



棉花纤维特异表达基因 *GhF1* 的分离及鉴定

孙杰^{1,2}, 李艳军², 李园莉¹, 汪若海³

(1. 中国科学院微生物研究所, 北京 100080; 2. 新疆石河子大学农学院, 新疆兵团绿洲生态农业重点实验室, 新疆石河子 832000; 3. 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: 采用 mRNA 荧光差异显示结合 cDNA 末端快速扩增技术, 克隆了一个棉纤维特异表达基因的全长 cDNA, 命名为 *GhF1*, 该 cDNA 全长 622 bp, 含有一个编码 66 个氨基酸蛋白的开放阅读框。Southern 杂交分析表明该基因在陆地棉 (*Gossypium hirsutum* L.) 中含有两个拷贝。Northern 杂交分析表明该基因在棉花纤维细胞特异表达, 在纤维发育过程中, *GhF1* 转录产物的累积主要发生在纤维细胞发育的早期阶段。尽管未发现该基因与已知基因有任何同源性, 但其分布的组织特异性和表达的发育阶段性暗示该基因在纤维伸长中起作用。

关键词: 棉花; 纤维特异基因; Northern 杂交

中图分类号: S562.035.3 **文献标识码:** A

文章编号: 1002-7807(2005)05-0259-05

Isolation and Identification of Specifically Expressed Gene *GhF1* in Cotton Fiber Cells

SUN Jie^{1,2}, LI Yan-jun², LI Yuan-li¹, WANG Ruo-hai³

(1. Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, China; 2. College of Agriculture, Key Laboratory of Oasis Ecology Agriculture of BINGTUAN, Shihezi University, Shihezi 832000, China; 3. Cotton Research Institute, Chinese Academy of Agricultural Sciences, Key Laboratory of Cotton Genetic Improvement, Ministry of Agriculture, Anyang 455000, China)

Abstract: Fluorescence differential display (FDD) technique was used to identify genes that are specifically or preferentially expressed in cotton fiber cells. A full-length cDNA of differential display cDNA fragments was cloned with rapid amplification of the cDNA ends. The cDNA contains an open reading frame of 622 bp encoding a protein of 66 amino acids. 3'-UTR of the cDNA is 85 bp and 5'-UTR is 389 bp. The gene was designated *GhF1*. Southern blot analysis showed that the gene had two copies in upland cotton genomic. PCR analysis using sequence of the cDNA as primer and genomic DNA as template showed the gene had no intron in cotton genomic. Northern blot analysis using the full-length cDNA as probe was performed to investigate the expressed level of *GhF1* in various organ and in the developing cotton fiber cells. The results showed that the gene was specifically expressed in fiber cells and its transcripts were abundant at the primary cell wall elongation stage and the expressed level decreased with the fiber cell development, inferring that *GhF1* may play a role in polar elongation of cotton fiber cells. Sequence comparison to GenBank indicates *GhF1* is a new gene and its function is unknown.

Key words: cotton; fiber-specific expression; Northern blot

收稿日期: 2005-03-25

作者简介: 孙杰(1969-), 男, 博士, 副教授, sunjiez@yahooh.com

基金项目: 国家高技术研究发展计划“863 计划”项目(2001AA222052)

棉纤维是由胚珠表皮细胞分化、发育而成的单细胞,其形成和发育过程可分为起始、伸长、次生壁合成和脱水成熟四个时期。与其它类型细胞相比,纤维细胞的发育有如下几个明显的特点:(1)在同一棉铃中,纤维细胞的发育高度同步化;(2)纤维细胞在较长的一段时间内(约20 d)快速伸长,长度可达20~30 mm,是其直径的1000~3000倍;(3)成熟棉纤维含95%以上的纤维素,在纤维发育中期和后期,其细胞内的生化过程几乎纯是纤维素的合成^[1]。由于这些特点,棉纤维细胞被认为是研究植物细胞伸长和纤维素生物合成的理想模式系统^[2]。

迄今为止,对于棉纤维发育的分子机理还知之甚少,故目前的研究工作主要是分离鉴定那些仅在或主要在纤维细胞内表达的基因,因为一般认为在纤维细胞内特异表达的基因可能对纤维发育起重要作用。虽然,近年来已经克隆了多个纤维特异表达基因,并对其中一些基因的功能进行了研究,但还有很多纤维特异表达的基因有待克隆和研究,这些工作不仅有助于深入了解纤维发育的分子机理,回答生物学理论上的一些重大问题,同时也可作为基因工程改良棉纤维品质提供目的基因和表达调控元件。本文利用mRNA荧光差异显示(FDD)技术^[3]分离棉纤维不同发育阶段差异表达基因的过程中获得了20多个阳性克隆^[4],把其中一个编号为26-2的cDNA克隆所代表的基因命名为*GhF1*, Northern杂交分析表明它为纤维特异表达,且主要在纤维发育的早期表达。进一步利用cDNA文库筛选克隆了*GhF1*的全长cDNA,序列分析表明这个cDNA为首次被克隆。

1 材料和方法

1.1 实验材料

棉花品种为陆地棉(*Gossypium hirsutum* L.)品种TM-1,由中国农业科学院棉花研究所品种资源室提供。2002年播种于中棉所试验田中,按常规进行农事作业和田间管理。在棉花开花当天挂牌标记棉铃,按6、12、18、24 DPA(开化后天数)采集棉铃,剥取子棉,再用灭过菌的镊子小心地从胚珠上剥取纤维细胞,液氮中速冻后保存于-80℃冰箱。

在实验室内TM-1种子经硫酸脱绒,0.1%氯化汞灭菌10 min,去除种壳后播种于MS盐培养

基中,在28~30℃光照培养箱中培养两周,分别取根、下胚轴和叶,花直接从田间植株上采集,上述材料在液氮中速冻后,保存于-80℃冰箱。上述材料用作RNA和DNA的提取。

1.2 总RNA的提取

取1 g材料在液氮中研碎,按1:2~5加入RNA提取缓冲液^[5],13000 g离心10 min,取3.5 ml上清转入含有9 ml 5 mol·L⁻¹ CsCl垫层的离心管中,113000 g(Beckman Ti90转子,约每分钟40300转)、22℃离心约16 h。离心结束后取出RNA沉淀,溶解于适量DEPC处理水中,测定OD值后于-20℃保存。

1.3 基因组DNA的提取

1 g棉花叶片,液氮研磨后用植物DNA抽提试剂盒(Amersham)提取基因组DNA。

1.4 荧光差异显示

荧光差异显示是在Genomymx(Beckman)仪上进行,用反转录酶Superscript II RT(Gibco)分别从9DPA、21DPA、27DPA纤维总RNA合成cDNA第一链,并以它为模板,选用8对引物进行PCR。锚定引物由17 bp的T7启动子序列,12 bp的polyT和两个可变碱基组成,随机引物由16 bp的M13反向序列(M13r)和10 bp的随机序列组成,锚定引物5'端标记荧光。本实验所用锚定引物为AP3(5' T7dT₁₂ GG 3')和AP7(5' T7dT₁₂ AC 3'),随机引物为APR17(5' M13r-CT-GCTAGGTA 3')、APR18(5' M13r-TGATGCTACC 3')、APR19(5' M13r-TTTTGGCTCC 3')和ARP20(5' M13r-TCGATACAGG 3'),引物组合为AP3/ARP17、AP3/ARP18、AP3/ARP19、AP3/APR20、AP7/ARP17、AP7/ARP18、AP7/ARP19、AP7/ARP20。PCR产物在高分辨率5.6%聚丙烯酰胺胶上电泳,干胶后扫描并分析差异显示cDNA条带^[4]。

1.5 cDNA末端序列的快速扩增(5'-RACE)

合成了两条*GhF1*基因特异反向引物:sp1:5' TCA TCA CCA GTG AGA GTG AA 3';sp2:5' CAC ACC AAC ATG GAA GAG CA 3',按照Gibco公司5'-RACE试剂盒说明进行操作。扩增片段从琼脂糖回收后连接到T-载体(Promga)上,酶切鉴定后测序。

1.6 Northern杂交

15 μg总RNA在1.2%琼脂糖甲醛变性胶上电泳,用毛细法将RNA转移到尼龙膜(Hy-

pond-N⁺, Amersham) 上, 65℃ 预杂交 2 h 后, 利用 Primer-a-gene Labeling System (Promga) 标记试剂盒, 对 GhF1 的 cDNA 进行标记, 用作 Northern 杂交探针, Church 缓冲液杂交 16 h 后洗膜, 压片。

1.7 Southern 杂交

20 μg 棉花 DNA 用限制性内切酶完全酶切后, 0.8% 琼脂糖凝胶电泳, 碱法转移到尼龙膜 (Hypond-N⁺, Amersham) 上, 65℃ 预杂交 2 h 后, 利用 Primer-a-gene Labeling System (Promga) 标记试剂盒, 对 GhF1 的全长 cDNA 进行标记, 用作 Southern 杂交探针, Church 缓冲液杂交 16 h 后洗膜, 压片。

1.8 DNA 序列测定与分析

DNA 序列由自动测序仪(上海基康生物技术有限公司)测序, 同源性比较在 NCBI Blast 上进行, ORF 及蛋白序列分析通过 DNAMAN 和

DNASIS 等软件完成。

2 结果与分析

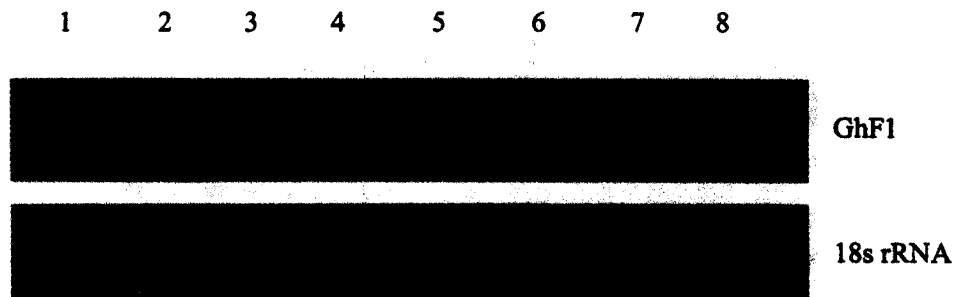
2.1 GhF1 全长 cDNA 的克隆

利用 mRNA 荧光差异显示技术, 从棉纤维中分离了一些在纤维发育不同阶段差异表达, 且仅在或优先在纤维细胞中表达的 cDNA 片段, 其中一个 cDNA 片段编号 26-2 长 450 bp, 将其核酸序列与 GenBank 中公布的已知基因序列无显著同源性, 推测所获得的序列可能代表一个未知新基因。利用 5'-RACE 技术克隆到其全长 cDNA 序列, 序列分析表明: 该 cDNA 长 622 个碱基, 3'非翻译区 389 bp, 5'非翻译区 85 bp, 含有一个开放阅读框架, 编码 66 个氨基酸残基组成的蛋白(图 1), 理论分子量 7.3 kD。我们将这个 cDNA 所代表的基因命名为 GhF1。

	ACAAACAACCTTTCTATTCTCAATT	25
TCCCCATCTCTCTCGTCTCTGGGTCTCGTAAGTTGGTGAATAGTGTGTAATTAGTGCCA		85
ATGGCTCGGATTGGGACCTCTGCAGCTCACATTGTGTGGCAATATTTGCTGTGGCCATG		145
M A R I G T S A A H I V L A I F A V A M		20
TTTGTGTGTCGGGACCATGGCACAGGATATTGCTCCTTCTCCTGCAATGGCTACCGGA		205
F V V S G T M A Q D I A P S P A M A T G		40
GCAGGCTCTGCTTTGCCGTTTCCGCTGTCTTATGCTCTTCCATGTTGGTCTCTTTA		265
A G S A L P V S A V F L C S S M L V S L		60
ATTGCTCTCTGGTGCATTGAATTCAAAGCTTTTCAAGACTTTATGACATTGGCTACCCCT		325
I A L L V H *		66
TAATTTCACTCTCACTGGTGATGAGGGGAGTAGCCTCTAATCTTCTCCGAGATAATATTT		385
GGGTGTATCAATTTTCAATTTCTCTAAAGTTTAAATATCTCTATATATATCATGTTAC		445
ATTAGTGACATAAATTTGATTTTGAATGTAATTGGTGTGATTTTCTTATATATCATAGA		505
TATATAAGGGTATTGATCCTATTCTTTTGTATTATATATGACAGTTTGCTTTGCTTTG		565
CCTTTTGAATAATAATTTTTTTTGGGTGGTTGAAAAATAAAAAAAAAAAAAAAAAAAAA		622

图 1 GhF1 基因的 cDNA 及蛋白序列

Fig. 1 Sequence of GhF1 cDNA and its protein



1 根, 2 下胚轴, 3 叶, 4 花, 5~8 分别为 6、12、18、24DPA 棉纤维, 18s rRNA 用作上样量对照

图 2 GhF1 在棉花不同器官中表达水平的 Northern 分析

Fig. 2 Northern blot analysis of GhF1 in different cotton tissue and developing cotton fibers

2.2 GhF1 基因在棉花不同器官中的表达水平分析

为了了解 *GhF1* 基因在棉花不同器官及纤维细胞不同发育阶段中的表达特性,以其 cDNA 的 3'-UTR 为基因特异探针,用 Northern 杂交方法,对该基因在棉花不同器官及不同发育阶段棉纤维中的转录水平进行了分析。如图 2 所示,在根、下胚轴、叶、花中均未检测到转录产物的表达,*GhF1* 基因在纤维中特异表达,且在棉花纤维发育的早期(6DPA)表达最强,并随着纤维发育延续表达量迅速下降。从 *GhF1* 基因分布的组织特异性和表达的发育阶段性暗示该基因可能在棉纤维发育的伸长期起作用。

2.3 GhF1 基因在棉花基因组中的特性分析

为了分析 *GhF1* 基因在(AD)₁ 基因组中的拷贝数和是否有同源基因存在,用 *GhF1* 全长 cDNA 为探针进行 Southern 杂交,如图 3 所示。在第 1、3、4、5 泳道各出现两个杂交带,而由于 *GhF1* cDNA 中存在第 2 泳道所用限制性内切酶

BamHI EcoRI XbaI KpnI SacI



图 3 *GhF1* 基因的 Southern 杂交分析

Fig. 3 Southern blot analysis of *GhF1*

的酶切点,出现 4 条杂交带。可见,*GhF1* 在陆地棉基因组中含有两个拷贝。除这些杂交带外,泳道中没有明显的其它条带,初步推断该基因不是以基因家族形式存在。我们分别在该 cDNA 的 +13~+23 和 +622~+644 bp 合成了一对引物,用基因组 DNA 为模板进行 PCR,结果与用该 cDNA 作模板为对照的 PCR 产物大小一致(结果未显示),该 PCR 产物克隆到 T-Vector (Promega)测序,发现与 cDNA 序列完全一致,表明该基因序列没有内含子。

3 结论

目前已利用 cDNA 文库差示筛选^[5-8]、mRNA 差异显示^[9-11]、减法杂交^[12]和同源序列^[13]等方法克隆了数十个在棉花纤维中特异或优先表达基因,这些基因在棉纤维的发育过程中发挥作用,但目前还没有发现对纤维品质形成起关键作用的基因。本研究利用 mRNA 荧光差异显示技术结合 5'-RACE 技术克隆了一个纤维特异表达基因,与 GenBank 数据库比对分析,没有发现与 *GhF1* 同源的已知蛋白,但从该基因表达产物的特异分布和发育的阶段性可以看出,它可能在纤维发育的早期(伸长阶段)起作用。目前,我们正利用 Sense 和 Anti-sense 技术研究 *GhF1* 在棉纤维发育中的功能,其上游调控序列的克隆也在进行中。

参考文献:

- [1] DELMER D P, Amor Y. Cellulose biosynthesis[J]. The Plant Cell, 1995, 7: 987-1000.
- [2] HEE J K, Barbara A. Cotton fiber growth in planta and in *vitro* models for plant cell elongation and cell wall biogenesis[J]. Plant Physiology, 2001, 127: 1361-1366
- [3] LIANG P, Pardee A B. Differential display of eukaryotic messenger RNA by means of chain reaction [J]. Science, 1992, 257(5072): 967-970.
- [4] 孙 杰,李园莉,汪若海,等. 利用 mRNA 荧光差异显示技术规模化筛选棉纤维特异表达基因[J]. 生物工程学报,2004,20(1):39-42.
- [5] JOHN M E, Crow L J. Gene expression in cotton (*Gossypium hirsutum* L.) fiber: cloning of the mRNAs[J]. Proc Natl Acad Sci, USA, 1992, 89: 5769-5773.
- [6] JOHN M E. Characterization of a cotton (*Gossypium hirsutum* L.) fiber mRNA(Fb-B6) [J]. Plant Physi-

- ol, 1995, 107: 1477-1478.
- [7] ORFORD S J, Timmis J N. Specific expression of an expansion gene during elongation of cotton fiber[J]. *Biochim Biophys Acta*, 1998, 1398: 342-346.
- [8] RINEHART J A, Petersen M W, John M E. Tissue-specific and developmental regulation of cotton gene *FbL2A*[J]. *Plant Physiol*, 1996, 112: 1331-1341.
- [9] KIM H J, Triplett B A. Cotton fiber germin-like protein. I. Molecular cloning and gene expression[J]. *Planta*, 2004, 218: 516-24.
- [10] SONG P, Allen R D. Identification of a cotton fiber-specific acyl carrier protein cDNA by differential display[J]. *Biochim Biophys Acta*, 1997, 1351: 305-312.
- [11] ZHAO G R, Liu J Y. Isolation of a cotton RGP gene: a homolog of reversibly glycosylated polypeptide highly expressed during fiber development[J]. *Biochim Biophys Acta*, 2002, 1574: 370-374.
- [12] FENG J X, Ji S J, Shi Y H, et al. Analysis of five differentially expressed gene families in fast elongating cotton fiber [J]. *Acta Biochim Biophys Sin (Shanghai)*, 2004, 36: 51-66.
- [13] RUAN Y L, Chourey P S, Delmer D P, et al. The differential expression of sucrose synthase in relation to diverse patterns of carbon partitioning in developing cotton seed[J]. *Plant Physiol*, 1997, 115(2): 375-385. ●