

## • 药材与资源 •

颠茄生长素应答因子 *ARF* 基因家族的全基因组鉴定及表达特征分析

周省委<sup>1,3</sup>, 陈璐<sup>1,3</sup>, 顾旭鹏<sup>1,3</sup>, 张韶珂<sup>1,3</sup>, 初雷霞<sup>1,3</sup>, 宋腾达<sup>1,3</sup>, 万婕<sup>2</sup>, 杨林林<sup>1,3,4,5\*</sup>, 董诚明<sup>1,3,5</sup>, 冯卫生<sup>1,4,5</sup>

1. 河南中医药大学药学院, 河南 郑州 450046
2. 河南中医药大学第一附属医院, 河南 郑州 450046
3. 河南省道地药材生态种植工程技术研究中心, 河南 郑州 450046
4. 河南省中药开发工程技术研究中心, 河南 郑州 450046
5. 呼吸疾病中医药防治省部共建协同创新中心, 河南 郑州 450046

**摘要:**目的 生长素应答因子 (auxin response factors, *ARFs*) 基因家族参与颠茄 *Atropa belladonna* 生长发育的调控, 对颠茄 *ARF* 基因 (*AbARF*) 进行全基因组鉴定及表达特征分析。方法 基于已发布的颠茄全基因组, 利用生物信息学方法对 *AbARF* 基因家族成员进行鉴定并对其理化性质、染色体分布、系统进化、基因结构、顺式作用元件进行分析, 利用实时荧光定量 PCR (qRT-PCR) 分析其在不同温度、光质、组织部位的表达特征。结果 共鉴定出 41 个 *AbARF* 基因 (*AbARF1*~*AbARF41*), 分为 4 个亚家族, 编码的氨基酸长度在 603~1 116 aa, 蛋白质相对分子质量为 68 709.35~125 100.60, 等电点介于 5.27~8.25, 亚细胞定位全部定位于细胞核。基因结构分析发现所有 *AbARF* 基因均含有 B3 保守结构域, 顺式元件预测结果表示光响应元件和脱落酸元件在 *AbARF* 基因家族启动子区域分布广泛。基因组内共线性分析表明, 全基因组复制和片段性复制在 *AbARF* 基因家族进化中发挥了重要作用, 复制过后经过了强烈的纯化选择。qRT-PCR 表达特征分析结果表明, 大部分 *AbARF* 在低温 (18 °C) 胁迫下表达量显著下调, 在白光调控下表达量显著高于红光和黄光, 多数 *AbARF* 基因在叶中表达量普遍高于主根、侧根、茎和花。结论 从全基因组水平对颠茄 *ARF* 基因家族进行鉴定和生物信息学分析, 为进一步阐明颠茄 *ARF* 基因的功能奠定基础。

**关键词:** 颠茄; *ARF* 转录因子; 生物信息学; 基因家族; 共线性分析; 表达特征

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2024)21-7419-16

DOI: 10.7501/j.issn.0253-2670.2024.21.021

## Genome-wide identification and expression profile analysis of *Atropa belladonna* auxin response factor *ARF* gene family

ZHOU Shengwei<sup>1,3</sup>, CHEN Lu<sup>1,3</sup>, GU Xupeng<sup>1,3</sup>, ZHANG Shaoke<sup>1,3</sup>, CHU Leixia<sup>1,3</sup>, SONG Tengda<sup>1,3</sup>, WAN Jie<sup>2</sup>, YANG Linlin<sup>1,3,4,5</sup>, DONG Chengming<sup>1,3,5</sup>, FENG Weisheng<sup>1,4,5</sup>

1. School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou 450046, China
2. The First Affiliated Hospital of Henan University of Traditional Chinese Medicine Zhengzhou 450046, China
3. Henan Provincial Ecological Planting Engineering Technology Research Center of Authentic Medicinal Materials, Zhengzhou 450046, China
4. Engineering and Technology Center for Chinese Medicine Development of Henan Province, Zhengzhou 450046, China
5. Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases by Henan Province and Ministry of Education of China, Zhengzhou 450046, China

**Abstract: Objective** The auxin response factors (ARFs) gene family is involved in the regulation of growth and development of

收稿日期: 2024-05-05

基金项目: 国家自然科学基金资助项目 (82104329, 32401226); 中国博士后科学基金特别资助项目 (2014T170252); 河南省科技攻关 (242102310549); 河南省重点研发专项 (231111312700, 241111310200); 河南省国际科技合作项目 (242102521056)

作者简介: 周省委, 男, 硕士研究生, 研究方向为药用植物生理与生态学。E-mail: zhou19991210@yeah.net

\*通信作者: 杨林林, 男, 博士, 副教授, 从事药用植物生态学研究。E-mail: yangll-hatcm@hactcm.edu.cn

*Atropa belladonna*. Genome-wide identification and expression profile analysis of the belladonna *ARF* gene (*AbARF*) can lay a scientific foundation for elucidating the growth and development functions and mechanism of the *AbARF* gene family. **Methods** Based on the published whole genome data of *A. belladonna*, bioinformatics methods were used to identify the *AbARF* gene family members and analyze their physical and chemical properties, chromosomal distribution, phylogenetic evolution, gene structure, and cis-acting elements, using real-time PCR (qRT-PCR) to analyze its expression characteristics at different temperatures, light quality, and tissue locations. **Results** A total of 41 *A. belladonna ARF* genes (*AbARF1—AbARF41*) were identified, divided into four subfamilies, The length of the encoded amino acids ranges from 603 to 1 116 aa, the relative molecular mass of the protein is 68 709.35 to 125 100.60, the protein isoelectric point ranges from 5.27 to 8.25, and all subcellular locations are located in the nucleus. Gene structure analysis found that all *ARF* proteins contain the B3 conserved domain, the prediction results of cis-elements indicate that light-responsive elements and abscisic acid elements are widely distributed in the promoter region of the *AbARF* gene family. Intra-genomic collinearity analysis showed that whole-genome duplication and segmental duplication played an important role in the evolution of the *AbARF* gene family, and that duplication underwent strong purifying selection. The results of qPCR expression profile analysis showed that the expression levels of most *AbARFs* were significantly down-regulated under low temperature (18 °C) stress. The expression levels under white light regulation were significantly higher than those under red and yellow light. The expression levels of most *AbARF* genes in leaves were generally higher than those in main roots, lateral roots, stems, and flowers. **Conclusion** Identification and bioinformatics analysis of the *ARF* gene family at the whole genome level will lay the foundation for further elucidating the function of the *A. belladonna ARF* genes.

**Keywords:** *Atropa belladonna* L.; *ARF* transcription factor; bioinformatics; gene family; collinearity analysis; expression characteristics

颠茄为茄科颠茄属植物颠茄 *Atropa belladonna* L. 的干燥全草<sup>[1]</sup>。原产于欧洲, 现在我国河南、山东、湖南等省有栽培<sup>[2]</sup>。颠茄主要药效成分是以莨菪碱和东莨菪碱为主的次生代谢产物, 主要用于胆、胃绞痛和晕动症以及迷走神经兴奋的导致的多汗、呕吐等<sup>[3-4]</sup>。植物次生代谢产物是植物为了抵御环境胁迫维持自身生长发育而产生的物质<sup>[5]</sup>。植物对环境胁迫的适应受细胞分子网络级联控制。它们激活应激反应机制以重新建立内稳态, 保护和修复受损的蛋白质和细胞膜<sup>[6]</sup>, 其中以转录因子 (transcription factors, TFs) 为首的调控是细胞内部调控网络中的重要一环, 转录因子能够通过自身合成蛋白质, 并与 DNA 模板相结合, 从而调控下游基因的表达<sup>[7]</sup>, 其具有 2 个基本功能: 识别和结合靶基因上的特定 DNA 序列和招募特定结构的蛋白质以促进下游反应<sup>[8-9]</sup>。因此转录因子可以参与调节植物的生长发育、非生物胁迫及对代谢网络的响应。

生长素应答因子 (auxin response factors, *ARFs*) 是植物中重要的一类转录因子, 具有 3 个主要结构域: 位于 N 端植物特异性 B3 型 DNA 结合域 (DBD), 具有高度保守性, 仅使用 DBD 蛋白序列构建的系统发育树与使用完整蛋白序列构建的系统发育树非常相似, 表明该结构域有助于功能特异性<sup>[10-11]</sup>, 此结构域能与生长素应答基因启动子中的 TGTCTC 生长素应答元件 (AuxREs) 结合<sup>[12-13]</sup>; 中间区域 (MR) 是由一个激活域 do-main (AD) 或一

个抑制域 (RD) 组成。MR 激活或抑制的功能是由其蛋白序列决定<sup>[14]</sup>; C 端二聚化结构域 (CTD) 与 Aux/IAA 转录因子中的基序 III 和 IV 十分相似, 通常与 *ARF* 蛋白形成同型二聚体, 或与 AUX/IAA 蛋白形成异源二聚体来行使相应功能<sup>[15-16]</sup>。到目前为止, 通过全基因组分析已经鉴定出拟南芥 (23 个 *ARF*, 其中 1 个为假基因)、水稻 (25 个 *ARF*)、大豆 (51 个 *ARF*)、桃子 (17 个 *ARF*) 和花生 (30 个 *ARF*) 的 *ARF* 基因<sup>[17-21]</sup>。*ARF* 蛋白对调节植物生长发育的各个方面至关重要, 包括根、叶、花的发育、果实的成熟、衰老和脱落<sup>[22]</sup>。拟南芥中 *AtARF1* 是一个转录抑制因子<sup>[23]</sup>, *AtARF2* 与植物花和叶的衰老和脱落有关, *AtARF1/AtARF2* 双突变体的表型与单突变体相似, 但 2 种单突变体表型性状均比双突变体进一步增强, 表明 *AtARF1* 和 *AtARF2* 有部分重复作用<sup>[13]</sup>; 正向和反向遗传学已经确定 *AtARF7* 和 *AtARF19* 是调节侧根形成的关键基因, *AtARF7* 和 *AtARF19* 单突变体的侧根数量略有减少, 而双突变体的侧根数量显著减少, 这表明它们在功能上是冗余的<sup>[24]</sup>; *OsARF12* 是水稻根系构型的关键基因, 敲除该基因导致拟南芥主根长度减少。此外, *OsARF12/25* 双突变体的主根比单突变体短得多, 表明它们在功能上亦是冗余的<sup>[25]</sup>; 在番茄中, *SlARF2* 的过表达加速了叶片衰老<sup>[26]</sup>, *SlIAA9*、*SlARF6A*、*SlARF8A*、*SlARF8B* 和 *SlARF24* 形成一个调控复合体, 控制叶片形态<sup>[27]</sup>。也有报道称, 在拟南芥中,

*AtARF6* 和 *AtARF8* 在花成熟过程中起冗余作用, *AtARF6* 和 *AtARF8* 单突变植株表现出雄蕊发育延迟和繁殖力下降, 而 *AtARF6/8* 双突变植株在花成熟过程中表现为花瓣短、雄蕊花丝短、花药不分裂(无花粉释放), 花的生长发育完全阻滞<sup>[28]</sup>。此外, *AtARF8* 可以与 BIGPETALp (BPEp) 相互作用, 从而影响细胞扩增, 限制拟南芥花瓣的生长<sup>[29]</sup>。

*ARF* 基因家族在植物生长发育中扮演重要角色, 影响植物对非生物胁迫的反应, 但尚未有研究鉴定和分析颠茄中的 *ARF* 基因 (*AbARF*)。因此, 我们对颠茄 *ARF* 基因家族进行了全基因组鉴定和分析, 运用最大似然法 (ML) 构建了系统发育树, 并使用系统发育方法将 *ARF* 基因分类为不同的分支。分析了 *AbARF* 基因的基本组成, 得到了 *AbARF* 基因与蛋白质在结构和组成上的异同。*AbARF* 在不同光照条件下全长转录组表达数据进行基因分析。此外, 进一步将鉴定的 *AbARF* 基因与课题组前期转录组数据结合, 挑选 12 个 FPKM 值差异明显的 *AbARF* 基因在不同光照、温度处理和不同组织部位的颠茄进行实时荧光定量 PCR 验证 (RT-qPCR), 初步探讨其对颠茄生长发育的影响作用, 为下一步研究 *AbARF* 基因在颠茄生长发育中的分子机制提供理论和实验参考。

## 1 材料与仪器

### 1.1 材料

实验用材料来自于河南省信阳市新县, 经河南中医药大学药学院董诚明教授鉴定为茄科颠茄属植物颠茄 *A. belladonna* L.。

### 1.2 主要仪器

QuantStudio5 实时荧光定量 PCR 仪 (美国 Thermo Fisher Scientific 公司)、411C-1 型多酚多糖植物总 RNA 提取试剂盒 (北京庄盟国际生物基因科技有限公司)、HiScript® III RT SuperMix for qPCR (gDNA wiper) 试剂盒 (南京诺唯赞生物科技股份有限公司)、PowerTrack™ SYBR Green Master Mix 试剂盒 (美国 Thermo Fisher Scientific 公司)。

## 2 方法

### 2.1 样品处理

**2.1.1 光质和温度处理** 取用颠茄幼苗, 在培养箱中进行光质培养, 使用塑料小花盆栽培颠茄幼苗, 温度 25 °C、湿度 45%、光强 20 000 lx, 分别设白光+正常温度 (25 °C)、白光+低温 (18 °C)、红

光+正常温度 (25 °C)、红光+低温 (18 °C)、黄光+正常温度 (25 °C)、黄光+低温 (18 °C) 6 个处理组, 每组 50 盆, 每盆 1 株。于第 21 天采集颠茄叶片, 将其存放于-80 °C 冰箱, 用于分析 12 个 *AbARF* 基因在颠茄不同光质和温度下的表达情况。

**2.1.2 不同组织部位采集** 取种植于本校植物驯化园的颠茄植株, 采集主根、须根、茎、叶、花 5 种组织部位, 迅速洗净擦干后在液氮迅速冷冻后放入-80 °C 冰箱中保存, 用于分析 12 个 *AbARF* 基因在颠茄不同组织部位的表达情况。

### 2.2 颠茄 *AbARF* 基因家族成员及理化性质分析

颠茄的全基因组数据<sup>[30]</sup>从国家生物信息中心下载 (<https://ngdc.cncb.ac.cn/>), 项目编号为 PRJCA012583, 登录号为 GWHBOWM00000000。拟南芥 24 条 *ARF* 蛋白序列数据集来自 tair 数据库 (<https://www.arabidopsis.org/>)。利用 TBtools<sup>[31]</sup>进行颠茄和拟南芥的双向 blast 筛选颠茄中的 *ARF* 蛋白序列, 然后通过 Pfam 数据库 (<http://pfam-legacy.xfam.org/>) 利用隐马尔可夫模型 (hidden markov model, HMM) 对颠茄基因组中的 *ARF* 蛋白序列进一步筛选。在 HMM 配置文件中, 参数为默认值, 截止值为 0.01, 最后去除 2 种方法中的重复转录本后获取 *AbARF* 蛋白序列。使用 TBtools 软件的 Protein Paramter Calc 工具对颠茄 *AbARF* 蛋白理化性质进行预测, 运用 WoLF PSORT (<https://wolfpsort.hgc.jp>) 网站对 *AbARF* 基因进行亚细胞定位。

### 2.3 颠茄 *AbARF* 基因家族系统进化分析

利用 TBtools 软件 One Step Build a ML Tre 工具对颠茄和拟南芥 *ARF* 蛋白序列进行比对, 搜索重复值设置为 5 000, 采用最大似然法 (ML) 构建发育树, 并利用 iTOL (<https://itol.embl.de/>) 在线工具进一步对系统发育树进行美化。

### 2.4 颠茄 *AbARF* 基因保守基序和基因结构分析及顺式作用元件预测

通过 MEME 数据库 (<https://meme-suite.org/meme/>) 对 *AbARF* 蛋白保守基序进行预测, 具体参数: any number of Repetitions (anr), maximum number of Motifs=10, minimum width≥6, maximum width≤50。使用 NCBI 中 CDD 在线网站 (<https://www.ncbi.nlm.nih.gov/cdd>) 对 *AbARF* 蛋白结构域进行预测。通过 TBtools 软件使用基因组序列和结构注释文件分别提取各个颠茄 *AbARF* 基因

上游各 2 000 bp 序列作为启动子区域, 然后通过 PlantCare 数据库 (<https://www.plantcare.co.uk/>) 对启动子区的顺式作用元件进行预测, 最后使用 TBtools 软件将蛋白保守基序、保守结构域、启动子预测结果和其系统进化树整合及可视化作图。

2.5 颠茄 *AbARF* 基因染色体定位和种内及种间共线性分析

使用 TBtools 软件分析颠茄的基因组注释文件, 提取每个 *AbARF* 基因的在染色体上的位置信息并进行可视化绘制。利用 TBtools 软件获得颠茄染色体 ID 及长度信息, 使用 TBtools 插件 One Step MCScanX-Super Fast 得到颠茄 *AbARF* 基因之间的共线性文件, 并在 Advanced Circos 工具中将 *AbARF* 基因的种内共线性关系可视化。从 NCBI 网站下载拟南芥全基因组数据, 使用 TBtools 将颠茄和拟南

芥全基因组进行比对获得种间共线性关系, 并进行可视化处理。

2.6 颠茄 *AbARF* 基因表达模式分析及实时荧光定量 PCR 验证

根据课题组前期不同光照条件下颠茄转录组数据, 选取 FPKM 值差异较大的 12 个 *AbARF* 基因进行荧光定量 PCR 验证, 分析其在不同温度和光质下及不同组织部位 (主根、侧根、茎、叶、花) 的颠茄中的表达量变化趋势。荧光定量反应体系 20  $\mu$ L: dd H<sub>2</sub>O 7  $\mu$ L; PowerTrack™ SYBR Green Master Mix 10  $\mu$ L; 上、下引物各 1  $\mu$ L; cDNA 1  $\mu$ L。反应程序: 95  $^{\circ}$ C 预变性 10 min, 95  $^{\circ}$ C 变性 10 s, 60  $^{\circ}$ C 退火 30 s, 72  $^{\circ}$ C 延伸 32 s, 40 个循环数。内参基因为 *PGK*, 引物序列信息见表 1。每个样品做 3 个生物学重复。

表 1 *AbARF* 基因引物序列  
Table 1 Primer sequences of *AbARF* genes

| 基因名称           | 正向引物 (5'→3')            | 反向引物 (5'→3')            |
|----------------|-------------------------|-------------------------|
| <i>AbARF1</i>  | ACAAGCAGCCTGCTGAGTTC    | CTCTTGAGCAGCAGGTTGCAT   |
| <i>AbARF7</i>  | CGGCGTGCAGCTGAGAAAAT    | AGGCTCCACCCTGTTGTTAGC   |
| <i>AbARF13</i> | AGCCGAGGGTCTATCGGAAAC   | AGGACCAGCACATGCATGC     |
| <i>AbARF17</i> | TGCTGAAACGGCGTTGTATACG  | TCTGCTGGTCAGCCTCTTGATT  |
| <i>AbARF18</i> | CGGAAAAGTTGATGCTGAGACGG | CTGCTGATCAGCCACTTGATTGG |
| <i>AbARF19</i> | TGCCGTGTGGTTAATGTCCG    | AGTGCTTGTATCAGAGGCCGT   |
| <i>AbARF20</i> | AGGAActCTGGCATGCCTGT    | CTCGGCCCCGAAGTTGAACATTC |
| <i>AbARF21</i> | TGAGTTGCTGGATCGTCCAACC  | ACTCGTGGGGAGACCTCATAGAT |
| <i>AbARF22</i> | AGAACGGGGAGCTTCGAGTT    | CTGACTGACAAGCCCTAGGCT   |
| <i>AbARF25</i> | CGACAACGAAGGAACCAATGCC  | ACCATAGGGGGAGGCATTTCATC |
| <i>AbARF29</i> | AGGCAGAGCAAGAGGAGCCTT   | AGTCATGTCCAGTTGCGGGAG   |
| <i>AbARF33</i> | AACTATGCCGGCGGTAGCAT    | ACAGAAAACGTGAGGAGGGAGG  |
| <i>AbARF41</i> | ACTCCGGAGTTCTGTGTGAAGG  | AGTAGATCAGGTTCTGTCACG   |
| <i>PGK</i>     | TCGCTCTTGAGAGAAGGTTGAC  | CTTGTCGCAATCACTACATCAG  |

2.7 数据处理

利用 Microsoft®Excel 2021 软件统计实时荧光定量 PCR 数据, 通过  $2^{-\Delta\Delta C_t}$  算法计算基因相对表达量, IBM SPSS Statistics 26 软件进行独立样本 *t* 检验和单因素方差分析, GraphPad 9.0 绘制柱状图, TBtools 软件绘制热图。

3 结果与分析

3.1 颠茄 *AbARF* 家族成员及理化性质分析

通过对颠茄和拟南芥的双向 blast 和 Pfam 数据库中 HMM 模型筛选, 去除冗余, 从颠茄全基因组中共鉴定到 41 个 *AbARF* 基因, 分别命名为 *AbARF1*~

*AbARF41*。使用 TBtools 对颠茄 *AbARF* 蛋白理化性质进行预测, 结果表明 (表 2), *AbARF1* 基因编码氨基酸最多, 为 1 116 个, *AbARF30* 基因编码氨基酸最少, 为 603 个, 相对分子质量分布区域为 68 709.35~125 100.60。除了 *AbARF33*、*AbARF37*、*AbARF38* 和 *AbARF40* 蛋白等电点大于 7.00, 其余蛋白均小于 7.00, 并且所有蛋白亲水性系数均小于零, 说明该家族蛋白大多数为酸性、亲水蛋白。同时, 所有 *AbARF* 蛋白不稳定系数大于 40, 表明该家族蛋白大多数为不稳定蛋白。亚细胞定位预测结果显示, *AbARF* 蛋白均定位在细胞核中。

表 2 颠茄 AbARF 蛋白理化性质分析

Table 2 Analysis of physical and chemical properties of AbARF proteins from *A. belladonna*

| 基因 ID   | 氨基酸数目 | 相对分子质量     | 等电点  | 不稳定系数 | 脂肪系数  | 亲水平均系数 | 亚细胞定位 |
|---------|-------|------------|------|-------|-------|--------|-------|
| AbARF1  | 1 116 | 125 100.57 | 6.37 | 71.67 | 69.47 | −0.716 | 细胞核   |
| AbARF2  | 1 093 | 122 392.50 | 6.23 | 69.68 | 68.45 | −0.693 | 细胞核   |
| AbARF3  | 1 115 | 123 520.70 | 6.16 | 62.02 | 73.38 | −0.545 | 细胞核   |
| AbARF4  | 1 108 | 122 640.73 | 6.25 | 63.76 | 73.21 | −0.540 | 细胞核   |
| AbARF5  | 1 107 | 122 404.37 | 6.61 | 64.01 | 71.80 | −0.545 | 细胞核   |
| AbARF6  | 1 102 | 121 777.19 | 5.89 | 56.35 | 73.08 | −0.569 | 细胞核   |
| AbARF7  | 1 103 | 122 073.51 | 6.22 | 55.53 | 71.95 | −0.601 | 细胞核   |
| AbARF8  | 1 049 | 116 774.51 | 6.13 | 58.90 | 73.25 | −0.620 | 细胞核   |
| AbARF9  | 890   | 98 907.17  | 6.19 | 62.90 | 73.93 | −0.453 | 细胞核   |
| AbARF10 | 832   | 92 948.50  | 5.85 | 59.27 | 75.36 | −0.444 | 细胞核   |
| AbARF11 | 826   | 92 392.98  | 6.01 | 56.11 | 73.05 | −0.480 | 细胞核   |
| AbARF12 | 891   | 99 020.22  | 6.19 | 62.10 | 73.96 | −0.456 | 细胞核   |
| AbARF13 | 882   | 98 874.90  | 5.99 | 56.47 | 80.23 | −0.318 | 细胞核   |
| AbARF14 | 905   | 100 296.37 | 5.35 | 52.58 | 75.38 | −0.442 | 细胞核   |
| AbARF15 | 907   | 100 325.50 | 5.27 | 53.76 | 79.63 | −0.389 | 细胞核   |
| AbARF16 | 806   | 89 767.68  | 6.07 | 62.15 | 74.03 | −0.473 | 细胞核   |
| AbARF17 | 826   | 92 466.80  | 6.92 | 54.41 | 65.59 | −0.656 | 细胞核   |
| AbARF18 | 850   | 94 164.72  | 6.44 | 55.45 | 69.15 | −0.578 | 细胞核   |
| AbARF19 | 847   | 94 346.99  | 6.35 | 50.10 | 67.10 | −0.598 | 细胞核   |
| AbARF20 | 656   | 73 101.71  | 6.16 | 59.96 | 70.27 | −0.507 | 细胞核   |
| AbARF21 | 656   | 73 377.85  | 5.99 | 60.65 | 69.10 | −0.544 | 细胞核   |
| AbARF22 | 656   | 72 965.36  | 5.93 | 60.69 | 70.14 | −0.525 | 细胞核   |
| AbARF23 | 796   | 88 404.31  | 5.89 | 55.50 | 73.18 | −0.466 | 细胞核   |
| AbARF24 | 705   | 78 832.94  | 5.79 | 51.52 | 77.80 | −0.479 | 细胞核   |
| AbARF25 | 785   | 88 448.30  | 6.29 | 52.53 | 75.87 | −0.492 | 细胞核   |
| AbARF26 | 798   | 88 863.29  | 6.24 | 55.72 | 73.48 | −0.448 | 细胞核   |
| AbARF27 | 681   | 76 490.21  | 5.75 | 52.34 | 78.40 | −0.481 | 细胞核   |
| AbARF28 | 787   | 87 242.17  | 6.00 | 55.51 | 71.68 | −0.445 | 细胞核   |
| AbARF29 | 692   | 77 761.79  | 5.77 | 54.03 | 78.70 | −0.457 | 细胞核   |
| AbARF30 | 603   | 68 709.35  | 6.84 | 57.78 | 77.10 | −0.493 | 细胞核   |
| AbARF31 | 750   | 82 470.50  | 6.46 | 53.52 | 71.77 | −0.422 | 细胞核   |
| AbARF32 | 623   | 69 743.53  | 5.86 | 51.11 | 72.42 | −0.513 | 细胞核   |
| AbARF33 | 754   | 82 610.20  | 8.25 | 50.06 | 72.71 | −0.400 | 细胞核   |
| AbARF34 | 693   | 76 353.41  | 6.79 | 48.40 | 72.31 | −0.395 | 细胞核   |
| AbARF35 | 697   | 76 694.92  | 6.79 | 48.16 | 73.86 | −0.377 | 细胞核   |
| AbARF36 | 679   | 75 380.49  | 5.66 | 45.52 | 81.53 | −0.285 | 细胞核   |
| AbARF37 | 778   | 86 289.14  | 8.06 | 48.59 | 80.33 | −0.254 | 细胞核   |
| AbARF38 | 697   | 76 840.00  | 7.86 | 45.41 | 78.34 | −0.278 | 细胞核   |
| AbARF39 | 673   | 74 587.83  | 5.98 | 49.01 | 80.53 | −0.252 | 细胞核   |
| AbARF40 | 686   | 75 285.50  | 7.56 | 47.04 | 80.87 | −0.236 | 细胞核   |
| AbARF41 | 680   | 75 063.87  | 6.48 | 47.13 | 73.97 | −0.417 | 细胞核   |

3.2 颠茄 *AbARF* 基因家族系统进化分析

采用 ML 构建系统发育树，结果如图 1 所示，

根据蛋白结构将其分为 4 个亚家族，颠茄 *AbARF* 基因在各个亚家族中均有分布，并且分布数量最多

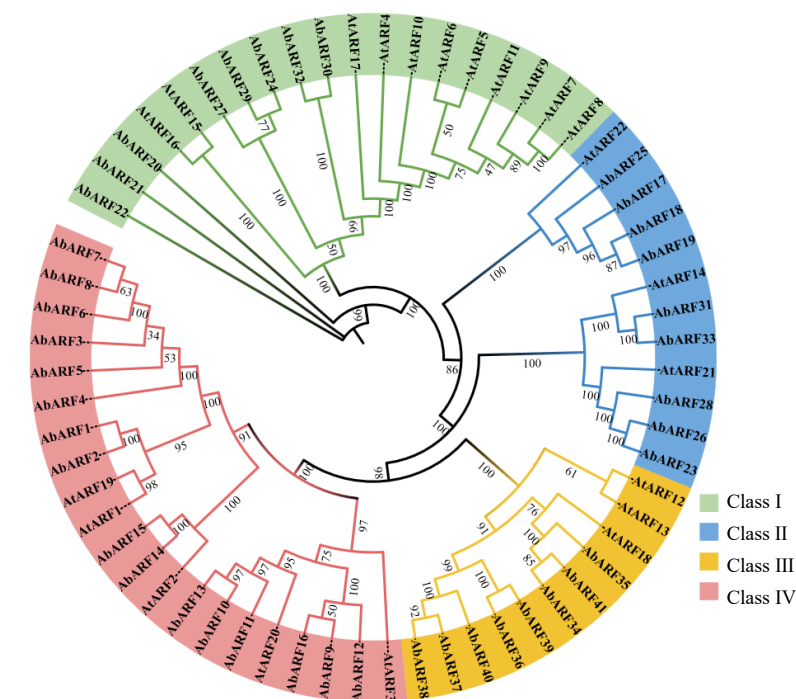


图 1 颠茄、拟南芥 ARF 蛋白系统进化树

Fig. 1 Phylogenetic tree of ARF proteins from *A. belladonna* and *Arabidopsis thaliana*

为亚家族IV，亚家族I包含 8 个成员，与拟南芥 *AtARF4/5/6/7/8/9/10/11/15/16* 基因同源，亚家族II有 9 个成员，与拟南芥 *AtARF14/21/22* 基因同源。亚家族III有 8 个成员，与拟南芥 *AtARF12/13/18* 蛋白同源，亚家族IV包括 16 个成员，与拟南芥 *AtARF1/2/3/19/20* 基因同源。

### 3.3 颠茄 *AbARF* 转录因子保守基序和基因结构分析及顺式作用元件预测

为了解析颠茄 *AbARF* 转录因子结构多样性和进化关系，结果如图 2 所示，*AbARF* 蛋白由 1~10 个 Motifs 组成（图 3 为 Motifs 氨基酸分布情况），均含有 B3 结构域（Motif1/3/6 和 8），说明表明这 4 个基序在 *AbARF* 蛋白序列中分布最广、保守程度最高，也从侧面佐证了筛选出的 41 个 *AbARF* 转录因子均属于 *ARF* 基因家族成员。*AbARF* 蛋白（除去 *AbARF16* 不含 Motif2 和 Motif7 和 *AbARF8* 不含 Motif2）均含有 Motif2 和 Motif7，除去 *AbARF25* 不含 Motif10 和 *AbARF31/33* 不含 Motif4/9 其余皆含有 Motif4/9 和 10。这些 *AbARF* 蛋白从保守基序组成不同从而被聚在不同亚家族，这可能与不同亚家族功能不同相关。

筛选的 41 个 *AbARF* 基因结构，所有 *AbARF* 基因 UTR 数量为 2~4 个，*AbARF1~AbARF31* 外显

子、内含子和 CDS 数量在 10~16 个，而 *AbARF32~AbARF41* 外显子、内含子和 CDS 数量在 3~5 个，差异明显，这说明 *AbARF32~AbARF41* 基因是颠茄最原始的 ARF 基因，在随着颠茄系统进化过程中，*AbARF1~AbARF31* 基因逐渐被分化出来，以适应新的环境变化，行使新的蛋白功能。

保守结构域结果如图 4 所示，所有 *AbARF* 蛋白均含有 B3 保守结构域，这说明 B3 是 *AbARF* 蛋白组成中最核心的保守结构域，除 *AbARF41* 蛋白外，所有 *AbARF* 蛋白均含有 *Auxin\_resp* 保守结构域，说明 *Auxin\_resp* 结构域是 *AbARF* 蛋白的重要组成部分。部分 *AbARF* 蛋白不含有 *AUX\_IAA superfamily* 保守结构域，它们被聚在同一分支，可能行使相同的蛋白功能。另有几个蛋白有特殊的保守结构域，*AbARF6* 蛋白含有 *PRK14948 superfamily*、*AbARF7* 蛋白含有 *Med15 superfamily*、*AbARF10* 蛋白含有 *Activator\_LAG-3 superfamily*、*AbARF41* 蛋白含有 *Auxin\_resp superfamily* 保守结构域。

启动子预测结果显示（图 4），*AbARF* 转录因子参与了 10 个非生物胁迫过程，所有 *AbARF* 转录因子共包括 419 个顺式元件，包括生长素响应元件、水杨酸响应元件、光响应元件等。每个 *AbARF* 启动子中都有大量的光响应性元件，占有所有顺式元



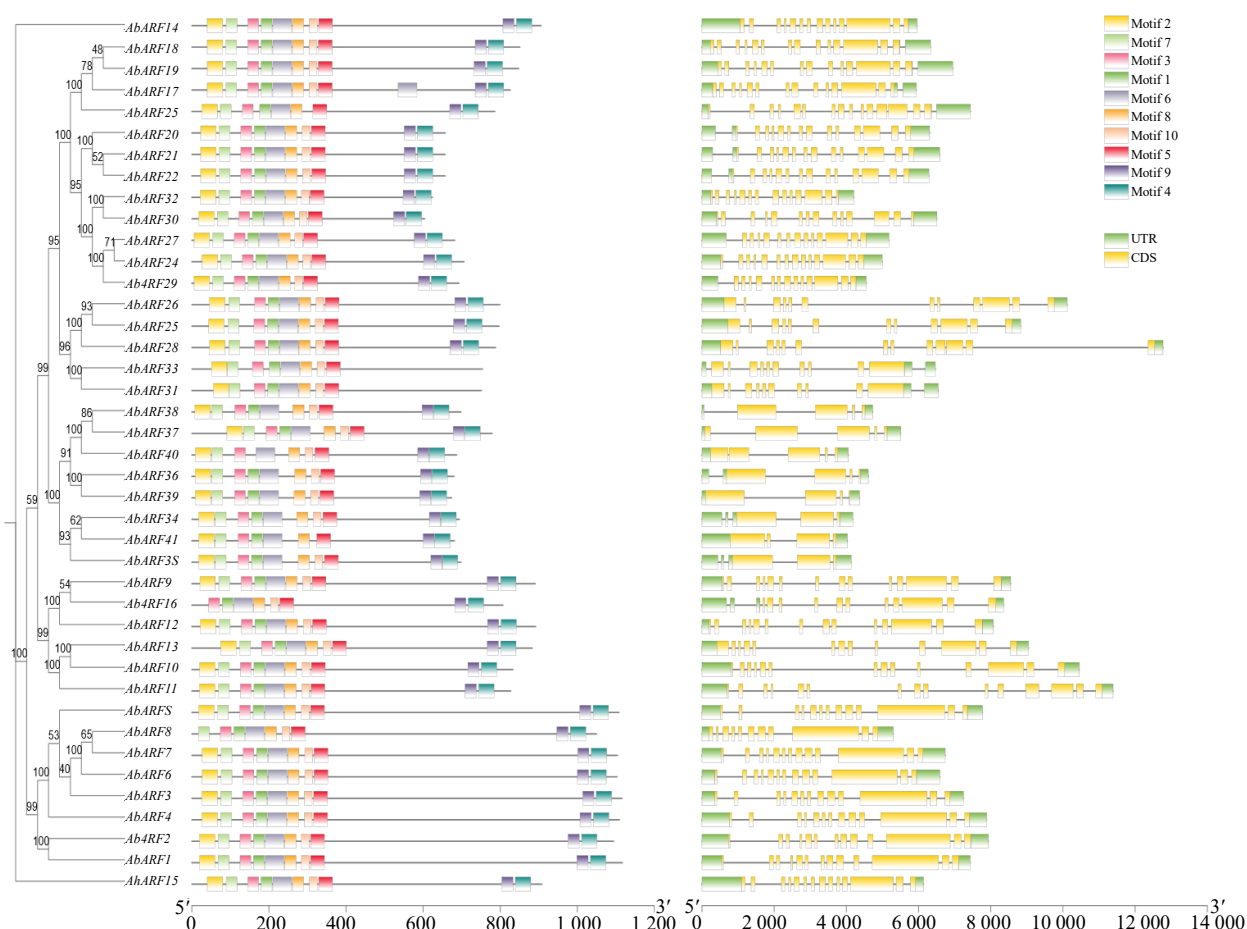


图 2 *AbARF* 转录因子保守基序和基因结构分析

Fig. 2 Analysis of conserved motifs and gene structures of *AbARF* transcription factors

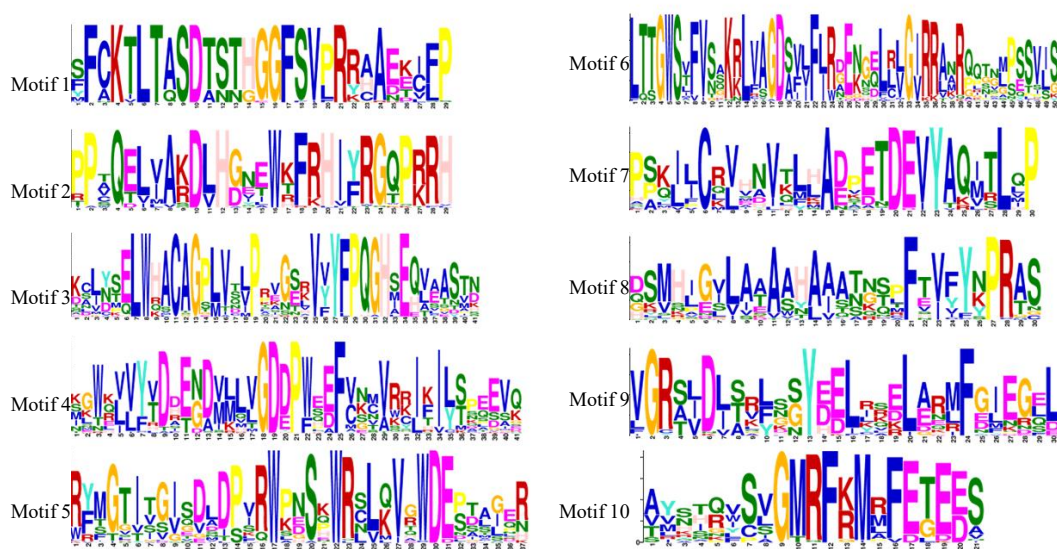
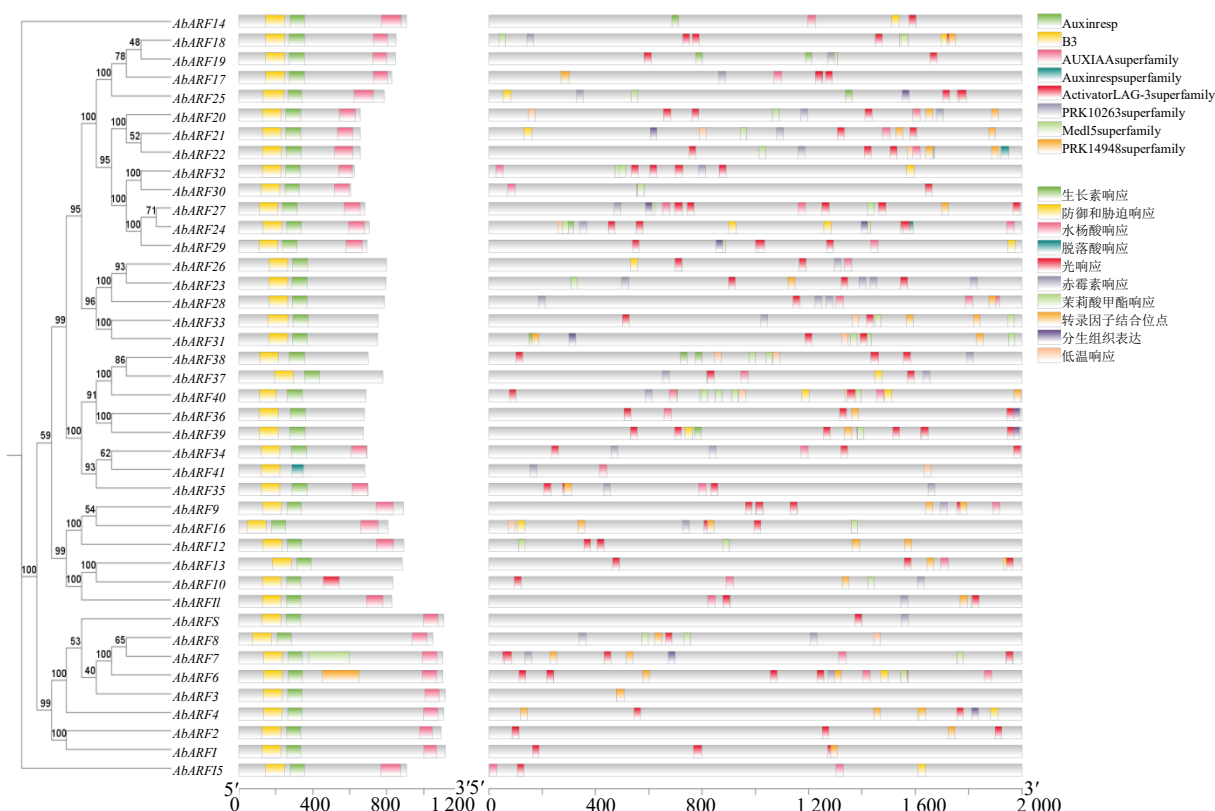
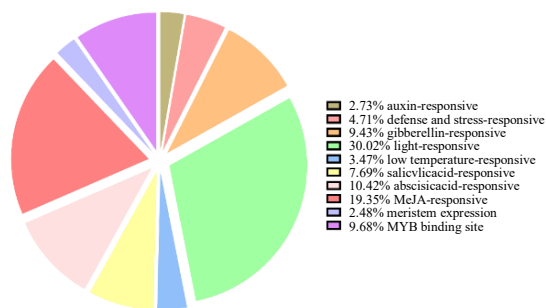


图 3 Motifs 氨基酸分布情况

Fig. 3 Motifs amino acid distribution

图 4 *AbARF* 转录因子保守结构域分析及顺式作用元件预测Fig. 4 Elements analysis of conserved domain of *AbARF* transcription factors and prediction of cis-acting elements

件的 30.02%，其数量远远超过其他响应元件，其次是茉莉酸甲酯响应元件，占有顺式元件的 19.35%，在所有顺式元件中分生组织表达调控元件占比最少，仅为 2.48%（图 5），有 9.68%为 MYB 转录因子结合元件，这可能说明 *AbARF* 和 *AbMYB* 转录因子能相互结合从而行使不同的蛋白功能。此外，由于含有光、茉莉酸甲酯、赤霉素和水杨酸响应元件的 *AbARF* 基因构成了这些基因的很大一部分，推测这些基因可能在响应这些胁迫方面发挥重要作用，可以作为未来验证这些基因功能的理论基础。

图 5 *AbARF* 转录因子总顺式作用元件占比Fig. 5 Proportion of total cis-acting elements of *AbARF* transcription factors

### 3.4 颠茄 *AbARF* 基因染色体定位和种内及种间共线性分析

根据颠茄基因组数据对 *AbARF* 基因家族成员进行染色体定位分析（图 6），发现 41 个 *AbARF* 基因不均匀的分布在 26 条染色体上，10 条染色体上没有 *AbARF* 基因，LG14 染色体上分布的 *AbARF* 基因最多为 3 个，其余染色体多分布为 1~2 个，多数 *AbARF* 基因分布在染色体两端，*AbARF26* 和 *AbARF37*、*AbARF23* 和 *AbARF38*、*AbARF28* 和 *AbARF40* 彼此在染色体上距离最近，但是整个 *AbARF* 基因家族在染色体上并无基因簇。

为检测颠茄中 *AbARF* 中的基因重复事件，我们使用 TBtools 软件进行共线性分析，在 *AbARF* 中共鉴定出 46 对基因间存在共线性关系（图 7），皆为全基因组复制和片段性复制事件，并无串联重复事件（表 3），这与决明的结果一致<sup>[32]</sup>。非同义取代率/同义取代率 ( $K_a/K_s$ ) 比通常是一个重要的进化中选择压力的指标<sup>[33]</sup>。对 46 个 *AbARF* 基因对共线性分析显示，所有共线性基因对  $K_a/K_s$  比值均小于 1，说明它们在复制后经过了强烈的纯化选择。另外，*AbARF35* 和 *AbARF40* 的  $K_s$  值远大于其他基因对，



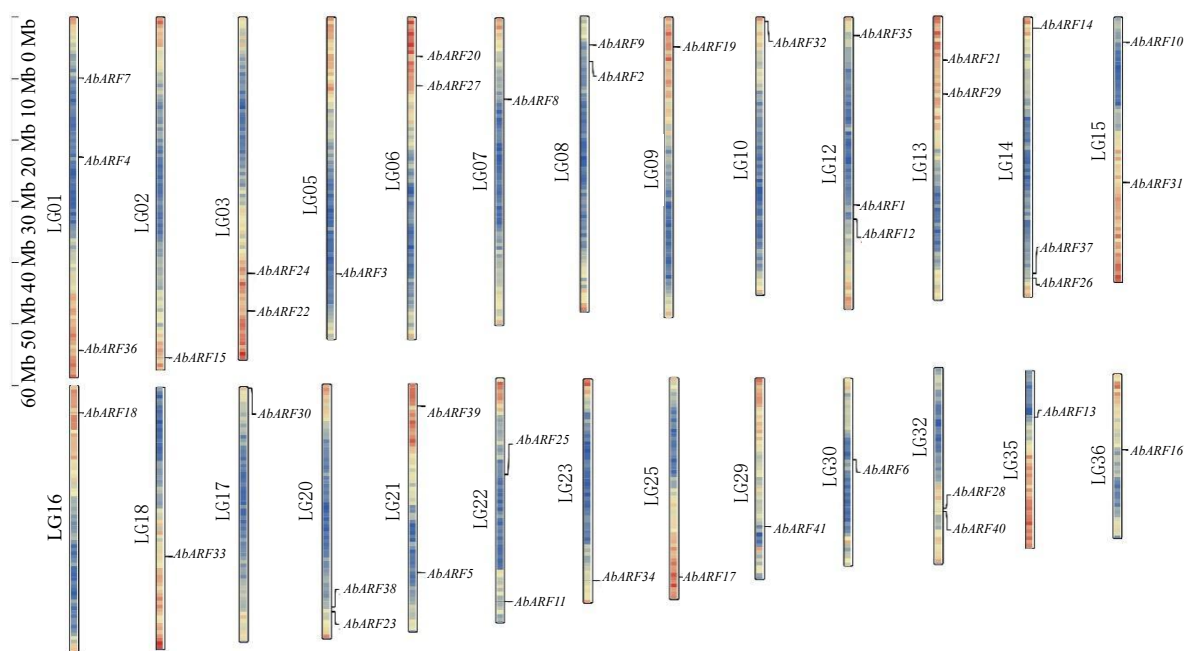
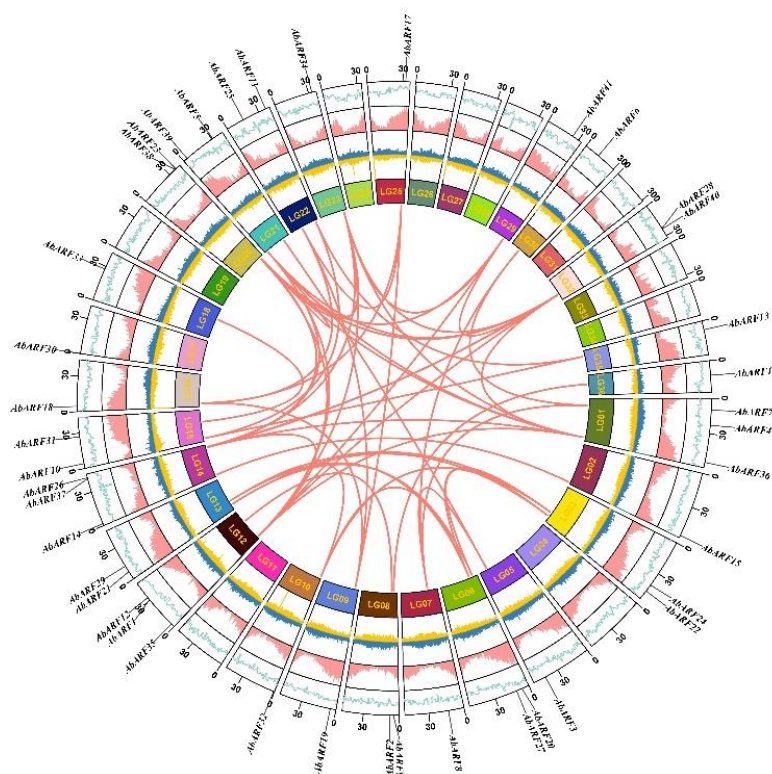


图 6 *AbARF* 基因家族的染色体定位

Fig. 6 Chromosomal location of *AbARF* gene family



红线表示 *AbARF* 有共线性；蓝色和黄色表示 GC 含量的偏移量，G 大于 C 为蓝色，G 小于 C 则黄色；粉色直方图为基因密度；青色为 GC 含量占比；LG01-36 为染色体名称。

The red line indicates that *AbARF* has collinearity; Blue and yellow indicate the offset of GC content. If G is greater than C, it is blue, if G is less than C, it is yellow; The pink histogram is the gene density; Cyan is the proportion of GC content; LG01-36: Chromosome name.

图 7 *AbARFs* 基因在基因组上的分布及共线性分析

Fig. 7 Distribution and collinearity analysis of *AbARFs* genes on genome

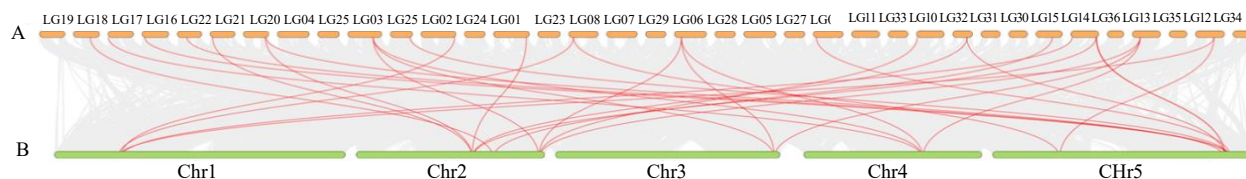
表3 *AbARF* 基因物种内共线性关系

Table 3 Collinear relationship among *AbARF* genes within species

| 基因对            |                | 复制类型         | <i>K<sub>a</sub></i> | <i>K<sub>s</sub></i> | <i>K<sub>a</sub>/K<sub>s</sub></i> | 选择压力类型 |
|----------------|----------------|--------------|----------------------|----------------------|------------------------------------|--------|
| <i>AbARF15</i> | <i>AbARF14</i> | 全基因组复制、片段性复制 | 0.020 381            | 0.138 404            | 0.147 260                          | 纯化选择   |
| <i>AbARF2</i>  | <i>AbARF1</i>  | 全基因组复制、片段性复制 | 0.023 136            | 0.088 573            | 0.261 204                          | 纯化选择   |
| <i>AbARF9</i>  | <i>AbARF16</i> | 全基因组复制、片段性复制 | 0.022 398            | 0.079 020            | 0.283 444                          | 纯化选择   |
| <i>AbARF19</i> | <i>AbARF18</i> | 全基因组复制、片段性复制 | 0.036 885            | 0.078 698            | 0.468 683                          | 纯化选择   |
| <i>AbARF19</i> | <i>AbARF25</i> | 全基因组复制、片段性复制 | 0.142 274            | 0.508 381            | 0.279 858                          | 纯化选择   |
| <i>AbARF19</i> | <i>AbARF17</i> | 全基因组复制、片段性复制 | 0.041 557            | 0.104 401            | 0.398 057                          | 纯化选择   |
| <i>AbARF10</i> | <i>AbARF11</i> | 全基因组复制、片段性复制 | 0.083 050            | 0.534 212            | 0.155 462                          | 纯化选择   |
| <i>AbARF10</i> | <i>AbARF13</i> | 全基因组复制、片段性复制 | 0.028 174            | 0.106 297            | 0.265 047                          | 纯化选择   |
| <i>AbARF18</i> | <i>AbARF25</i> | 全基因组复制、片段性复制 | 0.137 645            | 0.514 069            | 0.267 757                          | 纯化选择   |
| <i>AbARF18</i> | <i>AbARF17</i> | 全基因组复制、片段性复制 | 0.039 470            | 0.092 496            | 0.426 724                          | 纯化选择   |
| <i>AbARF25</i> | <i>AbARF17</i> | 全基因组复制、片段性复制 | 0.128 307            | 0.540 190            | 0.237 522                          | 纯化选择   |
| <i>AbARF11</i> | <i>AbARF13</i> | 全基因组复制、片段性复制 | 0.107 811            | 0.560 262            | 0.192 429                          | 纯化选择   |
| <i>AbARF24</i> | <i>AbARF27</i> | 全基因组复制、片段性复制 | 0.034 282            | 0.166 038            | 0.206 472                          | 纯化选择   |
| <i>AbARF22</i> | <i>AbARF20</i> | 全基因组复制、片段性复制 | 0.017 794            | 0.086 702            | 0.205 233                          | 纯化选择   |
| <i>AbARF24</i> | <i>AbARF32</i> | 全基因组复制、片段性复制 | 0.274 153            | 2.269 381            | 0.120 805                          | 纯化选择   |
| <i>AbARF24</i> | <i>AbARF29</i> | 全基因组复制、片段性复制 | 0.028 023            | 0.091 197            | 0.307 281                          | 纯化选择   |
| <i>AbARF22</i> | <i>AbARF21</i> | 全基因组复制、片段性复制 | 0.020 174            | 0.080 630            | 0.250 202                          | 纯化选择   |
| <i>AbARF27</i> | <i>AbARF32</i> | 全基因组复制、片段性复制 | 0.257 512            | 1.960 176            | 0.131 372                          | 纯化选择   |
| <i>AbARF20</i> | <i>AbARF21</i> | 全基因组复制、片段性复制 | 0.025 592            | 0.097 889            | 0.261 436                          | 纯化选择   |
| <i>AbARF27</i> | <i>AbARF29</i> | 全基因组复制、片段性复制 | 0.047 201            | 0.149 730            | 0.315 242                          | 纯化选择   |
| <i>AbARF32</i> | <i>AbARF29</i> | 全基因组复制、片段性复制 | 0.271 992            | 1.834 682            | 0.148 250                          | 纯化选择   |
| <i>AbARF26</i> | <i>AbARF23</i> | 全基因组复制、片段性复制 | 0.020 710            | 0.071 596            | 0.289 260                          | 纯化选择   |
| <i>AbARF26</i> | <i>AbARF28</i> | 全基因组复制、片段性复制 | 0.031 008            | 0.132 254            | 0.234 457                          | 纯化选择   |
| <i>AbARF23</i> | <i>AbARF28</i> | 全基因组复制、片段性复制 | 0.029 048            | 0.133 306            | 0.217 903                          | 纯化选择   |
| <i>AbARF36</i> | <i>AbARF37</i> | 全基因组复制、片段性复制 | 0.160 796            | 0.624 668            | 0.257 410                          | 纯化选择   |
| <i>AbARF36</i> | <i>AbARF38</i> | 全基因组复制、片段性复制 | 0.159 857            | 0.637 677            | 0.250 686                          | 纯化选择   |
| <i>AbARF36</i> | <i>AbARF39</i> | 全基因组复制、片段性复制 | 0.036 405            | 0.135 483            | 0.268 708                          | 纯化选择   |
| <i>AbARF36</i> | <i>AbARF40</i> | 全基因组复制、片段性复制 | 0.164 140            | 0.609 120            | 0.269 470                          | 纯化选择   |
| <i>AbARF3</i>  | <i>AbARF8</i>  | 全基因组复制、片段性复制 | 0.107 082            | 0.647 165            | 0.165 463                          | 纯化选择   |
| <i>AbARF3</i>  | <i>AbARF5</i>  | 全基因组复制、片段性复制 | 0.034 444            | 0.128 093            | 0.268 900                          | 纯化选择   |
| <i>AbARF35</i> | <i>AbARF39</i> | 全基因组复制、片段性复制 | 0.287 934            | 2.875 177            | 0.100 145                          | 纯化选择   |
| <i>AbARF35</i> | <i>AbARF34</i> | 全基因组复制、片段性复制 | 0.018 035            | 0.096 657            | 0.186 588                          | 纯化选择   |
| <i>AbARF35</i> | <i>AbARF41</i> | 全基因组复制、片段性复制 | 0.041 699            | 0.121 485            | 0.343 249                          | 纯化选择   |
| <i>AbARF35</i> | <i>AbARF40</i> | 全基因组复制、片段性复制 | 0.249 427            | 5.403 669            | 0.046 159                          | 纯化选择   |
| <i>AbARF37</i> | <i>AbARF38</i> | 全基因组复制、片段性复制 | 0.032 659            | 0.112 256            | 0.290 934                          | 纯化选择   |
| <i>AbARF37</i> | <i>AbARF39</i> | 全基因组复制、片段性复制 | 0.159 982            | 0.568 931            | 0.281 198                          | 纯化选择   |
| <i>AbARF37</i> | <i>AbARF40</i> | 全基因组复制、片段性复制 | 0.043 777            | 0.134 424            | 0.325 663                          | 纯化选择   |
| <i>AbARF31</i> | <i>AbARF33</i> | 全基因组复制、片段性复制 | 0.039 345            | 0.118 454            | 0.332 150                          | 纯化选择   |
| <i>AbARF38</i> | <i>AbARF39</i> | 全基因组复制、片段性复制 | 0.153 182            | 0.587 011            | 0.260 953                          | 纯化选择   |
| <i>AbARF38</i> | <i>AbARF40</i> | 全基因组复制、片段性复制 | 0.036 428            | 0.119 42             | 0.305 041                          | 纯化选择   |
| <i>AbARF39</i> | <i>AbARF34</i> | 全基因组复制、片段性复制 | 0.274 184            | 3.168 391            | 0.086 537                          | 纯化选择   |
| <i>AbARF39</i> | <i>AbARF40</i> | 全基因组复制、片段性复制 | 0.158 472            | 0.561 749            | 0.282 105                          | 纯化选择   |
| <i>AbARF34</i> | <i>AbARF41</i> | 全基因组复制、片段性复制 | 0.037 203            | 0.145 276            | 0.256 084                          | 纯化选择   |
| <i>AbARF7</i>  | <i>AbARF6</i>  | 全基因组复制、片段性复制 | 0.023 483            | 0.124 438            | 0.188 716                          | 纯化选择   |
| <i>AbARF8</i>  | <i>AbARF6</i>  | 全基因组复制、片段性复制 | 0.031 581            | 0.149 992            | 0.210 554                          | 纯化选择   |
| <i>AbARF7</i>  | <i>AbARF8</i>  | 全基因组复制、片段性复制 | 0.027 936            | 0.151 486            | 0.184 412                          | 纯化选择   |
| <i>AbARF15</i> | <i>AbARF14</i> | 全基因组复制、片段性复制 | 0.020 381            | 0.138 404            | 0.147 260                          | 纯化选择   |

表明这 2 个基因发生同位突变最多,突变时间越长。

为探索 *AbARF* 基因的潜在进化线索,将颠茄的 *AbARF* 基因和拟南芥基因组进行共线性比较。结果如图 8 所示,共有 32 个同源基因对, *AbARF3~13/16/20~22/34/35/41* 基因与拟南芥无同源基因对,这可能说明这几个基因是在颠茄演化过程中逐渐形成的,行使新的蛋白功能以适应环境变化。



红线表示 *AbARF* 同源基因对, 灰线表示其他同源基因对。

The red line indicates the *AbARF* homologous gene pair, and the gray line indicates other homologous gene pairs.

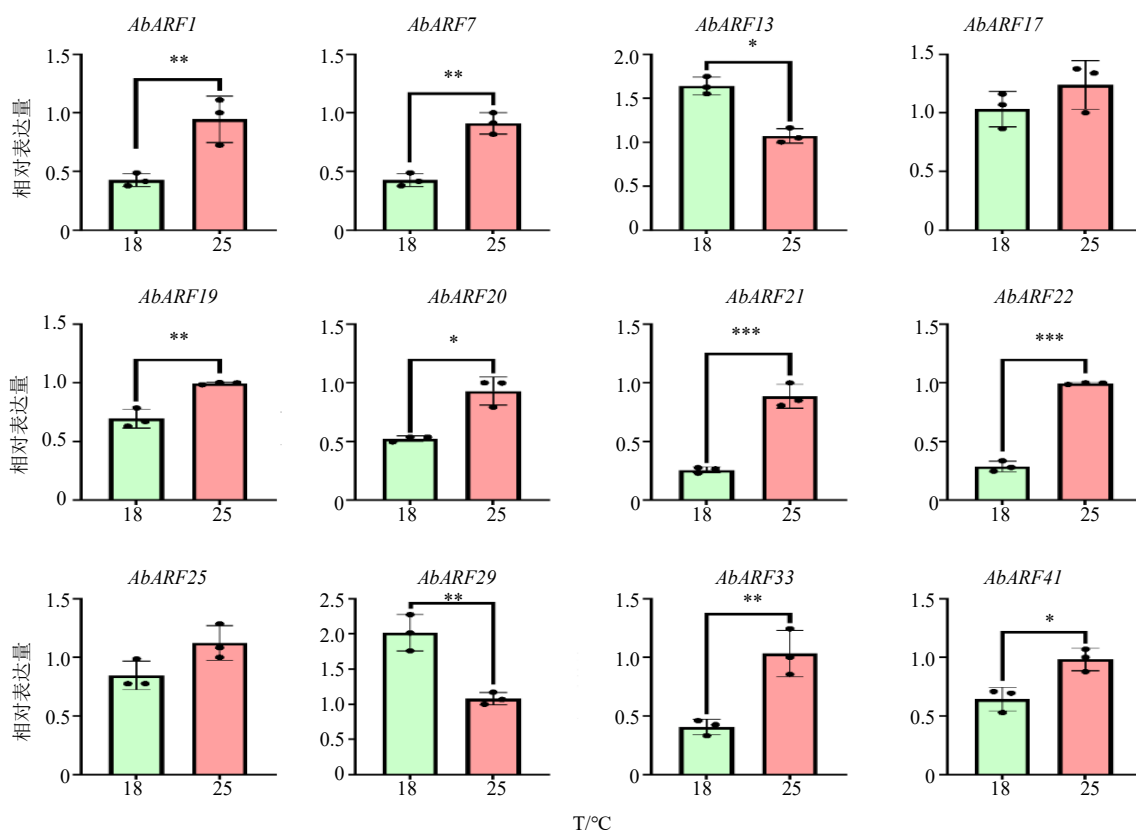
图 8 颠茄 (A) 和拟南芥 (B) 基因共线性分析

Fig. 8 Collinearity analysis of *A. Belladonna* (A) and *A. thaliana* genes (B)

### 3.5 颠茄 *AbARF* 基因表达模式分析

为进一步探究 *AbARF* 基因在不同温度, 不同光质及不同颠茄组织部位的表达模式, 选取了本课题组前期转录组中 12 个 FPKM 值差异最大的 *AbARF* 基因, 进行实时荧光定量 PCR 验证。*AbARF17* 和 *AbARF25* 基因在低温 (18 °C) 和对照 (25 °C) 条件下相对表达量无明显差异 (图 9), 说明它们表达模式不受低温影响; 在低温下, *AbARF20* 和 *AbARF41* 基因的相对表达量低于对照, 具有显著性差异 ( $P <$

0.05)。 *AbARF1*、*AbARF7*、*AbARF19* 和 *AbARF33* 基因相对表达量下调, 显著低于对照 ( $P < 0.01$ )。 *AbARF21* 和 *AbARF22* 基因的相对表达量极显著下调 ( $P < 0.001$ )。 *AbARF21* 和 *AbARF22* 基因在系统发育树中被聚在同一支, 具有相似的基因结构, 其基因表达模式也相似。另有 *AbARF13* 和 *AbARF29* 基因在低温条件下其相对表达量显著高于对照, 说明低温诱导有利于其表达。在不同光质条件下, 不同 *AbARF* 基因表达模式亦不相同 (图 10)。 *AbARF1*、*AbARF7*、

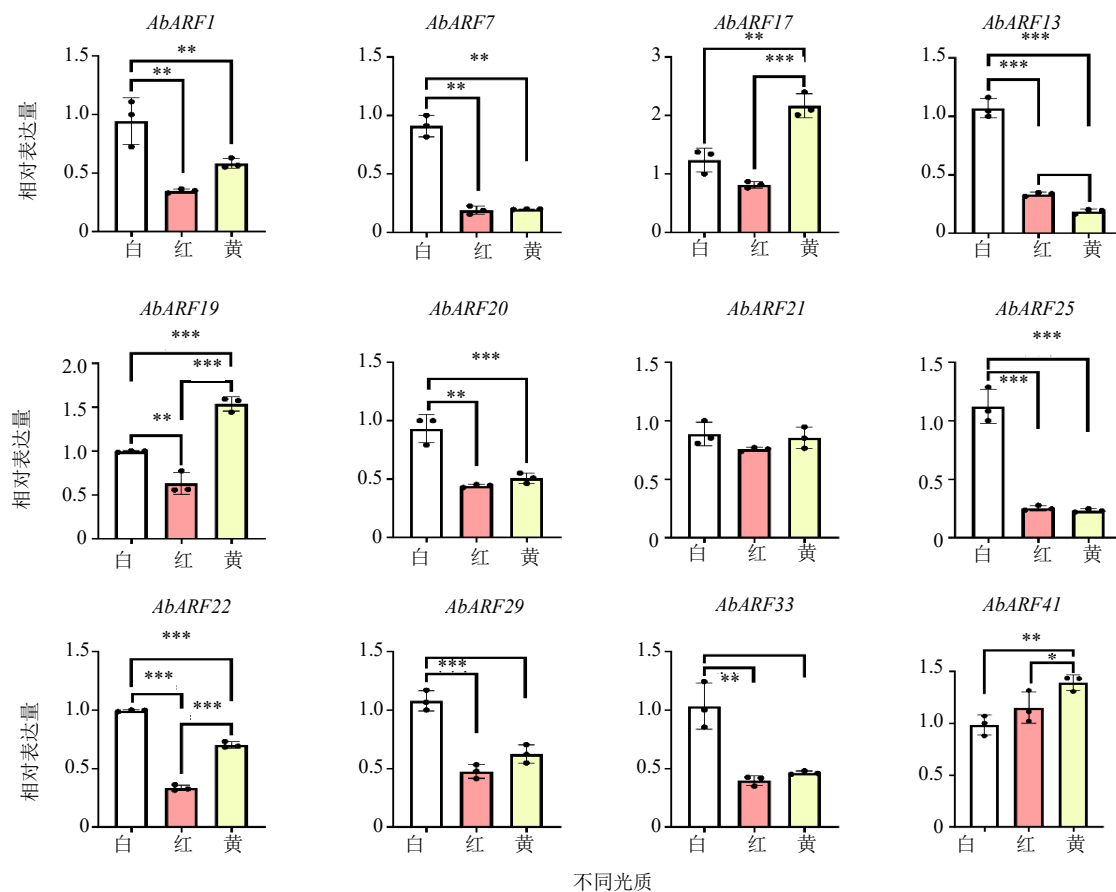


\*表示有显著性差异 ( $P < 0.05$ ), \*\*表示有非常显著性差异 ( $P < 0.01$ ), \*\*\*表示有极显著性差异 ( $P < 0.001$ )。

\*indicates an extremely significant difference ( $P < 0.05$ ), \*\*indicates an extremely significant difference ( $P < 0.01$ ), \*\*\*indicates an extremely significant difference ( $P < 0.001$ ).

图 9 *AbARF* 基因在低温条件下的表达模式

Fig. 9 Expression pattern of *AbARF* genes under low temperature conditions



\*表示显著性差异 ( $P < 0.05$ ), \*\*表示非常显著性差异 ( $P < 0.01$ ), \*\*\*表示极显著性差异 ( $P < 0.001$ ).

\*indicates an extremely significant difference ( $P < 0.05$ ), \*\*indicates an extremely significant difference ( $P < 0.01$ ), \*\*\*indicates an extremely significant difference ( $P < 0.001$ ).

图 10 *AbARF* 基因在不同光质条件下的表达模式

Fig. 10 Expression pattern of *AbARF* genes under different light quality conditions

*AbARF13*、*AbARF20*、*AbARF22*、*AbARF25*、*AbARF29* 和 *AbARF33* 基因在红光和黄光条件下相对表达量均显著下调, *AbARF17*、*AbARF19* 和 *AbARF41* 基因在黄光条件下相对表达量最高, 显著上调, *AbARF21* 基因在不同光质条件下其相对表达量变化无明显差异, 这说明 *AbARF21* 基因表达不受光质影响。

不同 *AbARF* 基因在颠茄不同组织的表达模式具有组织特异性。结果显示 (图 11), 所有 *AbARF* 基因在花中其相对表达量比其他部位低, *AbARF21* 和 *AbARF29* 基因在颠茄的茎、叶、主根和侧根表达量高, 在花中表达量低; *AbARF7*、*AbARF33* 和 *AbARF41* 基因在叶和茎中高表达, 在花、主根和侧根中低表达; 其余 *AbARF* 基因均在叶中相对表达量高, 在花、茎、主根和侧根中相对表达量低。综上所述, *AbARF* 基因在不同环境条件和不同组织部位中其表达模式有所不同。同一基因家族在虽然在

基因结构上有相似性, 但在不同条件下其表达模式有差异, 基因相似性并不代表基因功能相似性。

#### 4 讨论

颠茄是一种含有莨菪碱和东莨菪碱的药用植物, 但关于颠茄发育的生物学机制相关研究较少, *ARF* 基因作为植物的特异转录因子家族, 对调节植物根、叶、花的发育、果实的成熟、衰老和脱落等发育过程至关重要, 此外它们还响应生物和非生物胁迫与参与植物激素信号的串扰。随着越来越多的植物全基因组测序完成, 许多植物中的 *ARF* 基因家族在全基因组范围内进行鉴定, 不同物种中鉴定的 *ARF* 基因数目也不尽相同, 在拟南芥<sup>[13]</sup>、水稻<sup>[17]</sup>、玉米<sup>[34]</sup>、番茄<sup>[35]</sup>和小麦<sup>[36]</sup>中分别鉴定出 22、23、36、22 和 23 个 *ARF* 基因家族成员。本次研究从全基因组水平对颠茄 *ARF* 基因家族进行鉴定, 共鉴定出 41 个 *ARF* 基因, 这与其他物种中 *ARF* 基因数目



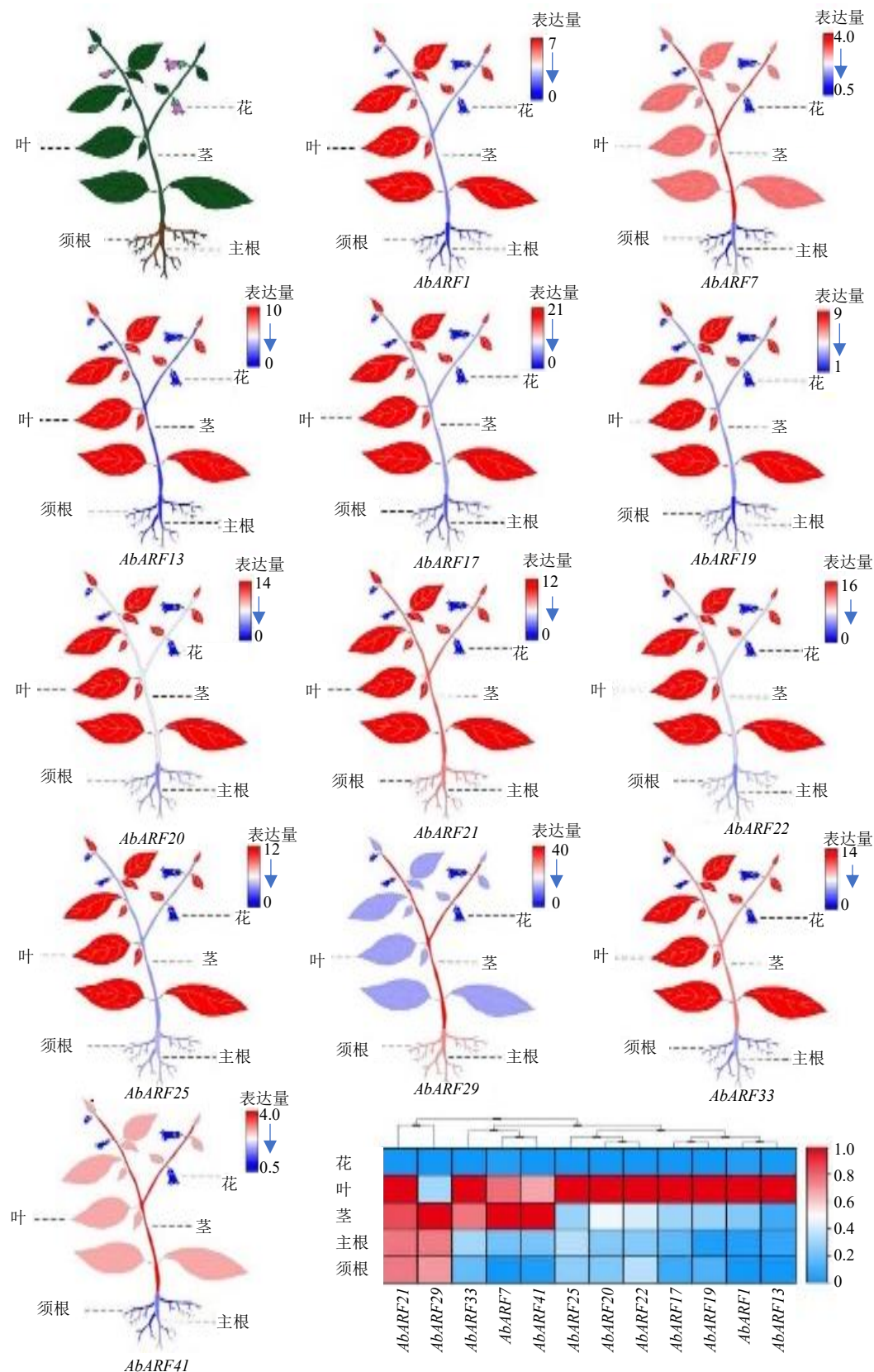


图 11 *AbARF* 基因在不同颠茄组织部位的表达模式

Fig. 11 Expression pattern of *AbARF* genes in different *A. belladonna* tissue parts

有所差异,说明相同基因家族在不同物种间其基因数量存在一定差异。颠茄基因组大小为 1.45 Gb<sup>[37]</sup>,分别为水稻(441 Mb)<sup>[38]</sup>、拟南芥(164 Mb)、番茄(850 Mb)的 3 倍、8 倍和 1.5 倍左右,但是各 ARF 基因数量与其基因组大小不成正比,这可能与颠茄和双子叶共同祖先发生分化后,进行全基因组复制(WGD)、基因丢失事件以及基因扩张程度不同有关。*AbARF1~AbARF41* 不均匀的分布在颠茄 26 条染色体上,*AbARF26* 和 *AbARF37*、*AbARF23* 和 *AbARF38*、*AbARF28* 和 *AbARF40* 彼此在染色体上距离最近,但是整个 *AbARF* 基因家族在染色体上并无基因簇。没有一个 *AbARF* 基因定位在 contig 上,说明此次用于鉴定的颠茄基因组测序质量好。在植物基因组中,基因复制和分化是基因家族扩张和新功能演化的必要步骤<sup>[39]</sup>。基因组内共线性分析结果表明 *AbARF* 基因家族复制类型主要为全基因组复制和片段性复制,除了 *AbARF4* 和 *AbARF35* 外,其余 *AbARF* 都参与了全基因组复制和片段复制事件,有 46 个基因间对存在共线性关系,但未发现串联重复,这与决明<sup>[32]</sup>、人参<sup>[40]</sup>的结果一致,表明全基因组复制和片段性复制在 *AbARF* 基因家族扩张过程中起了至关重要的作用。另外所有共线性关系基因对  $K_a/K_s$  比值均小于 1,说明它们在复制后经过了强烈的纯化选择。

序列同源分析表明,颠茄 ARF 基因家族基因在保守结构域等特征上与其他物种 ARF 基因家族相似,颠茄 ARF 基因家族成员按照是否在 C 端含有 Aux/IAA (CTD) 保守结构域可以分为 2 大类、4 个亚家族,保守结构域分析表明,所有颠茄 ARF 家族成员的 N 末端都具有一个植物特异性 B3 型高度保守结构域,这与拟南芥<sup>[13]</sup>和高粱<sup>[41]</sup>的结果一致。*AbARF23/26/28/31/33/36/37/38/39/40/41* 缺失 CTD 保守结构域,这可能导致它们不能通过和 Aux/IAA 蛋白的异二聚化或自身蛋白的同二聚化促进蛋白-蛋白的相互作用。*AbARF32~41* 外显子、内含子和 CDS 数量均为在 3~5 个左右,这说明 *AbARF32~AbARF41* 基因是颠茄最原始的 ARF 基因,在最原始的 *AbARF* 基因中缺乏 CTD 保守结构域,在随着颠茄系统进化过程中,CTD 保守结构域被演化出来成为 *AbARF* 基因的一部分,以适应新的环境变化,行使新的蛋白功能。

启动子顺式作用元件的预测分析发现 *AbARF* 基因含有大量有关生长发育、响应生物胁迫与非生

物胁迫的相关元件,表明 *AbARF* 基因可能参与相关途径从而调控颠茄的生长发育。在水稻中,低温会诱导 *OsARF4/14/18/19* 基因相对表达量上调从而抵御冷胁迫<sup>[42]</sup>;*AbARF13* 和 *AbARF29* 基因在颠茄低温条件下其相对表达量显著上调,*AbARF1*、*AbARF7*、*AbARF19*、*AbARF21*、*AbARF20*、*AbARF22*、*AbARF33*、*AbARF41* 基因相对表达量显著下调;在植物中 ARF 基因家族的启动子区域有大量的光响应元件,但是光照如何调控 ARF 基因的表达模式鲜有研究。本次研究发现红光和黄光会显著抑制 *AbARF1*、*AbARF7*、*AbARF13*、*AbARF20*、*AbARF22*、*AbARF25*、*AbARF29* 和 *AbARF33* 基因的表达,黄光促进 *AbARF17*、*AbARF19*、*AbARF41* 基因表达。这表明这些 *AbARF* 基因在不同环境条件下发挥不同的作用。同一基因家族的不同基因在植物不同组织部位具有不同的表达模式。在杜仲中,*EuARF19.2* 在幼叶中的表达量最高,*EuARF6.1*、*EuARF6.2* 和 *EuARF8.1* 在叶和茎中的表达水平较高<sup>[43]</sup>;*HsARF9*、*HsARF10* 和 *HsARF21* 在蛇足石杉根中表达量最高,而 *HsARF5* 和 *HsARF7* 在所有组织中均低表达<sup>[44]</sup>;李海波等<sup>[45]</sup>发现在油桐中,ARF 基因在花中的表达水平明显高于其在根、茎、叶和种子中的表达水平;石榴中也出现 *PgARF6a*、*PgARF6b*、*PgARF6c* 在花中较高表达现象<sup>[46]</sup>。但是在本次研究中显示所有 *AbARF* 基因在花中其相对表达量比其他部位低,这和上述植物中的 ARF 基因表达模式截然相反,说明相同基因家族在不同植物中表达模式亦不相同。*AbARF21* 和 *AbARF29* 基因在颠茄的茎、叶、主根和侧根表达量高,在花中表达量低;*AbARF7*、*AbARF33* 和 *AbARF41* 基因在叶和茎中高表达,在花、主根和侧根中低表达;其余 *AbARF* 基因均在叶中相对表达量高,在花、茎、主根和侧根中相对表达量低,表明不同 *AbARF* 基因在颠茄不同组织的表达模式具有组织特异性。颠茄的主要药用成分为莨菪碱和东莨菪碱,尽管 *AbARF* 基因在不同环境条件下及不同组织部位的呈现出不同的表达模式,但是否可能影响生物碱合成途径或者对这些信号做出反应、并调节下游基因表达的机制并不清楚,后续对于这些基因的功能及其与响应通路的关系还需要进一步的实验证实。

综上所述,本研究从全基因组层面对颠茄 ARF 基因家族进行了全面的生物信息学分析,共鉴定出 41 个 *AbARF* 基因,分为 2 大类,4 个亚家族。编码



的氨基酸长度在 603~1 116 aa, 蛋白质相对分子质量为 68 709.35~125 100.60, 蛋白等电点介于 5.27~8.25, 亚细胞定位全部定位于细胞核。所有 *AbARF* 基因均含有 B3 保守结构域。表达模式分析显示 *AbARF* 基因在不同环境条件和不同组织部位中其表达模式特征有所不同。本研究为深入探索 *AbARF* 基因功能奠定理论基础。

**利益冲突** 所有作者均声明不存在利益冲突

## 参考文献

- [1] 中国药典 [S]. 一部. 2020: 395.
- [2] 付维, 郑涛, 姜育松, 等. 颠茄 *AbCYP80F1* 基因的克隆与功能分析 [J]. 中草药, 2020, 51(16): 4291-4299.
- [3] 刘灿黄, 张继, 康帅, 等. 颠茄草的生药学研究 [J]. 中国中药杂志, 2014, 39(9): 1589-1592.
- [4] 汤景翔, 王青, 雷鹏. 东莨菪碱的中枢药理机制研究进展 [J]. 南京医科大学学报: 自然科学版, 2023, 43(8): 1161-1171.
- [5] 郭兰萍, 周良云, 康传志, 等. 药用植物适应环境胁迫的策略及道地药材“拟境栽培” [J]. 中国中药杂志, 2020, 45(9): 1969-1974.
- [6] 安昌, 陆琳, 沈梦千, 等. 植物 *bHLH* 基因家族研究进展及在药用植物中的应用前景 [J]. 生物技术通报, 2023, 39(10): 1-16.
- [7] Lee T I, Young R A. Transcription of eukaryotic protein-coding genes [J]. *Annu Rev Genet*, 2000, 34: 77-137.
- [8] Xiong H B, Lu D D, Li Z Y, et al. The DELLA-ABI4-HY5 module integrates light and gibberellin signals to regulate hypocotyl elongation [J]. *Plant Commun*, 2023, 4(5): 100597.
- [9] Lu L N, Zhai X G, Li X L, et al. Met1-specific motifs conserved in OTUB subfamily of green plants enable rice OTUB1 to hydrolyse Met1 ubiquitin chains [J]. *Nat Commun*, 2022, 13(1): 4672.
- [10] Szemenyei H, Hannon M, Long J A. TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis [J]. *Science*, 2008, 319(5868): 1384-1386.
- [11] Ulmasov T, Hagen G, Guilfoyle T J. Dimerization and DNA binding of auxin response factors [J]. *Plant J*, 1999, 19(3): 309-319.
- [12] Chandler J W. Auxin response factors [J]. *Plant Cell Environ*, 2016, 39(5): 1014-1028.
- [13] Guilfoyle T J, Hagen G. Auxin response factors [J]. *Curr Opin Plant Biol*, 2007, 10(5): 453-460.
- [14] Shen C J, Wang S K, Bai Y H, et al. Functional analysis of the structural domain of *ARF* proteins in rice (*Oryza sativa* L.) [J]. *J Exp Bot*, 2010, 61(14): 3971-3981.
- [15] Okushima Y, Overvoorde P J, ARIMA K, et al. Functional genomic analysis of the auxin response factor gene family members in *Arabidopsis thaliana*: Unique and overlapping functions of *arf7* and *arf19* [J]. *Plant Cell*, 2005, 17(2): 444-463.
- [16] Kim J, Harter K, Theologis A. Protein-protein interactions among the Aux/IAA proteins [J]. *Proc Natl Acad Sci USA*, 1997, 94(22): 11786-11791.
- [17] Wang D K, Pei K M, Fu Y P, et al. Genome-wide analysis of the auxin response factors (*ARF*) gene family in rice (*Oryza sativa*) [J]. *Gene*, 2007, 394(1/2): 13-24.
- [18] Ha C V, Le D T, Nishiyama R, et al. The auxin response factor transcription factor family in soybean: Genome-wide identification and expression analyses during development and water stress [J]. *DNA Res*, 2013, 20(5): 511-524.
- [19] Diao D H, Hu X, Guan D, et al. Genome-wide identification of the ARF (auxin response factor) gene family in peach and their expression analysis [J]. *Mol Biol Rep*, 2020, 47(6): 4331-4344.
- [20] Liu R, Guo Z F, Lu S Y. Genome-wide identification and expression analysis of the aux/IAA and auxin response factor gene family in *Medicago truncatula* [J]. *Int J Mol Sci*, 2021, 22(19): 10494.
- [21] Luo L, Wan Q, Yu Z P, et al. Genome-wide identification of auxin response factors in peanut (*Arachis hypogaea* L.) and functional analysis in root morphology [J]. *Int J Mol Sci*, 2022, 23(10): 5309.
- [22] Li Y H, Han S, Qi Y H. Advances in structure and function of auxin response factor in plants [J]. *J Integr Plant Biol*, 2023, 65(3): 617-632.
- [23] Ellis C M, Nagpal P, Young J C, et al. Auxin response factor1 and auxin response factor2 regulate senescence and floral organ abscission in *arabidopsis thaliana* [J]. *Development*, 2005, 132(20): 4563-4574.
- [24] Wilmoth J C, Wang S C, Tiwari S B, et al. *NPH4/ARF7* and *ARF19* promote leaf expansion and auxin-induced lateral root formation [J]. *Plant J*, 2005, 43(1): 118-130.
- [25] Qi Y H, Wang S K, Shen C J, et al. OsARF12, a transcription activator on auxin response gene, regulates root elongation and affects iron accumulation in rice (*Oryza sativa*) [J]. *New Phytol*, 2012, 193(1): 109-120.
- [26] Alexander L, Grierson D. Ethylene biosynthesis and action in tomato: A model for climacteric fruit ripening [J]. *J Exp Bot*, 2002, 53(377): 2039-2055.
- [27] Wu L J, Tian Z D, Zhang J H. Functional dissection of auxin response factors in regulating tomato leaf shape

- development [J]. *Front Plant Sci*, 2018, 9: 957.
- [28] Nagpal P, Ellis C M, Weber H, *et al.* Auxin response factors *ARF6* and *ARF8* promote jasmonic acid production and flower maturation [J]. *Development*, 2005, 132(18): 4107-4118.
- [29] Vauroid E, Brioudes F, Szécsi J, *et al.* AUXIN RESPONSE FACTOR8 regulates *Arabidopsis* petal growth by interacting with the *bHLH* transcription factor BIGPETALp [J]. *Plant Cell*, 2011, 23(3): 973-983.
- [30] Zhang F Y, Qiu F, Zeng J L, *et al.* Revealing evolution of tropane alkaloid biosynthesis by analyzing two genomes in the Solanaceae family [J]. *Nat Commun*, 2023, 14(1): 1446.
- [31] Chen C J, Chen H, Zhang Y, *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data [J]. *Mol Plant*, 2020, 13(8): 1194-1202.
- [32] 冯昭, 刘世鹏, 吕蕊花, 等. 决明 *Aux/IAA* 和 *ARF* 基因家族鉴定、表达及蛋白互作分析 [J]. *药学报*, 2024, 59(03): 751-763.
- [33] Roth C, Liberles D A. A systematic search for positive selection in higher plants (Embryophytes) [J]. *BMC Plant Biol*, 2006, 6: 12.
- [34] Wang Y J, Deng D X, Shi Y T, *et al.* Diversification, phylogeny and evolution of auxin response factor (*ARF*) family: Insights gained from analyzing maize *ARF* genes [J]. *Mol Biol Rep*, 2012, 39(3): 2401-2415.
- [35] Zouine M, Fu Y Y, Chateigner-Boutin A L, *et al.* Characterization of the tomato *ARF* gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing [J]. *PLoS One*, 2014, 9(1): e84203.
- [36] Qiao L Y, Zhang W P, Li X Y, *et al.* Characterization and expression patterns of auxin response factors in wheat [J]. *Front Plant Sci*, 2018, 9: 1395.
- [37] Zhang F Y, Qiu F, Zeng J L, *et al.* Revealing evolution of tropane alkaloid biosynthesis by analyzing two genomes in the Solanaceae family [J]. *Nat Commun*, 2023, 14(1): 1446.
- [38] Burr B. Mapping and sequencing the rice genome [J]. *Plant Cell*, 2002, 14(3): 521-523.
- [39] 谭政委, 郭水柱, 苏小雨, 等. 全基因组水平金银花 *TCP* 基因家族的鉴定及表达模式分析 [J]. *中草药*, 2024, 55(5): 1665-1676.
- [40] Yan M, Yan Y, Wang P, *et al.* Genome-wide identification and expression analysis of auxin response factor (*ARF*) gene family in *Panax ginseng* indicates its possible roles in root development [J]. *Plants*, 2023, 12(23): 3943.
- [41] Wang S K, Bai Y H, Shen C J, *et al.* Auxin-related gene families in abiotic stress response in *Sorghum bicolor* [J]. *Funct Integr Genomics*, 2010, 10(4): 533-546.
- [42] Du H, Liu H B, Xiong L Z. Endogenous auxin and jasmonic acid levels are differentially modulated by abiotic stresses in rice [J]. *Front Plant Sci*, 2013, 4: 397.
- [43] 刘闵豪, 李龙, 叶靖, 等. 杜仲 *ARF* 基因家族全基因组鉴定和表达分析 [J]. *林业科学*, 2021, 57(3): 170-180.
- [44] 李海波, 史济东, 张恺, 等. 基于全长转录组的蛇足石杉 *ARF* 基因家族鉴定分析 [J]. *广西植物*, 2024, 44(6): 1091-1104.
- [45] 王占军, 李静静, 周子寒, 等. 油桐 *ARF* 基因家族的鉴定及表达分析 [J]. *江苏农业学报*, 2024, 40(6): 984-992.
- [46] Wu M F, Tian Q, Reed J W. *Arabidopsis* microRNA167 controls patterns of *ARF6* and *ARF8* expression, and regulates both female and male reproduction [J]. *Development*, 2006, 133(21): 4211-4218.

[责任编辑 时圣明]