

拟南芥不定芽发生早期的数字基因表达谱分析

王兴春^{1,2}, 杨致荣², 张树伟¹, 李红英^{2,3}, 李生才⁴

1 山西农业大学生命科学学院, 山西 太谷 030801

2 山西农业大学农业生物工程研究所, 山西 太谷 030801

3 农业部黄土高原作物基因资源与种质创制重点实验室, 山西 太原 030031

4 山西农业大学农学院, 山西 太谷 030801

王兴春, 杨致荣, 张树伟, 等. 拟南芥不定芽发生早期的数字基因表达谱分析. 生物工程学报, 2013, 29(2): 189-202.
Wang XC, Yang ZR, Zhang SW, et al. Digital gene expression profiling analysis of the early adventitious shoot formation in *Arabidopsis thaliana*. Chin J Biotech, 2013, 29(2): 189-202.

摘要: 目前, 有关不定芽发生的研究主要集中在单基因的调控方面, 缺乏转录组方面的系统研究。利用 RNA-seq 高通量测序技术在全基因组范围内检测了不定芽发生早期的基因表达谱, 共检测到 2 457 个差异表达基因。这些基因参与了激素代谢和信号转导、愈伤组织和侧根的形成、茎顶端分生组织的发育和光合作用等过程。进一步的途径富集分析表明, 不定芽发生早期苯丙氨酸代谢和苯丙胺素合成等途径相关的基因显著富集。并且苯丙氨酸可以显著抑制不定芽的发生, 暗示了苯丙氨酸代谢和苯丙胺素的合成可能在不定芽发生过程起着重要的作用。

关键词: 拟南芥, 离体器官发生, 不定芽形成, 转录组, 数字基因表达谱

Received: September 17, 2012; **Accepted:** November 6, 2012

Supported by: National Natural Science Foundation of China (Nos. 31100235, 31171181), Natural Science Foundation of Shanxi Province (No. 2010021030-1), Postdoctoral Research Foundation (No. 80839).

Corresponding author: Xingchun Wang. Tel: +86-354-6287191-306; Fax: +86-354-6289318; E-mail: wxingchun@163.com

Shengcai Li. Tel/Fax: +86-354-6288324; E-mail: sxaulisc@126.com

国家自然科学基金 (Nos. 31100235, 31171181), 山西省青年科技研究基金 (No. 2010021030-1), 中国博士后研究经费 (No. 80839) 资助。

Digital gene expression profiling analysis of the early adventitious shoot formation in *Arabidopsis thaliana*

Xingchun Wang^{1,2}, Zhirong Yang², Shuwei Zhang¹, Hongying Li^{2,3}, and Shengcai Li⁴

1 College of Life Sciences, Shanxi Agricultural University, Taigu 030801, Shanxi, China

2 Institute of Agricultural Bioengineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China

3 Key Laboratory of Crop Gene Resources and Germplasm Enhancement on Loess Plateau, Ministry of Agriculture, Taiyuan 030031, Shanxi, China

4 College of Agriculture, Shanxi Agricultural University, Taigu 030801, Shanxi, China

Abstract: Most current research in the field of adventitious shoot formation is focused on the regulatory function of a single gene. However, a systematic transcriptomic analysis of the early adventitious shoot formation is still lacking. Here, we analyzed the transcriptome profiling of the early adventitious shoot formation in *Arabidopsis* by RNA-seq high throughput sequencing technology, and identified 2 457 differentially expressed genes. Detailed categorization revealed that these genes were mainly involved in hormone homeostasis or signal transduction, callus and lateral root formation, shoot apical meristem development and photosynthesis. Further pathway enrichment analysis showed that genes involved in phenylalanine metabolism and phenylpropanoid biosynthesis were significantly enriched. Moreover, exogenous phenylalanine could repress adventitious shoot formation, indicating that phenylalanine metabolism and phenylpropanoid biosynthesis might be important for adventitious shoot formation.

Keywords: *Arabidopsis thaliana*, *in vitro* organogenesis, adventitious shoot formation, transcriptome, digital gene expression profiling

作为植物再生完整植株的最主要方式之一，不定芽发生不仅广泛应用于植物发育调控机制等基础理论问题的研究，而且是植物快速繁殖和利用生物技术进行作物遗传改良的前提和基础^[1-2]。因此，深入研究不定芽发生的机制既具有重要的理论价值又有广阔的应用前景。

植物激素尤其是细胞分裂素和生长素是影响不定芽发生的关键因素，因此调控激素代谢、转运和信号转导的基因通常也参与不定芽发生过程^[3-6]。在无外源细胞分裂素的情况下，拟南芥功能获得性突变体植物生长激活子 *plant growth activator 22* (*pga22*) 和细胞分裂素非依赖突变体 *cytokinin independent 1* (*cki1*) 可以分化出大量不定芽^[5]。*PGA22* 基因编码一个异戊烯

基转移酶，该酶催化细胞分裂素生物合成的一个限速步骤；而 CKII 曾被认为是细胞分裂素的受体，但由于在体外不能结合细胞分裂素，故该推测还有待进一步证实。细胞分裂素信号转导的关键因子 B 型细胞分裂素响应子 RESPONSE REGULATOR (ARR) 过量表达通常促进不定芽的再生，而 A 型 ARR 通常会抑制不定芽的再生^[6-7]。

在模式植物拟南芥中，不定芽形成主要通过两步培养法进行：首先，将拟南芥的根或下胚轴在含有高浓度 2,4-二氯苯氧乙酸 (2,4-dichlorophenoxyacetic acid, 2,4-D) 的愈伤诱导培养基 (Callus induction medium, CIM) 中进行预培养；然后，再转移到含有高浓度细胞分裂

素的芽诱导培养基 (Shoot induction medium, SIM) 继续培养, 即可分化出不定芽^[8]。深入探讨上述步骤相关基因的表达及其关系, 对于阐明不定芽发生的分子机制具有重要的意义。Che 等^[9-10]利用基因芯片技术分析了拟南芥愈伤组织、不定芽和不定根再生过程中基因表达情况, 发现了一系列在不同的发育过程中差异表达的基因。最近, Xu 等^[11]利用基因芯片技术研究了愈伤组织形成早期的基因表达情况, 发现了一大批新的调控基因。其中转录因子 *HB52* 和 *CRF3* 的过量表达使外植体在无生长素的培养基上产生愈伤组织。这些研究虽然在一定程度上揭示了植物离体器官发生的基因调控网络, 但由于所用芯片的探针数量有限, 且基因芯片检测的灵敏度相对较低, 不能对所有基因进行检测。与基因芯片相比, 近年来发展起来的高通量测序技术有着诸多优势, 已经广泛应用于基因表达调控等领域的研究^[12]。但到目前为止, 高通量测序技术仍未应用于不定芽发生早期转录组研究。为此, 我们利用 RNA-Seq 高通量测序技术在全基因组范围内检测了在不定芽发生早期差异表达的基因, 为揭示不定芽发生的分子机制奠定了基础。

1 材料与方法

1.1 植物材料和培养基

本研究使用的野生型拟南芥 *Arabidopsis thaliana* 为 Wassilewskija (WS) 生态型。

萌发培养基: 1/2×MS (PhytoTechnology Laboratories, 货号 M519)、20 g/L 蔗糖和 8 g/L 琼脂粉, pH 5.8。

CIM 培养基: 1×MS (PhytoTechnology Laboratories, 货号 M519), 10 g/L 蔗糖, 0.5 g/L

MES, 100 mg/L 肌醇, 10 mg/L VB1, 1 mg/L VB3, 1 mg/L VB6, 0.5 mg/L 2,4-D, 0.05 mg/L 激动素 (Kinetin, KT), 和 2.5 g/L 植物凝胶, pH 5.8。

SIM 培养基: 与 CIM 成分相似, 只是将 CIM 中的 2,4-D 和 KT 替换为 0.2 mg/L 吲哚-3-乙酸 (Indole-3-acetic acid, IAA) 和 0.5 mg/L 反式玉米素 (*trans*-zeatin)。

1.2 方法

1.2.1 愈伤诱导和不定芽发生

拟南芥种子消毒后成簇播种在萌发培养基上, 注意不能太分散, 以便后期取材。4 °C 避光春化 2~3 d 后, 置于温室培养。培养条件为 22 °C, 16 h 光照/8 h 黑暗, 光照强度为 20~30 $\mu\text{mol}/(\text{m}^2\cdot\text{s})$ 的弱光, 以利于下胚轴的伸长。7 d 后, 取下胚轴放在富含生长素的 CIM 上培养 7 d; 然后再转移到 SIM 继续培养。

1.2.2 总 RNA 的提取和 RNA-Seq

分别取 CIM 培养 7 d (CIM7) 和 SIM 培养 2 d (SIM2) 的材料提取总 RNA。100 mg 植物材料在液氮中研磨成粉末, 然后利用康为世纪生物科技有限公司的 RNA 提取试剂盒 (含有 DNase I, 货号 CW0559) 提取总 RNA。总 RNA 利用 Agilent 2100 生物芯片分析系统检测合格后, 利用带有 Oligo (dT) 的磁珠富集 mRNA。富集的 mRNAs 随机打断成 200 nt 的片段, 再以片段化后的 mRNA 为模板, 用六碱基随机引物合成 cDNA 第一链; 加入缓冲液、dNTPs、RNase H 和 DNA polymerase I 合成 cDNA 第二链, 经过 QiaQuick PCR 试剂盒纯化并加 EB 缓冲液洗脱经末端修复、加碱基 A, 加测序接头, 再经琼脂糖凝胶电泳回收目的大小片段, 并进行 PCR 扩增。建好的文库用 Illumina HiSeq™ 2000 进行测序。

1.2.3 数据分析

基因表达量的计算使用 RPKM 法 (Reads Per Kb per Million reads)^[13], 该方法能消除基因长度和测序深度差异对计算基因表达的影响, 因此计算得到的基因表达量可直接用于比较不同样品间的基因表达差异。差异表达基因为 False Discovery Rate (FDR) ≤ 0.001 且差异倍数不低于 2 倍 (即 \log_2 的绝对值 ≥ 1) 的基因。

Gene Ontology (GO) 功能显著性富集分析采用 Gene Ontology 数据库 (<http://www.geneontology.org/>) 进行, 将差异表达的基因向 gene ontology 数据库的各 term 映射。计算每个 term 的基因数目。 $P \leq 0.05$ 为显著富集的 GO term。途径显著性富集分析 (Pathway enrichment analysis) 采用 KEGG 数据库^[14]进行, 将差异表达基因向 KEGG 数据库映射, 并统计基因在每个途径中的富集程度 (P value)。

2 结果与分析

2.1 RNA-Seq 高通量测序和评估

拟南芥下胚轴在富含生长素的 CIM 培养基培养 7 d 后 (CIM7), 中柱鞘部位膨大, 形成大量愈伤组织 (图 1A)。将其转移到富含细胞分裂素的 SIM 培养基继续培养 2 d (SIM2), 愈伤组织继续膨大, 颜色由白色变成淡黄色 (图 1B)。分别提取 CIM7 和 SIM2 的总 RNA, 构建 cDNA 文库, 进行 RNA-Seq 实验。CIM7 和 SIM2 样品 RNA-Seq 测序分别得到 11 001 875 和 1 054 546 raw reads, 其中 clean reads 分别为 10 956 306 和 10 956 306, 分别占 raw reads 的 99.55% 和 99.59%。去除杂质后的 clean reads 用作后续分析。

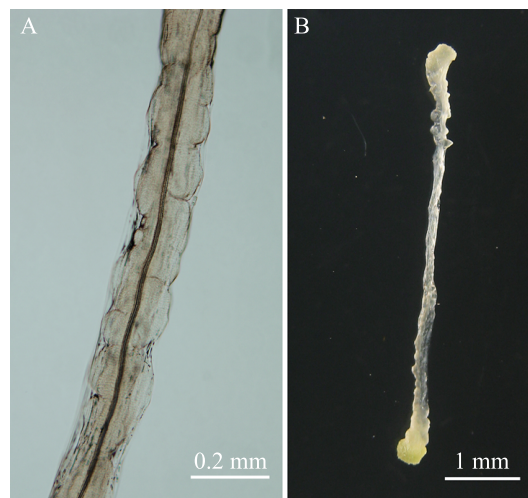


图 1 RNA-Seq 用外植体 CIM7 和 SIM2

Fig. 1 Explants of CIM7 and SIM2 for RNA-Seq. (A) Callus formation in CIM7. (B) Callus formation in SIM2.

我们使用短 reads 比对软件 SOAPaligner/soap2^[15]将得到的 clean reads 分别比对到拟南芥参考基因序列 (ftp://ftp.arabidopsis.org/Sequences/blast_datasets/TAIR10_blastsets/TAIR10_cdna_20101214)。结果表明, 样品 CIM7 和 SIM2 得到的 clean reads 中, 分别有 92.96% 和 93.13% 的 clean reads 匹配到拟南芥的 2 3878 个参考基因上 (表 1 和表 2)。每个基因平均被 426.52 (CIM7) 和 393.64 (SIM2) 个 clean reads 覆盖。

表 1 样品 CIM7 和拟南芥参考基因比对的统计结果
Table 1 Alignment statistics of CIM7 to *Arabidopsis* reference genes

Map to Gene	Number of reads	Percentage (%)
Total mapped reads	10184656	92.96
Perfect match	6327406	57.75
≤ 2 bp mismatch	3857250	35.21
Unique match	6437726	58.76
Multi-position match	3746930	34.20
Total unmapped Reads	771650	7.04

表 2 样品 SIM2 和拟南芥参考基因比对的统计结果

Table 2 Alignment statistics of SIM2 to *Arabidopsis* reference genes

Map to gene	Number of reads	Percentage (%)
Total mapped reads	9399484	93.13
Perfect match	5815303	57.62
≤2 bp mismatch	3584181	35.51
Unique match	5937721	58.83
Multi-position match	3461763	34.30
Total unmapped Reads	693874	6.87

测序饱和度分析表明, 当 clean reads 在 2 M 以下时, 检测到的基因数与测序量成正比; 当达到 2 M 以上时, 趋于平缓; 超过 8 M 时检测到的基因数已经接近饱和 (图 2A 和 2B)。样品 CIM7 和 SIM2 的测序量分别为 10.18 M 和 9.40 M, 已经接近饱和。随后, 我们又对测序的随机性进行了统计, 结果表明样品 CIM7 和 SIM2 的 reads 在参考基因上基本成随机分布 (图 2C 和 2D)。上述结果表明本次测序结果能够较好地反应细胞中基因表达的真实情况, 可以用于后续基因差异表达分析。

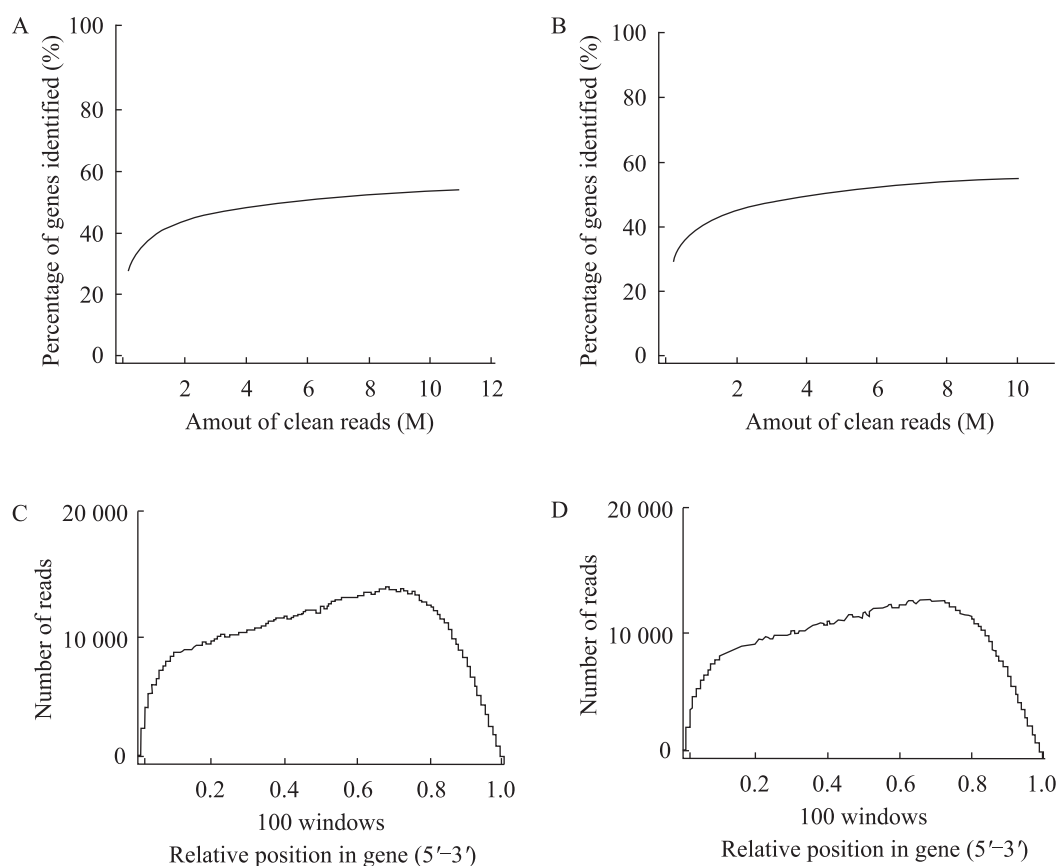


图 2 测序饱和度和随机性评估

Fig. 2 Sequencing saturation and randomness assessment. (A,B) Sequencing saturation analysis of CIM7 (A) and SIM2 (B). (C,D) Sequencing randomness analysis of CIM7 (C) and SIM2 (D).

2.2 不定芽发生早期差异表达的基因

在拟南芥不定芽发生早期,我们共检测到 2 457 个差异表达基因,约占拟南芥整个基因组的 10%左右。其中,有 1 761 个基因表达量上调,696 个基因表达量下调。这些基因参与了植物激素的平衡和信号转导、愈伤组织的形成、侧根的发育、光合作用和细胞周期等。

2.2.1 激素平衡和信号转导与不定芽发生

在拟南芥不定芽发生过程中,很多与细胞分裂素和生长素合成、代谢和信号转导相关的基因的表达量发生了显著改变。异戊烯基转移酶 (Isopentenyl transferase, IPT) 是细胞分裂素合成的关键酶,催化细胞分裂素合成的一个限速步骤。在拟南芥中过量表达 *IPT* 基因导致植物体内细胞分裂素含量上升,并促进了不定芽的再生^[5,16]。然而,我们发现在不定芽发生早期 *IPT5* 基因的表达量下降了近 1 500 倍 (表 3),而其他 *IPT* 基因的表达量没有显著变化,这暗示了 *IPT5* 基因可能在离体器官再生过程起着特殊的功能。作为细胞分裂素信号途径的下游组分,ARR 是执行细胞分裂素生理功能的效应元件。在拟南芥不定芽发生早期,有 6 个 *ARR* 基因的表达量显著升高 (表 3)。这 6 个 *ARR* 全部为 A 型 *ARR*,而 B 型 *ARR* 的表达量却没有明显变化。这与 A 型 *ARR* 基因的表达能被细胞分裂素快速诱导是相符的^[16]。

外植体从含有高浓度生长素的 CIM 中转到含有低浓度生长素的 SIM 培养基中后,一些参与生长素信号转导的基因如 *IAA*、*GH3*、*SAUR* 和 *ARF* 等的表达量也发生了变化 (表 3)。在拟南芥 29 个 *Aux/IAA* 中,有 4 个 *Aux/IAA* 基因 (*IAA3*、*IAA17*、*IAA18* 和 *IAA28*) 的表达量发生了上调;4 个 *Aux/IAA* 基因 (*IAA14*、*IAA19*、*IAA20* 和

IAA30) 的表达量发生了下调。

有趣的是,虽然 CIM 和 SIM 都不含有乙烯、赤霉素和油菜素内酯等其他激素,但有些与这些激素平衡和信号转导相关基因的表达量发生了显著变化 (表 3)。乙烯反应元件结合蛋白 (ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR, ERF) 是乙烯信号转导过程中的关键调控因子,在植物胁迫反应中起着重要的作用^[17]。在拟南芥不定芽发生早期,有 3 个 *ERF* 基因的表达量发生显著上升,其中 *ERF13* 的表达量上升了 21.21 倍。这些结果暗示了乙烯等其他激素也可能在拟南芥不定芽发生过程起着间接的调控作用。

2.2.2 愈伤形成和侧根发育相关基因的表达

在拟南芥不定芽发生早期,有 4 个 LATERAL ORGAN BOUNDARIES DOMAIN (LBD) 家族的基因 (*LBD16*、*LBD18*、*LBD19* 和 *LBD29*) 的表达量显著降低。其中 *LBD29* 基因的表达量降低幅度最大,达 1 445.81 倍,在拟南芥不定芽发生早期完全关闭 (表 3)。这 4 个 *LBD* 基因中,有 3 个 (*LBD16*、*LBD18* 和 *LBD29*) 可以被 CIM 快速诱导,且它们中的任何一个异位表达都可以使外植体在无激素的情况下形成愈伤组织^[18]。这表明,细胞由脱分化向再分化状态转变时,一些参与脱分化的基因将要被关闭。另外,由于愈伤组织和侧根起源相同^[19],所以一些参与侧根形成或发育基因的表达量也有所改变 (表 3)。

2.2.3 茎顶端分生组织相关基因的表达

外植体在 SIM 培养基培养 2 d 时,虽然还未出现明显的茎顶端分生组织 (SHOOT APICAL MERISTEM, SAM) (图 1),但 SAM 相关基因的表达已经发生了变化 (表 3)。其中,*CLAVATA3/ESR-RELATED 2 (CLE2)* 基因表达

表 3 不定芽发生早期差异表达的基因

Table 3 Differential expressed genes at the early stage of adventitious shoot formation

Gene ID	Log2 ratio (SIM2/CIM7)	<i>p</i> -value	FDR	Descriptions
Cytokinin-related gene expression patterns				
AT1G74890	4.08	8.95E-27	1.95E-25	RESPONSE REGULATOR 15 (ARR15)
AT5G21482	1.85	1.69E-12	1.78E-11	CYTOKININ OXIDASE 7 (CKX7)
AT2G17820	1.60	2.88E-183	6.82E-181	ARABIDOPSIS HISTIDINE KINASE 1 (AHK1)
AT3G48100	1.56	3.94E-114	4.56E-112	RESPONSE REGULATOR 5 (ARR5)
AT5G62920	1.47	1.02E-30	2.58E-29	RESPONSE REGULATOR 6 (ARR6)
AT1G19050	1.29	3.58E-34	1.02E-32	RESPONSE REGULATOR 7 (ARR7)
AT2G41310	1.27	3.80E-66	2.23E-64	RESPONSE REGULATOR 8 (ARR8)
AT1G59940	1.03	1.52E-08	1.16E-07	RESPONSE REGULATOR 3 (ARR3)
AT5G39340	1.02	9.91E-11	9.14E-10	HISTIDINE-CONTAINING PHOSPHOTRANSMITTER 3 (AHP3)
AT5G19040	-10.53	5.75E-05	2.88E-4	ISOPENTENYLTRANSFERASE 5 (IPT5)
AT4G11140	-2.18	5.39E-08	3.89E-07	CYTOKININ RESPONSE FACTOR 1 (CRF1)
AT4G23750	-1.84	6.13E-19	9.27E-18	CYTOKININ RESPONSE FACTOR 2 (CRF2)
AT5G10720	-1.15	6.68E-30	1.64E-28	ARABIDOPSIS HISTIDINE KINASE 5 (AHK5)
AT3G61630	-1.09	3.35E-08	2.47E-07	CYTOKININ RESPONSE FACTOR 6 (CRF6)
Auxin-related gene expression patterns				
AT1G48660	4.58	2.71E-12	2.80E-11	AUXIN-RESPONSIVE GH3 FAMILY PROTEIN
AT5G51470	4.14	1.27E-16	1.71E-15	AUXIN-RESPONSIVE GH3 FAMILY PROTEIN
AT5G25890	1.94	2.67E-77	1.89E-75	INDOLE-3-ACETIC ACID INDUCIBLE 28 (IAA28)
AT1G04250	1.82	7.14E-213	2.25E-210	INDOLE-3-ACETIC ACID INDUCIBLE 17 (IAA17)/ AUXIN RESISTANT 3 (AXR3)
AT1G72430	1.69	4.15E-14	4.82E-13	SAUR-LIKE AUXIN-RESPONSIVE PROTEIN FAMILY
AT2G46690	1.38	2.31E-21	3.96E-20	SAUR-LIKE AUXIN-RESPONSIVE PROTEIN FAMILY
AT1G19840	1.34	2.62E-05	1.38 E-04	SAUR-LIKE AUXIN-RESPONSIVE PROTEIN FAMILY
AT1G04240	1.28	1.76E-09	1.46E-08	INDOLE-3-ACETIC ACID INDUCIBLE 3 (IAA3)/SHORT HYPOCOTYL 2 (SHY2)
AT1G51950	1.19	1.02E-14	1.24E-13	INDOLE-3-ACETIC ACID INDUCIBLE 18 (IAA18)
AT1G23160	1.12	8.00E-07	5.11E-06	AUXIN-RESPONSIVE GH3 FAMILY PROTEIN
AT2G46990	-2.89	3.08E-11	2.96E-10	INDOLE-3-ACETIC ACID INDUCIBLE 20 (IAA20)
AT2G24400	-2.37	1.78E-10	1.61E-09	SAUR-LIKE AUXIN-RESPONSIVE PROTEIN FAMILY

续表 3

AT4G27260	-2.04	0	0	AUXIN-RESPONSIVE GH3 FAMILY PROTEIN
AT3G15540	-1.94	5.26E-59	2.77E-57	INDOLE-3-ACETIC ACID INDUCIBLE 19 (IAA19)
AT3G6210	-1.90	6.53E-64	3.65E-62	INDOLE-3-ACETIC ACID INDUCIBLE 30 (IAA30)
AT4G31320	-1.84	2.07E-12	2.16E-11	SAUR-LIKE AUXIN-RESPONSIVE PROTEIN FAMILY
AT4G14550	-1.47	1.91E-12	2.00E-11	INDOLE-3-ACETIC ACID INDUCIBLE 30 (IAA14)
AT2G14960	-1.44	1.59E-65	9.23E-64	AUXIN-RESPONSIVE GH3 FAMILY PROTEIN
AT3G61830	-1.19	1.12E-27	2.53E-26	AUXIN RESPONSE FACTOR 18 (ARF18)
Ethylene-related gene expression patterns				
AT2G44840	4.60	4.19E-52	1.89E-50	ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 13 (ERF13)
AT5G47220	1.99	2.81E-40	9.65E-39	ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 2 (ERF2)
AT4G17500	1.21	6.81E-62	3.72E-60	ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 1 (ERF1)
Lateral root and callus formation related gene expression patterns				
AT3G24300	2.58	5.01E-08	3.62E-07	AMMONIUM TRANSPORTER 1;3 (AMT1;3)
AT5G57740	1.65	2.69E-18	3.95E-17	XB3 ORTHOLOG 2 in ARABIDOPSIS THALIANA (XBAT32)
AT1G77690	1.23	2.52E-31	6.47E-30	LIKE AUX1 3 (LAX3)
AT3G58190	-13.82	5.71E-26	1.19E-24	LATERAL ORGAN BOUNDARIES DOMAIN 29 (LBD29)/ASYMMETRIC LEAVES 2-like 16 (ASL16)
AT2G45410	-3.05	2.77E-35	8.17E-34	LATERAL ORGAN BOUNDARIES DOMAIN 19 (LBD19)
AT2G42430	-2.53	4.88E-39	1.60E-37	LATERAL ORGAN BOUNDARIES DOMAIN 16 (LBD16)/ASYMMETRIC LEAVES 2-like 9 (ASL9)
AT2G45420	-2.24	7.06E-41	2.46E-39	LATERAL ORGAN BOUNDARIES DOMAIN 18 (LBD18)
AT3G59420	-1.44	1.39E-36	4.29E-35	CRINKLY4 (ACR4)
Shoot meristem related gene expression patterns				
AT4G18510	13.11	5.41E-10	4.70E-09	CLAVATA3/ESR-RELATED 2 (CLE2)
AT3G15170	3.16	1.72E-07	1.18E-06	CUP-SHAPED COTYLEDON 1 (CUC1)
AT5G53950	2.35	1.07E-09	9.06E-09	CUP-SHAPED COTYLEDON 2 (CUC2)
AT4G24540	2.22	1.61E-08	1.22E-07	AGAMOUS-like 24 (AGL24)
AT3G24770	1.70	8.59E-21	1.43E-19	CLAVATA3/ESR-RELATED 41 (CLE41)
AT1G46480	1.36	1.76E-23	3.32E-22	WUSCHEL RELATED HOMEBOX 4 (WOX4)
AT3G15510	-3.08	2.15E-10	1.93E-09	NAC DOMAIN CONTAINING PROTEIN 2 (NAC2)
AT3G11260	-1.80	2.12E-38	6.87E-37	WUSCHEL RELATED HOMEBOX 5 (WOX5)
AT2G45660	-1.54	3.81E-07	2.52E-06	AGAMOUS-like 20 (AGL20)

上升幅度最大, 上升了近 9 000 倍。其次是 *CUP-SHAPED COTYLEDON 1 (CUC1)* 和 *CUC2* 基因, 表达量分别提高了 8.92 倍和 5.07 倍。

除了上述基因外, 一些与细胞分化和细胞周期相关的基因的表达量也发生了显著变化 (数据未显示)。

2.3 不定芽发生早期代谢的变化

为了对差异表达基因的生物学功能有更清晰的了解, 我们进行了途径富集分析。结果表明, 差异表达的 2 457 个基因中有 1 364 个基因可以映射到 111 个不同的生物途径, 显著富集的路径 (Q value ≤ 0.05) 有 25 个 (表 4)。其中, 苯丙胺素合成和苯丙氨酸代谢相关途径的基因显著富集, 分别富集了 3.38 倍和 3.84 倍。为了进一步研究苯丙氨酸在不定芽发生过程的功能, 我们将脱分化的下胚轴外植体放在含有不同浓度的苯丙氨酸的 CIM 培养基培养。结果表明, 1 mmol/L 的苯丙氨酸即可以完全抑制不定芽的再生 (图 3), 暗示了苯丙氨酸和苯丙胺素的生物合成可能在不定芽发生过程中起着重要的作用。

除了苯丙胺素和苯丙氨酸途径外, 次生物质合成、植物激素信号转导和光合作用等相关途径的基因也显著富集了。进一步的分析表明, 差异表达的基因中光合途径 (Photosynthesis pathway)、黄酮和黄酮醇生物合成途径 (Flavone and flavonol biosynthesis) 的基因的表达量全部升高了, 暗示了不定芽发生早期外植体开始由异养向自养转变。而其余 23 个途径中既有上调基

因又有下调基因, 表明不定芽发生早期转录水平调节的复杂性。

2.4 不定芽发生早期叶绿体相关组分显著富集

GO 功能显著性细胞组分 (Cellular component) 分析表明, 在差异表达的 2 457 个基因中有 1 365 个基因可以映射到 101 个不同的 term。其中富集的有 20 个 term (表 5)。进一步的分析发现, 富集 term 中有一半以上 (11 个) term 位于叶绿体/质体中 (表 5), 这暗示了在不定芽出现之前, 外植体已经具备了部分光合作用的能力。

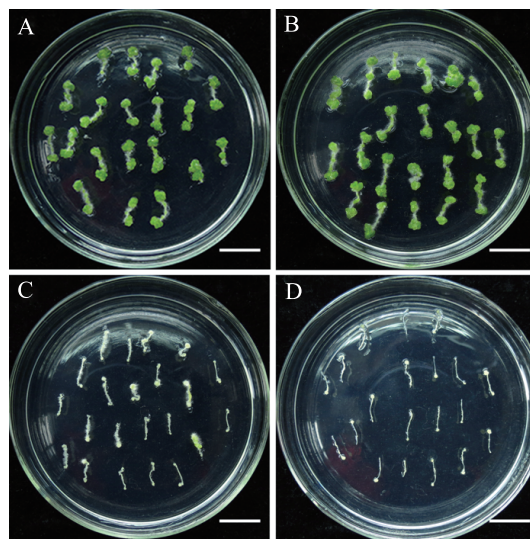


图 3 苯丙氨酸抑制不定芽发生

Fig. 3 Adventitious shoot formation was suppressed by phenylalanine. Hypocotyl explants were pre-cultured on CIM for 7 days and then cultured on SIM containing different concentration of phenylalanine for 12 days. (A) 0 mmol/L. (B) 0.1 mmol/L. (C) 1 mmol/L. (D) 10 mmol/L. Bar= 1 cm.

表 4 不定芽发生早期显著富集的途径

Table 4 Pathways significantly enriched at the early stage of adventitious shoot formation

Pathway	DEGs with pathway annotation (1 364)	All genes with pathway annotation (21 038)	<i>P</i> value	<i>Q</i> value
Phenylpropanoid biosynthesis	81 (5.94%)	371 (1.76%)	1.34E-22	1.49E-20
Phenylalanine metabolism	45 (3.3%)	180 (0.86%)	2.03E-15	1.12E-13
Biosynthesis of secondary metabolites	245 (17.96%)	2 454 (11.66%)	1.12E-12	4.13E-11
Stilbenoid, diarylheptanoid and gingerol biosynthesis	46 (3.37%)	261 (1.24%)	4.97E-10	1.38E-08
Pentose and glucuronate interconversions	42 (3.08%)	240 (1.14%)	3.49E-09	7.74E-08
Starch and sucrose metabolism	67 (4.91%)	490 (2.33%)	5.34E-09	9.89E-08
Flavonoid biosynthesis	36 (2.64%)	205 (0.97%)	4.11E-08	6.52E-07
Plant hormone signal transduction	123 (9.02%)	1 178 (5.6%)	8.30E-08	1.15E-06
Ascorbate and aldarate metabolism	26 (1.91%)	125 (0.59%)	9.91E-08	1.22E-06
Glucosinolate biosynthesis	19 (1.39%)	90 (0.43%)	4.04E-06	4.49E-05
Cyanoamino acid metabolism	25 (1.83%)	144 (0.68%)	5.75E-06	5.80E-05
Photosynthesis	21 (1.54%)	116 (0.55%)	1.63E-05	1.50E-04
Plant-pathogen interaction	120 (8.8%)	1 285 (6.11%)	2.93E-05	2.50E-04
Limonene and pinene degradation	30 (2.2%)	211 (1%)	4.20E-05	3.33E-04
Tryptophan metabolism	25 (1.83%)	164 (0.78%)	5.66E-05	4.19E-04
Metabolic pathways	359 (26.32%)	4 641 (22.06%)	6.70E-05	4.65E-04
ABC transporters	28 (2.05%)	202 (0.96%)	1.17E-04	7.62E-04
Flavone and flavonol biosynthesis	11 (0.81%)	56 (0.27%)	8.14E-04	5.02E-03
Sulfur metabolism	12 (0.88%)	69 (0.33%)	1.48E-03	8.27E-03
Brassinosteroid biosynthesis	8 (0.59%)	35 (0.17%)	1.49E-03	8.27E-03
Indole alkaloid biosynthesis	8 (0.59%)	42 (0.2%)	5.00E-03	2.64E-02
Photosynthesis-antenna proteins	7 (0.51%)	34 (0.16%)	5.43E-03	2.66E-02
Circadian rhythm-plant	27 (1.98%)	247 (1.17%)	5.51E-03	2.66E-02
Carotenoid biosynthesis	23 (1.69%)	205 (0.97%)	7.30E-03	3.28E-02
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	21 (1.54%)	182 (0.87%)	7.38E-03	3.28E-02

表 5 不定芽发生早期差异表达基因的 GO 富集分析

Table 5 GO analysis of the differential expressed genes at the early stage of adventitious shoot formation

Gene Ontology term	Cluster frequency	Genome frequency of use	<i>P</i> value
Extracellular region	117 out of 1 615 genes, 7.2%	778 out of 22 523 genes, 3.5%	1.18E-12
External encapsulating structure	166 out of 1 615 genes, 10.3%	1 271 out of 22 523 genes, 5.6%	1.54E-12
Cell periphery	176 out of 1 615 genes, 10.9%	1 444 out of 22 523 genes, 6.4%	1.20E-10
Cell wall	85 out of 1 615 genes, 5.3%	562 out of 22 523 genes, 2.5%	4.20E-09
Photosystem	23 out of 1 615 genes, 1.4%	102 out of 22 523 genes, 0.5%	6.89E-05
Photosynthetic membrane	27 out of 1 615 genes, 1.7%	142 out of 22 523 genes, 0.6%	2.70E-04
Photosystem I	8 out of 1 615 genes, 0.5%	20 out of 22 523 genes, 0.1%	3.99E-03
Thylakoid part	31 out of 1 615 genes, 1.9%	201 out of 22 523 genes, 0.9%	4.36E-03
Anchored to membrane	24 out of 1 615 genes, 1.5%	159 out of 22 523 genes, 0.7%	4.22E-02
Chloroplast stroma	17 out of 1 615 genes, 1.1%	97 out of 22 523 genes, 0.4%	5.04E-02
Chloroplast thylakoid	62 out of 1 615 genes, 3.8%	563 out of 22 523 genes, 2.5%	5.06E-02
Plastid thylakoid	64 out of 1615 genes, 4.0%	591 out of 22523 genes, 2.6%	6.45E-02
Intrinsic to membrane	206 out of 1 615 genes, 12.8%	2326 out of 22 523 genes, 10.3%	6.93E-02
Organelle sub-compartment	64 out of 1 615 genes, 4.0%	595 out of 22 523 genes, 2.6%	7.67E-02
Thylakoid	66 out of 1 615 genes, 4.1%	636 out of 22 523 genes, 2.8%	1.60E-01
Vesicle	92 out of 1 615 genes, 5.7%	945 out of 22 523 genes, 4.2%	1.68E-01
Cytoplasmic vesicle	90 out of 1 615 genes, 5.6%	928 out of 22 523 genes, 4.1%	2.09E-01
Chloroplast part	65 out of 1 615 genes, 4.0%	657 out of 22 523 genes, 2.9%	5.28E-01
Chloroplast	65 out of 1 615 genes, 4.0%	662 out of 22 523 genes, 2.9%	6.23E-01
Photosystem II	9 out of 1 615 genes, 0.6%	51 out of 22 523 genes, 0.2%	9.72E-01

3 讨论

植物不定芽发生是一个众多基因参与的复杂过程,在这一过程中许多基因的表达量发生了改变。Che 等^[9-10]曾利用基因芯片技术检测了不定芽发生时基因的表达情况。但由于基因芯片技术本身的局限性,检测的基因数目有限。另外, Che 等用的是在 SIM 培养基培养 3 d 的材料,此

时外植体已呈浅绿色,叶绿体已经开始分化。而我们利用的是在 SIM 培养基中培养 2 d 的材料,更有利于早期差异表达基因的研究。Che 等^[9-10]利用基因芯片进行的研究中,发现 *ARR5* 和 *RAP2.6L* 等基因的表达量显著上调。我们 RNA-Seq 技术检测结果表明,这两个基因分别上调了 2.96 和 2.25 倍。此外,本研究表明 *ARR15*、*CUC1*、*CUC2* 和 *LBD29* 等基因的表达量也明显

上调 (表 3), 这些基因已经被证明与不定芽发生有关。这些结果从侧面证明本研究的 RNA-Seq 检测结果是可信的。除了这些已知基因外, 我们也新发现了一大批以前在不定芽发生过程未研究过的基因 (如 *IPT5*、*CLE2*、*ERF13* 和 *GH3* 家族的基因等), 深入研究这些基因的功能将有助于我们对不定芽发生过程有更清晰的认识。同时, 我们还发现苯丙氨酸和苯丙胺素的合成与代谢可能在不定芽发生过程起着极其重要的作用。

细胞分裂素是影响不定芽发生的关键因子, 也是 SIM 培养基中添加的主要外源激素。当将外植体从富含生长素的 CIM 培养基转移到富含细胞分裂素的 SIM 培养基后, *IPT5* 基因的表达量大幅下降, 培养 2 d 时已经检测不到该基因的表达, 而其他 8 个 *IPT* 基因的表达却没有明显变化 (表 3)。细胞分裂素在 4 h 内即可负反馈抑制 *IPT1*、*IPT3*、*IPT5* 和 *IPT7* 基因的表达^[20]。因此, 我们推测不定芽分化早期 *IPT5* 基因的表达量下降可能是由于 SIM 培养基上高浓度细胞分裂素负反馈调节的结果。但我们不排除 *IPT5* 基因在脱分化时起着重要作用的可能性, 因此在脱分化时高表达, 在再分化时表达量降低。为此, 我们将深入研究 *IPT5* 基因在外植体脱分化阶段的功能。除了 *IPT* 基因外, 细胞分裂素的氧化分解是调节细胞分裂素含量动态平衡的一种重要方式。这一反应受到细胞分裂素氧化酶 (Cytokinin oxidase/dehydrogenase, CKX) 的催化。在不定芽发生早期, *CKX7* 基因的表达量略有上升 (表 3), 这可能是由于 SIM 中高浓度的细胞分裂素负反馈调节的结果。A 型 ARR 是细胞分裂素途径的负调控因子, 其中 *ARR15* 基因的表达可以被细

胞分裂素快速诱导, 且该基因可以作为细胞获得全能性具备不定芽分化能力的标记基因^[21-22]。与该报道相一致, 本实验表明在 SIM 培养基培养 2 d 时, *ARR15* 的表达量提高了 16.90 倍。除了 *ARR15* 基因外, 在本研究中, 我们发现 *ARR3*、*ARR5*、*ARR6*、*ARR7* 和 *ARR8* 等 5 个其他的 A 型 ARR 基因的表达量也都有所升高。除了上述细胞分裂素相关的基因外, 生长素相关的一些基因的表达量也发生了显著变化 (表 3)。生长素调控着愈伤的形成, 在细胞脱分化过程起着关键的作用。最近的一项研究表明, *LBD16*、*LBD17*、*LBD18* 和 *LBD29* 等 4 个 *LBD* 基因位于生长素响应因子 ARF 的下游, 调控着愈伤组织的形成^[18]。我们发现这些基因在不定芽形成早期表达量都有所下降, 其中 *LBD29* 基因的表达量降幅最大为 14 415.81 倍 (表 3)。这表明, 这些基因虽然在愈伤组织形成过程起着重要作用, 但在不定芽发生早期却不再发挥功能。

SAM 的发育和分生能力的维持与不定芽的再生关系密切, 许多 SAM 发育相关的基因突变也会影响不定芽的再生^[2,23]。不定芽发生早期有 9 个和 SAM 相关的基因的表达量发生了改变, 其中上调的有 6 个, 下调的 3 个。上调幅度较大的 3 个基因分别为 *CLE2*、*CUC1* 和 *CUC2*。有关 *CUC1* 和 *CUC2* 基因在不定芽发生过程的功能已有报道^[24], 但 *CLE2* 基因在不定芽发生过程的功能还不清楚。这些结果表明, 虽然在 SIM 培养基培养 2 d 时, 还未出现不定芽, 但此时一些芽顶端分生组织相关的基因已经开始表达。

途径富集分析表明, 不定芽发生早期苯丙氨酸代谢和苯丙胺素合成的基因显著富集 (表 4)。

进一步的实验表明, 苯丙氨酸可以抑制不定芽的形成 (图 3)。苯丙胺素是一类由苯丙氨酸合成的重要的次生代谢物质, 在植物生长发育过程起着极其重要的作用^[25]。但苯丙氨酸和苯丙胺素在不定芽发生过程的功能还不清楚, 深入研究二者合成和代谢相关突变体不定芽的再生情况将有助于这一问题的解决。

REFERENCES

- [1] Guan CM, Zhang XS. Advances in the molecular mechanism of *in vitro* plant organogenesis. *Chin Sci Bull*, 2006, 23(5): 595–602 (in Chinese).
关春梅, 张宪省. 植物离体器官发生控制机理研究进展. *植物学通报*, 2006, 23(5): 595–602.
- [2] Duclercq J, Sangwan-Norreel B, Catterou M, et al. *De novo* shoot organogenesis: from art to science. *Trends Plant Sci*, 2011, 16(11): 597–606.
- [3] Inoue T, Higuchi M, Hashimoto Y, et al. Identification of CRE1 as a cytokinin receptor from *Arabidopsis*. *Nature*, 2001, 409(6823): 1060–1063.
- [4] Sun J, Hirose N, Wang X, et al. *Arabidopsis* *SOI33/AtENT8* gene encodes a putative equilibrative nucleoside transporter that is involved in cytokinin transport *in planta*. *J Integr Plant Biol*, 2005, 47(5): 588–603.
- [5] Sun J, Niu QW, Tarkowski P, et al. The *Arabidopsis* *AtIPT8/PGA22* gene encodes an isopentenyl transferase that is involved in *de novo* cytokinin biosynthesis. *Plant Physiol*, 2003, 131(1): 167–176.
- [6] Ren B, Liang Y, Deng Y, et al. Genome-wide comparative analysis of type-A *Arabidopsis* response regulator genes by overexpression studies reveals their diverse roles and regulatory mechanisms in cytokinin signaling. *Cell Res*, 2009, 19(10): 1178–1190.
- [7] Hwang I, Sheen J. Two-component circuitry in *Arabidopsis* cytokinin signal transduction. *Nature*, 2001, 413(6854): 383–389.
- [8] Valvekens D, Montagu MV, Van Lijsebettens M. *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc Natl Acad Sci USA*, 1988, 85(15): 5536–5540.
- [9] Che P, Gingerich DJ, Lall S, et al. Global and hormone-induced gene expression changes during shoot development in *Arabidopsis*. *Plant Cell*, 2002, 14(11): 2771–2785.
- [10] Che P, Lall S, Nettleton D, et al. Gene expression programs during shoot, root, and callus development in *Arabidopsis* tissue culture. *Plant Physiol*, 2006, 141(2): 620–637.
- [11] Xu K, Liu J, Fan M, et al. A genome-wide transcriptome profiling reveals the early molecular events during callus initiation in *Arabidopsis* multiple organs. *Genomics*, 2012, 100(2): 116–124.
- [12] Wang XC, Yang ZR, Wang M, et al. High-throughput sequencing technology and its application. *China Biotechnol*, 2012, 32(1): 109–114 (in Chinese).
王兴春, 杨致荣, 王敏, 等. 高通量测序技术及其应用. *中国生物工程杂志*, 2012, 32(1): 109–114.
- [13] Mortazavi A, Williams BA, McCue K, et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Meth*, 2008, 5(7): 621–628.
- [14] Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*, 2008, 36(Database issue): D480–484.
- [15] Li R, Yu C, Li Y, et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*, 2009, 25(15): 1966–1967.
- [16] Deng Y, Wang XC, Yang SH, et al. New insights into cytokinins: metabolism, signal transduction, cross talks and potentials in agricultural applications. *Chin Sci Bull*, 2006: 23(5): 478–498 (in Chinese).

- 邓岩, 王兴春, 杨淑华, 等. 细胞分裂素: 代谢、信号转导、交叉反应与农艺性状改良. 植物学通报, 2006, 23(5): 478–498.
- [17] Onate-Sanchez L, Singh KB. Identification of *Arabidopsis* ethylene-responsive element binding factors with distinct induction kinetics after pathogen infection. *Plant Physiol*, 2002, 128(4): 1313–1322.
- [18] Fan M, Xu C, Xu K, et al. LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in *Arabidopsis* regeneration. *Cell Res*, 2012, 22(7): 1169–1180.
- [19] Sugimoto K, Jiao Y, Meyerowitz EM. *Arabidopsis* regeneration from multiple tissues occurs via a root development pathway. *Dev Cell*, 2010, 18(3): 463–471.
- [20] Miyawaki K, Matsumoto-Kitano M, Kakimoto T. Expression of cytokinin biosynthetic isopentenyltransferase genes in *Arabidopsis*: tissue specificity and regulation by auxin, cytokinin, and nitrate. *Plant J*, 2004, 37(1): 128–138.
- [21] Che P, Lall S, Howell SH. Developmental steps in acquiring competence for shoot development in *Arabidopsis* tissue culture. *Planta*, 2007, 226(5): 1183–1194.
- [22] Brenner WG, Romanov GA, Kollmer I, et al. Immediate-early and delayed cytokinin response genes of *Arabidopsis thaliana* identified by genome-wide expression profiling reveal novel cytokinin-sensitive processes and suggest cytokinin action through transcriptional cascades. *Plant J*, 2005, 44(2): 314–333.
- [23] Tamaki H, Konishi M, Daimon Y, et al. Identification of novel meristem factors involved in shoot regeneration through the analysis of temperature-sensitive mutants of *Arabidopsis*. *Plant J*, 2009, 57(6): 1027–1039.
- [24] Cary AJ, Che P, Howell SH. Developmental events and shoot apical meristem gene expression patterns during shoot development in *Arabidopsis thaliana*. *Plant J*, 2002, 32(6): 867–877.
- [25] Weisshaar B, Jenkins GI. Phenylpropanoid biosynthesis and its regulation. *Curr Opin Plant Biol*, 1998, 1(3): 251–257.

(本文责编 郝丽芳)