



# 大豆 *miR164* 家族的生物信息学分析

滕 露, 于月华, 何茹月, 陈全家, 倪志勇

(新疆农业大学 农学院, 新疆 乌鲁木齐 830052)

**摘 要:** *microRNA164* 是一个高度保守的 *miRNA* 家族, 广泛参与植物的细胞和生理过程。本论文旨在采用生物信息学方法, 了解大豆 *miR164* 基因家族在大豆生长发育及逆境胁迫应答过程中扮演的重要角色, 为其在大豆的分子育种和品种改良方面的研究提供理论基础。利用 *miRNA*、Plant MicroRNA Database、PLACE 和 NCBI 数据库以及 Clustal X2.0、MEGA 5.0 和 DNAMAN 软件对 *gma-miR164* 基因家族的序列情况、染色体定位、靶基因调控功能、启动子元件、二级结构和系统发育树进行生物信息学分析。在 miRBase 中搜索到 11 条 *gma-miR164* 基因同源序列, 分别分布在 7 条染色体上, 其中以位于 Chr3 上的序列最多, 共有 3 条, 位于 Chr10 和 Chr19 上各有 2 条, 而在 Chr02、Chr09、Chr18 和 Chr20 上各有 1 条。11 个 *gma-miR164* 基因家族成员共预测到 6 个靶基因, 均为 NAC 转录因子。*gma-miR164* 基因家族成熟序列的碱基保守性很高, 其前体序列均可形成稳定的二级茎环结构。启动子中顺式调控元件分析表明, ABRE、DRE、MYB 和 MYC 这 4 个元件在 *gma-miR164* 基因家族启动子序列中分布不均, 其中以 MYB 和 MYC 元件数目居多。本研究为探寻 *miR164* 基因家族在逆境胁迫应答中的调控作用提供了数据基础和理论依据。

**关键词:** 大豆; *miR164* 家族; 生物信息学; 启动子分析

## Bioinformatics Analysis of Soybean *miR164* Family

TENG Lu, YU Yue-hua, HE Ru-yue, CHEN Quan-jia, NI Zhi-yong

(College of Agriculture, Xinjiang Agricultural University, Urumqi 831300, China)

**Abstract:** *MicroRNA164* is a highly conserved *miRNA* family, widely involved in plant cell and physiological processes. The purpose of this study was to investigate the important role of *miR164* gene family in soybean growth and stress response by bioinformatics, and to provide a theoretical basis for molecular breeding and variety improvement in soybean. Using Plant MicroRNA Database, PLACE and NCBI Database, Clustal X2.0, MEGA5.0 and DNAMAN to analyze the sequence, chromosome location, target gene regulation function, promoter element, secondary structure and the phylogenetic tree was analyzed by bioinformatics. In the miRBase, 11 *gma-miR164* gene homologous sequences were searched and distributed on 7 chromosomes. There were 3 sequences in Chr3, two in Chr10 and two in Chr19, and one in Chr02, Chr09, Chr18 and Chr20. 11 members of *gma-miR164* gene family predicted 6 target genes, all of which were NAC transcription factors. The mature sequence of *gma-miR164* gene family was highly conserved, and its precursor sequence can form a stable secondary stem ring structure. The analysis of cis-regulatory elements in promoter showed that the four elements of ABRE, DRE, MYB and MYC were found in *gma-miR164* gene. The family promoter sequences were unevenly distributed, in which the number of MYB and MYC elements was the majority. This study provides a data basis and theoretical basis for exploring the regulatory role of *miR164* family in stress response.

**Keywords:** Soybean; *MiR164* family; Bioinformatics; Promoter analysis

MicroRNA (miRNA) 是一类由内源基因编码的单链 RNA 分子, 能够形成茎环结构, 长度约为 21 ~ 25 个氨基酸, 它们不仅调控植物的生长发育, 而且在生物形态建成, 信号传导, 逆境胁迫响应等方面发挥重要作用, 在动植物中参与转录后基因表达调控<sup>[1-2]</sup>。植物 miRNA 对靶基因的转录后调控是实

现其功能的主要方式, 在植物中 miRNA 通过与靶基因 mRNA 序列互补方式介导 mRNA 降解或抑制其翻译<sup>[3]</sup>, 从而精细调控某个基因的表达。分析 miRNA 的进化, 上游启动子调控元件, 预测其靶基因对揭示 miRNA 的调控机制具有重要意义。

*miR164* 主要靶基因是植物 NAC 转录因子, 而

收稿日期: 2018-06-04

基金项目: 国家自然科学基金(31660295); 中国博士后科学基金(2015M582741); 新疆维吾尔自治区天山英才计划(201720085); 新疆农业大学校级大学生创新项目(dxscx2017001, dxscx2018001)。

第一作者简介: 滕露(1995 -), 女, 硕士, 主要从事植物分子生物学研究。E-mail: 894546653@qq.com。

通讯作者: 倪志勇(1981 -), 男, 博士, 副教授, 主要从事植物分子生物学研究。E-mail: nizhiyong@126.com。

NAC 转录因子家族也同样是植物特有的<sup>[4]</sup>, NAC 转录因子参与细胞的分裂以及次生壁合成等生物发育过程<sup>[5]</sup>, 并且与高盐、低温、干旱等逆境胁迫响应过程相关<sup>[6]</sup>。miRNA 数据库 PMRD (plant microRNA database) 共收录了 43 种植物的 *miR164* 基因, 部分植物 *miR164* 基因家族成员的功能已得到解析。例如, 番茄 *miR164* 能影响花的形态建成及果实发育, 研究发现 *miR164* 超表达的番茄开花后期花瓣不脱落而是随果实继续生长, 果实形状奇特且不形成种子, 表明 *miR164* 可能调控花器官形成及果实发育过程<sup>[7]</sup>。杨树 *miR164* 在低氮胁迫前 12 h 诱导表达, 而后下调表达, 至 48 h 几乎检测不到, 靶基因 NAC1 显示出相反表达模式, 并通过调控下游关键基因的表达, 进而调控根系生长以适应低氮胁迫<sup>[8]</sup>。在盐胁迫下耐盐棉花中 *miR164a*、*miR164c*、*miR164d* 的表达量较敏感品种表达量上调<sup>[9]</sup>。拟南芥中, *miR164* 调控营养生长以及花器官的形成<sup>[10]</sup>, 随着拟南芥叶片的衰老, *miR164* 的表达量会下降<sup>[11]</sup>。甜橙 *miR164* 在球形胚中特异性表达, 与球形胚的形成密切相关<sup>[12]</sup>。

利用生物信息学方法分析目的基因的启动子中含有的顺式作用元件, 能够预测目的基因可能参与的生物进程。Bazzini 等<sup>[13]</sup> 分析了拟南芥 *miR164a* 的启动子, 发现含有赤霉素等激素和抗病响应相关元件, 通过将该启动子转化拟南芥, 与生物信息学分析结果一致, 该启动子受油菜花叶病毒和赤霉素诱导上调表达。Li 等<sup>[14]</sup> 分析了拟南芥中 11 个 *miR169* 家族成员的启动子序列, 发现部分启动子含有 ABRE 和 HSE 胁迫相关元件, 通过 qPCR 方法证明了启动子元件预测结果, 发现一些 *miR169* 受 ABA 和高温诱导。Zhao 等<sup>[15]</sup> 通过生物学信息学方法分析发现水稻 *miR169g* 含有干旱和盐胁迫相关的 DRE 元件, 通过 RNA gel 杂交分析发现 *miR169* 受盐胁迫诱导表达。

罗中钦<sup>[16]</sup> 通过 miRNA 芯片发现 *gma-miR164* 受干旱、盐和低温胁迫的诱导。Sun 等<sup>[17]</sup> 通过小 RNA 测序分析发现 *gma-miR164b/d* 在大豆根中受盐胁迫上调表达。许硕等<sup>[18]</sup> 通过小 RNA 测序发现 *gma-miR164* 在大豆叶片中受盐胁迫下调表达, 在低盐条件下, 干扰 *gma-miR164* 的转基因拟南芥侧根数量多于野生型。以上研究结果表明 *gma-miR164* 与逆境胁迫相关, 有必要对其进行生物信息学分析以进一步了解其表达调控机制和功能。然而, 目前对大豆 *miR64* 家族启动子的分析尚未开展。

本研究利用 miRNA、NCBI 等数据库及 Clustal X 2.0、MEGA 5.0 和 DNAMAN 在线软件对 *gma-miR164* 基因家族成员进行前体序列、成熟序列、二级结构、系统发育进化树和靶基因等生物信息学分析, 为明确其在大豆生长发育及逆境胁迫应答过程中的作用和后续大豆分子育种及抗逆性研究提供参考依据。

## 1 材料与方法

### 1.1 数据来源

通过 miRNA 数据库 miRBase (<http://www.mirbase.org/>) 检索 *gma-miR164* 基因的相关数据, 获得 *gma-miR164* 基因的前体序列, 成熟序列和染色体定位信息。

### 1.2 系统发育进化树分析

采用 Clustal X 2.0 对 *gma-miR164* 基因家族前体与拟南芥 (*Arabidopsis thaliana*)、水稻 (*Oryza sativa*) 和蒺藜苜蓿 (*Medicago truncatula*) 这 3 个物种的前体序列进行多序列比对分析, 比对结果采用 MEGA 5.0 软件通过邻接法 (Neighbor-Joining method) 构建系统发育进化树, 参数设为默认值。

### 1.3 前体二级结构预测

根据在 miRNA 数据库 miRBase 下载的前体序列, 利用 DNAMAN 软件预测 *gma-miR164* 基因家族前体序列的二级茎环结构, 系统参数设为默认值。

### 1.4 *miR164* 靶基因预测

利用 Plant MicroRNA Database (<http://bioinformatics.cau.edu.cn/PMRD/>) 网站, 在线预测 *gma-miR164* 的靶基因, 参数采用默认值。

### 1.5 *miR164* 启动子中顺式调控元件的分析

为进一步了解 *miR164* 基因表达的调控机制, 通过 PLACE (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>) 在线分析大豆的 *miR164* 家族成员上游 3 000 bp 启动子序列中的与非生物胁迫应答相关的顺式作用元件。*gma-miR164* 家族成员的全基因组序列下载于 NCBI (<https://www.ncbi.nlm.nih.gov/>) 数据库。

## 2 结果与分析

### 2.1 *gma-miR164* 基因家族序列分析及基因定位

miRBase 中搜索到 11 条 *gma-miR164* 基因同源序列 (表 1), *gma-miR164b*、*gma-miR164c* 和 *gma-miR164d* 基因成熟序列含有 20 个碱基, 其它 8 条 *gma-miR164* 基因成熟序列含有 21 个碱基。

表 1 *gma-miR164* 基因家族成熟序列及其基因定位结果

Table 1 Mature sequence of *gma-miR164* gene family and their locations

miRNA	成熟 miRNA Mature miRNA	基因位置 Gene location
<i>gma-miR164a</i>	UGGAGAAGCAGGGCACGUGCA	chr19; 49678662..49684756 [ + ]
<i>gma-miR164b</i>	UGGAGAAGCAGGGCACGUGC	chr10; 1555481..1561570 [ + ]
<i>gma-miR164c</i>	UGGAGAAGCAGGGCACGUGC	chr3; 5194846..5200935 [ - ]
<i>gma-miR164d</i>	UGGAGAAGCAGGGCACGUGC	chr10; 38615021..38621110 [ + ]
<i>gma-miR164e</i>	UGGAGAAGCAGGGCACGUGCA	chr2; 1532426..1538520 [ + ]
<i>gma-miR164f</i>	UGGAGAAGCAGGGCACGUGCA	chr3; 43530426..43536528 [ + ]
<i>gma-miR164g</i>	UGGAGAAGCAGGGCACGUGCA	chr3; 44892640..44898731 [ + ]
<i>gma-miR164h</i>	UGGAGAAGCAGGGCACGUGCA	chr9; 49457287..49463425 [ - ]
<i>gma-miR164i</i>	UGGAGAAGCAGGGCACGUGCA	chr18; 49503475..49509613 [ + ]
<i>gma-miR164j</i>	UGGAGAAGCAGGGCACGUGCA	chr19; 48276156..48282243 [ + ]
<i>gma-miR164k</i>	UGGAGAAGCAGGGCACGUGCA	chr20; 46916054..46922162 [ - ]

利用 *gma-miR164* 基因家族成员的前体序列做基因家族内部进化树,结果分为两大类,*miR164c*、*miR164h* 和 *miR164i* 在同一支上,其他成员在另一分支上,其中 *miR164a* 和 *miR164g* 两条前体序列的相似性较高,达到 95.79%。采用 DNAMAN 软件

对 *gma-miR164* 基因家族成熟序列进行比对分析,结果发现具有相同长度的 *miR164* 成员的成熟序列完全一致,说明大豆的 *miR164* 是非常保守的基因家族(图 1)。

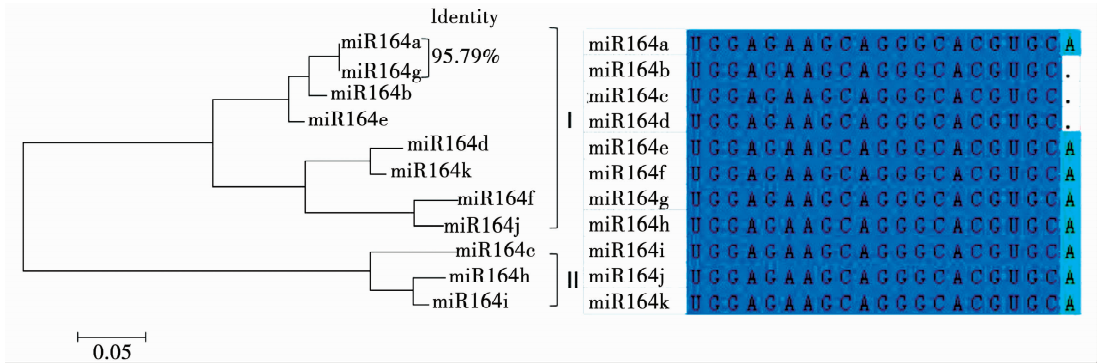


图 1 *gma-miR164* 基因家族成员成熟序列碱基保守性分析结果

Fig. 1 Conservation analysis for base of *gma-miR164* gene family mature sequence

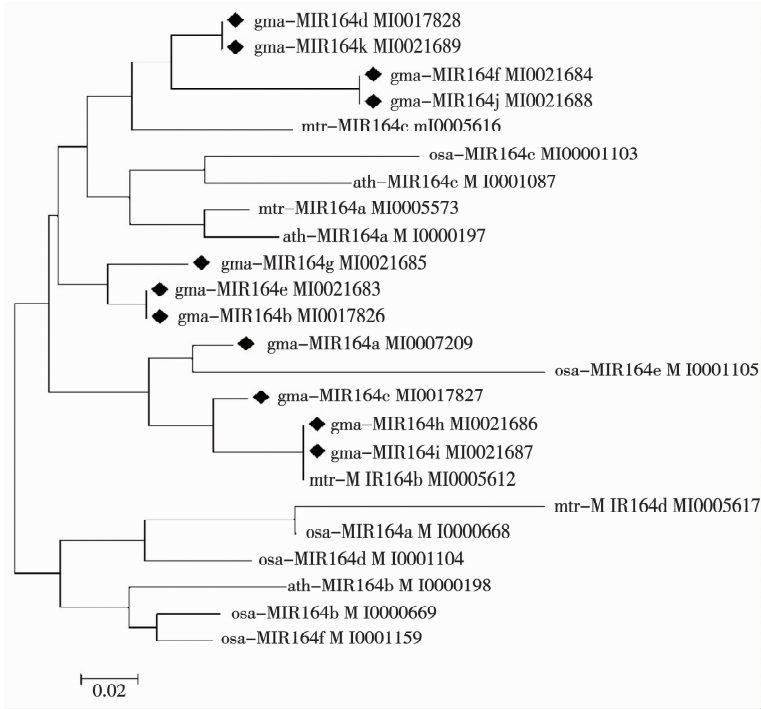
2.2 *miR164* 前体序列的系统发育进化树

分析结果表明:*osa-miR164a/b/d/f* 与 *ath-miR164b*、*mtr-miR164d* 聚为一簇,其余 18 条聚为另一簇,拟南芥 *ath-miR164b* 和蒺藜苜蓿 *mtr-miR164d* 相对于其他几个家族成员相对较远,即来自同一物种的家族不同成员可能被聚类在系统发育进化树的不同分支上,表明植物 *miR164* 的前体序列具有

比较丰富的多样性。

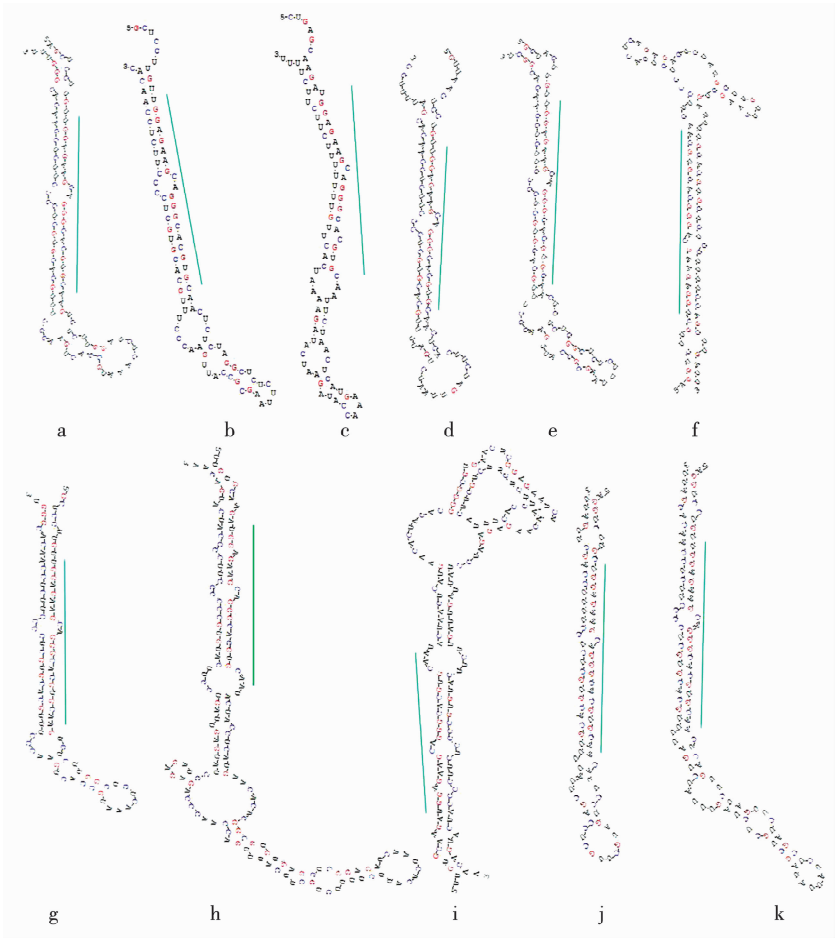
2.3 *gma-miR164* 基因家族前体序列的二级结构

对 *gma-miR164* 基因家族前体序列的二级结构进行预测,发现 11 条 *gma-miR164* 基因前体序列均可形成稳定的二级茎环结构,且成熟序列均产生于基因前体序列的 5'端臂上(图 3)。



◆标注的为大豆中的 *miR164* 序列。  
◆ indicate the *gma-miR164*.

图2 *gma-miR164* 前体序列系统发育进化树  
Fig. 2 Phylogenetic tree for *gma-miR164* precursor sequence



a-k: *gma-miR164a* - *gma-miR164k*.

图3 大豆 *miRNA164* 家族基因前体序列的二级结构预测  
Fig. 3 Secondary structure prediction for precursor sequence of genes *gma-miR164*

2.4 gma-miR164 的靶基因预测

11 个 gma-miR164 基因家族成员共预测到 6 个靶基因,均为 NAC 转录因子(表 2),这些靶基因的功能涉及转录因子调控和结合蛋白,不同 miRNA 具有相同的靶基因,说明 NAC 转录因子是 gma-miR164 基因家族的主要靶基因。

表 2 gma-miR164 基因家族靶基因预测结果

Table 2 Prediction for target genes of gma-miR164 gene family

miRNA	靶位点 Targets name	得分 Expection	靶位点类型 Target description
gma-miR164 a/h/c/d/e/f g/h/i/j/k	Glyma06g35660.1	1	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
	Glyma08g18470.1	1	NAC domain containing protein 1
	Glyma12g26190.1	1	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
	Glyma12g35530.1	1	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
	Glyma13g34950.1	1	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
	Glyma15g40510.1	1	NAC domain containing protein 1

2.5 miR164 启动子中顺式作用元件分析

分析大豆 miR164 启动子中顺式作用元件,为了解 miR164 表达调控机制提供基础。顺式作用元件在不同的 miR164 启动子中分布不均匀,在 gma-miR164 家族基因成员的启动子区域均含有 MYB、MYC 元件,且数量较多,而 11 个 gma-miR164 基因中有 10 个含有 ABRE 元件,说明在干旱应答机制中,它们参与 ABA 依赖的信号通路。而且 gma-miR164c、gma-miR164j 和 gma-miR164k 分别含有 4,2 和 1 个 DRE 元件,说明在干旱应答机制中,它们参与 ABA 依赖和 ABA 不依赖的信号通路。

表 3 miR164 启动子中顺式调控基序预测结果

Table 3 Prediction results of cis-regulatory motif in miR164 promoter

基因 Gene	ABRE	DRE	MYB	MYC
gma-miR164a	3	0	34	32
gma-miR164b	0	0	31	14
gma-miR164c	4	4	20	16
gma-miR164d	2	0	24	24
gma-miR164e	6	0	31	38
gma-miR164f	8	0	20	22
gma-miR164g	6	0	25	50
gma-miR164h	2	0	28	30
gma-miR164i	1	0	22	28
gma-miR164j	3	2	18	8
gma-miR164k	3	1	22	10

3 讨论

miRNA 的发现是生命科学领域一项重要的突破,揭示了基因的非编码区参与调控生命活动的机制。有研究表明 miRNA 不仅参与植物的生长发育过程,在植物应对非生物胁迫,如干旱<sup>[19]</sup>、冻害<sup>[20]</sup>、高盐<sup>[21]</sup>等过程中,也发挥着重要的调控作用。其

中,miR164 家族对植物组织器官生长分化及抗逆性具有重要的调节作用<sup>[10]</sup>。前人研究发现大豆 miR164 家族成员参与生物和非生物胁迫应答<sup>[17-18]</sup>,我们通过生物信息学分析发现 miR164 家族成员的启动子中含有相应的顺式作用元件,说明它们启动子中的顺式作用元件和相应的转录因子参与调控生物和非生物胁迫的应答。miR164 家族是植物特有的 miRNA 家族,其与靶基因之间形成极为复杂的调控网络,通过与对应靶基因之间的互作参与到植物器官形态建成以及逆境胁迫响应过程中。例如在拟南芥中过表达 miR164a 和 miR164b 时导致植株子叶和花器官出现融合等现象<sup>[22]</sup>。值得思考的是,不同植物中的 miR164 为何会表现出组织特异性;同一种植物中 miR164 家族各个成员之间的功能怎样协调;以及能否通过改造 miR164 作用的靶位点来直接提高植物的抗逆性等,相信随着研究的日益全面深入,更多的疑问都将得到解决。

4 结论

本文主要利用生物信息学方法鉴定了 11 条 gma-miR164 基因同源序列,分别分布在 7 条染色体上。11 个 gma-miR164 基因家族成员共预测到 6 个靶基因,均为 NAC 转录因子。gma-miR164 基因家族成熟序列的碱基保守性很高,其前体序列均可形成稳定的二级茎环结构。启动子中顺式调控元件分析表明,ABRE、DRE、MYB 和 MYC 这 4 个元件在 gma-miR164 基因家族启动子序列中分布不均,其中以 MYB 和 MYC 元件数目居多。

参考文献

[1] Liu H, Yu H, Tang G, et al. Small but powerful: Function of microRNAs in plant development[J]. Plant Cell Reports, 2018, 37 (3):515-528.

- [2] D'Ario M, Griffiths Jones S, Kim M. Small RNAs: Big impact on plant development[J]. Trends in Plant Science, 2017, 22(12): 1056-1068.
- [3] Kumar V, Khare T, Shriram V, et al. Plant small RNAs: The essential epigenetic regulators of gene expression for salt-stress responses and tolerance[J]. Plant Cell Reports, 2018, 37(1): 61-75.
- [4] Ohbayashi I, Sugiyama M. Plant nucleolar stress response, a new face in the nac-dependent cellular stress responses[J]. Frontiers in Plant Science, 2017, 8: 2247.
- [5] Afa S, Sajad M, Nazaruddin N, et al. MicroRNA and transcription factor: Key players in plant regulatory network[J]. Frontiers in Plant Science, 2017, 8: 565.
- [6] Kim H J, Nam H G, Lim P O. Regulatory network of NAC transcription factors in leaf senescence[J]. Current Opinion in Plant Biology, 2016, 33: 48-56.
- [7] 杨春文. 番茄 *miR164* 对花器官形成和果实发育的调控研究[D]. 重庆: 重庆大学, 2012;1-62. (Yang C W. Studies on the regulation of *miR164* on flower organ formation and fruit development in tomato[D]. Chongqing: Chongqing University, 2012;1-62. )
- [8] Wei H, Yordanov Y S, Georgieva T, et al. Nitrogen deprivation promotes populus root growth through global transcriptome reprogramming and activation of hierarchical genetic networks[J]. New Phytologist, 2013, 200(2):483-497.
- [9] 李春贺, 阴祖军, 刘玉栋, 等. 盐胁迫条件下不同耐盐棉花 miRNA 差异表达研究[J]. 山东农业科学, 2009(7): 12-17. (Li C H, Ying Z J, Liu Y D, et al. Study on miRNA differential expression of different salt tolerant cotton under salt stress[J]. Shandong Agricultural Science, 2009(7):12-17. )
- [10] 牟桂萍, 纪春艳, 许东林, 等. 植物 *miR164* 家族研究进展[J]. 生命科学, 2013, 25(5): 525-531. (Mu G P, Ji C Y, Xu D L, et al. Research progress of plant *miR164* family[J]. Chinese Bulletin of Life Sciences, 2013, 25(5):525-531. )
- [11] Kim J H, Woo H R, Kim J, et al. Trifurcate feed-forward regulation of age-dependent cell death involving *miR164* in *Arabidopsis* [J]. Science, 2009, 323(5917):1053-1057.
- [12] Wu X M, Liu M Y, Ge X X, et al. Stage and tissue-specific modulation of ten conserved miRNAs and their targets during somatic embryogenesis of Valencia sweet orange[J]. Planta, 2011, 233(3):495-505.
- [13] Bazzini A A, Almasia N I, Manacorda C A, et al. Virus infection elevates transcriptional activity of *miR164a* promoter in plants[J]. BMC Plant Biology, 2009, 9(1):1-12.
- [14] Li Y J, Fu Y, Ji L S, et al. Characterization and expression analysis of the *Arabidopsis* miR169 family[J]. Plant Science, 2010, 178(3):271-280.
- [15] Zhao B, Ge L, Liang R, et al. Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor[J]. BMC Molecular Biology, 2009, 10(1):1-10.
- [16] 罗中钦. 大豆逆境胁迫相关 microRNA 的发掘与验证[D]. 北京: 中国农业科学院, 2012. (Luo Z Q. Discovery and validation of microRNA related to stress in Soybean[D]. Beijing: Chinese Academy of Agricultural Sciences, 2012. )
- [17] Sun Z, Wang Y, Mou F, et al. Genome-wide small RNA analysis of soybean reveals auxin-responsive microRNAs that are differentially expressed in response to salt stress in root apex[J]. Frontiers in Plant Science, 2016, 6:1273.
- [18] 许硕. 野生大豆盐胁迫相关 microRNA 的功能分析[D]. 北京: 中国农业科学院, 2011. (Xu S. Functional analysis of microRNA related to salt stress in wild soybean[D]. Beijing: Chinese Academy of Agricultural Sciences, 2011. )
- [19] Zhao B, Liang R, Ge L, et al. Identification of drought-induced microRNAs in rice[J]. Biochemical & Biophysical Research Communications, 2007, 354(2): 585-590.
- [20] Zhou X, Wang G, Su T K, et al. Identification of cold-inducible microRNAs in plants by transcriptome analysis[J]. Biochimica et Biophysica Acta, 2008, 1779(11): 780-788.
- [21] Llave C, Xie Z, Kasschau K D, et al. Cleavage of scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA[J]. Science, 2002, 297(5589): 2053-2056.
- [22] Nikovics K, Blein T, Peaucelle A, et al. The balance between the *miR164a* and *CUC2* genes controls leaf margin serration in *Arabidopsis*[J]. Plant Cell, 2006, 18(11): 2929-2945.