

## cDNA 芯片筛选亚洲棉短绒分化发育相关基因

赵国红, 王 晟, 贾银华, 孙君灵, 王 杰, 杜雄明

(中国农业科学院棉花研究所/农业部棉花遗传改良重点开放实验室, 河南安阳 455000)

**摘要:** 【目的】筛选亚洲棉短绒分化发育相关基因。【方法】利用 cDNA 芯片技术在棉纤维发育前期开花前 3 天 (-3DPA)、开花当天 (0DPA)、开花后 3 天 (+3DPA)、5 天 (+5DPA)、7 天 (+7DPA) 筛选野生型 DPL971 与其短绒突变体 DPL972 的差异基因。通过与模式植物拟南芥数据库相结合, 找到与棉纤维起始发育相关基因, 并通过 RT-PCR 和 Real-Time PCR (QRT-PCR) 验证部分候选基因。【结果】筛选出与亚洲棉纤维起始分化发育相关基因 8 个, 其中相对应的与拟南芥同源性较高的蛋白有 7 个, 分别为 KAK (DR461366)、MYB5 (ES812048)、TTG1 (ES811600)、MYB23 (DR453866)、CSLD3 (DR459646)、RHD2 (DR461821) 和 ZWI (ES791383), 这些基因在拟南芥的表皮毛起始发育过程中都起着重要的作用。RT-PCR 和 QRT-PCR 结果表明 MYB23、MYB5、TTG1、CSLD3、RHD2 这 5 个基因在短绒突变体 DPL972 与其野生型 DPL971 的 +3DPA 胚珠中表达差异都有差异。【结论】上述结果表明这 5 个基因可能与短绒分化发育相关, 第 1 个基因可能抑制短绒分化和发育, 后 4 个则促进短绒分化和发育。

**关键词:** 亚洲棉; cDNA 芯片; 棉纤维; 光籽

## Genes Related to Fuzz Initiation and Development in *Gossypium arboreum* Identified by cDNA Microarray

ZHAO Guo-hong, WANG Sheng, JIA Yin-hua, SUN Jun-ling, WANG Jie, DU Xiong-ming

(Cotton Research Institute, Chinese Academy of Agricultural Sciences/Key Laboratory for Cotton Genetic Improvement, Ministry of Agriculture, Anyang 455000, Henan)

**Abstract:** 【Objective】 The objective of this study is to isolate genes related to fuzz initiation and development in *Gossypium arboreum*. 【Method】 The filter arrays of cDNAs were performed to identify the differentiation expressed genes between the wild-type DPL971 and its fuzz mutant line DPL972 in the early developing stages of -3DPA, 0DPA, +3DPA, +5DPA, and +7DPA. Compared with the database of model plant *Arabidopsis thaliana*, some genes related to the initiation of the cotton fiber were isolated, and part of them have been verified by RT-PCR and QRT-PCR. 【Result】 Eight genes related to fiber initiation and development in *Gossypium arboreum* were identified, and seven of them, KAK(DR461366), MYB5(ES812048), TTG1(ES811600), MYB23(DR453866), CSLD3(DR459646), RHD2(DR461821), ZWI(ES791383), were highly homologous to the Arabidopsis proteins which played important roles in the initiation of the trichomes development in *Arabidopsis*. The result of RT-PCR and QRT-PCR indicated that the 5 genes of MYB23, MYB5, TTG1, CSLD3, and RHD2 had significantly different expressions in +3DPA ovules between the fuzz mutant line DPL972 and its wild line DPL971. 【Conclusion】 It implied that these genes may be related to the initiation of the fuzz fiber, the first gene may inhibit the fuzz differentiation and the latter four may promote the fuzz development.

**Key words:** *Gossypium arboreum*; cDNA microarray; cotton fiber; fuzzless

## 0 引言

【研究意义】棉花是世界上最重要的纺织原料之

一。棉纤维的起始发育决定着纤维的产量, 而短绒与长纤维的发育存在着一定的互作关系。因此, 对棉花短绒分化发育相关基因的筛选, 不仅有利于棉花短绒发

收稿日期: 2009-03-18; 接受日期: 2009-05-07

基金项目: 国家“十一五”科技支撑计划项目(2006BAD13B04)、国家“973”计划项目(2004CB117301)

作者简介: 赵国红, 硕士研究生。Tel: 15836346708; E-mail: zhaoguohong2009@yahoo.cn。通信作者杜雄明, 研究员, 博士。Tel: 0372-2525352; E-mail: duxm@cricaas.com.cn

育分子机理的研究,也有利于揭示棉纤维分化发育机理,进而提高棉纤维的产量与品质。【前人研究进展】棉纤维是由棉花胚珠外珠被表皮层单细胞分化发育而成的,是棉花种子表皮毛的极度伸长,大约仅有30%的胚珠表皮细胞发育形成纤维<sup>[1]</sup>。棉花种子表皮毛细胞的发育经历了2次重要的分化,第1次突起形成长纤维,第2次形成短绒<sup>[2]</sup>。迄今为止,已有大量与棉花表皮毛发育相关的基因被克隆分离,例如, *E6*、*LTP6*、*SuSy*、*Expansin* 等<sup>[3-6]</sup>,而拟南芥作为双子叶植物的模式作物,已经建立起完备的功能基因组研究平台,可以极大地促进拟南芥本身及其近缘基因植物的功能基因组学研究<sup>[7]</sup>。表皮毛分化发育在拟南芥中已经有了十分深入的了解,棉花作为双子叶植物,也已有的一些与拟南芥表皮毛发育相关的基因被克隆。例如, Loguercio 等<sup>[8]</sup>通过 PCR 的方法从棉纤维 cDNA 文库中分离得到6个 R2R3MYB 家族的基因 (*GhMYB1*—*GhMYB6*),其中 *GhMYB1*、*GhMYB2*、*GhMYB3*、*GhMYB6* 表达的几个高峰期分别出现在开花后9d (用9DPA表示,下同),开花前1d (-1DPA)和开花后5d (5DPA),这些基因可能对棉纤维的生长发育起重要的调控作用;与拟南芥 *WERWOLF* (*WER*) 和 *GL1* 同源的棉花基因 *GhMYB109* 也属于 R2R3MYB 基因家族,并且在棉花基因组中是单拷贝,可能在棉纤维的起始发育中起重要的作用<sup>[9]</sup>; *GhTTG1*—*GhTTG4* 于2005年从陆地棉中分离出来,它们与拟南芥 *TTG1* 具有很高的同源性,其中有2个基因对拟南芥 *ttg1* 表皮毛突变体具有补偿作用<sup>[10]</sup>。【本研究切入点】虽然现在已有大量与棉纤维起始发育相关的基因被克隆分离,但与棉花短绒起始发育有关的基因至今尚未见报道。而20世纪90年代以来兴起的基因芯片技术,为从整体水平上探索基因表达差异提供了良好的试验手段。四倍体棉种具有大量的重复基因,为研究纤维分化发育增加了难度, Yang 等<sup>[11]</sup>提出陆地棉和海岛棉等异源四倍体的纤维性状很可能来源于A染色体组的棉种(亚洲棉或草棉)。本研究以亚洲棉野生型 DPL971 及其光籽突变体 DPL972 为材料,利用 cDNA 微阵列在纤维的不同发育时期基因群体水平检测与棉纤维表皮毛分化相关的基因,并利用 RT-PCR 和 QRT-PCR 进行部分基因的验证。【拟解决的关键问题】以期找到与棉花短绒分化发育相关的基因,为进一步揭示棉纤维的分化和发育机理,大幅度提高棉纤维产量和品质提供理论基础。

## 1 材料与方法

### 1.1 材料

1.1.1 植物材料 cDNA 微阵列检测材料为亚洲棉 DPL971 及其光籽突变体 DPL972,二者遗传背景十分相似,除在短绒的发育上存在着明显的差异外,在植物学形态特征上几乎没有差别。可以认为是近等基因系。材料种植于中国农业科学院棉花研究所实验农场。材料在开花前1d进行自交并做标记。分别收取-3DPA, 0DPA, +3DPA, +5DPA, +7DPA的胚珠,存于-70℃。

1.1.2 引物与试剂 cDNA 芯片所用的36 000个靶基因克隆由北京大学朱玉贤教授提供,芯片试验由北京博奥生物有限公司帮助完成。T<sub>7</sub>-oligo(dT)<sub>15</sub>, (5'-AAACGACGGCCAGTGAATTGTAATACGACTCACTATAG GCGCTTTTTTTTTTTTTTTT TTTT-3', V 可以是 G、C 或 A) 和 9 Random Primer 购于上海博亚生物技术有限公司。Klenow 酶及 RNA PCR Kit 购于大连宝生物公司。RT-PCR (表1) 和 QRT-PCR 所用特异引物由大连宝生物公司合成。QRT-PCR 中, DR453866 引物扩增片段长度为 94 bp, 内标基因为  $\beta$ -actin。反转录酶 SuperScriptIII (Invitrogen) 购于北京三博远志生物技术有限公司。SYBR Green PCR Master MIX (ABGene, Epsom, UK) 购于北京欧比特仪器有限公司。

### 1.2 方法

1.2.1 RNA 的提取 总 RNA 的提取使用 CTAB 方法<sup>[12]</sup>。用 DNase I 处理去除 DNA 污染。

1.2.2 表达谱芯片的制备 采用通用 PCR 引物扩增, PCR 反应及产物纯化用标准方法进行,通过琼脂糖凝胶电泳监控 PCR 质量。将上述制备的靶基因溶于 3×SSC 溶液,然后点制在基片上。经水合、室温干燥、UV 交联,然后用 0.2% SDS, 水及 0.2% 硼氢化钠溶液处理 10 min 后晾干备用。

1.2.3 cDNA 探针合成及杂交 总 RNA 以 T<sub>7</sub>-oligo(dT)<sub>15</sub> 为引物,用 cDNA Synthesis Kit 合成双链 cDNA; 双链 cDNA 进行体外转录合成 cRNA; 用 M-MLV (Invitrogen) 反转录酶和 9 Random Primer 进行反转录。产物以 9 Random Primer 为引物进行 KLENOW 酶标记,其中 Cy5-dCTP、Cy3-dCTP 终浓度为 40  $\mu\text{mol}\cdot\text{L}^{-1}$ 。标记的 DNA 溶于 30  $\mu\text{L}$  杂交液中 (3×SSC, 0.2% SDS, 5×Denhart's, 25% 甲酰胺), 42℃ 杂交过

表 1 部分差异表达基因 RT-PCR 序列引物

Table 1 Part of PCR primers with differentiation expressed genes used in RT-PCR amplification

| 登录号<br>Accession No. | 引物序列<br>Primer sequence (5'→3')                                |
|----------------------|--|
| DR453866             | GGT TGA GAT GGA AGA AAG AGT G<br>ATT GTG AAG AAT GAG CAG TAG G |
| ES812048             | GTG AAG GTT TGA GTG ATG GT<br>CCC CTT CTA AAG TTC TCT CG       |
| ES811600             | TAT TTG GGA GTT GCC TAC G<br>ACT CTC TTT CCT GCC CTT CT        |
| DR459646             | TTA AGT GGT CTG GCA TTG AGT T<br>CCC AAA GGA GTG ATA TGG TGA T |
| DR461821             | AAC TGA ACA TGG CAC TGA TAC T<br>GCA CTT CCA TAG AGG TCG TAA C |

夜。杂交结束后,在 42℃ 含 0.2% SDS, 2×SSC 的液体中洗 5 min, 然后在 0.2×SSC 中室温洗 5 min。玻片甩干后用于扫描。

1.2.4 芯片图像采集与数据分析 芯片用 LuxScan 10 KA 双通道激光扫描仪(CapitalBio 公司)进行扫描。采用 GenePix Pro 4.0 图像分析软件(Axon Instruments 公司)对芯片图像进行分析,把图像信号转化为数字信号;然后对芯片上的数据用 Lowess 方法进行归一化;最后以差异为两倍的标准来确定差异表达的基因。Ratio>2 为表达明显上调, Ratio<0.5 为表达明显下调,根据 Ratio 值的变化情况可以推断基因在不同时期的表达变化特征。

1.2.5 RT-PCR 和 QRT-PCR 验证差异基因 对芯片分析获得的亚洲棉 DPL971 及其光籽突变体 DPL972 间有差异的部分基因进行 RT-PCR 和 QRT-PCR 验证。用 RNA PCR Kit 将 1 μL 总 RNA 反转录成 cDNA 第一链,反应体系为 2 mol·μL<sup>-1</sup> 左右引物, 0.2 mmol·μL<sup>-1</sup> dNTP, 0.5 U *Ex Taq* DNA 聚合酶, 1/10 体积的 10× *Ex Taq* buffer, 加水调节到 20 μL。由于不同基因的扩增需要不同的循环数,优化扩增条件之后,每个反应进行 2 次重复。

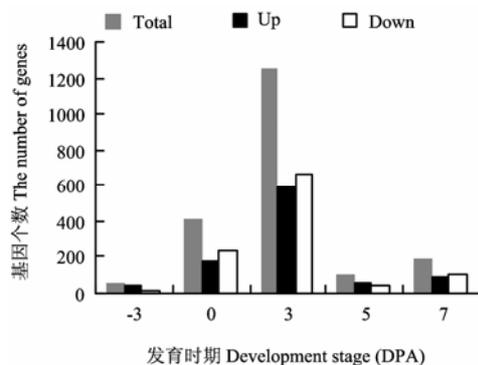
QRT-PCR 反应用 Lightcycler480 (Roche, Switzerland)进行,从不同时期的胚珠中分别取出 2 μg RNA 用于反转录,反转录酶 SuperScriptIII,反应体积为 20 μL。加水 80 μL 稀释之后用于 QRT-PCR。QRT-PCR 反应体系按照 SYBR Green PCR Master MIX,扩增程序如下, 95℃ 10 min; 95℃ 15 s, 60℃

1 min, 40 个循环,每个试验重复 3 次,最后取平均值。

## 2 结果

### 2.1 基因表达谱分析

试验分别收取亚洲棉 DPL971 及其光籽突变体 DPL972 的-3DPA, 0DPA, +3DPA +5DPA, +7DPA 时期的胚珠 RNA 制备了 5 张芯片。野生型 DPL971 及其短绒突变体 DPL972 分别用 Cy3 和 Cy5 标记杂交于同一张芯片上,用 Lowess 方法进行回归分析,以 Ratio>2 和 Ratio<0.5 为标准找到在亚洲棉 DPL971 及其光籽突变体 DPL972 所选定 5 个时期的表达差异基因一共有 1 779 个。参照与 NCBI 拟南芥数据库 BLASTx 比对结果,402 个差异表达基因已经有明确的功能注释,未知功能基因有 1 377 个。从芯片结果(图 1)可以看出基因的差异表达量在 0DPA 较高,在+3DPA 达到最高。在 0DPA 上调基因为 171 个,下调基因 235 个。在+3DPA 上调基因为 584 个,下调基因为 663 个。这个结果与前人研究的棉花表皮毛在+4DPA 开始进行第 2 次突起的结果相符。



Up: 上调基因; Down: 下调基因; Total: 上调基因和下调基因总和  
Up: Up-regulated genes; Down: Down-regulated genes; Total: The total number of up-regulated genes and down-regulated genes

图 1 cDNA 芯片中各个胚珠发育时期差异表达基因数量图

Fig. 1 The number of differentially expressed genes in ovules at different developmental stages by cDNA microarray

### 2.2 短绒起始发育相关基因的确定

将挑选的 1 779 个差异基因与博奥生物网站拟南芥数据库 (<http://bioinfo.capitalbio.com/mas>) 相结合,发现在芯片差异表达基因中存在 8 个与拟南芥表皮毛分化发育相关的同源基因(表 2),其中有 7 个在拟

南芥中编码有相应的蛋白, KAK (DR461366)、MYB5 (ES8120 48)、TTG1 (ES811600)、MYB23 (DR453866)、CSLD3 (DR459646)、RHD2 (DR461821)、ZWI (ES791383)。这些基因中有 3 个, MYB23、ATMYB5 和 TTG1, 曾被报道在拟南芥表

皮毛分化发育中发挥着十分重要的作用。RHD2、钙调蛋白 ZWI 和纤维素合成蛋白 CSLD3 在纤维发育中也有重要的作用。这 8 个差异表达基因均在 0DPA 及 +3DPA 有显著差异。其中有 6 个在 +3DPA 具有显著的差异, 有 3 个基因在 0DPA 具有显著的差异(表 2)。

表 2 cDNA 芯片筛选的可能与棉花短绒起始分化发育相关的基因

Table 2 Genes related to the initiation of the cotton fuzz fiber identified by cDNA microarray

| 登录号*          | 拟南芥编号       | 注释  | E 值       | 表达情况 **          |
|---------------|-------------|---|-----------|------------------|
| Accession No. | Ath_Gene ID | Description   | E-value   | -3、0、+3、+5、+7DPA |
| CM096F05      | AT4G38600   | HECT 类型泛素连接蛋白 Putative HECT ubiquitin-protein ligase 3 ( <i>Glycine max</i> )                                       | 7.00E-20  | 0 0 - 0 0        |
| UFL062F06     | AT3G13540   | MYB2 类似蛋白 MYB-like protein 2 ( <i>Gossypium barbadense</i> )  | 3.00E-48  | 0 + 0 0 +        |
| UFL047B02     | AT5G24520   | WD 重复蛋白 GhTTG3 WD-repeat protein GhTTG3 ( <i>Gossypium hirsutum</i> )   | 1.00E-56  | 0 0 - 0 0        |
| CM017G06      | AT5G40330   | DNA 结合蛋白/转录蛋白 DNA binding / transcription factor ( <i>Arabidopsis thaliana</i> )                                    | 7.00E-53  | 0 0 + 0 0        |
| CM005E08      | AT5G55480   | 甘油磷酸胆碱二脂酶磷酸二酯酶家族蛋白 Putative glycerophosphoryl diester phosphodiesterase family Protein ( <i>Nicotiana tabacum</i> ) | 3.00E-71  | 0 0 - 0 0        |
| CM078E04      | AT3G03050   | 纤维素合酶类似蛋白 D4 Cellulose synthase-like protein D4 ( <i>Populus tremuloides</i> )                                      | 3.00E-74  | 0 0 - 0 0        |
| CM101D12      | AT5G51060   | 氧化酶蛋白 C Respiratory burst oxidase protein C ( <i>Arabidopsis thaliana</i> )   | 2.00E-93  | 0 - 0 0 0        |
| UFL144A01     | AT5G65930   | 钙调素结合蛋白 Kinesin-like calmodulin binding protein ( <i>Gossypium hirsutum</i> )                                       | 3.00E-142 | 0 + - 0 0        |

\*: CM096F05、UFL062F06、UFL047B02、CM078E04、CM101D12 已经经过 RT-PCR 的验证。\*\*: 基因在-3、0、+3、+5、+7DPA cDNA 芯片中的表达情况, “0”表示在 cDNA 芯片中差异不显著, “+”表示上调表达, “-”表示下调表达

\*: The microarray results of genes CM096F05,UFL062F06,UFL047B02,CM078E04,CM101D12 have been verified by RT-PCR. \*\*: The expression of microarray results in -3、0、+3、+5、+7DPA, “0” shows no difference expression between wild type and its fuzz mutant, “+” means gene expression in fuzz mutant > 2 fold in wild type, “-” shows gene expression in fuzz mutant < 0.5 fold in wild type

### 2.3 RT-PCR 和 QRT-PCR 验证部分芯片差异表达基因

为证实从芯片中获得的结果, 将已经挑选出来的 8 个基因选取 5 个(图 2)在 0DPA、+3DPA、+5DPA、+7DPA 进行验证, 根据 cDNA 芯片杂交序列的 EST 设计引物, 通过 RT-PCR 分析在棉花不同时期表达量的变化。不同引物在扩增时采用不同的扩增循环数(25—45 个循环)来使扩增差异效果最好, 每个基因都重复扩增验证差异的稳定性。RT-PCR 结果显示, MYB23 (DR453866) 在突变体 DPL972 中的表达量比野生型 DPL971 中普遍要高, 且 +3DPA 差异最显著, 该基因在野生型 +3DPA 中表达量很低, 几乎被抑制, 而在突变体中该时期表达量却较高, 在其它 3 个时期该基因的表达也有差异但并不十分显著; MYB5 (ES812048) 在这 4 个时期的胚珠中表达量较低, 野生型和突变体 +5DPA 和 +7DPA 差异显著, 野生型中该基因在 +5DPA 明显上调, 突变体中该时期明显下调, 而 +7DPA 该基因在野生型中下调, 在突变体中却上调。该基因在 0DPA 和 +3DPA 差异不大; TTG1 (ES811600) 在 +3DPA 野生型中明显上调, 而在突变体中下调, 其它时期差异不大; CSLD3 (DR459646)

表达量也较低, 且在 +3DPA 和 +5DPA 差异显著, 在这 2 个时期野生型中该基因表达量比突变体明显要高; RHD2 (DR461821) 在这 4 个时期表达差异都比较显著, 其中, 除在 +3DPA, 野生型该基因表达量比在突变体中高之外, 其它 3 个时期该基因在突变体中的表达量都比野生型要高。通过与芯片 Ratio 值的比较可以看出, 多数 RT-PCR 扩增结果与芯片杂交结果都有较好的一致性(图 2 和表 3)。但是, 也有一些基因并不完全与芯片结果一致, 例如, MYB23 在 +7DPA、MYB5 在 0DPA 以及 RHD2 在 0DPA 和 +5DPA 并不符合芯片的结果。但总的来说, RT-PCR 和芯片的结果符合性较好。为进一步证实芯片的准确性, 笔者以 MYB23 (DR453866) 为例深入进行了 RT-PCR 和 QRT-PCR 分析(图 3)。为更详尽地了解该基因的表达规律, 笔者在 RT-PCR 中增加了不同时期胚珠的 RNA。结果显示, 与突变体 DPL972 相比较, 该基因在野生型 DPL971 中表达量普遍较低, 在 +1DPA, +3DPA, +4DPA 这几个时期差异最显著, 尤其在 +3DPA 该基因在野生体中几乎被抑制, 与芯片 Ratio 值结果相符。在 QRT-PCR 中, 该基因在 0DPA 和

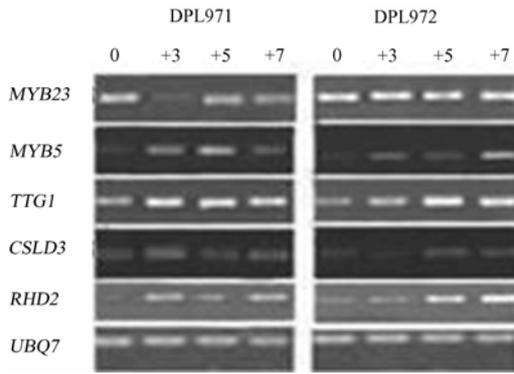


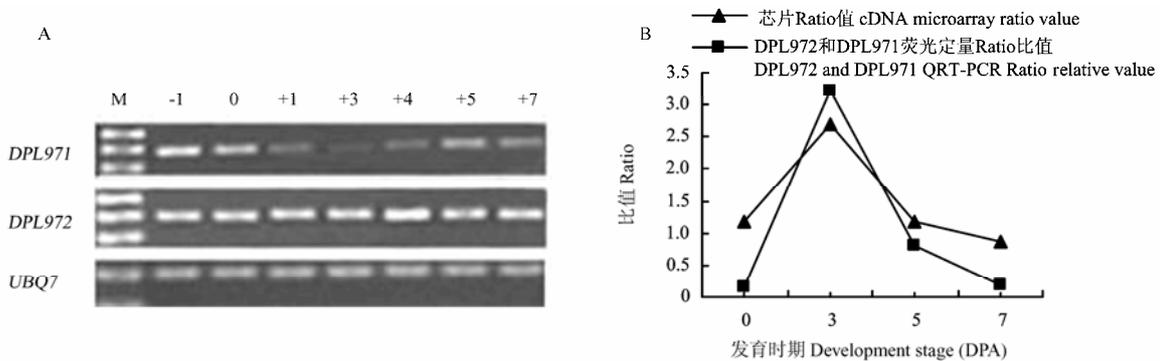
图 2 5 个差异基因在纤维发育不同时期的 RT-PCR  
Fig. 2 RT-PCR detection of five differentiation expressed genes at different fiber developing stages

+3DPA 差异最显著, 表达变化趋势和芯片结果在这几个时期也大致相同。

上述结果可以说明 *MYB23* (DR453866)、*MYB5* (ES812048)、*TTG1* (ES811600)、*CSLD3* (DR459646) 和 *RHD2* (DR461821) 这 5 个基因在短绒突变体 DPL972 与其野生型 DPL971 的+3DPA 胚珠中表达差异都很明显, 而+3DPA 正是短绒分化发育的关键时期, 可以说明这 5 个基因可能与短绒分化和发育相关。*MYB23* (DR453866) 在突变体 DPL972 中表现上调, 在野生型中表现下调, 而 *MYB5* (ES812048)、*TTG1* (ES811600)、*CSLD3* (DR 459646)、*RHD2* (DR461821) 在突变体 DPL972 中表现下调, 在野生型中表现上调。这说明 *MYB23* 可能抑制短绒分化和发

表 3 5 个差异表达基因在纤维发育不同时期 cDNA 芯片分析中的 Ratio 值  
Table 3 Ratio value of five selected genes from cDNA microarray analysis

| 登录号<br>Accession No. | 基因<br>Gene | E 值<br>E-value | Ratio           |                   |                   |                   |
|----------------------|------------|----------------|-----------------|-------------------|-------------------|-------------------|
|                      |            |                | 0DPA<br>R(M/W0) | +3DPA<br>R(M/W+3) | +5DPA<br>R(M/W+5) | +7DPA<br>R(M/W+7) |
| CM017G06             | MYB23      | 8.00E-51       | 1.1886          | 2.7015            | 1.1657            | 0.8766            |
| UFL062F06            | MYB5       | 2.00E-45       | 4.4666          | 0.5892            | 0.7508            | 11.8886           |
| UFL047B02            | TTG1       | 4.00E-51       | 0.6904          | 0.2781            | 0.8168            | 0.9737            |
| CM078E04             | CSLD3      | 7.00E-74       | 1.0988          | 0.3161            | 0.6532            | 0.9962            |
| CM101D12             | RHD2       | 2.00E-109      | 0.3884          | 0.5572            | 1.1165            | 1.4774            |



A: *MYB23* (DR453866)的 RT-PCR 验证; B: *MYB23* (DR453866)的 QRT-PCR 验证  
A: Verification of microarray results by RT-PCR; B: Verification of microarray results by QRT-PCR

图 3 RT-PCR 与 QRT-PCR 对 cDNA 芯片结果的验证

Fig. 3 Verification of cDNA microarray results by RT-PCR and QRT-PCR

育, 而 *MYB5* (ES812048)、*TTG1* (ES811600)、*CSLD3* (DR459646)、*RHD2* (DR461821) 基因则可能促进短绒分化和发育。

### 3 讨论

棉纤维胚珠表皮毛的分化发育历经 4 个阶段, 纤

维起始、细胞伸长、次生壁沉积和纤维成熟<sup>[13]</sup>。这 4 个时期各有特点, 但并没有明显的界限, 它们之间存在着重叠。在这 4 个时期中, 纤维细胞形态结构改变, 伴随着重要的生理生化过程, 有大量的基因参与纤维发育过程的调控。长纤维发育的起始期较短, 在开花前后的 1—2 d 外珠被的许多表皮细胞就开始发生增大突起, 最终发育成为棉纤维。短绒的突起是在开花后 4 d<sup>[3]</sup>, 也有人认为在开花后 5—10 d<sup>[14]</sup>。赖童飞<sup>[15]</sup>用石蜡切片观察到在野生型 DPL971 的+4DPA 胚珠部分表皮细胞排列已经出现不规则现象, 极个别的胚珠表皮细胞出现了突起, DPL971 短绒突起主要出现在+5DPA, 一般来说, 纤维和短绒的分化发生在其突起之前 2—3 d, 因此 DPL971 胚珠表皮细胞的第 2 次分化(短绒起始)应该发生在+3DPA 之前。运用基因芯片技术对亚洲棉野生型 DPL971 及其短绒突变体 DPL972 纤维发育前期进行比较的过程中, 发现在 0—3DPA, 尤其是+3DPA 差异表达的基因数量突然增加, 这可能是由于纤维细胞形态结构要发生改变造成的, 这个现象很可能与棉花表皮毛第 2 次突起有重要的关系。

拟南芥表皮毛和棉纤维细胞都是单细胞分化发育而来的, 二者起始时期均在表皮表面, 拟南芥表皮毛已经被作为研究表皮毛发育的一种模式。通过与拟南芥数据库的比较, 挑选出 8 个可能与棉纤维发育相关的基因。在这 8 个基因中, 有 4 个基因具有相似的表达模式, 在突变体+3DPA 具有显著下调趋势, 这不仅显示了+3DPA 在棉纤维第 2 次突起中的重要性, 也表明正向调控因子可能在纤维的起始发育中占据了很大的比例。

利用 cDNA 芯片技术从亚洲棉野生型 DPL971 及其短绒突变体 DPL972 挑选的 8 个与纤维发育相关的差异表达基因中, DR461366 与拟南芥 KAK 蛋白具有较高的同源性, KAK 蛋白编码大量的 HECT 泛素蛋白连接酶基因, 这个家族的基因与表皮毛细胞形态发育相关, 该基因的突变会增加表皮毛的分支<sup>[16]</sup>。DR453866 和 ES812048 也分别与拟南芥 MYB23 和 MYB5 具有较高的同源性, 这 2 个基因属于 R2R3MYB 家族基因, *AtMYB23* 在拟南芥野生型中的过表达可导致叶片表皮毛数量的减少, *AtMYB23* 可能与 *GLI/WER* 有部分的功能重叠, 控制着拟南芥叶片表皮毛的发育启动<sup>[17]</sup>; *ATMYB5* 可能在拟南芥表皮毛分叉的过程中起负调控的作用<sup>[18]</sup>。QRT-PCR 分析发现 *MYB23* (DR453866) 在野生型+3DPA 达到最低, 而在突变

体中该基因表达量一直都比较低, 可能该基因在棉纤维突变体中过量表达导致短绒不能突起, 是短绒突起的抑制因子。RT-PCR 显示, *MYB5* (ES812048) 在突变体中表达量较低, 尤其在 0DPA、+3DPA 和+5DPA, 而在野生型+3DPA 和+5DPA 中该基因表达量较高, 可能由于该基因在短绒生长发育的过程中有一定的促进作用。目前已证实拟南芥表皮毛的启动至少由 3 个基因 *GLI*、*GL3*、*TTG1* 的产物组成的复合体 MYB-bHLH-WD40 所控制。ES811600 与拟南芥 *TTG1* 有较高的相似性, *TTG1* 编码含有 WD40 重复序列的蛋白, 这个结构域的功能可以提高与其它蛋白的相互作用<sup>[19]</sup>。*TTG1* (ES811600) 在这 4 个时期的 RT-PCR 中, 突变体中 0DPA 和+3DPA 表达量较低, 而在野生型中表达量一直较高, 可能由于该基因在短绒发育的关键时期表达量过低影响了关键蛋白的结合, 从而导致短绒不能正常发育。DR459646 和 DR461821 分别与拟南芥 *CSLD3* 和 *RHD2* 蛋白同源, 这 2 个蛋白对拟南芥表皮毛的形态建成具有重要的作用, 在表皮毛突起生长这段时期是必须的, *AtCSLD3* 基因编码纤维素合酶蛋白, 在拟南芥根毛早期发育生长过程中起重要的作用, 该基因突变体植株在拟南芥根毛完成起始分化后停止生长<sup>[20]</sup>。而基因 *RHD2* 的突变也可以导致表皮毛停止生长<sup>[21]</sup>。在 RT-PCR 分析中, 发现这 2 个基因在+3DPA 突变体中表达量突然下降, 有可能影响短绒的正常发育。与 ES791383 同源的拟南芥 *ZWI* 基因编码钙调节蛋白, 在表皮毛形态建成以及细胞分化方面具有重要的作用<sup>[22]</sup>, *ZWI* 基因的突变可以导致表皮毛分支的减少<sup>[23]</sup>。上述分析再一次说明 *KAK*、*MYB5*、*MYB23*、*TTG1*、*CSLD3*、*RHD2*、*ZWI* 等基因不但对拟南芥表皮毛起始发育有关, 对棉纤维的分化和发育也有重要的影响作用。

一般来讲, cDNA 芯片需要重复, 但对于一些特殊的试验, 通过精心的设计, 即使没有重复, 芯片数据也是可信的。本试验将同一发育时期的野生型和突变体点制在一张芯片上, 试验材料遗传背景一致, 同时, 通过对多个发育时期野生型和突变体胚珠的差异基因进行比较, 有利于背景的消除, 减小芯片的误差; 笔者此前的研究已经确定了野生型和突变体短绒发育起始时期(+3DPA), 本试验通过对这个时期前后的野生型和突变体胚珠取样研究, 有类似重复的效果, 将这些试验结果从表达趋势上进行分析, 也可以排除只从一个时期一个数值选择目的基因的弊端, 这样做也可以减小芯片的误差。此外, 笔者进行了多个基因

片段的半定量 PCR, 用于验证芯片的结果, 所以从总体上来说, 本试验设计可以避免 cDNA 芯片没有重复产生的假阳性问题, 芯片结果还是比较可信的。

芯片能从整体的角度, 去寻找与短绒发育相关的基因及基因网络。但由于芯片数据的庞大, 很难对大量基因进行一一分子验证, 很可能遗漏了不少与短绒分化发育相关的基因。本研究找到 8 个可能与亚洲棉短绒发育相关的基因, 仅仅局限于在拟南芥中已经有明确功能注释的 402 个差异表达基因, 还有 1 377 个未知功能基因没有涉及, 其中也可能还含有很多与短绒发育相关的基因, 这将是下一步研究工作的重点。显然, 棉花表皮毛的分化生长发育是一个复杂的过程, 需要大量基因的调控, 本试验虽然将 cDNA 芯片与拟南芥数据库相结合仅仅筛选出少量与棉花短绒起始发育相关的基因, 但将模式作物表皮毛的研究应用于棉纤维生长发育机理的研究中也是一种可行的途径。

## 4 结 论

本研究以 cDNA 芯片技术为基础, 筛选出 8 个与棉花纤维和短绒初始发育相关的基因 *KAK* (DR461366)、*MYB5* (ES812048)、*TTG1* (ES811600)、*MYB23* (DR453866)、*CSLD3* (DR459646)、*RHD2* (DR461821)、*ZWI* (ES79 1383) 和 DR452709, 经 RT-PCR 和 QRT-PCR 初步验证有 5 个基因可能与亚洲棉短绒分化发育相关, *MYB23* (DR453866) 基因可能抑制短绒分化和发育, 而 *TTG1* (ES811600)、*CSLD3* (DR459646)、*RHD2* (DR461821)、*ZWI* (ES791383) 4 个基因则可能促进短绒分化和发育。

## References

- [1] Ramsey J C, Berlin J D. Ultrastructure of early stages of cotton fiber differentiation. *Botanical Gazette*, 1976, 137(1): 11-19.
- [2] Joshi P C, Wadhvani A M, Johri B M. Morphological and embryological studies of *Gossypium* L. *Proceedings of the National Institute of Sciences of the India*, 1967, 33: 37-93.
- [3] John M E, Crow L J. Gene expression in cotton (*Gossypium hirsutum* L.) fiber: Cloning of the mRNAs. *Proceedings of the National Academy of Sciences of the USA*, 1992, 89: 5769-5773.
- [4] Ma D P, Tan H, Si Y, Creech R G, Jenkins J N. Differential expression of lipid transfer protein gene in cotton fiber. *Biochimica et Biophysica Acta*, 1995, 1257: 81-84.
- [5] Ruan Y L, Llewellyn D J, Furbank R T, Chourey P S. The delayed initiation and slow elongation of fuzz-like short fiber cells in relation to altered patterns of sucrose synthase expression and plasmodesmata gating in a lintless mutant of cotton. *Journal of Experimental Botany*, 2005, 56: 977-984.
- [6] Harmer S E, Orford S J, Timmis J N. Characterisation of six alpha-expansin genes in *Gossypium hirsutum* (upland cotton). *Molecular Genetics and Genomics*, 2002, 268: 1-9.
- [7] 常凤启, 刘选明, 李银心, 贾庚祥, 马晶晶, 刘公社, 朱至清. 低能 N<sup>+</sup> 辐射拟南芥诱导基因组 DNA 碱基变异分析. *中国科学(C 辑)*, 2003, 33(2): 117-124.  
Chang F Q, Liu X M, Li Y X, Jia G X, Ma J J, Liu G S, Zhu Z Q. The analysis of bases Variation of genome DNA in *Arabidopsis thaliana* irradiated by low energy N<sup>+</sup> ion beam. *Science in China (Series C)*, 2003, 33 (2): 117-124. (in Chinese)
- [8] Loguercio L L, Zhang J Q, Wilkins T A. Differential regulation of six novel MYB-domain genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L.). *Molecular Genetics and Genomics*, 1999, 261: 660-671.
- [9] Suo J F, Liang X, Pu L, Zhang Y S, Xue Y B. Identification of GhMYB109 encoding a R2R3 MYB transcription factor that expressed specifically in fiber initials and elongating fibers of cotton (*Gossypium hirsutum* L.). *Biochimica et Biophysica Acta*, 2003, 1630: 25-34.
- [10] Humphries J A, Walker A R, Timmis J N, Orford S J. Two WD-repeat genes from cotton are functional homologues of the *Arabidopsis thaliana* *TRANSPARENT TESTA GLABRA1* (*TTG1*) gene. *Plant Molecular Biology*, 2005, 57: 67-81.
- [11] Yang S S, Cheung F, Lee J J, Ha M, Wei N E, Sze S H, Stelly D M, Thaxton P, Triplett B, Town C D, Chen Z J. Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. *The Plant Journal*, 2006, 47: 761-775.
- [12] 蒋建雄, 张天真. 利用 CTAB/酚法提取棉花组织总 RNA. *棉花学报*, 2003, 15(3): 166-167.  
Jiang J X, Zhang T Z. Extraction of total RNA in cotton tissues with CTAB-acidic phenolic method. *Cotton Science*, 2003, 15(3): 166-167. (in Chinese)
- [13] Graves D A, Stewart J M. Chronology of the differentiation of cotton (*Gossypium hirsutum* L.) fiber cells. *Planta*, 1988, 175: 254-258.
- [14] Beasley C A. *Ovule Culture: Fundamental and Pragmatic Research for the Cotton Industry*. Springer-Verlag, New York, NY. 1977.
- [15] 赖童飞. 亚洲棉短绒纤维发育突变体蛋白质差异分析[D]. 北京: 中国农业科学院, 2007.  
Lai T F, Analysis of different proteins of fuzz mutant in *Gossypium*

- arboretum*[D]. Beijing: Chinese Academy of Agricultural Sciences, 2007. (in Chinese)
- [16] Refy A E, Perazza D, Zekraoui L, Valay J G, Bechtold N, Brown S, Hülskamp M, Herzog M, Bonneville J M. The *Arabidopsis* *KAKTUS* gene encodes a HECT protein and controls the number of endoreduplication cycles. *Molecular Genetics Genomics*, 2003, 270: 403-414.
- [17] Gonzalez A, Mendenhall J, Huo Y J, Lloyd A. TTG1 complex MYBs, MYB5 and TT2, control outer seed coat differentiation. *Developmental Biology*, 2009, 325(2): 412-421.
- [18] Kirik V, Bouyer D, Schöbinger U, Bechtold N, Herzog M, Bonneville J M, Hülskamp M. *CPR5* is involved in cell proliferation and cell death control and encodes a novel transmembrane protein. *Current Biology*, 2001, 11: 1891-1895.
- [19] Szymanski D B, Lloyd A M, Marks M D. Progress in the molecular genetic analysis of trichome initiation and morphogenesis in *Arabidopsis*. *Trends in Plant Science*, 2000, 5: 214-219.
- [20] Favery B, Ryan E, Foreman J, Linstead P, Boundonck K, Steer M, Shaw P, Dolan L. *KOJAK* encodes a cellulose synthase-like protein required for root hair cell morphogenesis in *Arabidopsis*. *Genes and Development*, 2001, 15: 79-89.
- [21] Foreman J, Demidchik V, Bothwell J H, Mylona P, Miedema H, Torres M A, Linstead P, Costa S, Brownlee C, Jones D G, Davies J M, Dolan L. Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. *Nature*, 2003, 422: 442-446.
- [22] Reddy V S, Day I S, Thomas T, Reddy A S. KIC, a novel Ca<sup>2+</sup> binding protein with one EF-hand motif, interacts with a microtubule motor protein and regulates trichome morphogenesis. *The Plant Cell*, 2004, 16: 185-200.
- [23] Schnittger A, Hülskamp M. Trichome morphogenesis: A cell-cycle perspective. *The Royal Society*, 2002, 357: 823-826.

(责任编辑 毕京翠, 李 莉)