

毛竹生长素反应因子基因的生物信息学分析及差异表达

程占超, 侯丹, 马艳军, 高健

(国际竹藤中心 国家林业局竹藤科学与技术重点开放实验, 北京 100102)

摘要: 生长素反应因子(ARF)基因家族在植物的生长发育中起至关重要的作用。关于毛竹 *Phyllostachys edulis* ARF 基因家族在花器官中生物信息学分析未见报道。以毛竹花器官为材料, 采用生物信息学的方法, 对毛竹中 ARF 基因进行筛选, 并对其系统进化关系、保守基序及在花器官中的差异表达模式进行了初步的分析。结果表明: 毛竹全基因组中含有 44 个 ARF 基因, 分为 I, II, III 类。与拟南芥 *Arabidopsis thaliana* 和水稻 *Oryza sativa* 分别比较, 发现毛竹与水稻存在 11 个姐妹同源基因对。此外, *PheARF13*, *PheARF14* 和 *PheARF35* 在花芽、雌蕊和幼胚中高量表达, 同时 *PheARF2*, *PheARF30*, *PheARF14*, *PheARF13*, *PheARF35*, *PheARF7*, *PheARF37*, *PheARF38* 在雌蕊和幼胚中高量表达, 推测这些基因可能在花发育和种子发育中发挥重要的作用。图 3 表 1 参 32

关键词: 林木育种学; 毛竹; 生长素应答因子; 花发育

中图分类号: S722 **文献标志码:** A **文章编号:** 2095-0756(2017)04-0574-07

Bioinformatic analysis and differential expression of auxin response factor (ARF) gene in *Phyllostachys edulis*

CHENG Zhanchao, HOU Dan, MA Yanjun, GAO Jian

(Key Laboratory of Bamboo and Rattan Science and Technology of State Forestry Administration, International Center for Bamboo and Rattan, Beijing 100102, China)

Abstract: The auxin response factor (ARF) gene family plays a key role in plant growth and developmental processes, such as root and shoot development as well as flower and fruit development. To provide a theoretical basis for flower and seed development of *Phyllostachys edulis*, a genome-wide analysis of the previously undocumented ARF gene family for *Ph. edulis* was conducted. In this study a whole-genome survey of *Ph. edulis* was performed and a detailed analysis of the gene motif and phylogenetic classification was provided. Results showed 44 ARF genes which were classified into three groups. A comparative analysis of the ARF genes among *Ph. edulis*, rice, and *Arabidopsis* suggested a total of 11 sister pairs (OsARF-PheARF) providing insights into various orthologous relationships between OsARFs and PheARFs. For ARF expression patterns of diverse floral organs, *PheARF13*, *PheARF14*, and *PheARF35* showed the highest expression in the flower bud, pistil, and young embryo; whereas, *PheARF2*, *PheARF30*, *PheARF14*, *PheARF13*, *PheARF35*, *PheARF7*, *PheARF37*, and *PheARF38* were highly expressed in the pistil and the young embryo. This study suggested that ARF genes may play a very critical role during flower and fruit development of *Ph. edulis*. [Ch, 3 fig. 1 tab. 32 ref.]

Key words: forest tree breeding; *Phyllostachys edulis*; auxin response factors; floral development

植物生长激素能够调节或者影响植物不同的生理过程, 比如顶端优势、侧根萌发、导管分化、胚胎发育和芽的伸长, 也可以促进细胞的分裂、延伸和分化^[1]。在分子水平上, 生长素能够特异性地调节基

收稿日期: 2016-06-13; 修回日期: 2016-07-14

基金项目: 国家高技术研究发展计划(“863”计划)项目(2013AA102607-4); 国家自然科学基金资助项目(31570673)

作者简介: 程占超, 博士, 从事林木遗传育种研究。E-mail: chengzhan_chao@126.com。通信作者: 高健, 研究员, 博士, 博士生导师, 从事林木遗传育种研究。E-mail: gaojian@icbr.ac.cn

因的表达^[2]。生长素信号转导相关的 3 类主要蛋白是 Aux/IAAs, ARFs 和 SCF 复合体^[3-4]。作为生长素信号途径一个重要的成员, ARFs 基因能够通过结合生长素启动子上的 AuxRE 来调控生长素反应基因的表达^[5]。这些 ARFs 基因调控不同的发育过程, 比如顶芽形成, 导管组织的形成, 胚胎、花和果实的形成^[6]。一个经典的 ARF 蛋白, 结构相对简单, 大部分蛋白的分子量为 70~130 kD, 含有 3 个保守的结构域, 它们分别是 N 末端 DNA 结合结构域(DBD), 中间区域(MR)和 C 末端二聚体结构域(CTD)。不同的结构域也决定 ARF 不同的功能, DBD 在信号转导中起着鉴别作用, CTD 可以使 ARF 之间形成二聚体, 而 MR 则起着转录激活或抑制的功能^[7]。ULMASOV 等^[8]鉴定出第 1 个 ARF 基因, 即 *AtARF1*, 先后在拟南芥 *Arabidopsis thaliana* 和水稻 *Oryza sativa* 中鉴定出 23 和 25 个 ARF 基因^[9-10]。ARFs 大部分在植物中发现, 比如在双子叶、单子叶、裸子和蕨类植物中都有发现, 但是在动物和微生物中至今没有发现。因此, ARFs 是植物中特有的一类转录因子^[11]。ARF 的生物学功能主要来自拟南芥 ARF 基因功能缺失的表型, 到目前为止, 已经发现很多 ARF 基因及其功能, 比如 *ARF3* 和 *ARF7*^[12-14]。研究表明: 拟南芥 *arf3* 的突变体会出现花蕊基部和顶端的发育不良, 说明 *ARF3* 在调节花器官发育上起作用; 缺失 *ARF7*, 导致上胚轴的向光性和下胚轴的向地性功能的消失。这 2 个 ARF 基因有着不同的功能, 很少在功能上出现冗余现象。毛竹 *Phyllostachys edulis* 与水稻、玉米 *Zea mays* 等同属于禾本科 Gramineae 单子叶植物, 但是, 毛竹营养生长周期长, 开花时期不确定, 开花后死亡, 导致竹林面积减少, 对经济发展和生态环境造成重大损失和破坏, 竹子开花的调控机制一直是竹类植物研究中的难点和热点。目前, ARF 家族基因在模式植物拟南芥、水稻、玉米中已有研究, 而毛竹 ARF 家族基因在花器官和幼胚发育上鲜有报道。本研究通过生物信息学的方法, 根据毛竹的基因组, 鉴定 ARF 基因家族, 进行进化树分析、基序分析、基因差异表达模式分析, 为研究 ARF 基因在毛竹花和种子发育过程中的功能奠定基础。

1 材料与方 法

1.1 实验材料

毛竹开花实验地位于广西壮族自治区桂林市南岭山系的西南部。该毛竹林属于自然生长状态, 基本无人干扰。以毛竹的花器官为材料, 进行解剖, 分离出花芽、苞片、颖片、稃片、雄蕊、雌蕊和幼胚以及未开花的成熟叶片, 建立 8 个样本进行转录组高通量测序。

1.2 方 法

1.2.1 毛竹中 PheARF 家族成员的检索 从拟南芥基因组数据库(<http://www.arabidopsis.org/>)和水稻基因组数据库(<http://rice.plantbiology.msu.edu/index.html>)中分别检索拟南芥和水稻中 ARF 蛋白序列; 毛竹 ARF 相关蛋白序列从毛竹基因组数据库(<http://www.bamboogdb.org/index.jsp>)中获得。毛竹 PheARF 家族蛋白的序列号, 开放阅读框的长度、氨基酸数目、分子质量和等电点在表 1 中提供。毛竹 ARF 分子量及等电点数据通过 ExPASy(http://web.expasy.org/compute_pi/)获得。

1.2.2 进化树分析 用 ClustalX 1.83(<http://www.clustal.org/>)^[15] 软件对蛋白全长的多序列比对进行分析, 进化树分析前去掉比对序列的差异和不明确序列。用 no-rooted neighbor-joining 方法通过 MEGA 6.0 (<http://www.megasoftware.net/mega.html>)^[16] 构建系统进化树。

1.2.3 基序分析 MEME version 4.11.2(<http://meme-suite.org/>)^[17] 在线工具鉴定候选蛋白序列的保守区域, 公式为 any, maximum number of motifs = 20, minimum width ≥ 6 和 maximum width ≤ 200 。

1.2.4 表达量分析 将花芽、苞片、颖片、稃片、雄蕊、雌蕊和幼胚以及未开花的成熟叶片的 FPKM 值输入到 Cluster 3.0, 用 Java TreeView 生成热点图^[18]。

2 结 果

2.1 毛竹 PheARF 家族序列鉴定

通过 23 个拟南芥和 25 个水稻 ARF 蛋白序列检索毛竹基因组数据库, 共得到 44 个 ARF 蛋白(表 1)。从表 1 可以看出, ARF 家族的蛋白的等电点为 5.22~9.23, 蛋白序列的长度为 192~1 257 个氨基酸, 分子量的大小为 31.70 ~139.07 kD。

表1 毛竹中 ARF 成员数量及其属性

Table 1 Properties and numbers of ARF identified from *Phyllostachys edulis*

ARF 蛋白 名称	毛竹 ARF 序列号	氨基酸 长度	开放阅 读框/bp	等电 点	分子 量/kD	ARF 蛋白 名称	毛竹 ARF 序列号	氨基酸 长度	开放阅 读框/bp	等电 点	分子 量/kD
PheARF1	PH01013505G0010	278	837	7.61	31.67	PheARF23	PH01000483G0220	866	2 601	5.74	96.15
PheARF2	PH01008675G0010	260	780	6.70	28.56	PheARF24	PH01000384G0170	763	2 292	5.36	84.80
PheARF3	PH01007612G0010	672	2 019	5.72	74.77	PheARF25	PH01000305G0690	635	1 908	7.25	69.46
PheARF4	PH01005322G0010	792	2 379	6.09	87.49	PheARF26	PH01000277G0820	591	1 776	6.90	64.69
PheARF5	PH01003631G0130	832	2 499	5.84	92.70	PheARF27	PH01000259G1120	904	2 715	6.60	100.71
PheARF6	PH01003096G0200	921	2 766	5.31	102.29	PheARF28	PH01000227G0020	908	2 727	5.87	100.72
PheARF7	PH01002857G0060	708	2 127	5.99	78.96	PheARF29	PH01000222G0180	750	2 253	6.11	83.00
PheARF8	PH01002850G0020	476	1 431	9.23	54.32	PheARF30	PH01000183G0570	897	2 694	6.03	99.11
PheARF9	PH01002806G0200	840	2 523	6.97	93.01	PheARF31	PH01000176G0540	1 121	3 366	6.21	124.73
PheARF10	PH01002685G0120	605	1 818	6.29	65.89	PheARF32	PH01000116G0910	817	2 454	7.61	90.11
PheARF11	PH01002498G0280	658	1 977	6.46	72.04	PheARF33	PH01000114G0050	1 257	3 774	6.23	139.07
PheARF12	PH01002160G0180	757	2 274	6.16	84.64	PheARF34	PH01000093G0670	750	2 253	6.82	83.28
PheARF13	PH01001899G0250	728	2 187	5.68	80.88	PheARF35	PH01000087G1340	738	2 217	5.85	82.09
PheARF14	PH01001690G0310	701	2 106	6.50	76.99	PheARF36	PH01000057G1420	928	2 787	5.54	102.92
PheARF15	PH01001555G0390	878	2 637	6.16	97.81	PheARF37	PH01000046G0220	734	2 205	6.70	80.51
PheARF16	PH01001285G0430	603	1 812	7.02	65.37	PheARF38	PH01000044G0540	439	1 320	5.87	47.45
PheARF17	PH01001212G0190	364	1 095	7.25	39.91	PheARF39	PH01000018G0940	667	2 004	6.96	74.36
PheARF18	PH01001026G0300	744	2 235	7.55	81.07	PheARF40	PH01000014G0610	427	1 284	9.00	47.32
PheARF19	PH01000667G0020	637	1 914	5.22	71.47	PheARF41	PH01000011G0660	554	1 665	5.66	61.35
PheARF20	PH01000626G0040	693	2 082	5.34	77.55	PheARF42	PH01000002G3110	828	2 487	6.41	92.55
PheARF21	PH01000623G0440	1 186	3 561	8.87	132.01	PheARF43	PH01000093G0690	192	579	5.44	21.02
PheARF22	PH01000548G0300	985	2 958	6.05	110.11	PheARF44	PH01000237G0420	514	1 545	8.52	56.44

2.2 毛竹 PheARF 家族进化分析

系统进化结果显示：44 个基因之间相似性比较高(图 1)。将 44 个毛竹的 ARF 蛋白列同 23 个拟南芥和 25 个水稻的蛋白序列同时比对，发现可以将这 3 个物种的 ARF 分为三大类，分别为 I，II 和 III。其中有 13 个基因分布在 I 类，17 个基因分布在 II 类，还有 14 个基因属于 III 类。大部分的 PheARFs 包含 3 个经典的结构域：DBD，结构域 II 和 AUX/IAA 家族结构域。有 11 个与水稻的同源关系较近的基因对：*OsAR25/PheARF30*(II)，*OsARF16/PheARF40*(II)，*OsARF19/PheARF31*(II)，*OsARF15/PheARF37*(III)，*OsARF2/PheARF14*(III)，*OsARF14/PheARF39*(III)，*OsARF3/PheARF26*(III)，*OsARF13/PheARF38*(III)，*OsARF8/PheARF16*(III)，*OsARF10/PheARF10*(III)和 *OsARF18/PheARF18*(III)。这些基因大部分出现在 II 类和 III 类中，说明这 2 类的基因相对比较保守。

2.3 毛竹 PheARF 基序分析

为了进一步了解毛竹 ARF 保守区域的结构，通过 MEME 在线工具构建基序分析图，结果如图 2 所示。毛竹 ARF 成员含有的基序结构不一，大概有 20 个不同的基序组成。其中每个基因都包含不同种类的基序，大部分的基序为 4~17 个。大部分的 PheARF 基因都包含 Motif1~8，Motif10，Motif11，Motif14，Motif18 和 Motif20 等，这些基序出现的次数较多，而 Motif13，Motif16 和 Motif17 等基序在这些基因中不常见。从进化角度来看，同源关系较近的 I，II，III 类的基因在基序的长度和种类较为相似，比如 I 类的 *PheARF4*，*PheARF9* 和 *PheARF32* 等，II 类的 *PheARF22*，*PheARF31* 和 *PheARF33* 等，III 类的 *PheARF17*，*PheARF18* 和 *PheARF37* 等。以上结果说明大部分的毛竹 ARF 家族基因是相当保守的。

2.4 毛竹 PheARF 基因在花器官中差异表达分析

为了研究毛竹 PheARF 基因在花器官发育中的作用和调控机制，进一步分析 PheARF 基因在花芽、苞片、颖片、稃片、雌蕊、雄蕊、幼胚和未开花的叶中表达分析。结果表明：*PheARF12*，*PheARF13*，*PheARF14*，*PheARF15*，*PheARF24*，*PheARF35*，*PheARF43*，*PheARF23* 和 *PheARF32* 在花芽中高量表

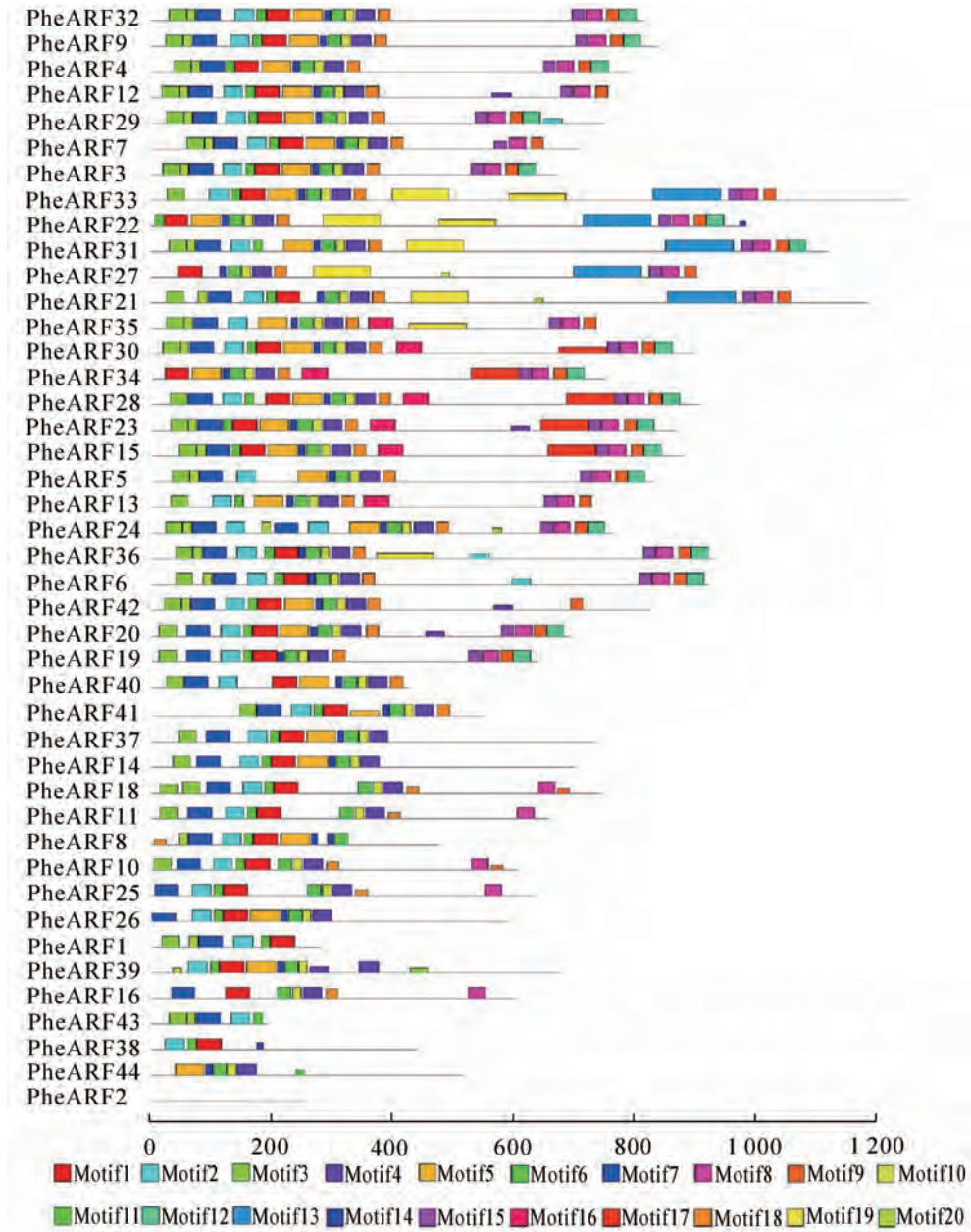


图2 毛竹 ARF 保守基序分析

Figure 2 Distribution of conserved motifs in *Phyllostachys edulis* ARF proteins identified using MEME search tool

组。以上分析表明：毛竹 ARF 基因复制事件的频繁发生有可能导致毛竹 ARF 家族成员数量增多。通过系统进化分析，毛竹 ARF 家族基因大致分为 3 个亚组(图 1)，亚组内的基因相似性比较高，但是亚组之间相似性不高，或许在基因功能上也有差异。毛竹与拟南芥和水稻的 ARF 家族的系统进化关系表明，毛竹与水稻之间有 11 对同源基因对，这些基因对都属于 II 类和 III 类，一方面说明毛竹的 ARF 家族与水稻有很高的同源性(图 1)，另一方面说明毛竹 ARF 家族 II 类和 III 类的基因保守性很高。

研究人员对水稻和拟南芥的 ARF 基因进行了广泛地研究^[29-31]。ARF 基因在花和种子发育中起着非常重要的作用，为研究毛竹的开花及种子发育提供了许多有用的信息。据报道，拟南芥的 *AtARF3* 和 *AtARF4* 参与花的发育，在花中高量表达^[32]。与 *AtARF3* 和 *AtARF4* 同源的毛竹 *PheARF14* 和 *PheARF37* 在幼胚和雌蕊中都高量表达，说明 2 个毛竹 ARF 基因可能与 *AtARF3* 和 *AtARF4* 基因功能相似，推测它们在毛竹花和种子发育过程中起着关键的作用。

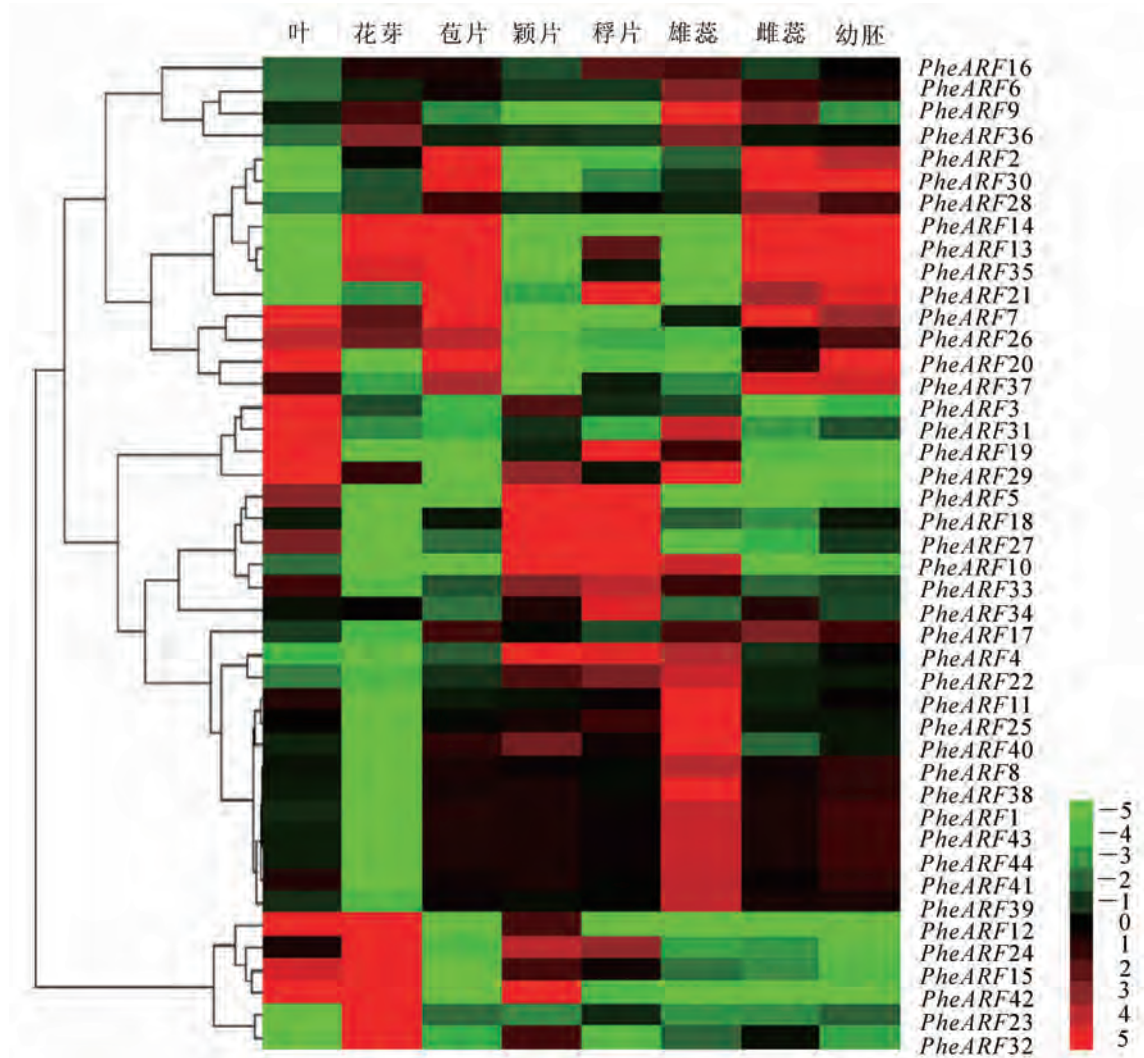


图 3 毛竹 *PheARF* 基因在不同花器官中表达模式分析

Figure 3 Expression analysis of *PheARF* genes among different flower organs of *Phyllostachys edulis*

4 参考文献

- [1] DAVIES P J. *Plant Hormones: Physiology, Biochemistry and Molecular Biology* [M]. Dordrecht: Kluwer Academic Publishers, 2013.
- [2] THEOLOGIS A. Rapid gene regulation by auxin [J]. *Ann Rev Plant Physiol*, 1986, **37**(1): 407 – 438.
- [3] GUILFOYLE T J, ULMASOV T, HAGEN G. The ARF family of transcription factors and their role in plant hormone-responsive transcription [J]. *Cell Mol Life Sci*, 1998, **54**(7): 619 – 627.
- [4] SHEN Chenjia, YUE Runqing, SUN Tao, *et al.* Genome-wide identification and expression analysis of auxin response factor gene family in *Medicago truncatula* [J]. *Front Plant Sci*, 2015, **6**(11): 3932 – 3935.
- [5] TIWARI S B, HAGEN G, GUILFOYLE T. The roles of auxin response factor domains in auxin-responsive transcription [J]. *Plant Cell*, 2003, **15**(2): 533 – 543.
- [6] GUILFOYLE T J, HAGEN G. Auxin response factors [J]. *Curr Opin Plant Biol*, 2007, **10**(5): 453 – 460.
- [7] RAMOS J A, ZENSER N, LEYSER O, *et al.* Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain II and is proteasome dependent [J]. *Plant Cell*, 2001, **13**(10): 2349 – 2360.
- [8] ULMASOV T, HAGEN G, GUILFOYLE T J. ARF1, a transcription factor that binds to auxin response elements [J]. *Science*, 1997, **276**(5320): 1865 – 1868.
- [9] HAGEN G, GUILFOYLE T. Auxin-responsive gene expression: genes, promoters and regulatory factors [J]. *Plant Mol Biol*, 2002, **49**(3/4): 373 – 385.
- [10] REMINGTON D L, VISION T J, GUILFOYLE T J, *et al.* Contrasting modes of diversification in the Aux/IAA and

- ARF gene families [J]. *Plant Physiol*, 2004, **135**(3): 1738 – 1752.
- [11] LISCUM E, REED J W. Genetics of Aux/IAA and ARF action in plant growth and development [J]. *Plant Mol Biol*, 2002, **49**(3/4): 387 – 400.
- [12] SESSIONS R. *Arabidopsis* (Brassicaceae) flower development and gynoecium patterning in wild type and ettin mutants [J]. *Am J Bot*, 1997, **84**(9): 1179 – 1179.
- [13] HARDTKE C S, BERLETH T. The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development [J]. *EMBO J*, 1998, **17**(5): 1405 – 1411.
- [14] HARPER R M, STOWE-EVANS E L, LUESSE D R, *et al.* The NPH4 locus encodes the auxin response factor ARF7, a conditional regulator of differential growth in aerial *Arabidopsis* tissue [J]. *Plant Cell*, 2000, **12**(5): 757 – 770.
- [15] THOMPSON J D, GIBSON T J, PLEWNIAC F, *et al.* The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. *Nucl Acids Res*, 1997, **25**(24): 4876 – 4882.
- [16] TAMURA K, STECHER G, PETERSON D, *et al.* MEGA6: molecular evolutionary genetics analysis version 6.0 [J]. *Mol Biol Evol*, 2013, **30**(12): 2725 – 2729.
- [17] BAILEY T L, ELKAN C. The value of prior knowledge in discovering motifs with MEME [C]. *Proc Third Int Conf Intel Syst Mol Biol*, California: AAAI Press, 1995: 21 – 29.
- [18] SALDANHA A J. Java Treeview-extensible visualization of microarray data [J]. *Bioinformatics*, 2004, **20**(17): 3246 – 3248.
- [19] FARCOT E, LAVEDRINE C, VERNOUX T. A modular analysis of the auxin signalling network [J]. *PloS One*, 2015, **10**(3): 864 – 865.
- [20] ULMASOV T, MURFETT J, HAGEN G, *et al.* Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements [J]. *Plant Cell*, 1997, **9**(11): 1963 – 1971.
- [21] 王垒, 陈劲枫, 娄丽娜, 等. 黄瓜果实中 ARF 和 Aux/IAA 基因对外源激素的应答[J]. 西北植物学报, 2011, **31**(6): 1127 – 1131.
- WANG Lei, CHEN Jinfeng, LOU Lina, *et al.* Expression analysis about some ARF and Aux/IAA family members in fruits of cucumber with exogenous hormones [J]. *Acta Bot Boreal-Occident Sin*, 2011, **31**(6): 1127 – 1131.
- [22] 史梦雅, 李阳, 张巍, 等. 生长素反应因子作用机制研究进展[J]. 生物技术通报, 2012(8): 24 – 28.
- SHI Mengya, LI Yang, ZHANG Wei, *et al.* Progress in mechanism of auxin response factors [J]. *Biotechnol Bull*, 2012(8): 24 – 28.
- [23] AUDRAN-DELALANDE C, BASSA C, MILA I, *et al.* Genome-wide identification, functional analysis and expression profiling of the Aux/IAA gene family in tomato [J]. *Plant Cell Physiol*, 2012, **53**(4): 659 – 672.
- [24] CHO H, RYU H, RHO S, *et al.* A secreted peptide acts on BIN2-mediated phosphorylation of ARFs to potentiate auxin response during lateral root development [J]. *Nat Cell Biol*, 2014, **16**(1): 66 – 76.
- [25] PAULL R E, IRIKURA B, WU Pingfang, *et al.* Fruit development, ripening and quality related genes in the papaya genome [J]. *Trop Plant Biol*, 2008, **1**(3/4): 246 – 277.
- [26] KUMAR R, TYAGI A K, SHARMA A K. Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development [J]. *Mol Gen Genomics Mgg*, 2011, **285**(3): 245 – 260.
- [27] OHNO S. *Evolution by Gene Duplication* [M]. Heidelberg: Springer Verlag, 2013.
- [28] PENG Zhenhua, LU Ying, LI Lubin, *et al.* The draft genome of the fast-growing non-timber forest species moso bamboo (*Phyllostachys heterocycla*) [J]. *Nat Gen*, 2013, **45**(4): 456 – 461.
- [29] QI Yanhua, WANG Suikang, SHEN Chenjia, *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.
- [30] COLE M, CHANDLER J, WEIJERS D, *et al.* DORNROSCHE is a direct target of the auxin response factor MONOPTEROS in the *Arabidopsis* embryo [J]. *Development*, 2009, **136**(10): 1643 – 1651.
- [31] LI Jisheng, DAI Xinhua, ZHAO Yunde. A role for auxin response factor 19 in auxin and ethylene signaling in *Arabidopsis* [J]. *Plant Physiol*, 2006, **140**(3): 899 – 908.
- [32] HUNTER C, WILLMANN M R, WU Gang, *et al.* Trans-acting siRNA-mediated repression of Ettin and ARF4 regulates heteroblasty in *Arabidopsis* [J]. *Development*, 2006, **133**(15): 2973 – 2981.