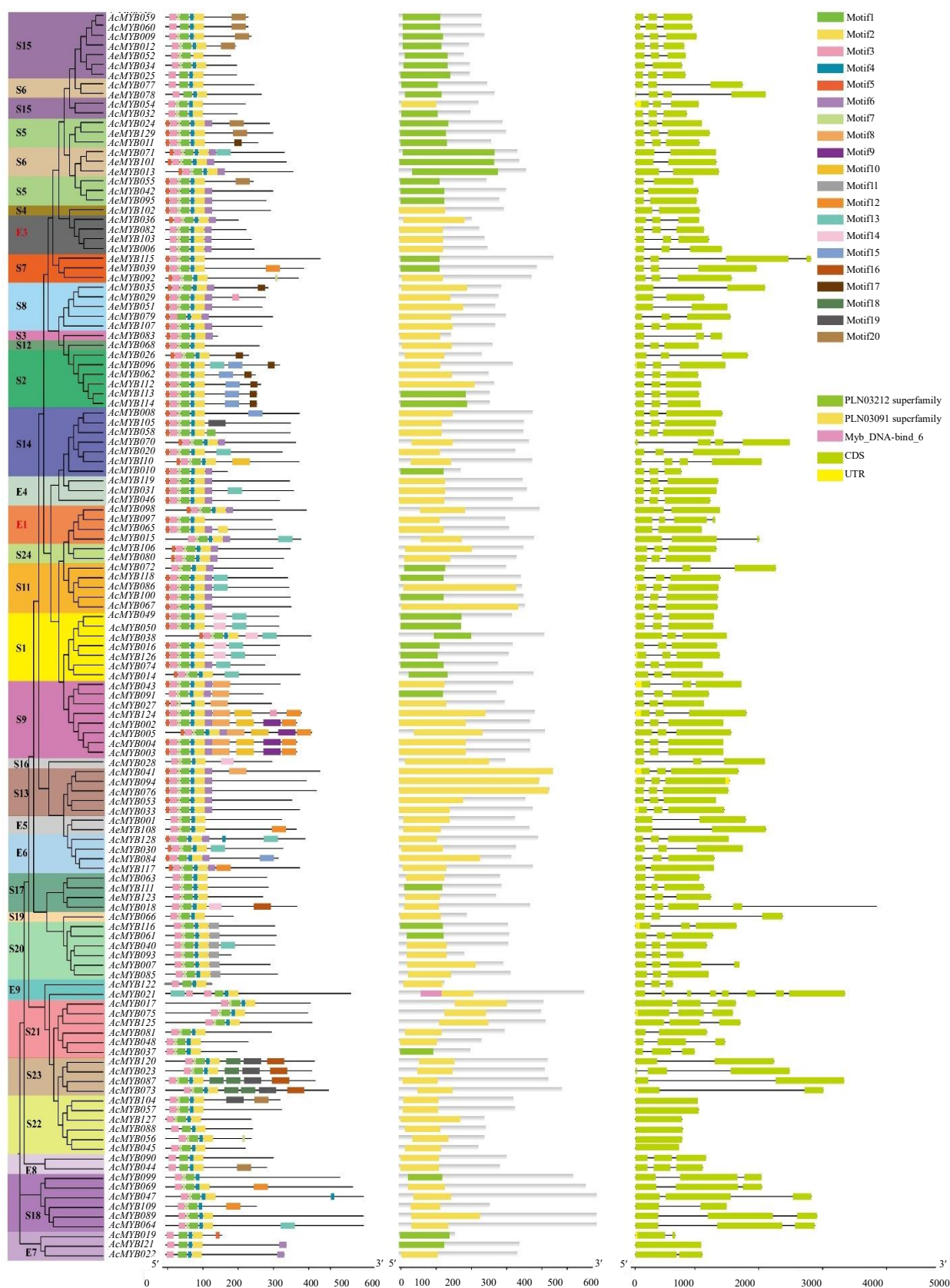


图 5 七叶树 *R2R3-MYB* 基因的进化关系、基序、基因结构

Fig. 5 Evolutionary relationship, motifs and genetic structure of *R2R3-MYB* genes in *A. chinensis*

图 6 七叶树 *R2R3-MYB* 基因家族成员顺式作用元件分析Fig. 6 Analysis of cis-acting elements in *R2R3-MYB* gene family members of *A. chinensis*

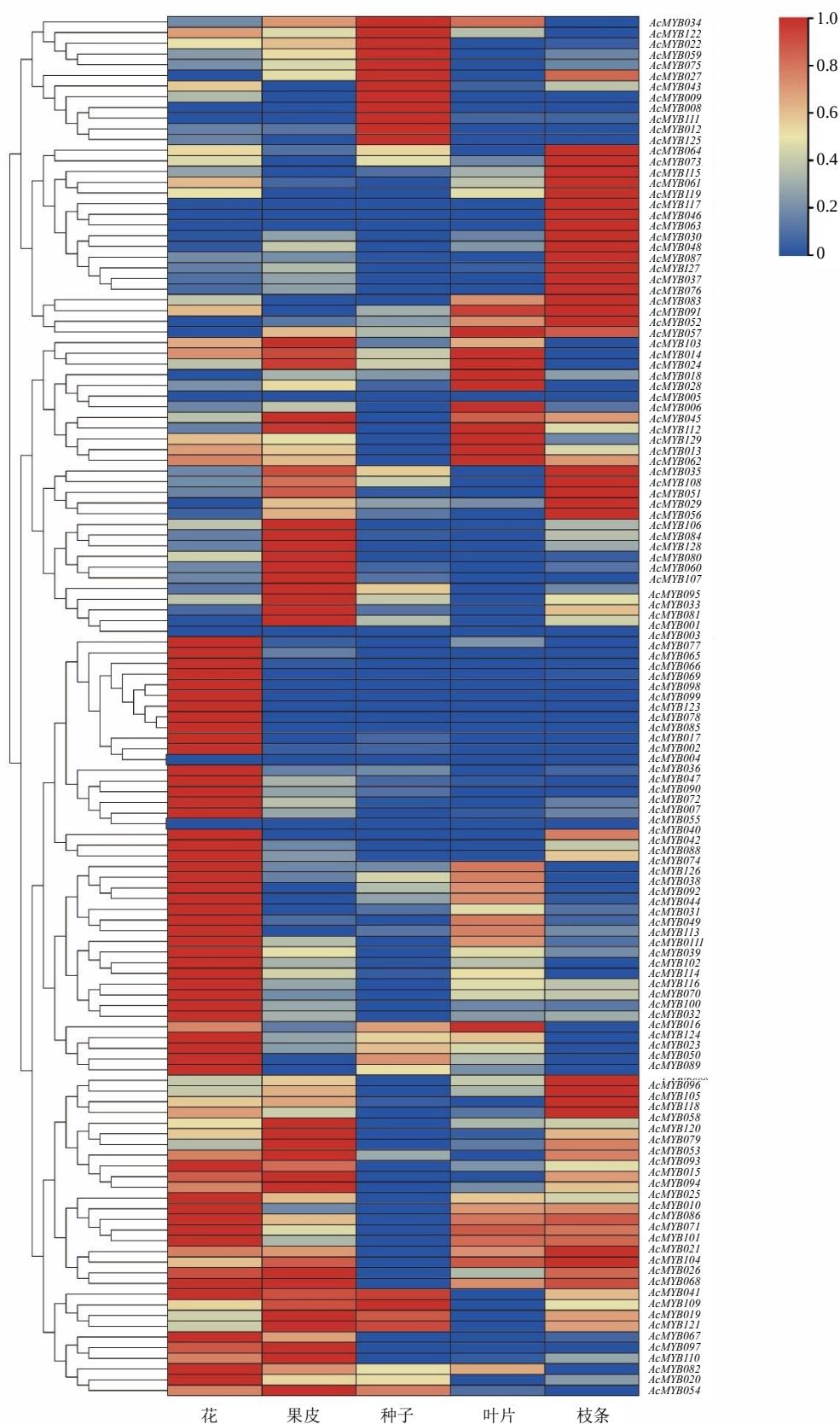


图 7 七叶树 *R2R3-MYB* 基因家族成员的表达模式分析

Fig. 7 Expression pattern analysis of *R2R3-MYB* gene family members in *A. chinensis*

组织中的表达模式进行分析。结果显示, *AcMYB003*、*AcMYB004*、*AcMYB055* 在七叶树各个组织部位均不表达, 126 个七叶树 *R2R3-MYB* 家族基因在根、茎、叶、花、果实中均表达, 但具有组织表达特异性且各成员间的表达水平存在差异。25 个 *R2R3-MYB* 基因在花瓣中的相对表达量显著高于其他组织, 其中 *AcMYB066* (*AC7A207T92*) 在花中相对表达量最高为 1045.894, 而在其他组织中的相对表达量几乎为 0。11 个 *R2R3-MYB* 基因在果皮中的相对表达量相对较高, 其中 *AcMYB121*、*AcMYB095* 和 *AcMYB019* 在叶片中的相对表达量均在 200 以上; 6 个 *R2R3-MYB* 基因 *AcMYB057*、*AcMYB062*、*AcMYB129*、*AcMYB006*、*AcMYB013*、*AcMYB018* 在叶中的表达量相对较高, *AcMYB057* 在叶片中的表达量最高为 628.012; 13 个 *R2R3-MYB* 基因在枝条中的相对表达量显著高于其他组织, *AcMYB104*、*AcMYB096* 在枝条中的相对表达量分别为 465.34 和 235.03; 7 个 *R2R3-MYB* 基因 *AcMYB043*、*AcMYB022*、*AcMYB012*、*AcMYB059*、*AcMYB075*、*AcMYB111* 和 *AcMYB125* 在种仁中的表达量显著高于其他组织, *AcMYB043*、*AcMYB059*、

AcMYB022 3 个基因在种子中相对表达量较高, 分别为 187.39、112.33、55.18, 推测这些基因可能与娑罗子中的化合物合成有关, 值得进一步研究。

基于转录组中 *R2R3-MYB* 基因在七叶树不同组织部位中的表达模式, 筛选 10 个在花瓣、果皮、种子、叶和枝条中差异表达的七叶树 *R2R3-MYB* 基因进行实时荧光定量 PCR 检测, RT-qPCR 结果如图 8 所示, *AcMYB043*、*AcMYB022*、*AcMYB012* 在七叶树种仁中的表达量显著高于其他组织; *AcMYB066*、*AcMYB036* 在七叶树花中的表达量显著高于其他组织; *AcMYB096* 在七叶树枝条中的表达量显著高于其他组织; *AcMYB095*、*AcMYB121* 在七叶树果皮中的表达量显著高于其他组织; *AcMYB129* 在七叶树叶片中的表达量显著高于其他组织; 10 个基因的表达趋势与转录组数据基本一致, 进一步验证了转录组数据的准确性。

4 讨论

本研究通过生物信息学分析方法, 首次对七叶树 *R2R3-MYB* 基因家族成员进行鉴定, 从基因组中共鉴定到 129 个 *R2R3-MYB* 家族基因, 这与部分木

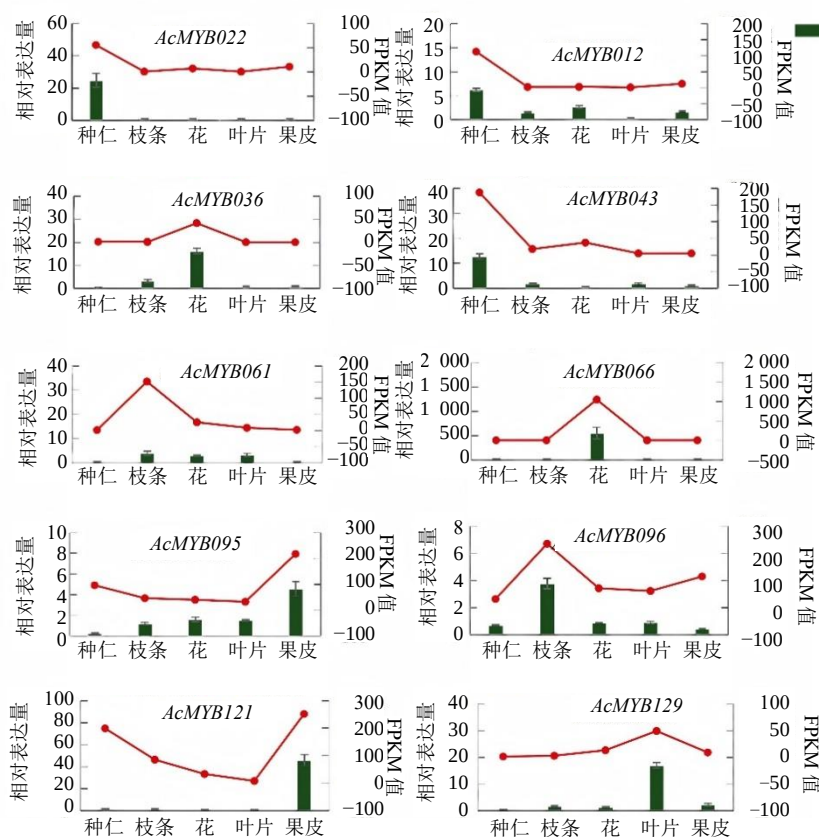


图 8 七叶树 *R2R3-MYB* 差异表达基因在不同组织部位的相对表达量

Fig. 8 Relative expressions of differentially expressed *R2R3-MYB* genes in different tissues of *A. chinensis*

本植物中报道的 *R2R3-MYB* 基因数量相似,如在药用茶树中有 153 个^[28],在梨树 *Pyrus hopeiensis* 中有 122 个^[29],在可可树 *Theobroma cacao* 中有 116 个^[30],在椰子 *Cocos nucifera* 中有 124 个^[31]等,而在银杏中仅筛选到 78 个 *R2R3-MYB* 基因,不同物种基因组中包含的基因家族成员数量相差较大可能是基因组大小存在较大差异导致的。129 个 *AcR2R3-MYB* 基因不均等的分布在七叶树 20 条染色体上,其中 Chr1、Chr2、Chr3 和 Chr12 染色体上基因分布比较密集,AcMYB002、AcMYB003、AcMYB004、AcMYB005 在 Chr1 上呈簇状分布,可能是因为基因的串联复制导致的。基因复制是新基因和新功能的主要来源,基因组内共线性分析可知七叶树 *R2R3-MYB* 基因具有较高的共线相关性,20 条七叶树染色体上均检测到共线基因对,这可能是导致 *R2R3-MYB* 转录因子家族在进化过程中发生扩张的主要原因。七叶树 *R2R3-MYB* 蛋白由高度保守的 MYB 结构域组成,R2 结构域和 R3 结构域分别含有 3 个和 2 个高度保守的色氨酸残基(W),这些保守的氨基酸残基可能对维持 MYB 结构域和功能起重要作用。系统进化分析表明,129 个七叶树 *R2R3-MYB* 蛋白和 125 个拟南芥 *R2R3-MYB* 蛋白被划分为 34 个亚家族,其中 30 个亚组都含有七叶树和拟南芥成员,说明了 2 个物种间的 *R2R3-MYB* 成员之间具有一定的同源性;有 2 个分支的七叶树 MYB 分别单独聚类(A1、A3),推测这 2 个分支上的成员在进化上基因结构发生了较大改变。

AcMYB 基因表达模式分析发现 *AcMYB066* (*AC7A207T92*) 在花中表达量最高,而在其他组织中的表达量几乎为 0,结合系统发育分析结果,AcMYB066 蛋白与拟南芥 AtMYB021、AtMYB024 蛋白聚在一起,研究表明 AtMYB21 和 AtMYB024 这 2 种转录因子在拟南芥雄蕊和花粉发育中起重要作用^[32-33],推测 AcMYB066 可调控七叶树花粉及雄蕊的发育。在果皮中表达量相对较高的 *AcMYB095* (*AC12C31T16*) 与拟南芥纤维发育相关基因 AtMYB123 聚在一组,推测 AcMYB095 调控了七叶树种子外果皮的发育。在叶片中表达量相对较高的 *AcMYB057* (*AC6C6T16*) 与 AtMYB044 聚在一组,AtMYB044 是拟南芥中 ABA 信号转导通路的重要转录因子^[34-35],推测 AcMYB057 在 ABA 受体激活启动下游响应的过程中发挥重要作用,进而调控叶片的发育与防卫反应。*AcMYB104* (*AC14C29T7*) 与

拟南芥干旱胁迫相关转录因子 AtMYB077 聚在一组,AcMYB061 (*AC7A2T154*) 与 AtMYB108 聚在一起,AtMYB108 可调节机械损伤及 ABA 依赖的细胞死亡,进而影响植物的抗病能力^[36],推测 AcMYB104 与 AcMYB061 分别在七叶树抗干旱和生物胁迫中起重要作用。

转录因子对三萜皂苷的生物合成起到重要的调控作用,本研究在七叶树基因组中鉴定到 129 个 *R2R3-MYB* 转录因子,然而其功能及调控机制的报道相对匮乏。研究发现三七中 PnMYB31 和 PnMYB78 与 PnbHLH31 的调控作用影响了植物激素和三七皂苷的积累,从而影响了三七的发育和品质^[37]。PnMYB4 与 PnMYB1 竞争与 PnbHLH 结合,通过调控 PnMYB1-PnbHLH 复合物诱导的皂苷结构基因启动子的激活抑制皂苷的积累^[38]。植物生长调节剂 PBZ 处理可下调忍冬桑黄 *Sanghuangporus lonicericola* SIMYB 的表达,SIMYB 通过负调控 ACAT、MVD、IDI 和 FDPS 的表达,进而促进忍冬桑黄的三萜生物合成^[39]。通过枣树 *Ziziphus jujuba* 转录组及三萜含量的相关性分析及分子实验验证,ZjMYB39 和 ZjMYB4 通过直接结合和激活 ZjFPS 和 ZjSQS 启动子调控三萜生物合成^[40],这说明 *R2R3-MYB* 转录因子可以通过直接调控萜类合成通路中结构基因的转录表达从而调节萜类化合物的合成,还可以与其他转录因子形成蛋白复合物共同参与萜类的合成调控。目前报道较多的调控植物三萜生物合成的转录因子类型有 bHLH、AP2/ERF、bZIP 和 WRKY^[41],MYB 转录因子调控三萜生物合成的研究相对较少。对七叶皂苷生物合成途径中关键酶基因 *AcOSC6*、*AcCYP716A275*、*AcCYP716A278* 的转录因子结合位点进行预测,发现 MYB 转录因子结合位点在各个关键酶基因中均有存在且数量相对较高(图 9),其中 *AcOSC6*、*AcCYP716A275*、*AcCYP716A274*、*AcCYP716A278* 4 个基因中 MYB 转录因子结合位点数量最多,这些基因均为七叶皂苷母核原七叶皂苷元形成的关键基因,且均在种子中特异性高表达,因此可推测 MYB 转录因子可能在七叶皂苷的生物合成过程中发挥着重要的调控作用。由于七叶树的主要药用部位为其种子娑罗子,且七叶皂苷仅在七叶树的种仁中产生,基于七叶树花瓣、果皮、种子、叶片和枝条 5 个不同组织中的转录组数据分析发现,AcMYB043、AcMYB022、AcMYB012、AcMYB059、AcMYB075、AcMYB111

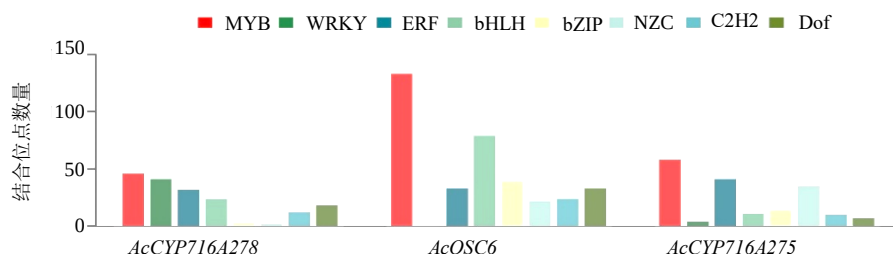


图 9 七叶皂苷合成关键酶基因转录因子结合位点数量统计

Fig. 9 Statistics on number of transcription factors binding sites in key enzyme genes for aescin synthesis

和 *AcMYB125* 在种仁中的表达量显著高于其他组织, 结合转录因子结合位点预测结果, 推测在七叶树种子中相对高表达的 *R2R3-MYB* 基因可能在七叶皂苷的生物合成过程中发挥重要调控作用。

本研究在全基因组水平上鉴定了七叶树 *R2R3-MYB* 基因家族, 分析了其理化性质、染色体定位分布、基因结构特征和表达模式, 并通过预测七叶皂苷生物合成关键酶基因转录因子结合位点, 结合转录组学和代谢组学, 推测 *AcMYB043*、*AcMYB022*、*AcMYB012*、*AcMYB059*、*AcMYB075*、*AcMYB111* 和 *AcMYB125* 7 个基因可能参与娑罗子主要功效物质七叶皂苷的代谢调控, 这些发现为七叶皂苷生物合成的调控机制奠定基础。后续需要结合转录调控相关分子生物学实验进一步验证这些基因的功能, 深入解析其调控机制, 进而通过合成生物学及分子育种手段扩大七叶皂苷原材料来源, 推动七叶皂苷类功效成分资源的综合开发和应用, 保障七叶树资源的可持续利用和中药产业的可持续发展。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 方文培. 中国植物志 (46 卷) [M]. 北京: 科学出版社, 1981: 276.
- [2] 石召华. 七叶树属植物资源及品质研究 [D]. 武汉: 湖北中医药大学, 2013.
- [3] 中国药典 [S]. 一部. 2025: 305.
- [4] Gallelli L. Escin: A review of its anti-edematous, anti-inflammatory, and venotonic properties [J]. *Drug Des Devel Ther*, 2019, 13: 3425-3437.
- [5] Wang Y Y, Han X W, Wan X H, *et al.* β -escin: An updated review of its analysis, pharmacology, pharmacokinetics, and toxicity [J]. *Am J Chin Med*, 2023, 51(8): 2095-2120.
- [6] 张辰露, 李新生, 梁宗锁. 七叶树属植物的分布特征及化学成分研究进展 [J]. 西北林学院学报, 2009, 24(6): 142-145.
- [7] Latchman D S. Transcription factors: An overview [J]. *Int*

J Biochem Cell Biol, 1997, 29(12): 1305-1312.

- [8] Wu X Y, Xia M, Su P, *et al.* MYB transcription factors in plants: A comprehensive review of their discovery, structure, classification, functional diversity and regulatory mechanism [J]. *Int J Biol Macromol*, 2024, 282(Pt 2): 136652.
- [9] Wang X P, Niu Y L, Zheng Y. Multiple functions of MYB transcription factors in abiotic stress responses [J]. *Int J Mol Sci*, 2021, 22(11): 6125.
- [10] Pratyusha D S, Sarada D V L. MYB transcription factors-master regulators of phenylpropanoid biosynthesis and diverse developmental and stress responses [J]. *Plant Cell Rep*, 2022, 41(12): 2245-2260.
- [11] Yu Y B, Zhang S, Yu Y, *et al.* The pivotal role of MYB transcription factors in plant disease resistance [J]. *Planta*, 2023, 258(1): 16.
- [12] Thakur S, Vasudev P G. MYB transcription factors and their role in Medicinal plants [J]. *Mol Biol Rep*, 2022, 49(11): 10995-11008.
- [13] Liu Z Y, Li J N, Li S, *et al.* The 1R-MYB transcription factor SIMYB1L modulates drought tolerance via an ABA-dependent pathway in tomato [J]. *Plant Physiol Biochem*, 2025, 222: 109721.
- [14] Liu L, Li S, Tang F J, *et al.* MaMYBR30, a novel 1R-MYB, plays important roles in plant development and abiotic stress resistance [J]. *Plants*, 2024, 13(13): 1794.
- [15] Feng G Q, Burleigh J G, Braun E L, *et al.* Evolution of the 3R-MYB gene family in plants [J]. *Genome Biol Evol*, 2017, 9(4): 1013-1029.
- [16] Yuan S, Jiang H Y, Wang Y T, *et al.* A 3R-MYB transcription factor is involved in methyl jasmonate-induced disease resistance in *Agaricus bisporus* and has implications for disease resistance in *Arabidopsis* [J]. *J Adv Res*, 2025, 73: 117-131.
- [17] Thiedig K, Weisshaar B, Stracke R. Functional and evolutionary analysis of the *Arabidopsis* 4R-MYB protein SNAPc4 as part of the SNAP complex [J]. *Plant Physiol*,

- 2021, 185(3): 1002-1020.
- [18] Stracke R, Werber M, Weisshaar B. The *R2R3-MYB* gene family in *Arabidopsis thaliana* [J]. *Curr Opin Plant Biol*, 2001, 4(5): 447-456.
- [19] Chen S, Ma F L, Chen J Y, *et al.* Function of R2R3-type myeloblastosis transcription factors in plants [J]. *Rice Sci*, 2025, 32(3): 307-321.
- [20] Liu M M, Liu G Z, Wang G D, *et al.* Identification and functional characterization of *AcMYB113* in anthocyanin metabolism of *Aesculus chinensis* Bunge var. *chinensis* leaves [J]. *Plant Physiol Biochem*, 2023, 199: 107709.
- [21] Yin J, Sun L, Li Y, *et al.* Functional identification of BpMYB21 and BpMYB61 transcription factors responding to MeJA and SA in birch triterpenoid synthesis [J]. *BMC Plant Biol*, 2020, 20(1): 374.
- [22] Zhang Z Z, Li S Y, Lian X Y. An overview of genus *Aesculus* L.: Ethnobotany, phytochemistry, and pharmacological activities [J]. *Pharm Crops*, 2010, 1(1): 24-51.
- [23] Drăghici L R, Hădărugă D I, Hădărugă N G. *Aesculus* species: a review on biologically active compounds and their possible applications [J]. *J Agroalimentary Processes Technolog*, 2020, 26(4): 422-428.
- [24] 秦振芬, 孟祥霄, 温东, 等. 乌拉尔甘草 TIFY 基因家族鉴定及调控分析 [J]. 世界科学技术—中医药现代化, 2022, 24(5): 1855-1864.
- [25] Sun W, Yin Q G, Wan H H, *et al.* Characterization of the horse chestnut genome reveals the evolution of aescin and aesculin biosynthesis [J]. *Nat Commun*, 2023, 14(1): 6470.
- [26] 任伟超, 米要磊, 孟祥霄, 等. 人参 NAC 基因家族成员的鉴定与表达分析 [J]. 世界科学技术—中医药现代化, 2021, 23(9): 3376-3387.
- [27] 蒋莉萍, 徐宏峰, 程璐, 等. 菊花脑 WRKY 基因家族鉴定与表达分析 [J]. 世界科学技术—中医药现代化, 2024, 26(9): 2374-2384.
- [28] Zhang X N, Xu Z W, Liu B M, *et al.* Identification of MYB gene family in medicinal tea tree *Melaleuca alternifolia* (Maiden and Betche) and analysis of members regulating terpene biosynthesis [J]. *Mol Biol Rep*, 2024, 51(1): 70.
- [29] Li Y T, Zhang J, Wang S J, *et al.* Genome-wide identification of the *Pyrus* R2R3-MYB gene family and PhMYB62 regulation analysis in *Pyrus hopeiensis* flowers at low temperature [J]. *Int J Biol Macromol*, 2024, 257(Pt 2): 128611.
- [30] Du J H, Zhang Q Q, Hou S J, *et al.* Genome-wide identification and analysis of the *R2R3-MYB* gene family in *Theobroma cacao* [J]. *Genes*, 2022, 13(9): 1572.
- [31] Li J, Guo S K, Min Htwe Y, *et al.* Genome-wide identification, classification and expression analysis of *MYB* gene family in coconut (*Cocos nucifera* L.) [J]. *Front Plant Sci*, 2024, 14: 1263595.
- [32] Shin B, Choi G, Yi H, *et al.* *AtMYB21*, a gene encoding a flower-specific transcription factor, is regulated by COP1 [J]. *Plant J*, 2002, 30(1): 23-32.
- [33] Huang H, Gao H, Liu B, *et al.* *Arabidopsis* MYB24 regulates jasmonate-mediated stamen development [J]. *Front Plant Sci*, 2017, 8: 1525.
- [34] Li D K, Li Y, Zhang L, *et al.* *Arabidopsis* ABA receptor RCAR1/PYL9 interacts with an R2R3-type MYB transcription factor, AtMYB44 [J]. *Int J Mol Sci*, 2014, 15(5): 8473-8490.
- [35] Nguyen N H, Cheong J J. The AtMYB44 promoter is accessible to signals that induce different chromatin modifications for gene transcription [J]. *Plant Physiol Biochem*, 2018, 130: 14-19.
- [36] Cui F Q, Brosché M, Sipari N, *et al.* Regulation of ABA dependent wound induced spreading cell death by MYB108 [J]. *New Phytol*, 2013, 200(3): 634-640.
- [37] Yu M Y, Ma C X, Tai B, *et al.* Unveiling the regulatory mechanisms of nodules development and quality formation in *Panax notoginseng* using multi-omics and MALDI-MSI [J]. *J Adv Res*, 2025, 69: 463-475.
- [38] Man J H, Shi Y, Huang Y Y, *et al.* PnMYB4 negatively modulates saponin biosynthesis in *Panax notoginseng* through interplay with PnMYB1 [J]. *Hortic Res*, 2023, 10(8): uhad134.
- [39] Zhang D X, Liu B Y, Xue F F, *et al.* Paclobutrazol induces triterpenoid biosynthesis via downregulation of the negative transcriptional regulator SlMYB in *Sanguangporus lonicericola* [J]. *Commun Biol*, 2025, 8(1): 551.
- [40] Wen C P, Zhang Z, Shi Q Q, *et al.* Transcription factors *ZjMYB39* and *ZjMYB4* regulate farnesyl diphosphate synthase-and squalene synthase-mediated triterpenoid biosynthesis in jujube [J]. *J Agric Food Chem*, 2023, 71(11): 4599-4614.
- [41] Li Y L, Wang J, Li L Y, *et al.* Natural products of pentacyclic triterpenoids: From discovery to heterologous biosynthesis [J]. *Nat Prod Rep*, 2023, 40(8): 1303-1353.

[责任编辑 时圣明]