

## 基于全基因组的栀子苷生物合成相关 MYB 转录因子系统分析

许文杰<sup>1</sup>, 黄远浩<sup>1</sup>, 韩蓉蓉<sup>2</sup>, 刘燕琴<sup>2</sup>, 曹敏<sup>2</sup>, 宋经元<sup>1,3\*</sup>

(1. 中国医学科学院北京协和医学院药用植物研究所, 国家中医药管理局中药资源保护重点研究室, 北京 100193;  
2. 重庆市药物种植研究所, 重庆 408435; 3. 中药资源教育部工程研究中心, 北京 100193)

**摘要:** MYB 转录因子参与调控多种次生代谢产物的生物合成。栀子是我国大宗常用中药材, 主要活性成分为栀子苷。本研究以栀子不同发育时期的叶片和花蕾器官为研究材料, 借助基因组和转录组分析, 全面挖掘栀子苷生物合成相关 MYB 转录因子。转录组数据分析表明, 不同于萜类合成共有途径基因表达模式, 栀子苷合成特异途径基因在花蕾中的表达量显著高于叶片, 表达模式与该成分的器官积累规律相同, 且香叶醇合酶、环烯醚萜合酶和栀子苷酸甲基转移酶编码基因的启动子区域均含有多个 MYB 转录因子结合位点。对栀子基因组编码基因进行注释, 共获得 105 个 MYB 转录因子, 根据保守结构域数量将其划分为 68 个 1R-MYB、33 个 R2R3-MYB、3 个 3R-MYB 和 1 个非典型 MYB 转录因子。进一步结合系统发育树和实时荧光定量 PCR 分析, 筛选出 3 个可能调控栀子苷生物合成的候选 MYB 转录因子, 包括潜在正调控转录因子 GjMYB23 及负调控转录因子 GjMYB31 和 GjMYB73。本研究结果将为栀子苷生物合成调控研究及进一步解析栀子品质形成机制奠定基础, 以推动栀子优良品种选育。

**关键词:** 栀子; 栀子苷; MYB; 转录因子; 基因组

中图分类号: R932 文献标识码: A 文章编号: 0513-4870(2023)08-2522-10

## Systematic analysis of MYB transcription factors related to the geniposide biosynthesis in *Gardenia jasminoides* Ellis based on whole genome

XU Wen-jie<sup>1</sup>, HUANG Yuan-hao<sup>1</sup>, HAN Rong-rong<sup>2</sup>, LIU Yan-qin<sup>2</sup>, CAO Min<sup>2</sup>, SONG Jing-yuan<sup>1,3\*</sup>

(1. Key Lab of Chinese Medicine Resources Conservation, State Administration of Traditional Chinese Medicine of the People's Republic of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China; 2. Chongqing Institute of Medicinal Plant Cultivation, Chongqing 408435, China; 3. Engineering Research Center of Chinese Medicine Resource, Ministry of Education, Beijing 100193, China)

**Abstract:** MYB transcription factors are involved in the regulation of various secondary metabolites biosynthesis. *Gardenia jasminoides* Ellis is the commonly used Chinese herbal medicine, and its main active ingredient is geniposide. Here, leaves and flower buds at different developmental stages of *G. jasminoides* were used to explore MYB transcription factors related to geniposide biosynthesis based on genome and transcriptome analysis. Transcriptome data analysis showed that, different from the expression of the common pathway genes for terpenoid biosynthesis, the expression level of genes in the specific pathway of geniposide biosynthesis was significantly higher in flower buds than in leaves, which was the same as the organ accumulation pattern of this component. And the promoter regions of geraniol synthase, iridoid synthase and geniposidic acid methyltransferase involved in the specific pathway all contained multiple MYB-binding sites. A total of 105 MYB transcription factors were obtained by annotating the coding genes of *G. jasminoides*, which were divided into 68 1R-MYB, 33 R2R3-MYB, 3 3R-MYB and 1 atypical MYB transcription factor according to the number of conserved domain.

收稿日期: 2023-04-26; 修回日期: 2023-06-30.

基金项目: 国家自然科学基金项目 (82073966); 川渝道地药材绿色种植技术研究与应用 (cstc2020jcsx-cylhX0008).

\*通讯作者 Tel: 86-10-57833199, E-mail: jysong@implad.ac.cn

DOI: 10.16438/j.0513-4870.2023-0516

Based on the analysis of phylogenetic tree and quantitative real-time PCR, three candidate MYB transcription factors related to geniposide biosynthesis were selected, including potential positive regulation factor GjMYB23 and negative regulation factors GjMYB31 and GjMYB73. The results of this study will lay a foundation for searching the regulation of geniposide biosynthesis and further analysis of the quality formation mechanism of *G. jasminoides*, so as to promote the breeding of excellent varieties of *G. jasminoides*.

**Key words:** *Gardenia jasminoides*; geniposide; MYB; transcription factor; genome

中药栀子来源于茜草科栀子属植物栀子 (*Gardenia jasminoides* Ellis) 的干燥成熟果实, 味苦, 性寒, 具有泻火除烦、清热利湿和凉血解毒等功效。现代药理研究表明, 栀子在保肝利胆、保护神经、抗炎和抗血栓等方面具有一定活性, 是 280 余种中药制剂和 2 330 余种治疗疾病经典方剂配方<sup>[1]</sup>, 如急救名药安宫牛黄丸和血脉清片等。栀子天然产物包括环烯醚萜、二萜、有机酸和黄酮等多种化合物, 主要活性成分环烯醚萜多以糖苷形式存在, 其中栀子苷 (geniposide) 含量最高, 《中华人民共和国药典》(2020 年版一部) 规定该成分是栀子的质量控制指标。近期多项研究发现, 栀子苷在阿尔茨海默症和抑郁症等疾病治疗中展现出多种生理活性, 是潜在的多种脑部疾病预防和治疗药物或先导化合物<sup>[2,3]</sup>。

栀子始记于《神农本草经》, 迄今已有千余年栽培历史, 主产于江西、四川、重庆和福建等长江以南多地<sup>[4]</sup>, 具有较高遗传多样性<sup>[5]</sup>。栀子年需求量在 6 000 吨左右, 这促使人们对栀子进行盲目引种与大规模种植, 品种混杂导致药材质量参差不齐<sup>[6]</sup>, 严重影响栀子疗效及产业化发展。栀子苷生源途径研究可以为栀子种质选育提供关键分子标志物, 广受关注。与众多萜类成分相同, 栀子苷的合成前体物是来源于甲基赤藓醇 4-磷酸途径 (MEP) 和甲羟戊酸途径 (MVA) 的异戊烯基二磷酸 (IPP) 和二甲基烯丙基二磷酸 (DMAPP)。IPP 和 DMAPP 在一系列酶的催化下合成环烯醚萜醇骨架, 其中香叶醇合酶 (GES) 是环烯醚萜类成分合成特异途径的第一个酶<sup>[7]</sup>, 环烯醚萜合酶 (IS) 则催化关键的环化过程<sup>[8]</sup>。上述产物进一步在糖基转移酶、甲基转移酶和氧化还原酶的作用下合成栀子苷等环烯醚萜苷成分, 其中催化栀子苷合成最后关键步骤的栀子苷酸甲基转移酶 (GAMT) 已被鉴定<sup>[9]</sup>, 为栀子苷代谢调控研究提供了重要靶基因。

植物 MYB (v-myb avian myeloblastosis viral oncogene homolog) 转录因子广泛参与胁迫响应、次生代谢调控和激素应答等过程<sup>[10]</sup>, 与黄酮、萜类和木素等天然产物合成密切相关<sup>[11]</sup>。根据保守结构域 SANT 的数量, MYB 转录因子可划分为 4 种类型: 1R-MYB、R2R3-MYB、3R-MYB 及非典型 MYB (4 个或 5 个 MYB

结构域)<sup>[12]</sup>。研究人员已在香雪兰 (*Freesia x hybrida*)<sup>[13]</sup>、丹参 (*Salvia miltiorrhiza*)<sup>[14]</sup> 和黄花蒿 (*Artemisia annua*)<sup>[15,16]</sup> 等中发现调控萜类天然产物合成的 MYB 转录因子, 但目前尚未见栀子苷生物合成调控相关 MYB 转录因子报道。前期转录组数据分析和靶向代谢物检测结果表明, 栀子叶片中的栀子苷含量较低, 而花蕾中该成分含量较高, 且栀子苷合成途径关键酶编码基因 *GAMT* 的表达模式与该成分的器官积累规律一致<sup>[9]</sup>。因此, 为探究栀子苷的生物合成是否受到 MYB 转录因子调控, 本研究同样以栀子不同发育时期叶片和花蕾器官为研究材料, 通过栀子基因组及转录组数据系统分析, 全面挖掘调控栀子苷生物合成特异途径基因表达的 MYB 转录因子, 为栀子优良种质选育及代谢工程研究提供潜在遗传资源。

## 材料与方法

**转录组数据分析** 栀子转录组数据为实验室前期积累数据<sup>[9]</sup>, 包括幼嫩叶片 (young leaf, YL)、成熟叶片 (mature leaf, ML)、小花蕾 (small flower bud, SFB, 长约 2 cm) 和大花蕾 (large flower bud, LFB, 长约 3 cm), 每个器官三个生物学重复。基于栀子参考基因组 (NCBI: ASM1310374v1), 借助 Hisat2 (v. 2.1.0) 和 Cufflinks (v. 2.1.1) 软件分别进行序列比对及基因表达值 FPKM (fragments per kilobase of exon per million reads mapped) 计算。从 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/>) 下载拟南芥 (*Arabidopsis thaliana*)、长春花 (*Catharanthus roseus*) 和栀子中环烯醚萜苷合成途径基因的蛋白序列 (表 1<sup>[9]</sup>), 以其为参考序列, 使用本地版 Blastp (v. 2.10.0) 对栀子相同功能蛋白进行注释。提取栀子中已验证功能基因及与其他物种高度同源基因的 4 个器官表达值, 使用 TBtools (v. 1.108) 软件, 以  $\text{Log}_{10}(\text{FPKM}+1)$  绘制表达热图。

**MYB 转录因子结合位点预测** 根据栀子参考基因组, 提取栀子中已表征功能的栀子苷合成特异途径基因上游 2 000 bp 为启动子序列, 包括 *GES*、*IS* 和 2 个 *GAMT*。使用 New PLACE 在线工具 (<https://www.dna.affrc.go.jp/PLACE>) 预测上述基因启动子区域的顺式

**Table 1** Genes involved in the biosynthetic pathway of iridoid glycosides. Sequences of *GAMT1* and *GAMT2* genes were used with permission from reference<sup>[9]</sup>, copyright (2021) Elsevier B. V.

Gene	Gene name	Accession ID/reference
<i>DXS</i>	1-Deoxy- <i>D</i> -xylulose-5-phosphate synthase	Q38854
<i>DXR</i>	1-Deoxy- <i>D</i> -xylulose-5-phosphate reductoisomerase	ARU08103
<i>CMS</i>	4-Diphosphocytidyl-2-C-methyl- <i>D</i> -erythritol synthase	Q5N8G1
<i>CMK</i>	4-Diphosphocytidyl-2-C-methyl- <i>D</i> -erythritol kinase	O81014
<i>MCS</i>	2-C-Methyl- <i>D</i> -erythritol-2,4-cyclodiphosphate synthase	Q9CAK8
<i>HDS</i>	4-Hydroxy-3-methylbut-2-enyl diphosphate synthase	F4K0E8
<i>HDR</i>	4-Hydroxy-3-methylbut-2-enyl diphosphate reductase	Q94B35
<i>AACT</i>	Acetoacetyl-CoA thiolase	AAM00280
<i>HMGS</i>	3-Hydroxy-3-methylglutaryl CoA synthase	P54873
<i>HMGR</i>	3-Hydroxy-3-methylglutaryl CoA reductase	P43256
<i>MVK</i>	Mevalonate kinase	P46086
<i>PMK</i>	5-Phosphomevalonate kinase	Q9C6T1
<i>MVD</i>	Mevalonate-5-diphosphate decarboxylase	O23722
<i>IS</i>	Iridoid synthase	MN120556
<i>IS2</i>	Iridoid synthase 2	MN120557
<i>IS4</i>	Iridoid synthase 4	MN120558
<i>GPPS</i>	Geranyl diphosphate synthase	AHA82035
<i>GES</i>	Geraniol synthase	QCT83300
<i>IO</i>	Iridoid oxidase	AHK60833
<i>8HGO</i>	8-Hydroxygeraniol oxidoreductase	AHK60836
<i>G10H</i>	Geraniol 10-hydroxylase	Q8VWZ7
<i>GAMT1</i>	Geniposidic acid methyltransferase 1	[9]
<i>GAMT2</i>	Geniposidic acid methyltransferase 2	[9]

元件结合位点,使用TBtools软件进行可视化。

**MYB转录因子家族分析** 基于本课题组已有的栀子基因组注释编码基因序列,使用plantTFDB数据库(<http://planttfdb.gao-lab.org/index.php>)进行转录因子预测。使用SMART在线工具(<http://smart.embl-heidelberg.de>)预测MYB转录因子的SANT保守结构域,过滤无完整结构域基因。使用MEME在线工具(<https://meme-suite.org/meme/tools/meme>)预测MYB转录因子保守基序,数量设置为20。使用TBtools软件对MYB转录因子系统发育树和保守基序进行可视化。分别使用SOPMA在线工具([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa%20\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa%20_sopma.html))和Swiss-Model(<https://swissmodel.expasy.org/>)预测MYB转录因子的二级结构和三级结构。

**候选MYB转录因子挖掘** 收集与萜类合成相关的MYB转录因子蛋白序列<sup>[13-15,17-21]</sup>,提取在栀子任一器官表达MYB基因(FPKM > 10)的蛋白序列,分别使用MAFFT和IQ-TREE2软件对其进行多序列比对和系统发育树构建,自动筛选最优建树模型。根据栀子4个器官的基因表达值FPKM,使用R软件(v. 3.4.3)Cor函数计算基因间表达相关系数,提取与栀子萜合成特异途径基因表达模式呈正、负相关的MYB转录因子基因。最终选取与已知萜类代谢相关MYB转录因子亲缘关系较近、编码基因与栀子萜合成特异途径基因表达模式相关的MYB转录因子作为栀子萜生物合

成调控候选转录因子。

**候选MYB基因qRT-PCR分析** 候选*GjMYB*基因实时荧光定量PCR(qRT-PCR)验证所用样本包括YL、ML、SFB和LFB,均为2021年5月采集于重庆市南山区,实验室-80℃冰箱保存。该材料与转录组测序样本为同批次采集,但为不同样本。使用多糖多酚植物总RNA提取试剂盒(TIANGEN, DP441)提取栀子样品总RNA。使用PrimerQuest在线工具(<https://sg.idtdna.com/Primerquest>)设计MYB基因的qRT-PCR引物(表2)。扩增片段长度约100 bp, Tm值为60℃左右。以栀子YL、ML、SFB和LFB样本质量合格的反转录cDNA为模板,每个器官2个生物学重复,每个基因3个技术重复,使用SYBR Green I染料(CWBIO, CW0955M)进行MYB基因表达相对定量分析。以栀子*Actin*为内参基因,以YL器官为对照组,根据 $2^{-\Delta\Delta Ct}$ 方法计算候选MYB基因相对表达量。

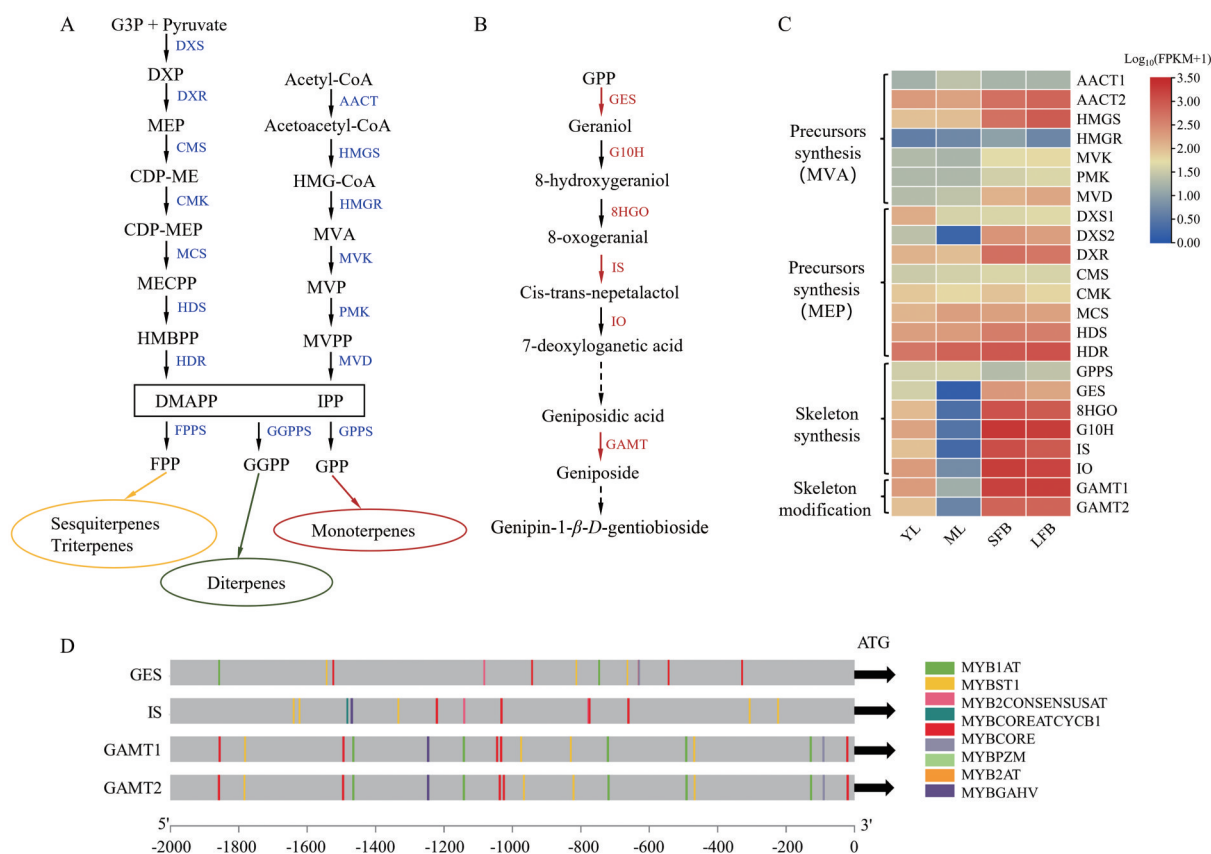
## 结果

### 1 栀子萜生物合成途径基因表达模式分析

根据栀子基因组注释文件,计算4个器官35 967个编码基因的表达水平,并提取已鉴定栀子萜生源途径基因及环烯醚萜苷合成途径同源基因的表达值。结果显示,与萜类合成共有途径(图1A)基因相比,除*GPPS*外的环烯醚萜醇骨架形成及修饰基因均在不同发育阶段的花蕾器官中特异性高表达( $P < 0.05$ ),而在

**Table 2** Primers for qRT-PCR of MYB transcription factor genes in *Gardenia jasminoides*. MYB: v-myb avian myeloblastosis viral oncogene homolog

Gene	Forward primer (5'→3')	Reverse primer (5'→3')
<i>Actin</i>	TCCTCTTCCAGCCTTCTATC	GCTCATACGGTCAGCAATAC
<i>GjMYB9</i>	TACTCGTACGAAGAGGTGAG	GTTTACGACCTGGAATTCCTCC
<i>GjMYB16</i>	GTGTTAGTGCAGGGTCTTAC	TGCGTTTCCAGACTCTAATG
<i>GjMYB19</i>	CAGTTCAATGCCACCTCTAA	AGGATGGCTACTGTTGATATTG
<i>GjMYB23</i>	GGACGAAAGTCAAGACTCTAAG	TTGTGTTCCAATACTCCTCAAT
<i>GjMYB31</i>	CATGGAAGCAGCAAGTTTAAAG	GACTGTGGAAGTCCAATAGATAG
<i>GjMYB73</i>	GCCCAATACGATAAGGAGAC	GTAGTTGGGATACGGAACAC
<i>GjMYB76</i>	CGAATGCCTCTTCTTCTTC	AGGCTAATGGATGGCTTTATAC
<i>GjMYB101</i>	CTCCTCATCATCGTCCATTC	CAGCTGGCTGCATTATCT



**Figure 1** Genes involved in the biosynthetic pathway of geniposide and their expression in *Gardenia jasminoides*. A: Terpene biosynthetic pathway. G3P: Glyceraldehyde 3-phosphate; DXP: 1-Deoxy-*D*-xylulose-5-phosphate; MEP: 2-C-Methyl-*D*-erythritol-4-phosphate; CDP-ME: 4-(Cytidine 5'-diphospho)-2-C-methyl-*D*-erythritol; CDP-MEP: 2-Phospho-4-(cytidine 5'-diphospho)-2-C-methyl-*D*-erythritol; MECPP: 2-C-Methyl-*D*-erythritol-2,4-cyclodiphosphate; HMBPP: 1-Hydroxy-2-methyl-2-butenyl-4-diphosphate; HMG-CoA: 3-Hydroxy-3-methylglutaryl CoA; MVA: Mevalonate; MVP: Mevalonate phosphate; MVPP: Mevalonate diphosphate; IPP: Isopentenyl pyrophosphate; DMAPP: Dimethylallyl pyrophosphate; B: Specific step in the biosynthetic pathway of iridoid glycoside in *Gardenia jasminoides*. Red arrows represent catalytic steps that have been elucidated in *Gardenia jasminoides*; C: Expression heatmap of the genes involved in the biosynthetic pathway of geniposide. YL: Young leaf; ML: Mature leaf; SFB: Small flower bud; LFB: Large flower bud; FPKM: Fragments per kilobase of exon per million; D: Distribution of MYB transcription factors binding sites in the promoter of the specific pathway genes for geniposide biosynthesis. The promoter sequence was 2 000 bp upstream of the transcription start site

不同发育时期叶片中表达量较低, 尤其是ML样本(图1B、C)。该结果说明栀子苷合成特异途径基因的表达均具有组织特异性, 且表达模式与栀子苷在不同器官中的积累规律一致。

## 2 栀子苷合成特异途径基因MYB结合位点分析

由于栀子苷合成特异途径基因均为器官特异性表达, 提取已验证功能的栀子苷合成特异途径基因上游2 000 bp序列, 包括 *GES*、*IS*、*GAMT1* 和 *GAMT2* (图

1D)。New PLACE 顺式作用元件分析结果显示, *GES*、*IS*和2个 *GAMT*基因启动子区域均含有16~18个MYB转录因子结合位点,如MYB1AT、MYBST1、MYBCORE和MYB2CONSENSUSAT等,提示栀子苷生物合成基因的器官特异性表达可能受到MYB转录因子调控。

### 3 栀子MYB转录因子家族鉴定

通过PlantTFDB数据库对栀子编码基因进行转录因子预测,共获得119个MYB转录因子。进一步使用SMART数据库筛选含有完整SANT结构域的MYB,最终确定105个MYB转录因子,根据其编码基因的染色体位置排序命名为GjMYB1~GjMYB105。基于SANT结构域数量将104个MYB转录因子划分为3类,包括68个1R-MYB、33个R2R3-MYB和3个3R-MYB(表3)。此外,GjMYB64含有5个SANT结构域,为非典型MYB转录因子。

对105个MYB转录因子进行序列分析发现,其平均蛋白长度为395个氨基酸。所有MYB转录因子基因不平均分布于栀子的11条染色体,其中9号染色体包含54个MYB转录因子编码基因,平均MYB基因数量最多(图2)。MEME预测结果显示,所有MYB转录因子的保守motif为1~12个(图3)。数量最多的1R-MYB均含有motif 1、motif 2、motif 3和motif 4中至少1个motif,除GjMYB81外的32个R2R3-MYB注释到motif 1、motif 2、motif 3和motif 4中至少2个motif,3个

3R-MYB均包含motif 1、motif 2和motif 3。该结果说明,motif 1、motif 2、motif 3和motif 4可能是MYB转录因子保守序列。

### 4 栀子苷生物合成调控相关MYB转录因子筛选

提取表达*GjMYB*基因(FPKM > 10)的蛋白序列,将53个GjMYB转录因子与已鉴定参与萜类合成调控的MYB转录因子构建系统发育树。结果显示,拟南芥、留兰香(*Mentha spicata*)、香雪兰、黄花蒿、丹参、白桦(*Betula platyphylla*)和番茄(*Solanum lycopersicum*)中的10个萜类合成调控MYB转录因子聚在一支,7个GjMYB与其亲缘关系较近,包括GjMYB9、GjMYB16、GjMYB19、GjMYB23、GjMYB31、GjMYB76和GjMYB101(图4A)。在7个*GjMYB*基因中,*GjMYB9*、*GjMYB23*和*GjMYB101*表达量较高,*GjMYB23*和*GjMYB31*与*GES*、*IS*、*GAMT1*和*GAMT2*的表达模式相关性较高,表达相关系数分别大于0.91和小于-0.91。此外,通过对53个*GjMYB*基因的表达模式分析发现,*GjMYB73*基因表达量极高,且其与上述功能基因的表达模式呈现高度负相关(相关系数小于-0.90)。因此,将上述8个GjMYB作为调控栀子苷生物合成的候选MYB转录因子。转录因子二维结构分析显示,筛选的8个GjMYB转录因子均包含 $\alpha$ -螺旋、 $\beta$ -转角、 $\beta$ -折叠和无规则卷曲构型,且以 $\alpha$ -螺旋和无规则卷曲为主,与三维结构同源建模结果相近(图4B)。

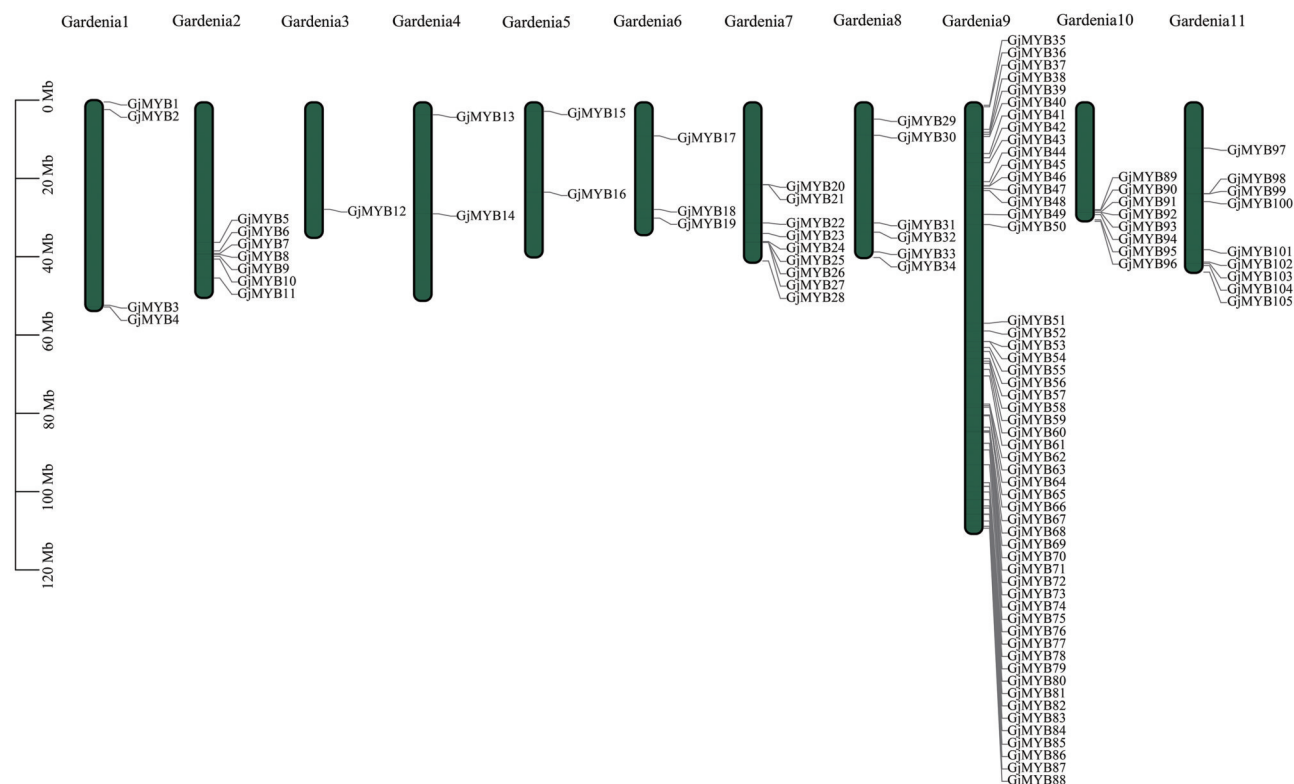


Figure 2 Chromosome location of MYB transcription factor genes in *Gardenia jasminoides*

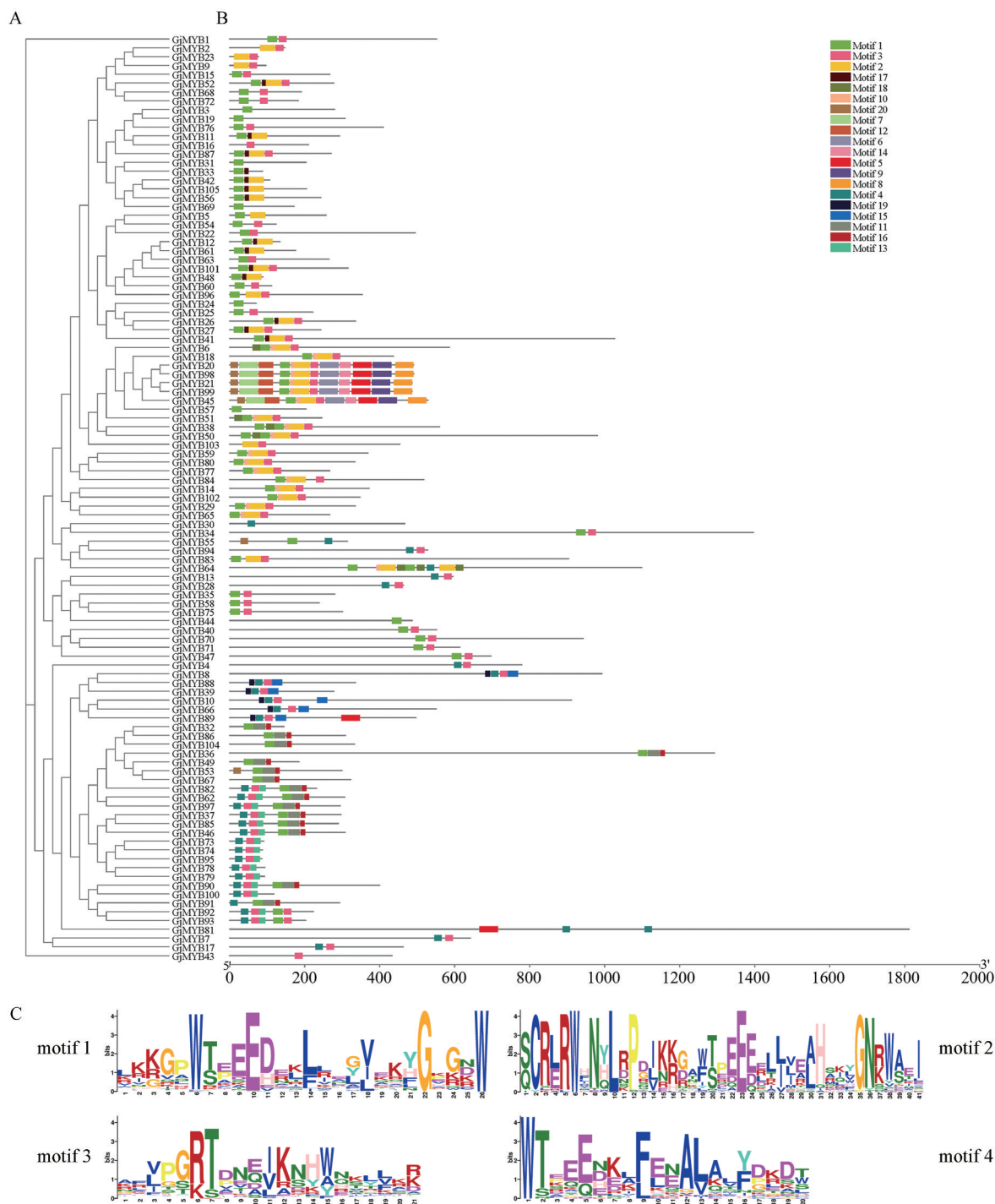
**Table 3** The classification of MYB transcription factor genes in *Gardenia jasminoides*. AA size: The number of amino acid

Gene name	Gene ID	Class	AA size	Gene name	Gene ID	Class	AA size
<i>GjMYB1</i>	Gj1P4T24	1R	552	<i>GjMYB54</i>	Gj9A610T144	1R	124
<i>GjMYB2</i>	Gj1A24T33	1R	147	<i>GjMYB55</i>	Gj9X626T25	1R	314
<i>GjMYB3</i>	Gj1P524T10	1R	280	<i>GjMYB56</i>	Gj9A636T100	1R	243
<i>GjMYB4</i>	Gj1P528T16	1R	779	<i>GjMYB57</i>	Gj9A654T110	1R	204
<i>GjMYB5</i>	Gj2A357T40	1R	257	<i>GjMYB58</i>	Gj9A661T106	1R	239
<i>GjMYB6</i>	Gj2P379T15	R2R3	586	<i>GjMYB59</i>	Gj9X666T65	R2R3	369
<i>GjMYB7</i>	Gj2A386T122	1R	642	<i>GjMYB60</i>	Gj9A682T134	1R	112
<i>GjMYB8</i>	Gj2P388T19	1R	993	<i>GjMYB61</i>	Gj9P698T16	1R	176
<i>GjMYB9</i>	Gj2A393T78	1R	97	<i>GjMYB62</i>	Gj9A770T55	R2R3	307
<i>GjMYB10</i>	Gj2A400T31	1R	912	<i>GjMYB63</i>	Gj9A774T63	1R	265
<i>GjMYB11</i>	Gj2P448T7	1R	293	<i>GjMYB64</i>	Gj9A778T112	Atypical	1 099
<i>GjMYB12</i>	Gj3A273T7	1R	134	<i>GjMYB65</i>	Gj9A798T79	R2R3	267
<i>GjMYB13</i>	Gj4A31T63	1R	596	<i>GjMYB66</i>	Gj9P800T24	1R	551
<i>GjMYB14</i>	Gj4A284T28	R2R3	372	<i>GjMYB67</i>	Gj9A829T62	1R	323
<i>GjMYB15</i>	Gj5A22T109	1R	267	<i>GjMYB68</i>	Gj9A839T43	1R	191
<i>GjMYB16</i>	Gj5A229T11	1R	210	<i>GjMYB69</i>	Gj9A839T49	1R	172
<i>GjMYB17</i>	Gj6P85T6	1R	463	<i>GjMYB70</i>	Gj9P839T9	1R	943
<i>GjMYB18</i>	Gj6A273T22	R2R3	437	<i>GjMYB71</i>	Gj9A842T63	1R	614
<i>GjMYB19</i>	Gj6P295T8	1R	308	<i>GjMYB72</i>	Gj9A842T65	1R	183
<i>GjMYB20</i>	Gj7A210T21	R2R3	491	<i>GjMYB73</i>	Gj9X870T42	1R	91
<i>GjMYB21</i>	Gj7A210T23	R2R3	487	<i>GjMYB74</i>	Gj9X871T92	1R	88
<i>GjMYB22</i>	Gj7P308T5	1R	495	<i>GjMYB75</i>	Gj9A887T113	1R	301
<i>GjMYB23</i>	Gj7A335T61	1R	77	<i>GjMYB76</i>	Gj9P925T22	1R	410
<i>GjMYB24</i>	Gj7P355T9	1R	71	<i>GjMYB77</i>	Gj9X971T52	R2R3	267
<i>GjMYB25</i>	Gj7P357T7	1R	222	<i>GjMYB78</i>	Gj9X980T34	1R	94
<i>GjMYB26</i>	Gj7A357T62	R2R3	336	<i>GjMYB79</i>	Gj9X981T83	1R	93
<i>GjMYB27</i>	Gj7P357T10	R2R3	243	<i>GjMYB80</i>	Gj9X995T40	R2R3	334
<i>GjMYB28</i>	Gj7A405T102	1R	464	<i>GjMYB81</i>	Gj9P1015T8	R2R3	1 812
<i>GjMYB29</i>	Gj8A42T41	R2R3	335	<i>GjMYB82</i>	Gj9A1031T85	R2R3	232
<i>GjMYB30</i>	Gj8A84T56	1R	467	<i>GjMYB83</i>	Gj9A1036T88	3R	904
<i>GjMYB31</i>	Gj8A308T22	1R	204	<i>GjMYB84</i>	Gj9P1051T18	R2R3	518
<i>GjMYB32</i>	Gj8E331T0	1R	145	<i>GjMYB85</i>	Gj9A1052T114	R2R3	290
<i>GjMYB33</i>	Gj8A382T92	1R	88	<i>GjMYB86</i>	Gj9P1069T25	1R	309
<i>GjMYB34</i>	Gj8P396T16	1R	1 397	<i>GjMYB87</i>	Gj9A1083T89	R2R3	271
<i>GjMYB35</i>	Gj9A7T75	1R	280	<i>GjMYB88</i>	Gj9A1088T92	1R	336
<i>GjMYB36</i>	Gj9P11T26	1R	1 293	<i>GjMYB89</i>	Gj10A273T124	1R	497
<i>GjMYB37</i>	Gj9A69T86	R2R3	297	<i>GjMYB90</i>	Gj10P276T12	R2R3	400
<i>GjMYB38</i>	Gj9A76T57	3R	560	<i>GjMYB91</i>	Gj10A280T93	1R	293
<i>GjMYB39</i>	Gj9A81T123	1R	278	<i>GjMYB92</i>	Gj10P280T27	R2R3	223
<i>GjMYB40</i>	Gj9A87T89	1R	552	<i>GjMYB93</i>	Gj10P280T28	R2R3	203
<i>GjMYB41</i>	Gj9A130T84	R2R3	1 027	<i>GjMYB94</i>	Gj10X286T42	1R	528
<i>GjMYB42</i>	Gj9X141T53	1R	107	<i>GjMYB95</i>	Gj10X300T28	1R	86
<i>GjMYB43</i>	Gj9P154T10	1R	433	<i>GjMYB96</i>	Gj10X304T40	R2R3	354
<i>GjMYB44</i>	Gj9A202T95	1R	487	<i>GjMYB97</i>	Gj11A117T23	R2R3	295
<i>GjMYB45</i>	Gj9A212T108	R2R3	529	<i>GjMYB98</i>	Gj11A233T32	R2R3	491
<i>GjMYB46</i>	Gj9A213T96	R2R3	308	<i>GjMYB99</i>	Gj11A233T34	R2R3	487
<i>GjMYB47</i>	Gj9P220T9	1R	697	<i>GjMYB100</i>	Gj11X253T23	1R	118
<i>GjMYB48</i>	Gj9A225T92	1R	89	<i>GjMYB101</i>	Gj11A376T83	R2R3	316
<i>GjMYB49</i>	Gj9A286T43	1R	185	<i>GjMYB102</i>	Gj11A408T77	R2R3	348
<i>GjMYB50</i>	Gj9A312T88	3R	981	<i>GjMYB103</i>	Gj11A411T50	1R	454
<i>GjMYB51</i>	Gj9A564T94	R2R3	246	<i>GjMYB104</i>	Gj11A417T81	1R	333
<i>GjMYB52</i>	Gj9A584T54	R2R3	278	<i>GjMYB105</i>	Gj11A433T110	1R	205
<i>GjMYB53</i>	Gj9A610T142	1R	300				

**5 候选MYB基因qRT-PCR定量分析**

根据基因注释CDS序列对筛选的8个*GjMYB*基因进行qRT-PCR验证。结果显示,除*GjMYB19*和*GjMYB101*外,其余6个*GjMYB*基因的表达模式与转

录组数据分析结果相近(图5),即*GjMYB23*在花蕾器官中表达量最高,*GjMYB31*和*GjMYB73*在叶片中的表达量高于花蕾器官,且成熟叶片中表达量最高。*GjMYB9*、*GjMYB16*和*GjMYB76*与栀子苷合成特异的途



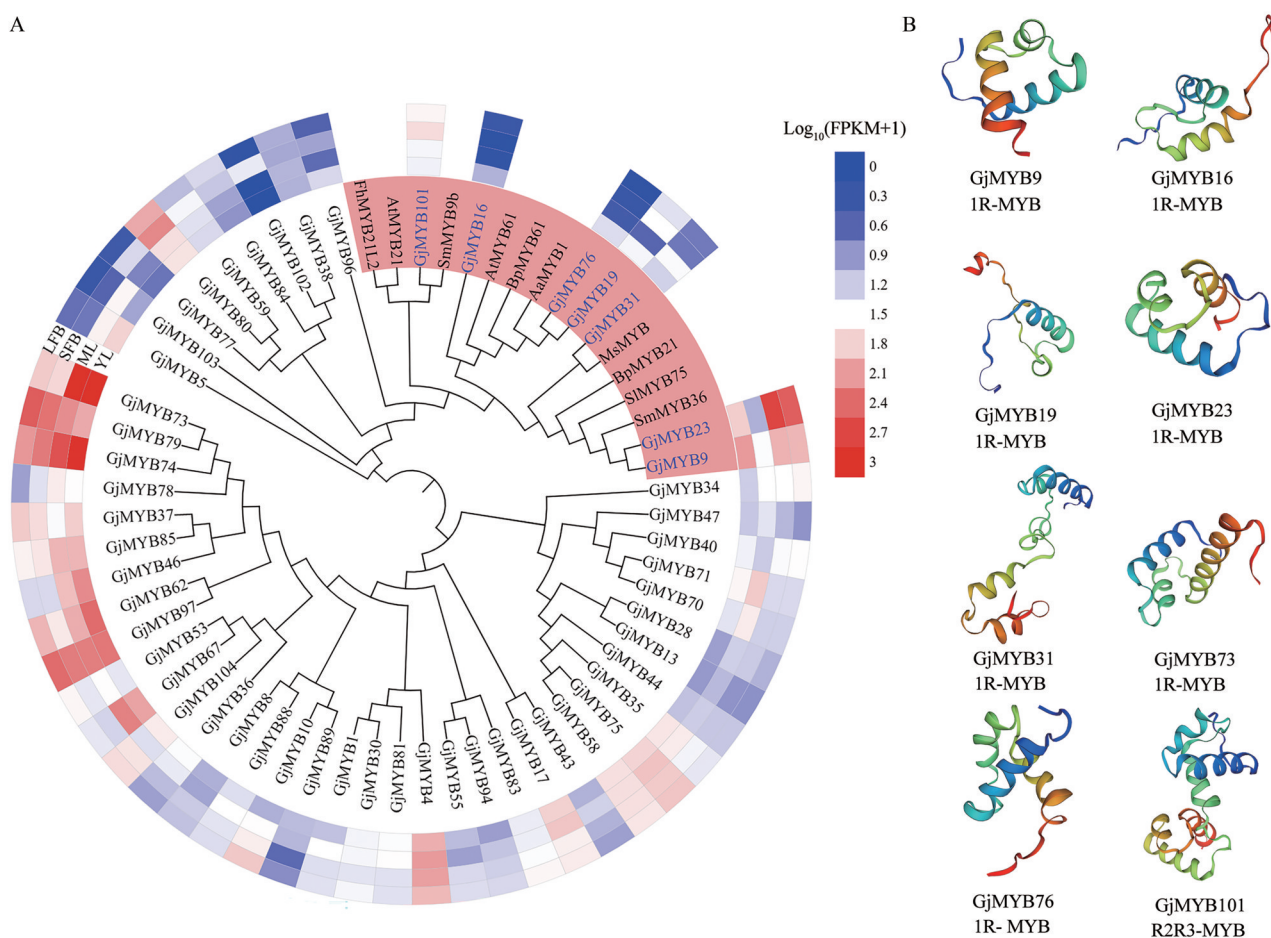
**Figure 3** Sequence analysis of MYB transcription factors in *Gardenia jasminoides*. A: The phylogenetic tree of GjMYBs; B: The motif distribution of GjMYBs; C: Four conserved motifs of GjMYBs

径基因的表达模式相关性相对较低, 且与梔子苷的器官积累规律无显著关联。据此, 推测 GjMYB23 是正向调控梔子苷生物合成的 MYB 转录因子, 而 GjMYB31 和 GjMYB73 可能参与该成分合成的负调控过程。

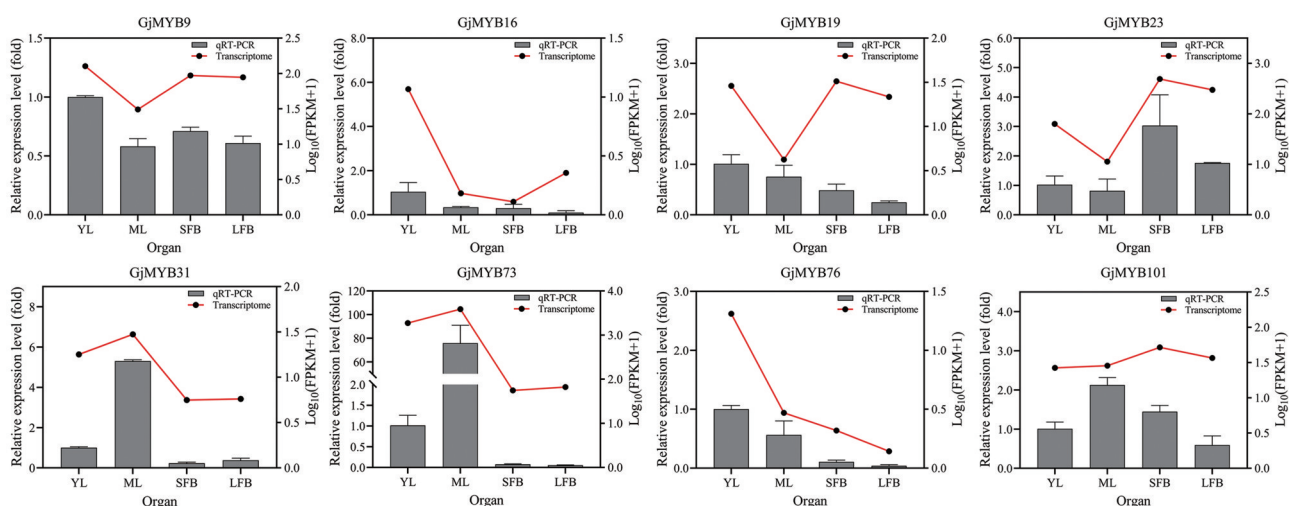
### 讨论

梔子是临床常用中药之一, 代表性活性成分为环

烯醚萜苷, 研究人员已从梔子中分离和鉴定出 48 种该类化合物<sup>[1]</sup>, 指标成分梔子苷是多种环烯醚萜苷的进一步修饰产物及合成底物。由于次生代谢产物合成基因及其表达调控因子是优良种质选育的重要遗传信息, 研究人员先后鉴定了梔子苷生物合成特异途径基因 *GES*、*IS*、*GAMT1* 和 *GAMT2*, 然而缺乏相应的基因表达调控机制研究。前期研究发现, 梔子的叶片和花蕾



**Figure 4** Analysis of MYB transcription factors related to geniposide biosynthesis. A: Phylogenetic tree and expression pattern of expressed *GjMYB* genes; B: Tertiary structures of selected eight *GjMYB* transcription factors



**Figure 5** Relative expression levels of selected MYB transcription factor genes related to geniposide biosynthesis in *Gardenia jasminoides*

器官中栀子苷含量存在较大差异, 活性成分积累呈现器官特异性分布。在此基础上, 本文进一步对栀子苷生源途径基因的表达模式进行分析发现, 所有栀子苷合成特异途径基因均为器官特异性表达, 且表达模式均与栀子苷积累规律相同。此外, *GES*、*IS* 及 2 个 *GAMT*

基因启动子区域包含多个 MYB 转录因子结合位点, 推测由于 MYB 转录因子调控栀子苷生物合成基因的器官特异性表达导致活性成分在特定部位合成。

多组学数据联合分析可全面且精准挖掘次生代谢产物合成基因及调控转录因子<sup>[22-24]</sup>, 栀子染色体水平

基因组<sup>[25]</sup>为系统分析栀子苷生物合成相关MYB转录因子提供良好的数据支撑。本研究通过转录组测序、多种生物信息学分析及qRT-PCR验证,共筛选到3个潜在的栀子苷生物合成相关MYB转录因子。首先,通过系统发育树分析发现,候选转录因子GjMYB23和GjMYB31与萜类成分合成相关MYB转录因子聚在一支,提示已知天然产物合成调控MYB转录因子序列存在一定保守性。qRT-PCR实验证实GjMYB23基因在花蕾中高表达,而GjMYB31基因在叶片中表达量相对较高,即两个基因的表达模式分别与栀子苷的器官积累规律呈现正相关与负相关,推测二者分别是栀子苷生物合成正调控和负调控转录因子。其次,鉴于栀子苷合成特异途径基因的表达及该成分的积累均呈现器官特异性,根据基因表达定量分析筛选出高表达的候选基因GjMYB73,其表达模式与栀子苷合成特异途径基因IS、GES和GAMTs的表达趋势及栀子苷含量分布为负相关,可能参与栀子苷生物合成负调控过程。此外,需要特别说明的是,GjMYB73蛋白序列与MYB转录因子家族中的RAD蛋白同源性较高,这导致其与萜类合成相关MYB转录因子亲缘关系较远。研究表明,MYB-RAD类蛋白参与对称花型发育过程<sup>[26,27]</sup>,但GjMYB73基因在花蕾中表达量极低,可能具有新的调控功能。

本研究筛选出的3个潜在栀子苷生物合成调控MYB转录因子均为1R-MYB转录因子。然而,多数研究发现R2R3-MYB型转录因子在植物天然产物合成中发挥重要作用<sup>[28]</sup>,其不仅可以直接调控靶基因表达,还可以与光和激素等协同参与次生代谢产物合成。Reddy等<sup>[17]</sup>在留兰香中发现了能直接与香叶基焦磷酸合酶(GPPS)编码基因的顺式作用元件结合,通过抑制GPPS表达负调控单萜合成的R2R3-MYB转录因子。Liu等<sup>[16]</sup>发现只有当黄花蒿中的R2R3型转录因子AaMYB108存在时,光和茉莉酸才能促进青蒿素合成。1R-MYB转录因子研究基础相对薄弱<sup>[29]</sup>,仅有少量研究发现其参与天然产物合成<sup>[30]</sup>。因此,本研究结果可能将进一步拓展1R-MYB转录因子的调控功能,下一步计划通过酵母单杂交或转基因技术等验证候选GjMYB转录因子是否参与栀子苷的生物合成调控过程。

**作者贡献:** 许文杰负责实验设计、数据分析、基因表达定量实验及文章撰写;黄远浩参与基因表达定量实验;韩蓉蓉、刘燕琴及曹敏负责栀子样品采集;宋经元负责总体实验设计及稿件修改等。

**利益冲突:** 本文无任何利益冲突。

## References

- [1] Chen L, Li M, Yang Z, et al. *Gardenia jasminoides* Ellis: ethnopharmacology, phytochemistry, and pharmacological and industrial applications of an important traditional Chinese medicine [J]. *J Ethnopharmacol*, 2020, 257: 112829.
- [2] Zhang W, Zhang F, Hu Q, et al. The emerging possibility of the use of geniposide in the treatment of cerebral diseases: a review [J]. *Chin Med*, 2021, 161: 86.
- [3] Zhou YX, Zhang RQ, Rahman K, et al. Diverse pharmacological activities and potential medicinal benefits of geniposide [J]. *Evid Based Complement Alternat Med*, 2019, 2019: 4925682.
- [4] Pan Y, Li L, Wang Y, et al. Investigation report on distribution and comprehensive utilization of *Gardenia jasminoides* cultivar resources in China [J]. *Nat Prod Res Dev* (天然产物研究与开发), 2019, 31: 1823-1830.
- [5] Cai Y, Zhang S, Zeng J, et al. Evaluation of germplasm resources of *Gardenia Fructus* based on principal component and hierarchical cluster analysis [J]. *Chin J Exp Tradit Med Form* (中国实验方剂学杂志), 2017, 23: 30-37.
- [6] Jiao YL, Tang ZZ, Tang YQ, et al. Analysis of geniposide and crocin contents in fruits of different leaf-fruit types of gardenia [J]. *Non-wood Forest Res* (经济林研究), 2017, 353: 193-198.
- [7] Ye P, Liang S, Wang X, et al. Transcriptome analysis and targeted metabolic profiling for pathway elucidation and identification of a geraniol synthase involved in iridoid biosynthesis from *Gardenia jasminoides* [J]. *Ind Crops Prod*, 2019, 132: 48-58.
- [8] Xu C, Ye P, Wu Q, et al. Identification and functional characterization of three iridoid synthases in *Gardenia jasminoides* [J]. *Planta*, 2022, 2553: 58.
- [9] Xu W, Lou Q, Hao L, et al. O-Methyltransferases catalyze the last step of geniposide biosynthesis in *Gardenia jasminoides* [J]. *Ind Crops Prod*, 2022, 177: 114438.
- [10] Wei X, Lan H. Advances in the regulation of plant MYB transcription factors in secondary metabolism and stress response [J]. *Biotechnol Bull* (生物技术通报), 2022, 38: 12-23.
- [11] Zhang C, Wang Y, Chen J, et al. Research advance on the regulation of second metabolism by plant MYB transcription factors [J]. *Genomics Applied Biol* (基因组学与应用生物学), 2020, 399: 4171-4177.
- [12] Katiyar A, Smita S, Lenka SK, et al. Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis* [J]. *BMC Genomics*, 2012, 13: 544.
- [13] Yang Z, Li Y, Gao F, et al. MYB21 interacts with MYC2 to control the expression of terpene synthase genes in flowers of *Freesia hybrida* and *Arabidopsis thaliana* [J]. *J Exp Bot*, 2020, 7114: 4140-4158.
- [14] Zhang J, Zhou L, Zheng X, et al. Overexpression of SmMYB9b enhances tanshinone concentration in *Salvia miltiorrhiza* hairy roots [J]. *Plant Cell Rep*, 2017, 36: 1297-1309.

- [15] Matias-Hernandez L, Jiang W, Yang K, et al. AaMYB1 and its orthologue AtMYB61 affect terpene metabolism and trichome development in *Artemisia annua* and *Arabidopsis thaliana* [J]. *Plant J*, 2017, 903: 520-534.
- [16] Liu H, Li L, Fu X, et al. AaMYB108 is the core factor integrating light and jasmonic acid signaling to regulate artemisinin biosynthesis in *Artemisia annua* [J]. *New Phytol*, 2023, 237: 2224-2237.
- [17] Reddy VA, Wang Q, Dhar N, et al. Spearmint R2R3-MYB transcription factor MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (MsGPPS.LSU) [J]. *Plant Biotechnol J*, 2017, 159: 1105-1119.
- [18] Ding K, Pei T, Bai Z, et al. SmMYB36, a novel R2R3-MYB transcription factor, enhances tanshinone accumulation and decreases phenolic acid content in *Salvia miltiorrhiza* hairy roots [J]. *Sci Rep*, 2017, 71: 5104.
- [19] Reeves PH, Ellis CM, Ploense SE, et al. A regulatory network for coordinated flower maturation [J]. *PLoS Genet*, 2012, 82: e1002506.
- [20] Jian W, Cao H, Yuan S, et al. SIMYB75, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits [J]. *Hortic Res*, 2019, 6: 22.
- [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, 201: 374.
- [22] Chu Y, Xiao S, Su H, et al. Genome-wide characterization and analysis of bHLH transcription factors in *Panax ginseng* [J]. *Acta Pharm Sin B*, 2018, 84: 666-677.
- [23] Ma F, Zhou H, Xu Y, et al. Comprehensive analysis of bZIP transcription factors in passion fruit [J]. *iScience*, 2023, 26: 106556.
- [24] Xu W, Lou Q, Chen Q, et al. Systematic analysis of TPS genes for the floral scent biosynthesis in *Gardenia jasminoides* based on whole genome [J]. *Plant Sci J (植物科学学报)*, 2023. DOI: 10.11913/PSJ.2095-0837.23017.
- [25] Xu Z, Pu X, Gao R, et al. Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants [J]. *BMC Biol*, 2020, 181: 63.
- [26] Corley S, Carpenter R, Copsey L, et al. Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum* [J]. *Proc Natl Acad Sci U S A*, 2005, 10214: 5068-5073.
- [27] Zhong J, Preston J, Hileman L, et al. Repeated and diverse losses of corolla bilateral symmetry in the Lamiaceae [J]. *Ann Bot*, 2017, 1197: 1211-1223.
- [28] Xie D, Ye Y, Yang D, et al. Identification and analysis of the R2R3-MYB transcription factor based on a transcriptome database in *Morinda officinalis* [J]. *Acta Pharm Sin (药学报)*, 2020, 55: 160-167.
- [29] Liu C, Yao P, Song H, et al. Identification and analysis of 1R MYB transcription factor subfamily in tobacco [J]. *Plant Physiol J (植物生理学报)*, 2022, 585: 904-918.
- [30] Yong Y, Zhang Y, Lyu Y. A MYB-related transcription factor from *Lilium lancifolium* L. (LlMYB3) is involved in anthocyanin biosynthesis pathway and enhances multiple abiotic stress tolerance in *Arabidopsis thaliana* [J]. *Int J Mol Sci*, 2019, 20: 3195.