

# 望春玉兰 WRKY 基因家族生物信息学分析

戴晓港, 刘景胜, 李淑娴\*

(南京林业大学林学院/南方杨树工程技术研究中心/江苏省杨树种质创新与品种改良重点实验室, 江苏 南京 210037)

**摘要:**WRKY 超级转录因子基因家族,参与植物生长、发育、非生物胁迫以及次生代谢产物的调节。该文鉴定了药用植物望春玉兰全基因组 WRKY 基因,并采用生物信息学方法,分析了该基因家族成员的蛋白理化性质、亚细胞定位、系统进化、保守结构域、顺式作用元件,以及在不同组织中的表达模式。在望春玉兰全基因组中共鉴定 56 个 WRKY 基因;根据进化分析的结果,将这些基因分为 3 个组(Group I—III),又将 Group II 进一步分为 5 个亚组(II a—II e)。该家族基因蛋白分子量在 12 215.15—191 326.88 Da 之间,理论等电点在 4.9—10.06 之间。亚细胞定位结果显示,55 个基因家族成员定位在细胞核,其中 1 个基因同时定位在细胞外间隙,还有 1 个基因定位在细胞质。顺式作用元件分析表明,望春玉兰 WRKY 基因除了广泛参与到调节非生物胁迫中,还参与次生代谢产物的合成调控。该文同时分析了 WRKY 基因在根、叶片和花中的表达模式,并鉴定了组织特异表达基因。通过对望春玉兰 WRKY 基因家族的生物信息学分析,为深入研究 WRKY 基因家族的功能提供了基础。

**关键词:**望春玉兰;WRKY 转录因子;进化分析;基因表达模式

**中图分类号:**Q756;S792.9.9 **文献标志码:**A **doi:**10.3969/j.issn.1001-7380.2021.06.001

## Bioinformatics analysis of WRKY gene family in *Yulania biondii*

Dai Xiaogang, Liu Jingsheng, Li Shuxian\*

(The Key Laboratory of Tree Genetic Improvement and Biotechnology of Jiangsu Province, The Southern Poplar Germplasm Engineering and Wood Processing Technology Center, College of Forestry, Nanjing Forestry University, Nanjing 210037, China)

**Abstract:**As a super transcription factor family, WRKY proteins were involved in regulating plant growth, development, abiotic stress response and secondary metabolism. In this study, genome-wide WRKY transcription factors were identified from a medical tree of *Yulania biondii*, and their physico-chemical properties, subcellular localization, phylogenetic evolution, conserved domain, cis-acting elements, gene expression pattern in different tissues were analyzed by bioinformatics. Fifty-six WRKY genes identified in *Y. biondii* genome were classified into 3 main groups (Group I—III), with 5 subgroups (IIa—IIe) in group II based on the evolution analysis result. The proteins varied ranging from 12 215.15 Da to 191 326.88 Da and theoretical pI varying from 4.9—10.06. Subcellular localization results showed that 55 genes were localized in nucleus, with one localized in extracellular space, another in cytoplasm simultaneously. Cis-acting element analysis showed that *Y. biondii* WRKY genes were not only found in being involved in abiotic stress response, but also regulating the synthesis of secondary metabolites. Gene expression pattern were analyzed with the RNA-seq data from root, leaf and flower, with the specific expression genes identified in different tissues. The bioinformatics analysis will provide the basis for further functional study of WRKY gene family in *Y. biondii*.

**Key words:***Yulania biondii*; WRKY transcription factor; Evolution analysis; Gene expression pattern

植物在整个生命活动过程中会不断受到各种生物胁迫(如食草动物和害虫的啃食、病原体感染)

收稿日期:2021-10-02;修回日期:2021-10-29

基金项目:国家自然科学基金项目“籼箕柳组培再生体系及农杆菌介导的遗传转化研究”(31500533)

作者简介:戴晓港(1984—),男,江苏新沂人,实验师,博士。主要从事林木遗传育种研究。E-mail: xgdai@njfu.edu.cn

\* 通信作者:李淑娴(1969—),女,山东潍坊人,教授,博士。主要从事观赏植物育种研究。E-mail: shuxianli@njfu.com.cn

和非生物胁迫(如干旱、低温、高温、土壤高盐碱等)的影响,然而植物不能自由移动来躲避这些不利环境,因此在漫长的进化过程中,植物逐渐从形态、生理、细胞以及分子水平发生进化以应对这些不利环境条件。当植物受到胁迫时会产生相应的应答反应,相关基因、蛋白质和代谢产物构成复杂调控网络以应答相应的胁迫。转录因子作为反式作用因子可以与靶标基因启动子区域相应的顺式作用元件特异结合,从而调控相关基因的转录,在植物应答生物胁迫和非生物胁迫中具有举足轻重的作用<sup>[1]</sup>。WRKY 转录因子是植物中的超级转录因子家族之一,其典型特征是在氨基酸 N-端有保守的 WRKYGQK 结构域,在 C-端含有 C<sub>2</sub>H<sub>2</sub>或 C<sub>2</sub>HC 型锌指结构<sup>[2]</sup>。根据上述保守结构域数量和锌指结构类型可将 WRKY 转录因子分为 3 组:第 I 组包含 2 个 WRKY 结构域和 C<sub>2</sub>H<sub>2</sub>型锌指结构;第 II 组含有 1 个 WRKY 结构域和 C<sub>2</sub>H<sub>2</sub>型锌指结构;而第 III 类是由 1 个 WRKY 保守结构域和 C<sub>2</sub>HC 型锌指结构组成<sup>[2]</sup>。大量研究表明,WRKY 在调控植物生长发育、生物胁迫和非生物胁迫<sup>[3-4]</sup>、次生代谢等多种生物过程中起着重要的作用<sup>[5-6]</sup>。如木本鸡脚棉中 GaWRKY1 转录因子通过调控卡丁烯合酶-A 的活性,从而参与棉酚等倍半萜烯类化合物的合成<sup>[5]</sup>;从黄花蒿分泌型腺毛中分离的 AaWRKY1 通过上调倍半萜环化酶基因 ADS 的表达,从而提高抗疟青蒿素的生物合成<sup>[6]</sup>。随着基因组测序的发展,药用植物越来越引起大家的关注,这也为鉴定更多的参与调控次生代谢的 WRKY 转录因子提供可能。

望春玉兰(*Yulania biondii*)是木兰科(Magnoliaceae)玉兰属多年生落叶乔木,是中国特有的珍稀濒危树种,原产湖北、河南、陕西、甘肃等地海拔 600—1 200 m 的阔叶林中<sup>[7]</sup>。望春玉兰先花后叶,种子鲜红,树形优美,被广泛应用于园林绿化<sup>[7]</sup>,其干燥的花蕾是中药材“辛夷”的正品,具有散风寒、通肺窍、降压、杀菌等作用,是一种用途广泛的药用植物<sup>[8]</sup>。望春玉兰具有一定的耐寒性,喜微酸性土壤,但碱性土壤对生长的影响较大。望春玉兰全基因组测序的完成为 WRKY 转录因子家族的鉴定和分析提供了基础,本研究对望春玉兰 WRKY 基因家族成员进行鉴定,同时对蛋白理化性质、亚细胞定位、系统进化、保守结构域、顺式作用元件、以及在不同组织中的表达模式进行了初步研究,以期为深入研究该基因家族在望春玉兰中应答非生物胁迫

以及次生代谢的调控奠定基础。

## 1 材料与方法

### 1.1 WRKY 转录因子查找及理化性质分析

从国家基因库生命大数据平台(<https://db.cngb.org/codeplot/>)中下载望春玉兰基因组和转录组原始数据(CNP0000884)<sup>[9]</sup>,利用 hmmsearch v3.3.2 将 WRKY 隐马尔可夫模型(Pfam: PF00847)和下载的望春玉兰蛋白序列进行比对,保留 E-value  $\leq 1e-5$  的蛋白序列,再采用 SMART(<http://smart.embl-heidelberg.de/>)进行保守结构域鉴定<sup>[10]</sup>,去除不含 WRKY 保守结构域的蛋白序列。采用 Expasy<sup>[11]</sup>(<https://web.expasy.org/protparam/>)在线预测蛋白分子量、等电点等理化性质,并利用 Cell-PLoc 2.0(<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>)对鉴定的 WRKY 基因进行亚细胞定位预测。

### 1.2 WRKY 转录因子进化树分析

从拟南芥基因组数据库(<https://www.arabidopsis.org/browse/genefamily/WRKY-Som.jsp>)下载拟南芥 WRKY 转录因子序列。根据 SMART 分析结果提取拟南芥和望春玉兰 WRKY 转录因子的保守结构域序列,并提交到 MEGA-6.0<sup>[12]</sup>软件中利用 ClustalW 进行多重比对,然后用邻接法(Neighbor Joining method)构建进化树,bootstrap 值设置为 1 000,最后采用 Figtree(<http://tree.bio.ed.ac.uk/software/figtree/>)对构建的进化树进行可视化,根据进化树中拟南芥 WRKY 家族的分类对望春玉兰 WRKY 家族进行类群的划分。

### 1.3 WRKY 转录因子保守基序和基因结构

利用 MEME(<https://meme-suite.org/meme/tools/meme>)在线预测望春玉兰 WRKY 蛋白的保守基序(Motif),Motif 个数设置为 10 个,基序长度为 6—70,其他参数均为默认值。根据 MEME 的分析结果,结合基因注释的 gff 文件,利用 TBtools<sup>[13]</sup>对保守基序、基因结构和进化树进行整合,绘制望春玉兰 WRKY 转录因子保守基序和基因结构图。

### 1.4 WRKY 转录因子顺式作用元件分析

利用 perl 脚本根据基因组注释的 gff 文件,从基因组中提取目的基因上游 2 000 bp 序列,用 Plant-CARE(<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)对启动子区域进行顺式作用元件分析,最后利用 TBtools<sup>[13]</sup>结合进化树并绘制顺式作用元件分布图。

1.5 WRKY 转录因子染色体定位及串联重复基因分析

根据基因组注释的 gff 文件,使用 TBtools<sup>[13]</sup> 绘制望春玉兰 WRKY 转录因子在染色体上的分布图。串联重复基因的鉴定参照 Lehti-Shiu 对毛果杨基因组串联重复基因的鉴定标准:2 个基因之间不超过 350 kb 且中间插入的基因不超过 10 个,进化树上它们还需要同属于 1 个亚家族<sup>[14]</sup>。

1.6 WRKY 转录因子表达分析

将从国家基因库生命大数据平台中下载的望春玉兰叶片(CNX0149886)和花(CNX0149883)的转录组,以及实验室测序的根系转录组,使用 Trimmomatic<sup>[15]</sup> 过滤低质量序列,然后利用 STAR<sup>[16]</sup> 默认参数将上述过滤后的序列分别比对到望春玉兰基因组,统计根、叶和花分别比对到每个 WRKY 基因的序列数量,采用 RPKM(reads per kilobase of transcripts per million mapped reads)对每个基因表达量进行均一化。

2 结果与分析

2.1 望春玉兰 WRKY 基因家族成员鉴定及理化性质分析

本研究共鉴定出望春玉兰 WRKY 基因家族有 56 个成员,各成员蛋白长度、理化性质、理论等电点和亚细胞定位等信息如表 1。从表 1 中结果可以看出,蛋白长度在 106—1 675 个氨基酸之间,不同基因氨基酸数量存在较大差异。通过 Expasy 在线分析了望春玉兰 WRKY 转录因子蛋白理化性质,该家族基因蛋白分子量在 12 215. 15—191 326. 88 Da 之间,理论等电点在 4. 9—10. 06 之间。亚细胞定位结果显示,55 个基因家族成员定位在细胞核,其中 MBI36866\_MAGBIO 成员同时定位在细胞外间隙和细胞核,而 MBI24213\_MAGBIO 还定位在细胞质。

表 1 望春玉兰 WRKY 家族基因基本信息及蛋白理化性质

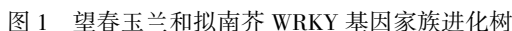
基因 ID	染色体	氨基酸数量	分子量/Da	等电点 pI	亚细胞定位
MBI12904_MAGBIO	Chr01	585	63 147. 24	6. 14	细胞核
MBI24029_MAGBIO	Chr01	187	21 159. 15	5. 44	细胞核
MBI07699_MAGBIO	Chr02	429	46 830	5. 65	细胞核
MBI11405_MAGBIO	Chr03	290	32 528. 75	5. 79	细胞核
MBI09114_MAGBIO	Chr03	375	40 796. 11	5. 63	细胞核
MBI11137_MAGBIO	Chr03	452	49 677. 58	5. 23	细胞核
MBI37592_MAGBIO	Chr04	328	36 595. 39	9. 18	细胞核
MBI39244_MAGBIO	Chr04	355	39 361. 62	9. 71	细胞核

续表 1					
基因 ID	染色体	氨基酸数量	分子量/Da	等电点 pI	亚细胞定位
MBI08947_MAGBIO	Chr04	106	12 690. 45	9. 82	细胞核
MBI13841_MAGBIO	Chr04	216	24 733. 38	7. 09	细胞核
MBI23915_MAGBIO	Chr04	201	22 612. 75	5. 84	细胞核
MBI05745_MAGBIO	Chr04	583	64 264. 68	6. 11	细胞核
MBI29696_MAGBIO	Chr05	243	27 023. 43	9. 45	细胞核
MBI46007_MAGBIO	Chr05	278	30 713. 39	6. 39	细胞核
MBI35288_MAGBIO	Chr06	348	39 374. 86	5. 85	细胞核
MBI18685_MAGBIO	Chr06	266	29 044. 98	5. 08	细胞核
MBI15279_MAGBIO	Chr07	314	34 790. 52	8. 24	细胞核
MBI37844_MAGBIO	Chr07	615	67 222. 51	5. 77	细胞核
MBI46335_MAGBIO	Chr07	163	18 410. 49	9. 66	细胞核
MBI24350_MAGBIO	Chr07	334	37 035. 95	4. 91	细胞核
MBI27415_MAGBIO	Chr10	313	34 920. 82	8. 19	细胞核
MBI35588_MAGBIO	Chr10	526	57 622. 25	5. 67	细胞核
MBI12228_MAGBIO	Chr11	789	85 650. 49	5. 79	细胞核
MBI05340_MAGBIO	Chr11	159	17 980. 3	9. 72	细胞核
MBI42321_MAGBIO	Chr11	401	45 047. 48	9. 52	细胞核
MBI36866_MAGBIO	Chr11	111	12 215. 51	8. 82	细胞外间隙/细胞核
MBI17784_MAGBIO	Chr14	473	51 808. 77	8. 64	细胞核
MBI10258_MAGBIO	Chr14	731	78 882. 78	6. 17	细胞核
MBI09154_MAGBIO	Chr14	300	32 965. 69	10. 04	细胞核
MBI34974_MAGBIO	Chr14	303	33 048. 61	9. 65	细胞核
MBI45550_MAGBIO	Chr14	229	25 920. 22	9. 05	细胞核
MBI37104_MAGBIO	Chr14	241	27 151. 67	9. 13	细胞核
MBI46238_MAGBIO	Chr14	246	27 915. 72	9. 61	细胞核
MBI32392_MAGBIO	Chr14	309	34 144. 25	7. 61	细胞核
MBI05415_MAGBIO	Chr14	281	32 412. 38	10. 06	细胞核
MBI36009_MAGBIO	Chr15	1 043	113 470. 85	5. 87	细胞核
MBI47426_MAGBIO	Chr15	770	83 026. 65	6. 09	细胞核
MBI36613_MAGBIO	Chr15	448	48 967. 8	7. 23	细胞核
MBI39387_MAGBIO	Chr15	314	34 869. 72	4. 98	细胞核
MBI39388_MAGBIO	Chr15	345	37 592. 92	6	细胞核
MBI01639_MAGBIO	Chr15	242	26 455. 98	5. 99	细胞核
MBI46105_MAGBIO	Chr15	316	34 583. 62	8. 35	细胞核
MBI19961_MAGBIO	Chr15	299	33 062. 34	8. 74	细胞核
MBI17574_MAGBIO	Chr15	177	20 263. 22	8. 74	细胞核
MBI10883_MAGBIO	Chr15	231	26 886. 92	9. 34	细胞核
MBI10872_MAGBIO	Chr15	304	34 250. 44	9. 54	细胞核
MBI36007_MAGBIO	Chr15	169	18 632. 42	5. 86	细胞核
MBI09342_MAGBIO	Chr16	301	32 655. 16	9. 73	细胞核
MBI09906_MAGBIO	Chr16	276	30 372. 54	7. 18	细胞核
MBI27140_MAGBIO	Chr18	428	46 772. 17	5. 22	细胞核
MBI10366_MAGBIO	Chr18	503	55 363. 99	7. 62	细胞核
MBI24213_MAGBIO	Chr18	1 675	191 326. 88	5. 84	细胞质
MBI17309_MAGBIO	Chr18	221	24 692. 53	5. 99	细胞核
MBI35846_MAGBIO	Chr19	425	45 869. 67	6. 09	细胞核
MBI03263_MAGBIO	Chr19	278	31 360. 08	6. 61	细胞核



b, II c, II d, II e 共 5 个亚组, 其成员数量分别为 9, 7, 10, 7, 4 个; Group III 虽然也只含有 1 个 WRKY 保守结构域, 但其锌指结构 ( $C_2HC$ ) 和其他 2 组不同, 该组中共有 7 个基因。还有 3 个基因 (*MBI36007\_MAGBIO*、*MBI46007\_MAGBIO* 和 *MBI29696\_MAGBIO*) 虽然含有 WRKY 保守结构域, 但序列与其他基因分化较大而未被分类到现有的 I, II 和 III 组中 (见图 1)。

b, II c, II d, II e 共 5 个亚组, 其成员数量分别为 9, 7, 10, 7, 4 个; Group III 虽然也只含有 1 个 WRKY 保守结构域, 但其锌指结构 ( $C_2HC$ ) 和其他 2 组不同, 该组中共有 7 个基因。还有 3 个基因 (*MBI36007\_MAGBIO*、*MBI46007\_MAGBIO* 和 *MBI29696\_MAGBIO*) 虽然含有 WRKY 保守结构域, 但序列与其他基因分化较大而未被分类到现有的 I, II 和 III 组中 (见图 1)。



### 2.3 WRKY 转录因子蛋白基序分析

利用 MEME 在线分析望春玉兰 WRKY 基因家族的 Motif,利用 TBTools 对 Motif 分析结果进行可视化(见图 2)。望春玉兰 WRKY 蛋白共包含 10 个 Motif,其中 Motif 1, Motif 2 和 Motif 3 是 WRKY 的保守结构域, Motif 1 含有保守序列 WRKYGQK,望春玉兰所有 WRKY 蛋白都含有这个保守结构域; Motif 1 尾端 2 个蛋白和 Motif 2 构成了 WRKY 结构域的锌指结构( $C_2H_2$ ),其中 48 个基因含有 Motif 2 结构域; Motif 3 是

WRKY 的完整结构域,其中 7 个基因同时含有 Motif 3 和 Motif 1,这类基因因含有 2 个保守的 WRKY 结构域而被划分为 Group I。Motif 4—Motif 10 虽然在数据库中暂无功能记录,但这些 Motif 也为望春玉兰 WRKY 基因分类提供了重要参考,如 Motif 8 和 Motif 10 是 Group I 特有的,而 Motif 5, Motif 6 和 Motif 9 是 Group IIa 和 Group IIb 中特有。总体而言,望春玉兰 WRKY 基因家族不同分组中 Motif 是相似的,这个结果也有利地支持了进化树对基因家族分类的可靠性。

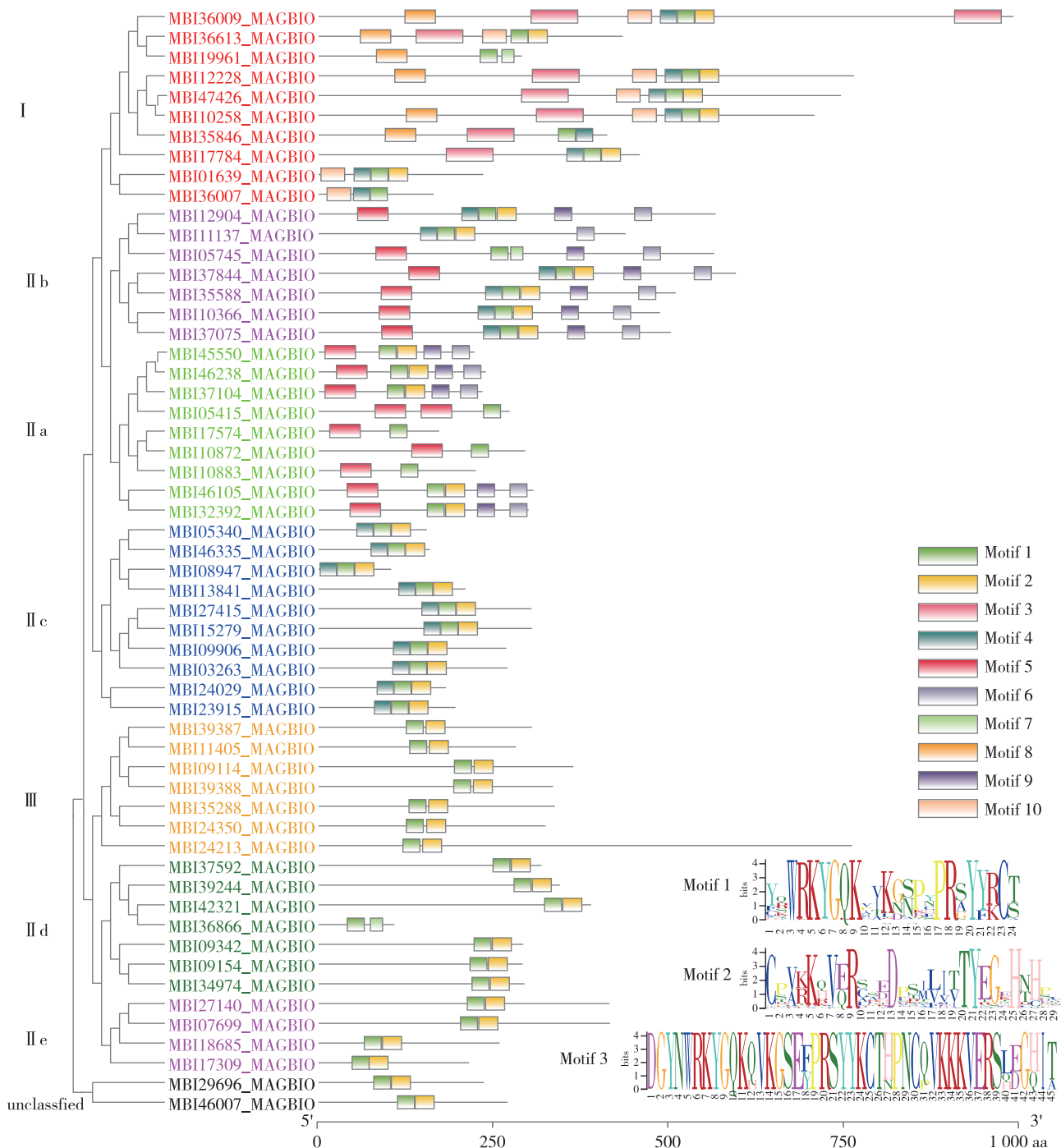


图 2 望春玉兰 WRKY 基因家族蛋白保守结构域

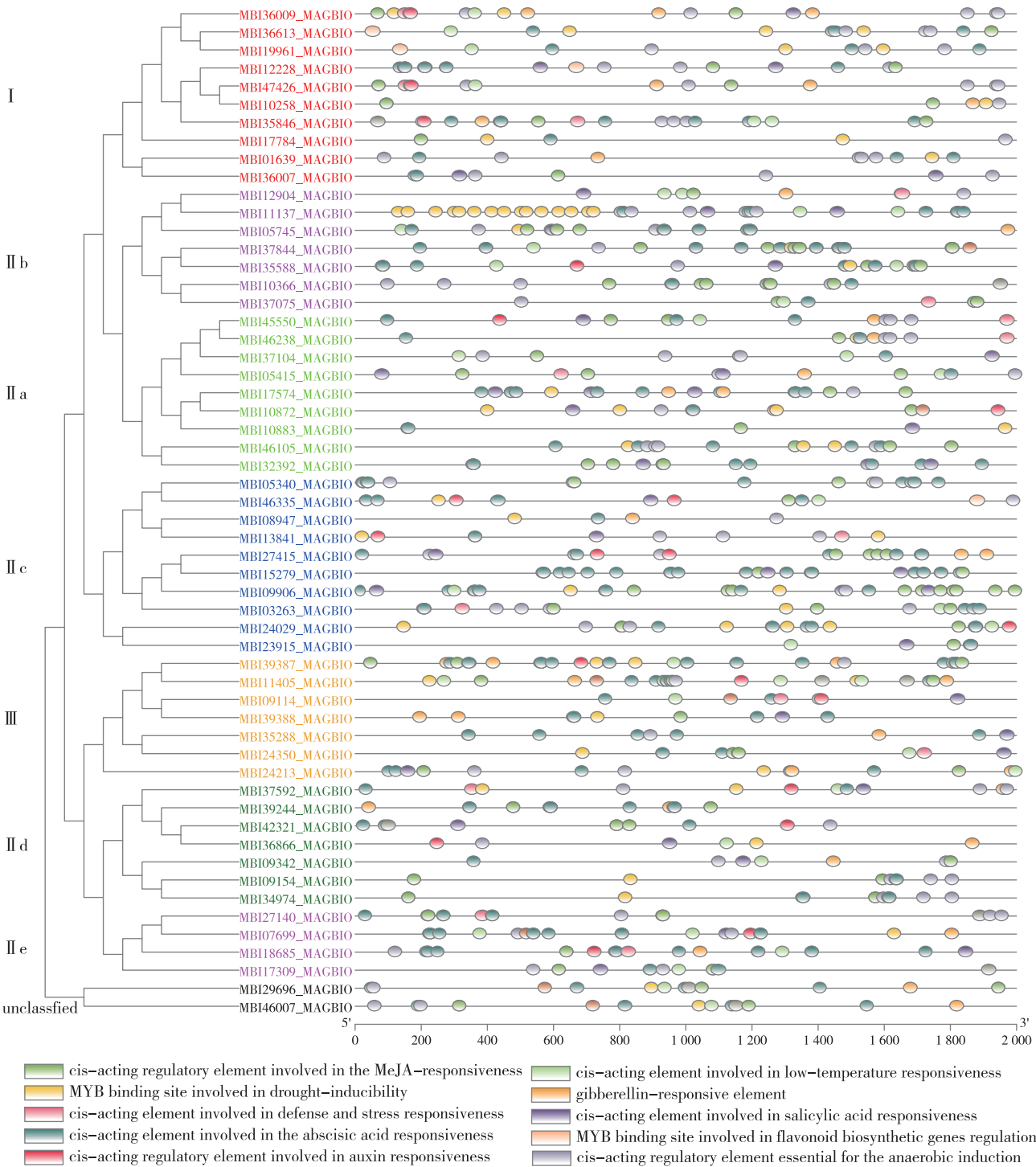


图 3 望春玉兰 WRKY 基因家族成员启动子区域顺式作用元件

2.4 WRKY 转录因子启动子区域顺式作用元件

利用 PlantCARE 注释望春玉兰 56 个 *WRKY* 基因上游 2 000 bp 启动子区域序列,得到 *WRKY* 基因家族顺式作用元件,除最基本的 TATA box,CAAT box 和光响应元件之外,还含有参与各种激素、调控生长、参与次生代谢和非生物胁迫等响应元件。在激素类

响应元件中,分别有 53 个和 47 个基因含有脱落酸反应元件(abscisic acid responsiveness)和茉莉酸甲酯反应元件(MeJA-responsiveness);应答水杨酸、赤霉素和生长素响应元件的基因相对较少,分别有 32,26,18 个。参与非生物胁迫相关的主要有无氧诱导(the anaerobic induction)、干旱诱导(drought-inducibility)、低

温响应 (low-temperature responsiveness) 和防御胁迫 (defense and stress responsiveness) 等所必需的调节元件。*MBI11137\_MAGBIO* 含有多个干旱诱导响应顺式作用元件,推测该基因可能调控望春玉兰响应干旱胁迫。值得注意的是望春玉兰 *WRKY* 基因除了参与激素响应和非生物胁迫的顺式作用元件,其中还有 9 个基因 (*MBI08947\_MAGBIO*, *MBI10258\_MAGBIO*, *MBI10883\_MAGBIO*, *MBI12228\_MAGBIO*, *MBI19961\_MAGBIO*, *MBI36613\_MAGBIO*, *MBI36866\_MAGBIO*, *MBI42321\_MAGBIO*, *MBI46335\_MAGBIO*) 含有类黄酮合成调节元件,参与植物的次生代谢。望春玉兰的花蕾和根系入药,表明了 *WRKY* 基因可能参与了次生代谢产物的合成调节。

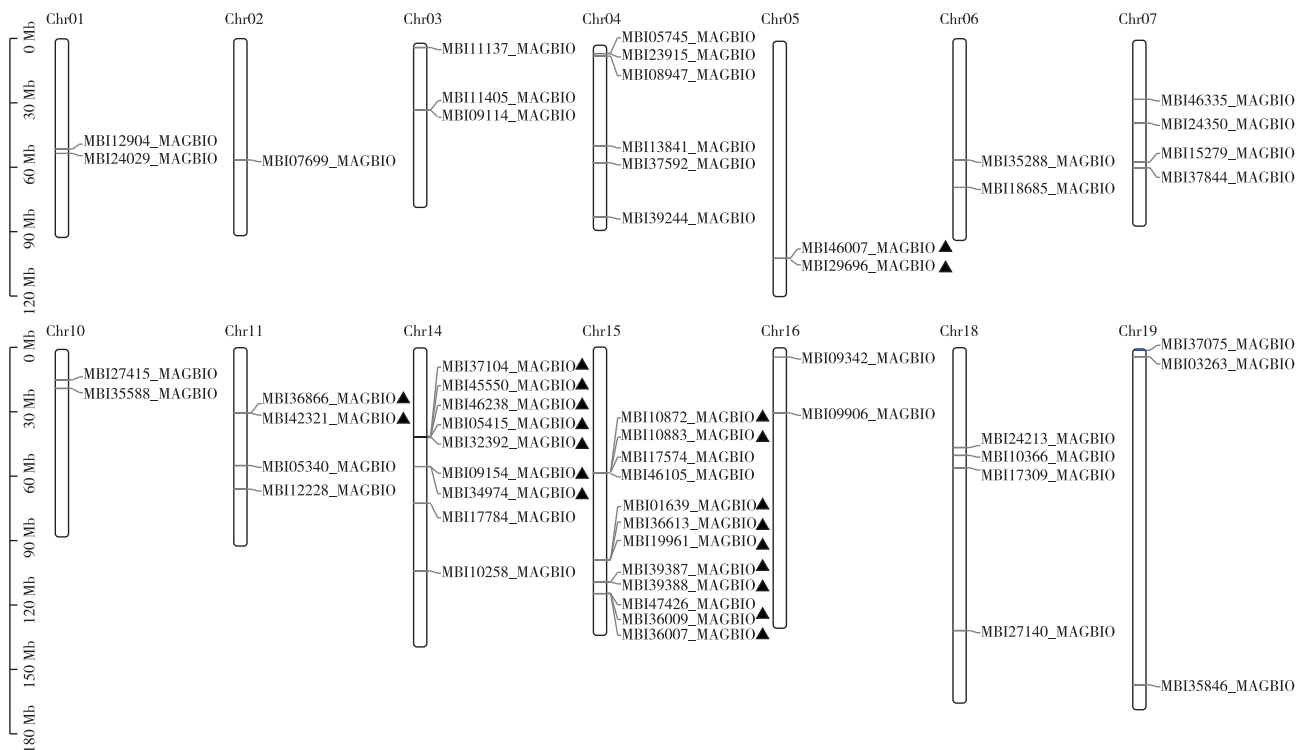


图4 望春玉兰 *WRKY* 基因家族在染色体上的分布(串联复制)

## 2.6 望春玉兰 *WRKY* 基因家族组织表达分析

基因表达模式可以揭示植物不同组织的生物学功能。将望春玉兰叶片、花和根系转录组测序序列分别比对到基因组,分析 *WRKY* 家族基因在不同组织中的表达模式,结果如图 5 所示。从图 5 中可以看出,Group I 中除了 *MBI36007\_MAGBIO* 在根系中不表达,Group IId 中除了 *MBI36007\_MAGBIO* 在根系和叶片中不表达外,其余基因在所有组织中均检测到表达,且 Group IId 中的基因整体表达量均较高。Group IIa 中只有 *MBI32392\_MAGBIO* 和

## 2.5 *WRKY* 基因家族染色体分布

望春玉兰基因组共有 19 条染色体,而 *WRKY* 基因只分布在其中 14 条染色体上(见图 4),Chr08, Chr09, Chr12, Chr13 和 Chr17 均不含有 *WRKY* 基因。在含有 *WRKY* 基因的染色体中,Chr15, Chr14 和 Chr04 这 3 条染色体含有 *WRKY* 基因的数量约占总数的 48.2%,分别有 12, 9, 6 个;其余 11 条染色体 *WRKY* 基因的数量在 1—3 个不等。串联重复是导致基因家族扩张的主要方式之一,对望春玉兰 *WRKY* 基因复制方式分析发现,这个基因家族中有 9 个串联复制事件,包含 20 个 *WRKY* 基因,其中 7 次串联复制发生在 Chr14 和 Chr15 上,这也是引起 *WRKY* 基因家族成员在染色体上分布不均匀的主要原因。

*MBI46105\_MAGBIO* 在 3 个组织中均能检测到表达,有 5 个基因在 3 个组织中均未检测到表达,而 *MBI37104\_MAGBIO* 和 *MBI05415\_MAGBIO* 是叶片组织特异性基因,只在叶片中有表达量。Group IIb 中有 1 个基因 *MBI37844\_MAGBIO* 在 3 个组织中表达量均较高,FPKM 值均超过 10,还有 3 个基因只在根系特异表达。Group IIc 中除了 *MBI24029\_MAGBIO* 在叶片中特异表达外,其余基因在不同的组织中也都能检测到表达。

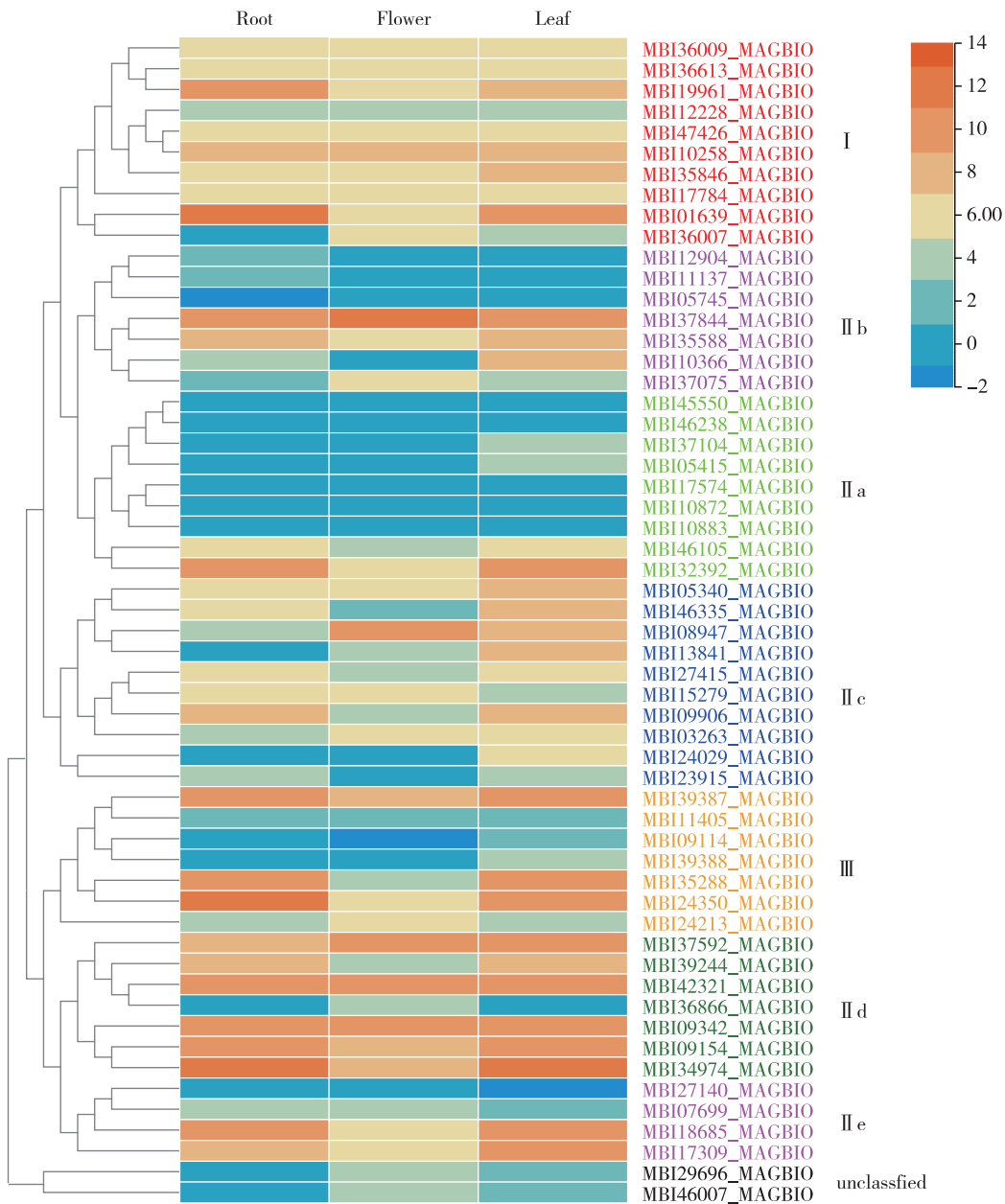


图 5 望春玉兰不同组织 WRKY 基因家族表达模式

3 讨论

WRKY 是植物中较大的基因家族,已被证实在调节植物生长、发育、非生物胁迫中起到重要作用。随着全基因组测序发展,已有几十种双子叶植物全基因组 WRKY 基因家族被鉴定和分析,而其基因家族中基因数量在 45—188 个不等<sup>[17]</sup>。本研究从望春玉兰基因组中鉴定了 56 个 WRKY 基因,比较分析发现,望春玉兰基因组(2.20 Gb)大小约是拟南芥(125 Mb)的 18 倍,但望春玉兰 WRKY 基因的数量(56 个)却只有拟南芥(74 个)的 3/4;同样,望春玉

兰基因组是木本植物毛果杨(0.55 Gb)的 4 倍,但 WRKY 基因的数量却不足毛果杨(122 个)的 1/2<sup>[17]</sup>。上述结果说明了基因组中,基因家族基因的数量并不是由基因组大小决定的。

根据 WRKY 基因保守结构域序列和锌指结构类型,同时结合拟南芥 WRKY 基因家族的分类,将望春玉兰 WRKY 基因家族分为 3 个组:Group I, Group II 和 Group III,分别含有 8,38 和 7 个基因。然而在拟南芥和毛果杨中,Group I 基因的数量占主导地位<sup>[18]</sup>,而望春玉兰中 Group II 基因数量占 67.9%,说明 Group II 在进化过程中可能发生了更



多的基因复制。这与对望春玉兰 WRKY 基因家族串联复制分析结果一致,在望春玉兰 WRKY 基因家族中共发现 9 个串联重复,其中 Group II 中有 5 个串联重复共产生了 11 个基因。望春玉兰 WRKY 基因家族 Group II 可进一步细分为 IIa, IIb, IIc, IId 和 IIe 共 5 个亚组,分别含有 9, 7, 10, 7, 4 个基因。对拟南芥、水稻、葡萄等 7 种植物基因组 WRKY 基因家族 Group II 基因数量分析发现, Group IIa 基因的比例为 8.8%—12.2% 之间<sup>[18]</sup>, 而望春玉兰 Group IIa 基因的比例为 23.7%, 约是其他物种的 2—3 倍, 说明 Group IIa 在望春玉兰中发生了大量扩张。已有研究表明, WRKY 基因家族的 Group IIa 基因可以提高拟南芥苗期和苹果愈伤的渗透胁迫能力和抗盐性<sup>[19]</sup>, 也是调节水稻先天免疫能力的重要转录因子<sup>[20]</sup>, 望春玉兰 Group IIa 基因的大量扩张, 可能是在长期进化过程中提高非生物胁迫的表现。

WRKY 不仅调节植物生长、发育和参与非生物胁迫的调节, 还参与调控植物的次生代谢。Xu 等在棉花中克隆的 GaWRKY1 转录因子可以调控卡丁烯合酶-A 的活性, 从而参与调节棉花倍半萜烯类化合物的合成<sup>[5]</sup>; Liu 等研究发现药用植物杜仲 WRKY 转录因子参与苯基丙氨酸和黄酮类物质合成代谢的调控<sup>[17]</sup>。望春玉兰转录因子顺式作用元件分析发现, WRKY 启动子区域不仅含有大量参与非生物胁迫的顺式作用元件, 还有 9 个基因含有参与黄酮类物质合成代谢的顺式作用元件, 其中 MBI19961\_MAGBIO 和 MBI08947\_MAGBIO 分别在根系和花中高表达, 而根系和花是望春玉兰药物合成的主要组织, 因此这 2 个基因可能影响望春玉兰根系和花的次生代谢。

#### 参考文献:

- [1] 李书粉, 孙富丛, 肖理慧, 等. 植物对非生物胁迫应答的转录因子及调控机制[J]. 西北植物学报, 2006, 26(6): 1295-1300.
- [2] EULGEM T, RUSHTON P J, ROBATZEK S, et al. The WRKY superfamily of plant transcription factors[J]. Trends in Plant Science, 2000, 5(5): 199-206.
- [3] EULGEM T, SOMSSICH I E. Networks of WRKY transcription factors in defense signaling[J]. Current Opinion in Plant Biology, 2007, 10(4): 366-371.
- [4] REN X, CHEN Z, LIU Y, et al. ABO3, a WRKY transcription factor, mediates plant responses to abscisic acid and drought tolerance in *Arabidopsis* [J]. The Plant Journal, 2010, 63(3): 417-429.
- [5] XU Y H, WANG J W, WANG S, et al. Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)- $\delta$ -cadinene synthase-A[J]. Plant Physiology, 2004, 135(1): 507-515.
- [6] MA D, PU G, LEI C, et al. Isolation and characterization of AaWRKY1, an *Artemisia annua* transcription factor that regulates the amorpho-4, 11-diene synthase gene, a key gene of artemisinin biosynthesis[J]. Plant and Cell Physiology, 2009, 50(12): 2146-2161.
- [7] 张庆宝, 申亚梅, 范义荣. 木兰属(*Magnolia*) 观赏植物育种现状及育种策略[J]. 江苏林业科技, 2008, 35(6): 46-48.
- [8] 中华人民共和国卫生部药典委员会. 中国药典: 第一版[M]. 北京: 人民卫生出版社, 1977.
- [9] DONG S, LIU M, LIU Y, et al. The genome of *Magnolia biondii* Pamp. provides insights into the evolution of Magnoliales and biosynthesis of terpenoids[J]. Horticulture Research, 2021, 8(1): 1-13.
- [10] LETUNIC I, DOERKS T, BORK P. SMART: recent updates, new developments and status in 2015 [J]. Nucleic Acids Research, 2015, 43(Database issue): 257-260.
- [11] ARTIMO P, JONNALAGEDDA M, ARNOLD K, et al. ExPASy: SIB bioinformatics resource portal[J]. Nucleic Acids Research, 2012, 40(W1): 597-603.
- [12] TAMURA K, STECHER G, PETERSON D, et al. MEGA6: Molecular evolutionary genetics analysis version 6.0 [J]. Molecular Biology & Evolution, 2013, 30(12): 1-12.
- [13] CHEN C, CHEN H, ZHANG Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. Molecular Plant, 2020, 13(8): 1194-1202.
- [14] LEHTI-SHIU M D, ZOU C, HANADA K, et al. Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes [J]. Plant Physiology, 2009, 150(1): 12-26.
- [15] BOLGER A M, LOHSE M, U SADEL B. Trimmomatic: a flexible trimmer for Illumina sequence data[J]. Bioinformatics. 2014; 30(15): 2114-2120.
- [16] DOBIN A, DAVIS C A, SCHLESINGER F, et al. STAR: ultrafast universal RNA-seq aligner[J]. Bioinformatics, 2013, 29(1): 15-21.
- [17] LIU J, WANG X, CHEN Y, et al. Identification, evolution and expression analysis of WRKY gene family in *Eucommia ulmoides* [J]. Genomics, 2021, 113: 3294-3309.
- [18] 宋 辉, 南志标. 蒺藜苜蓿全基因组中 WRKY 转录因子的鉴定与分析[J]. 遗传, 2014, 36(2): 152-168.
- [19] DONG Q, ZHENG W, DUAN D, et al. MdWRKY30, a group IIa WRKY gene from apple, confers tolerance to salinity and osmotic stresses in transgenic apple callus and *Arabidopsis* seedlings [J]. Plant Science, 2020, 299: 110-117.
- [20] PENG Y, BARTLEY L E, CANLAS P, et al. OsWRKY IIa transcription factors modulate rice innate immunity[J]. Rice, 2010, 3(1): 36-42.