

# CYP72B1 基因和 AUR3 基因响应光、生长素 和油菜素内酯的转录调控机制研究 \*

张长青<sup>1, 4)</sup> 王进<sup>1)\*\*</sup> 朱怀球<sup>3)</sup> 高翔<sup>2)</sup>

(<sup>1</sup>)南京大学生命科学学院, 医药生物技术国家重点实验室, 南京 210093; <sup>2</sup>南京大学模式动物研究所, 南京 210061;

<sup>3</sup>北京大学理论生物中心, 北京 100871; <sup>4</sup>金陵科技学院园艺学院, 南京 210038)

**摘要** 为研究光、生长素和油菜素内酯在基因层次上的互作机制, 开发了转录调控元件识别工具 OCMMat, 其中, 在对共表达基因信息和直系同源基因信息进行整合时, 利用了转录调控元件在直系同源基因启动子中的富集性。利用该方法发现, *CYP72B1* 基因和 *AUR3* 基因启动子含有 3 个相同的调控模序 GAGACA、AAGAAAAA、ATCATG, 它们分别承担了 AuxRE 元件、GT 元件和 GT 辅助元件的功能。其中, ATCATG 模序是目前尚未报道过的调控元件, 与 AAGAAAAA 模序的距离相对恒定。基于调控元件识别结果, 构建了 *CYP72B1* 基因和 *AUR3* 基因响应光、生长素和油菜素内酯的转录调控模型, 模型显示: 光信号和生长素、油菜素内酯信号在 *CYP72B1* 基因和 *AUR3* 基因的转录调控元件上相互交叠, 而生长素和油菜素内酯信号则在转录因子 ARF 水平上相交。

**关键词** *CYP72B1* 基因, *AUR3* 基因, 转录调控, 元件识别, GT 元件, AuxRE 元件

**学科分类号** Q756, Q341

**DOI:** 10.3724/SP.J.1206.2009.00088

植物诸多生长发育过程, 如种子萌发、光形态建成、向光性弯曲、开花结果等, 均受光、生长素和油菜素内酯调控。利用拟南芥、水稻等模式植物, 人们已确立了它们信号转导的基本组分和通路。但进一步研究发现, 这些通路并不独立, 它们在多个环节和步骤上相互交叉和联接<sup>[1]</sup>, 表现出不同层次上的互作(crosstalk)。基于突变体筛选、酵母双杂交、蛋白质芯片等技术, 人们已鉴定出多个光、生长素和油菜素内酯的共调控蛋白<sup>[2, 3]</sup>, 并揭示了它们的部分互作机制, 但在生物调控最基本的基因层面上, 相关工作却进展缓慢, 迄今仅停留在对共表达基因的识别上<sup>[4]</sup>, 如何鉴定其中的共调控基因依然是个难题<sup>[4]</sup>, 而这也正是揭示复杂信号通路间如何通过基因间相互作用以调控植物正常生长发育的关键。

共调控基因拥有相同的转录调控元件(*cis-element*), 因此, 通过识别共表达基因间相同的转录调控元件可以鉴定共调控基因<sup>[5]</sup>。目前, 转录调控元件识别的工具主要有两类: 基于共表达基

因启动子中转录调控元件高频特征开发的工具和基于直系同源基因启动子中转录调控元件保守特征开发的工具。这两类方法均有其不足之处, 前者往往假阳性率比较高, 而后者则仅适用于分离时间为 500~700 万年的物种<sup>[6]</sup>。对植物而言, 目前完成全基因组测序的物种较少, 而基因表达谱数据又相对丰富, 因此, 还仅局限于前一类方法的广泛应用。如何能既降低假阳性率、同时又摆脱对物种分离时间上的依赖是开发植物基因转录调控元件识别工具面临的主要问题。

基于转录调控元件的进化保守特征, 我们提出了转录调控元件在直系同源基因启动子中的富集思想, 将其与转录调控元件在共表达基因启动子中的高频特征相结合, 绕过了依赖物种分离时间的障

\* 国家自然科学基金资助项目(90208021)。

\*\* 通讯联系人。

Tel: 025-83686785, E-mail: jwang@nju.edu.cn

收稿日期: 2009-02-17, 接受日期: 2009-06-16

碍, 开发了 OCMat 工具. 对生长素和油菜素内酯诱导的共表达基因分析发现, *CYP72B1* 基因和 *AUR3* 基因具有相同的转录调控元件和调控模式, 鉴定出了它们的共调控关系, 构建了它们应答光、生长素和油菜素内酯信号的转录调控机制模型.

## 1 材料和方法

### 1.1 调控元件识别方法

通过组合单物种中的共表达基因信息和多物种间的直系同源基因信息, 提出了识别转录调控元件的一种新方法. 它包括 3 个基本步骤:

#### a. 共表达基因启动子中的高频元件识别.

表达谱相似或生化功能相似的基因通常被认为由相似的转录调控机制控制, 因此它们分享有相似的转录调控元件. 高频性是转录调控元件在该组启动子中最基本的统计学特征.

利用 Fisher's exact test, 我们对共表达基因启动子中所有 5~25 bp 的寡核苷酸序列进行高频性评估. *P* 值是实际的衡量指标<sup>[7]</sup>, 由公式(1)计算:

$$P = \frac{(n_1+n_2)!(m_1+m_2)!(n_1+m_1)!(n_2+m_2)!}{n_1!n_2!m_1!m_2!(n_1+n_2+m_1+m_2)!} \quad (1)$$

其中,  $n_1$  和  $n_2$  分别表示共表达基因启动子中含某寡核苷酸序列的启动子数目和不含该序列的启动子数目,  $m_1$  和  $m_2$  分别表示非共表达基因启动子(即对照组)中含该寡核苷酸的启动子数和不包含该元件的启动子数.

将  $P < 0.05$  作为寡核苷酸序列高频度判断的最低标准.

#### b. 寡核苷酸序列的保守度评估.

由于受进化选择压影响, 调控序列在直系同源基因启动子中具有保守特征, 因此推断, 它们在该组启动子序列中同样也富集(enrichment), 我们将这一现象称为富集保守性, 它能在一定程度上代替当前序列比对识别方法中的寡核苷酸序列保守性.

改进作者先前提出的富集度计算模型<sup>[8]</sup>, 得到如公式(2)所示的新模型:

$$S = k \times \frac{AO}{EO} \quad (2)$$

其中  $S$  表示富集系数.  $k$  为富集度( $AO/EO$ )的校正系数, 由训练数据集计算产生, 其倒数反映了待评估序列对中性启动子序列集(没有任何进化关系的启动子序列集)和对中性寡核苷酸序列(不具有调控功能的序列)的偏爱度.  $AO$  为待评估序列在直

系同源基因启动子序列集中的观察次数,  $EO$  为待评估序列在直系同源的启动子序列中的期望出现次数, 由公式(3)计算产生.

$$EO = \sum_i^{2N} \left( \sum_{j=1}^{L_{pj}-L_m+1} \prod_{i=1}^{L_m} P_i \right) \quad (3)$$

其中  $N$  为启动子序列数, 考虑反向互补序列, 故公式中为  $2N$ .  $L_{pj}$  为第  $j$  条启动子的序列长度,  $L_m$  是评估元件的长度,  $P_i$  是评估元件第  $i$  位置上的碱基概率, 由启动子对应碱基频度来估算.

$S > 1$  是序列富集的最低判断标准.

#### c. 识别结果的后续处理.

对在共表达基因启动子集中高频出现, 并且在直系同源基因启动子集中富集保守的寡核苷酸序列进行聚类, 产生调控元件的频率矩阵, 按照 Cavener<sup>[9]</sup>的方法产生频率矩阵的一致序列. 为方便分析, 程序自动将计算结果与植物调控元件数据库 PLACE<sup>[10]</sup> 中的已知序列进行比较, 建立链接. 最后, 使用网页技术对结果进行展示.

## 1.2 材料

本文中使用的共表达基因为文献[11]报道. 对照组为拟南芥全基因组中除共表达基因外的其余基因, 取自 Tair 数据库(<http://www.arabidopsis.org/>). 直系同源基因从 Inparanoid 数据库<sup>[12]</sup> 和 DOOP 数据库<sup>[13]</sup> 下载.

## 2 结果与分析

### 2.1 转录调控元件识别工具开发

基于上文描述的方法, 我们开发了拟南芥转录调控元件预测工具 OCMat(<http://www.nicemice.cn/OCMat/OCMat.html>), 其中预先存储了拟南芥全基因组的基因启动子和它们的直系同源序列, 用户只需要输入一套具有共表达关系的基因代码即可进行转录调控元件的预测和调控模块分析. 图 1 是 OCMat 的运行界面.

### 2.2 转录调控元件识别

使用 OCMat 对 Goda 等<sup>[11]</sup> 报道的共表达基因进行分析, 预测出 3 个调控模序(motif) (表 1). 与植物调控元件数据库 PLACE 中的已知序列相比, 模序 1 完全等同于生长素反应元件(AuxRE)<sup>[14]</sup>, 模序 2 包含有光应答元件——GT 元件<sup>[15]</sup> 序列, 模序 3 是一个全新序列, 我们将之定义为 EDIB (element detected in genes up-regulated by IAA and BL).

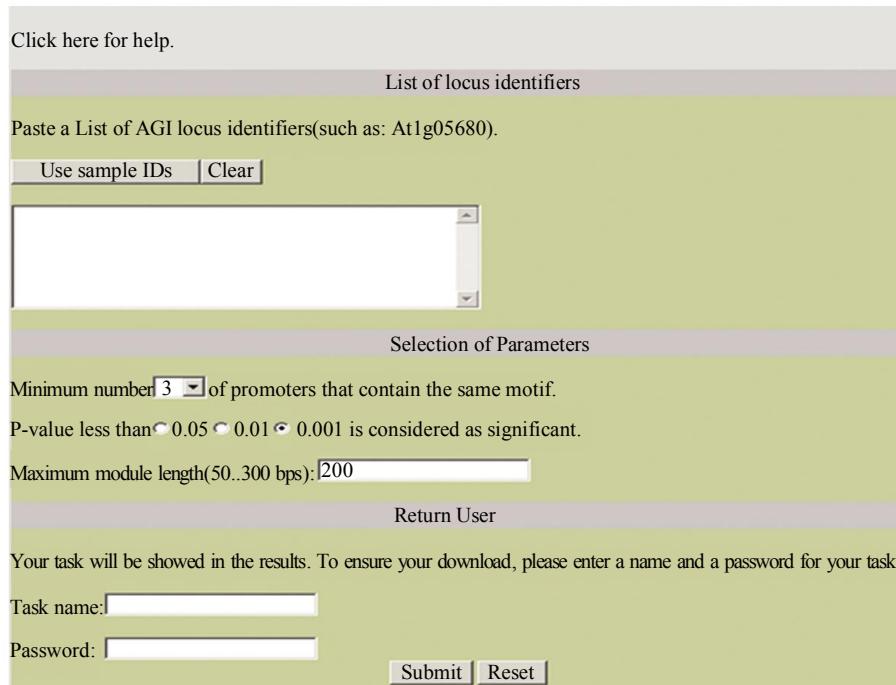


Fig. 1 OCMMat user interface

Table 1 Motifs predicted by OCMMat

No.	Name	Motifs	Consensus sequence	Known <i>cis</i> -elements	
				Name	Sequence
1	Motif1	bits 2 0 1 2 3 4 5 6 GAGACA	GAGACA	AuxRE	GAGACA
2	Motif2	bits 2 0 1 2 3 4 5 6 7 8 AAGAAAAA	AAGAAAAA	GT element	GRWAAW
3	EDIB	bits 2 0 1 2 3 4 5 6 ATCATG	ATCATG		

Parameters for this prediction: (1) At least half of the promoters contain the same motif; (2)  $P < 0.001$ ; (3) The maximum length of *cis*-module is 200 bp.

### 2.3 共调控模式和共调控基因识别

将识别的调控模序定位在其表达基因启动子序列中，发现，在 *CYP72B1*(At2g26710)和 *AUR3*(At4g37390)基因序列中，模序 2 与 EDIB 元件之间的距离相似，在 21~25 bp 或 145~152 bp 之间(图 2a)。*CYP72B1* 与其直系同源基因序列相比，模序 2 与 EDIB 元件的距离也接近，在 145~174 bp 之间(图 2b)。由此可推测，模序 2 与 EDIB 间的组合是一种共调控模式(*cis*-module)，而 *CYP72B1* 基因和 *AUR3* 基因通过这个调控模式构成了共调控关系。

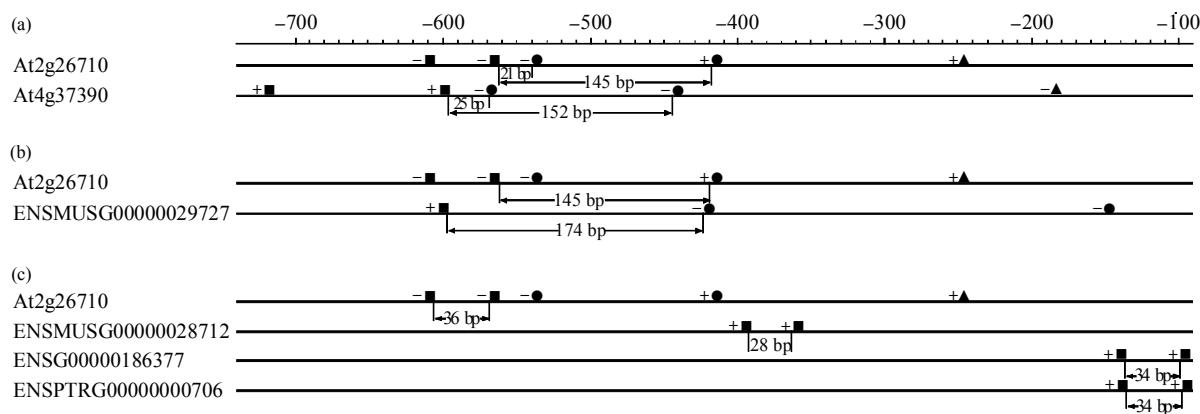
此外，模序 2 和其拷贝在 *CYP72B1* 直系同源

基因启动子中也具有间距恒定的特征，它们均维持在 28~36 bp 之间(图 2c)。这个结果与实验结果相吻合<sup>[15]</sup>。因此，模序 2 和其拷贝间的组合也可能是一种共调控模式。

### 2.4 转录因子预测

基于识别结果中所含的已知转录调控元件信息，我们预测了各调控模序的结合蛋白，即 *CYP72B1* 基因和 *AUR3* 基因的转录因子。

由于生长素反应元件是转录因子 ARF 识别和结合的唯一序列<sup>[14]</sup>，而模序 1 又完全等同于生长素反应元件的序列，因此推断模序 1 是 ARF 的识别结合位点。Takase 等<sup>[16]</sup>的实验支持这个结果。他们



**Fig. 2 cis-Element's patterns in promoters of two co-expressed genes and their orthologs**

The space between motif 2 (■AAGAAAAAA) and EDIB (●ATCATG) is similar in co-expressed genes At2g26710 and At4g37390 (a), and is relatively conservative between orthologs At2g26710 and ENSMUSG00000029727 (b). The space between two motif2 is also relatively conservative among orthologs At2g26710, ENSMUSG00000028712, ENSG00000186377 and ENSPTRG00000000706 (c). The sign +, - indicate the elements occur in the sense strand and the reverse complementary strand respectively. ▲: GAGACA; ■: AAGAAAAAA; ●: ATCATG.

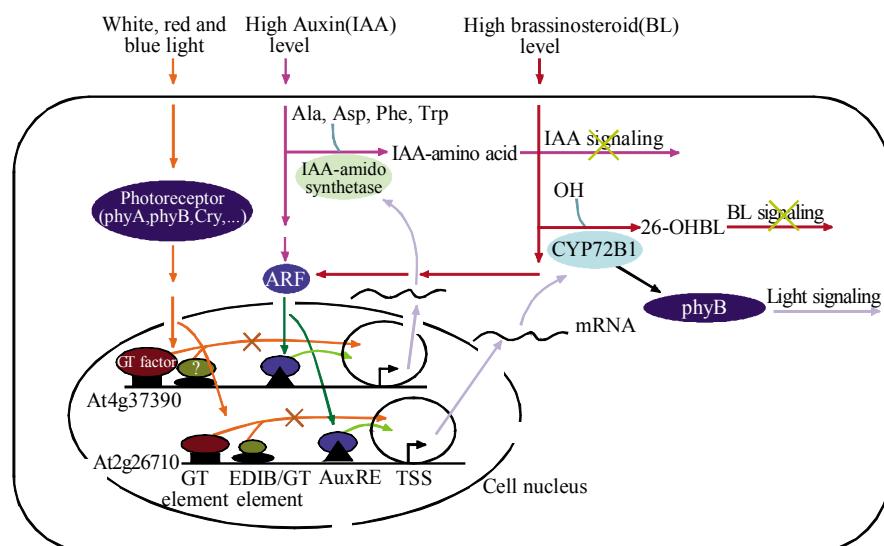
的研究发现，ARF7 基因突变的拟南芥植株中，*AUR3* 的转录不再受生长素影响。

Park 等<sup>[17]</sup>对豆(*Glycine max*)钙离子结合蛋白 SCaM-4 基因的研究表明，其转录受 GT 蛋白直接调控。我们注意到，拟南芥钙离子结合蛋白基因 *At4g27280* 和 *CYP72B1*、*AUR3* 具有共表达的关系<sup>[18]</sup>，而且它们的启动子中同样也存在与模序 2 完全相同的序列。由此推断，包含 GT 元件序列的模序 2 具有结合 GT 因子、调控 *CYP72B1* 和 *AUR3* 基因转录的功能。

EDIB 模序由于缺乏已知信息，无法推测其结合蛋白。鉴于它与模序 2 间的距离相对恒定(图 2a, b)，推测它为模序 2 的辅助元件，其结合蛋白具有与模序 2 的结合蛋白——GT 因子协同调控基因转录的功能。

## 2.5 CYP72B1 基因和 AUR3 基因的转录调控模型构建

通过文献挖掘和知识整合，我们构建了 *CYP72B1* 基因和 *AUR3* 基因的转录调控网络模型(图 3)。其中，光信号经光受体传导至细胞后，通



**Fig. 3 Transcriptional regulation mechanism of the genes At2g26710 and At4g37390.**

In responding to light, GT factor and an unknown protein bind respectively to the GT element and EDIB that locates in promoters of At2g26710 and At4g37390 to repress the gene expression. While in the stimulus of high auxin or brassinosteroid level, ARF binds to AuxRE and enhances the gene expression. The products of genes, CYP72B1 and IAA-amino synthetase, can inactivate the IAA or BL signaling respectively and rescues the light signal through the photoreceptor phyB. TSS indicates the transcriptional start site.

过 GT 因子和目前未知的某个 EDIB 结合蛋白与其对应的调控元件结合, 从而产生对 *CYP72B1* 基因、*AUR3* 基因的转录调控作用, 而生长素信号和油菜素内酯信号则借助 ARF 转录因子与其结合位点 motif 1 的结合进而起始 *CYP72B1* 和 *AUR3* 基因的转录。*CYP72B1* 基因和 *AUR3* 基因的编码产物——油菜素内酯 -26- 羟基化酶(*CYP72B1*)和吲哚乙酸 - 氨基酸合成酶(IAA-amino synthetase), 则分别承担对油菜素内酯和吲哚乙酸(IAA)的反馈抑制, 同时 *CYP72B1* 通过与 phyB 蛋白相互作用调控光信号传导<sup>[16,18]</sup>。该模型显示, 光、生长素和油菜素内酯分别拥有各自不同的信号转导通路, 但各信号通路间并不相互独立, 它们存在明显的互作现象, 基本可归纳为两点: a. 光对油菜素内酯和生长素信号转导通路的调控作用。光照情况下, GT 因子呈非活性状态<sup>[15]</sup>, 靶基因的转录被抑制, 油菜素内酯和生长素信号通路正常, 但在黑暗条件下, *CYP72B1* 和 *AUR3* 基因的转录被激活, 其蛋白质产物则分别催化油菜素内酯和生长素发生羟基化反应和氨基酸化反应, 减少它们的细胞内含量, 削弱它们的信号强度, 从而出现光信号对油菜素内酯和生长素信号转导通路的干扰。b. 油菜素内酯和生长素信号通路间的调控作用, 以及它们对光信号通路的干扰。当油菜素内酯和生长素之一浓度出现异常(过高)时, *CYP72B1* 基因和 *AUR3* 基因的转录将被同时激活, 其表达产物通过催化油菜素内酯和生长素发生羟基化反应和氨基酸化反应而抑制 2 个激素信号通路, 减少了细胞内激素含量, 形成 2 个激素信号通路间的互作。另一方面, 油菜素内酯和生长素还可通过 *CYP72B1* 基因的表达来调控光受体 phyB 的活性, 产生对光信号通路的干扰。鉴于油菜素内酯和生长素在诸多生物功能上的相似性, 推测它们信号通路间的互作可能与其功能上的自洽性相一致, 而光与油菜素内酯以及生长素信号间的互作在某种层面上可能反映了植物昼夜节律的调控机制。

### 3 讨 论

光、生长素和油菜素内酯是调控植物生长发育的 3 个基本信号, 广泛用于作物栽培的化学控制和环境控制之中。目前, 人们已经在蛋白质层次上鉴定出这些信号通路上的多个互作机制, 但在基因水平上却少有研究。本文通过对 *CYP72B1* 基因和 *AUR3* 基因的调控元件和转录因子预测, 在基因转

录水平上建立了光信号与油菜素内酯、生长素通路间的联系。我们发现, 光信号和生长素、油菜素内酯信号在 *CYP72B1* 基因和 *AUR3* 基因的转录调控元件上相互交叠, 而生长素和油菜素内酯信号的调控则在转录因子 ARF 水平上相交。

在 *CYP72B1* 基因和 *AUR3* 基因启动子中, 我们预测出生长素反应元件和光响应元件, 给出它们应答生长素信号和光信号的分子基础。同时, 我们还报道了一个全新的 EDIB 元件。由于 GT 元件具有协同调控基因转录的特性<sup>[15]</sup>, 而 EDIB 元件又与承担 GT 元件功能的模序 2 间的距离相对恒定(图 2a, b), 因此推测, 它承担了 GT 元件的辅助功能, 当然, EDIB 元件的调控特异性和结合蛋白还有待进一步研究。基于 *CYP72B1* 和 *AUR3* 基因的共调控关系, 可以通过其中一个基因的表达状态预测另一基因的表达情况。比如, *CYP72B1* 基因表达受白光、红光、蓝光的诱导而下调, 但对黑暗和远红光处理无变化<sup>[18]</sup>, 由此可以预知, *AUR3* 基因可能存在同样的表达特征。目前除红光外, 其他表现均已得到实验证实<sup>[16]</sup>。鉴于 GT 因子对红光、远红光、蓝光、紫外光等各类光受体无特异性, 因此推测 EDIB 元件的结合蛋白可能具有对不同光受体的专一性功能。

在 *CYP72B1* 基因和 *AUR3* 基因启动子中, 我们并未发现响应油菜素内酯信号的调控元件, 如 BZR1 结合位点(CGTGT/CG)和 BES1 的结合位点(CANNTG)等<sup>[19]</sup>, 这说明它们对油菜素内酯信号的响应可能不是直接的。结果提示, 这两个基因通过 ARF 响应油菜素内酯信号。在生长素诱导基因 DR5::GUS 启动子中发现, ARF 结合位点可承担响应油菜素内酯信号的功能<sup>[20]</sup>。

协同调控是真核生物转录调控元件发挥功能的主要形式, 而间距恒定则是其最为常见的属性。本研究发现, 模序 2 和其拷贝、模序 2 与 EDIB 元件均在进化上保持了间距相对恒定的特征, 这暗示了模序 2 的协同调控方式和这种组合模式在功能上的意义。目前, GT 元件和其拷贝间的协同性已经被证实, 但与 EDIB 元件间的协同功能还有待深入研究。

调控元件由于受到选择压的影响, 突变频率一般较低, 因此在直系同源基因启动子中, 它们具有一定的线性排布特征。然而, 对于分离时间很远的物种来讲, 其调控元件除发生碱基上的突变外, 还会因基因组的其他各种进化事件(如插入、缺失等)

而出现排布的方向、顺序、正负链、间距等多种改变，因此，目前基于直系同源性的转录调控元件识别工具均要求物种的分离时间不宜超过70万年<sup>[6]</sup>，否则将无法得到较为满意的结果。本文开发的OCMMat工具利用调控元件在直系同源基因启动子中的富集度作为参数，克服了上述缺陷，因而其用途将更为广泛<sup>[21]</sup>。

整合序列分析结果和文献知识，我们获得了CYP72B1和AUR3基因响应光信号和激素信号的转录调控模型，阐明了光信号通路、生长素通路和油菜素内酯信号间通过CYP72B1和AUR3基因进行相互作用的分子机理。将序列分析结果和已有知识进行整合，有助于在系统层面上揭示植物生长发育的分子机理，从而为作物栽培的环境控制和化学控制提供更为完善的科学依据。

### 参 考 文 献

- 1 Hardtke C S. Transcriptional auxin-brassinosteroid crosstalk: who's talking? *Bioessays*, 2007, **29**(11): 1115~1123
- 2 Vert G, Walcher C L, Chory J, et al. Integration of auxin and brassinosteroid pathways by Auxin Response Factor 2. *Proc Nat Acad Sci USA*, 2008, **105**(28): 9829~9834
- 3 Li H, Johnson P, Stepanova A, et al. Convergence of signaling pathways in the control of differential cell growth in *Arabidopsis*. *Developmental Cell*, 2004, **7**(2): 193~204
- 4 Gyenesi A, Wagner U, Barkow-Oesterreicher S, et al. Mining co-regulated gene profiles for the detection of functional associations in gene expression data. *Bioinformatics*, 2007, **23**(15): 1927~1935
- 5 Qin Z S, McCue L A, Thompson W, et al. Identification of co-regulated genes through Bayesian clustering of predicted regulatory binding sites. *Nature Biotechnology*, 2003, **21**(4): 435~439
- 6 Wasserman W W, Sandelin A. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Gene*, 2004, **5**(4): 276~287
- 7 Chen G, Hata N, Zhang M Q. Transcription factor binding element detection using functional clustering of mutant expression data. *Nucleic Acids Research*, 2004, **32**(8): 2362~2371
- 8 Xue W, Wang J, Shen Z, et al. Enrichment of transcriptional regulatory sites in non-coding genomic region. *Bioinformatics*, 2004, **20**(4): 569~575
- 9 Cavener D R. Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Research*, 1987, **15**(4): 1353~1361
- 10 Higo K, Ugawa Y, Iwamoto M, et al. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, 1999, **27**(1): 297~300
- 11 Goda H, Sawa S, Asami T, et al. Comprehensive comparison of auxin-regulated and brassinosteroid-regulated genes in *Arabidopsis*. *Plant Physiology*, 2004, **134**(4): 1555~1573
- 12 O'Brien K P, Remm M, Sonnhammer E L. Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Research*, 2005, **33**(Database issue): D476~480
- 13 Barta E, Sebestyen E, Palfy T B, et al. DoOP: Databases of Orthologous Promoters, collections of clusters of orthologous upstream sequences from chordates and plants. *Nucleic Acids Research*, 2005, **33**(Database issue): D86~90
- 14 Guilfoyle T, Hagen G, Ulmasov T, et al. How does auxin turn on genes?. *Plant Physiology*, 1998, **118**(2): 341~347
- 15 Zhou D X. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends in Plant Science*, 1999, **4**(6): 210~214
- 16 Takase T, Nakazawa M, Ishikawa A, et al. ydk1-D, an auxin-responsive GH3 mutant that is involved in hypocotyl and root elongation. *Plant J*, 2004, **37**(4): 471~483
- 17 Park H C, Kim M L, Kang Y H, et al. Pathogen- and NaCl-induced expression of the SCaM-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. *Plant Physiology*, 2004, **135**(4): 2150~2161
- 18 Tanaka K, Asami T, Yoshida S, et al. Brassinosteroid homeostasis in *Arabidopsis* is ensured by feedback expressions of multiple genes involved in its metabolism. *Plant Physiology*, 2005, **138**(2): 1117~1125
- 19 Vert G, Chory J. Downstream nuclear events in brassinosteroid signalling. *Nature*, 2006, **441**(7089): 96~100
- 20 Bao F, Shen J, Brady S R, et al. Brassinosteroids interact with auxin to promote lateral root development in *Arabidopsis*. *Plant Physiology*, 2004, **134**(4): 1624~1631
- 21 Zhang C Q, Wang J, Zhu H, et al. A Mutation Degree Model for Identifying Transcriptional Regulatory Elements. The Seventh Asia Pacific Bioinformatics Conference, Beijing, 2009

## The Transcriptional Regulatory Mechanism of *CYP72B1* and *AUR3* in Response to Light, Auxin and Brassinosteroid<sup>\*</sup>

ZHANG Chang-Qing<sup>1,4)</sup>, WANG Jin<sup>1)\*\*</sup>, ZHU Huai-Qiu<sup>3)</sup>, GAO Xiang<sup>2)</sup>

(<sup>1</sup>) State Key Laboratory of Pharmaceutical Biotechnology, School of Life Science, Nanjing University, Nanjing 210093, China;

(<sup>2</sup>) Model Animal Research Center, Nanjing University, Nanjing 210093, China;

(<sup>3</sup>) Center for Theoretical Biology, Peking University, Beijing 100871, China;

(<sup>4</sup>) College of Horticulture, Jinling Institute of Technology, Nanjing 210038, China)

**Abstract** Light, auxin and brassinosteroid play important roles in plant growth and development. Genetic analysis has demonstrated complex interactions between their signaling pathways, but the gene regulatory mechanisms connecting these pathways are poorly understood. *CYP72B1* and *AUR3* are two important genes responsive to light, auxin and brassinosteroid at transcription level. To understand the regulation mechanism of the two genes, a new tool called OCMMat was developed for identifying *cis*-elements, OCMMat combines both the over-representation property of regulatory elements in co-expressed genes and the conservation property in orthologous genes, for the latter, it was estimated by an enrichment score of regulatory element in orthologous promoter sequences. Using this tool, 3 regulatory motifs shared by genes *CYP72B1* and *AUR3* were reported, motif GAGACA which is the same as a known *cis*-element AuxRE, motif AAGAAAAA containing the sequence of GT element and the third ATCATG which is a new one named EDIB element. The space and the order of AAGAAAAA and EDIB show the same pattern in promoters of both the co-expressed genes and the orthologs of *CYP72B1*. Based on the sequence analysis and the literature knowledge to date, a model was proposed for describing the transcriptional regulatory mechanism of *CYP72B1* and *AUR3* in response to light, auxin and brassinosteroid. The model presents how the signaling pathways of light, auxin and brassinosteroid are interplaying at gene transcription level. In response to light, the transcription factors GT factor and an unknown protein repress the expression of *CYP72B1* and *AUR3*, the hormone pathways are not interfered and thus work in their own way. While in the absence of light stimulation, *CYP72B1* and *AUR3* are expressed and the products, in turn, inhibit both the auxin pathway and the brassinosteroid pathways. On the other side, at high hormone level, gene expression is up-regulated through ARF binding, the gene products inhibit the hormone pathways in a feedback manner, and meanwhile, rescues the light signal through the photoreceptor phyB.

**Key words** *CYP72B1*, *AUR3*, transcriptional regulation, *cis*-element identification, GT element, AuxRE

**DOI:** 10.3724/SP.J.1206.2009.00088

\*This work was supported by a grant from The National Natural Science Foundation of China (90208021).

\*\*Corresponding author.

Tel: 86-25-83686785, E-mail: jwang@nju.edu.cn

Received: February 17, 2009 Accepted: June 16, 2009