



棉花 *GhCO* 基因的克隆与表达分析

吴 嫔^{1,2}, 范术丽², 宋美珍², 庞朝友², 喻树迅^{2*}

(1.西北农林科技大学, 陕西 杨凌 712100; 2.中国农业科学院棉花研究所 / 农业部棉花遗传改良重点实验室, 河南 安阳 455000)

摘要: 以中棉所 36 均一化全长 cDNA 文库为基础, 利用 RT-PCR 技术从棉花中克隆了一个新的 CO 蛋白基因, 命名为 *GhCO* (GenBank: HM006910)。 *GhCO* cDNA 的 ORF 全长为 1017 bp, 编码 338 个氨基酸, 含有一个 CCT 域和两个 BBOX 域。 序列比较分析结果表明, *GhCO* 蛋白与蓖麻 *RcCO*、芒果 *MiCO* 具有较高的同源性, 是棉花 CO 蛋白家族中的新成员。 QRT-PCR 结果表明, *GhCO* 在棉花的花、蕾、胚珠等均有表达, 而且在蕾和花中优势表达。 *GhCO* 在花芽分化形态出现以前就已经高调表达, 推测可能与棉花的花芽分化有关。 *AtCO* 已经证明在花发育过程中对开花时间起正向调节因子的作用, 推测 *GhCO* 蛋白在花发育过程中可能起重要作用。 因此, 构建了 pBIGhCO 过量表达载体, 为进一步研究 *GhCO* 的功能奠定了重要的基础。

关键词: 棉花; *GhCO* 基因; QRT-PCR; 过量表达载体

中图分类号: S562.035.3 文献标识码: A

文章编号: 1002-7807(2010)05-0387-06

Cloning and Expression Analysis of *GhCO* Gene in *Gossypium hirsutum* L.

WU Man^{1,2}, FAN Shu-li², SONG Mei-zhen², PANG Chao-you², YU Shu-xun^{2*}

(1. College of Agronomy, Northwest Sci-Tech University of Agriculture and Forestry, Yangling, Shaanxi 712100, China; 2. Cotton Research Institute of CAAS, Key Laboratory of Cotton Genetic Improvement, Ministry of Agriculture, Anyang, Henan 455000, China)

Abstract: Based on a normalized full-length cDNA library of CCRI 36, a new CO protein gene was isolated from *Gossypium hirsutum* L. using RT-PCR methods, it was named *GhCO* (HM006910). The cDNA of *GhCO* contained 1017 bp and encoded a putative protein of 338 aa. The encoding protein had a typical CCT domain and two BBOX domains. Bioinformatics analysis indicated that the identities of the deduced *GhCO* amino acid were very high with *Ricinus communis* protein *RcCO*, and *Mangifera indica* protein *MiCO*. Sequence analysis indicated that the gene was a new member of the CO protein family. It was demonstrated that *GhCO* was expressed in flower, square, ovule, and other organs. The highest level of transcripts was accumulated in flower and square. *GhCO* was also strongly expressed prior to any morphological evidence of floral initiation, suggesting that cotton apex differentiation may be relevant. While the *AtCO* had been proved to be a positive on flowering time during the process of flower development, therefore, we speculate that *GhCO* protein play an important role in cotton flowering pathway. In order to study the function of *GhCO* protein furtherly, we constructed a pBIGhCO over-expression vector.

Key words: cotton (*Gossypium hirsutum* L.); *GhCO* gene; QRT-PCR; over-expression vector

棉花是典型的短日照作物, 在短日照条件下, 开花提前, 生育期缩短; 在长日照条件下, 开花推迟, 生育期延长。这种日照长短决定开花时间的现象称为光周期现象。Garner 等是最早对植物开花光周期现象进行研究的, 他们发现许多植

物的开花受日照长短的控制^[1]。近年来, 随着分子遗传学的发展, 尤其是对两种模式植物拟南芥和水稻开花光周期现象的研究, 使得人们对控制这一复杂生物过程的分子机制有了较为清晰的认识。植物将生物钟信号和光信号整合起来最终形

收稿日期: 2010-03-17

作者简介: 吴 嫔 (1980-), 女, 在读博士, wuman2004@163.com; * 通讯作者, yu@cricaas.com.cn

基金项目: 国家转基因重大专项 (2009ZX08005-020B)

成对开花时间的控制,是由 *CO*(*CONSTANS*)基因的转录丰度和 *CO* 蛋白的稳定性所共同决定^[2]。*CO* 基因编码的一种转录调控子受生物钟调控,表达量在一天之内呈节律性变化,它能够促进拟南芥在长日照条件下开花。

棉花开花途径是否也存在 *CO* 基因,开花时间是否受 *CO* 基因的调控,光周期反应的机制与水稻和拟南芥等植物是否有差异,都还未见报道。本实验利用本课题构建的一个中棉所 36 均一化全长 cDNA 文库克隆了一个新的棉花 *GhCO* 全长基因,并进行了初步表达分析,构建了 pBIGhCO 表达载体,为以后进一步研究其功能奠定了重要的基础。

1 材料和方法

1.1 材料处理

试验材料为中棉所 36,2008 年 4 月 20 日在中国农业科学院棉花研究所试验田播种。为了扩增得到与花发育相关的基因和观察基因的时空表达模式,到开花期时,采取花的各个组织(苞叶、花瓣、雄蕊、雌蕊、萼片),不同发育时期的花(-1 d, 0 d, +1 d, +2 d),叶、根、主茎生长点、蕾(长度 3 mm, 4~5 mm, 1 cm, 2 cm) 等器官立即浸于液氮中,-70℃保存备用。

1.2 RNA 提取和 cDNA 制备

总 RNA 分离用 CTAB 法^[3]。取 2~5 g 花、蕾、根、茎、叶等材料用液氮研磨后,加入 15 mL CTAB 提取缓冲液,震荡均匀后 65℃温浴 3~5 min,加入等体积的氯仿-异戊醇(24:1),剧烈震荡后,12000 r·min⁻¹,5 min,吸取上清。重复 1 次,1/4 体积 LiCl 沉淀过夜。溶于 400 μL DEPC-H₂O 中,风干后,加入等体积的酚和氯仿-异戊醇各抽提一次,3 mol·L⁻¹ 醋酸钠(pH=2.5)沉淀,风干。最后 RNA 溶于 20 μL DEPC-H₂O 中,用 Dnase I (promega) 处理后,用于 cDNA 合成。cDNA 第一链合成采用 Oligo dt (18) 和 M-MLV 逆转录酶 (Invitrogen),反应体系按照其说明书进行。

1.3 基因克隆和序列测定

根据生物信息学分析,在本实验室中棉所 36 均一化全长 cDNA 文库中筛选到一个推测与棉花花发育相关的 EST,通过序列拼接得到其全

长,以此序列设计引物。以大田正常生长条件下中棉所 36 植株上的根、茎、叶、花等材料提取 RNA,合成 cDNA 第一链,将第一链产物稀释 10 倍后作为模板,用于基因扩增。扩增 *GhCO* 全长的引物为 5'-AACCCCAGCAACTTGTGAA-3' 和 5'-TACCTTCATCTTCTTTACCTAT-3',反应体系为 25 μL,含第一链 cDNA 稀释产物 1 μL,5 μmol·L⁻¹ 引物各 2 μL, 10×PCR buffer 2.5 μL, MgCl₂ (1.5 mmol·L⁻¹) 1.5 μL, Taq DNA 聚合酶 1 mol·L⁻¹ (宝生物生物工程公司),灭菌 ddH₂O 17 μL。PCR 反应在 PTC-200 DNA Engine Cycler (MJ) 上进行,反应条件为 95℃ 5 min; 94℃ 1 min, 60℃ 1 min, 72℃ 2 min, 35 个循环; 72℃ 延伸 10 min。

基因克隆按照刘天明和宋国琦等方法^[4,5]进行。PCR 产物经琼脂糖凝胶电泳分离后,用刀片切下目标片段,用 DNA 凝胶回收试剂盒(宝生物工程有限公司)纯化,将纯化基因产物与 T-easy 载体(promega)连接。取 5 μL 连接产物转化 100 μL 大肠杆菌 DH5α 的感受态细胞,然后加入 900 μL LB 培养液,37℃、中低速 (150 r·min⁻¹) 振荡培养 1.5 h,之后取 200 μL 菌液涂于含氨苄 (60 μg·mL⁻¹) 的 LB/X-gal/ IPTG 培养板上,37℃培养 14 h,挑取白色克隆。以 T-easy 载体通用引物 M13 对挑取的菌落进行 PCR 扩增,检测阳性克隆插入片段大小。阳性克隆由上海生物工程技术有限公司采用 M13 正向或反向引物进行测序。

1.4 序列比较分析

得到的序列去除载体序列后用 BLASTX (<http://www.ncbi.nlm.nih.gov/BLAST/>) 进行同源序列分析,用 Clustal W (<http://www.ebi.ac.uk/Clustalw/>) 进行多重序列比对。

1.5 QRT-PCR 分析

分别提取棉花根、茎、叶等总 RNA 5 μg 进行反转录反应。实验采用 Invitrogen 公司的 Super ScriptTM First-Strand Synthesis System for RT-PCR 试剂盒,生成的第 1 链 cDNA 用作 QRT-PCR 的模板。QRT-PCR 使用 SYBR green PCR 试剂盒 (Applied Biosystems 公司) 标记反应产物,PCR 分析仪器为罗氏 LightCycler[®] 480 Real-time Cycler。棉花的 18S 基因为内标,使用基因特异性引物

5'-TCGGGTCTTGGTCTGTGAAGT-3' 和 5'-GAGGGTTAGCGGAGTGGATG-3'。

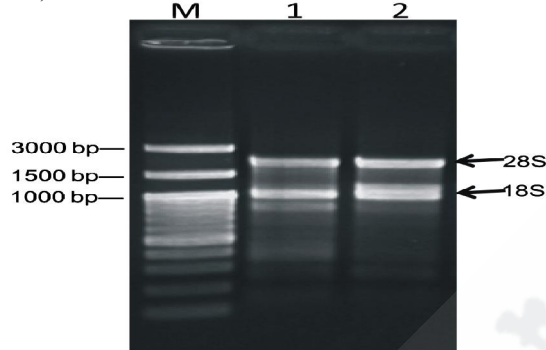
1.6 植物过量表达载体 pBI121-*GhCO* 的构建及鉴定

使用本实验室保存的 pBI121 质粒,构建 *GhCO* 基因的植物过量表达载体。

2 结果与分析

2.1 RNA 提取

中棉所 36 RNA 经琼脂糖凝胶电泳,从图 1 可看出,28S 和 18S 电泳条带清晰,无拖尾现象产生,表明所提取的 RNA 质量较好。



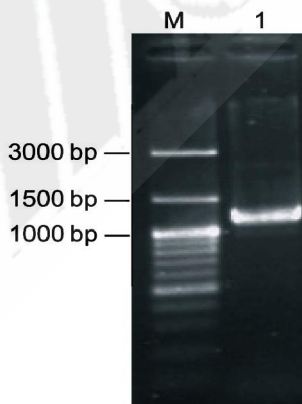
M: DNA 分子量标记(100 bp DNA ladder); 1-2: 总 RNA。

图 1 中棉所 36 总 RNA

Fig. 1 The RNA of CCRI 36

2.2 RT-PCR 扩增 *GhCO* 基因

通过 RT-PCR 扩增获得 *GhCO* 基因,反应物经 1.0% 琼脂糖凝胶电泳检测,结果如图 2 所示。由图 2 可知,扩增片段大小和预期大小基本一致,长度在 1100 bp 左右。对扩增得到的 *GhCO* 进行克隆测序,结果发现 *GhCO* 的 ORF 全长为 1017 bp,编码 338 个氨基酸。



M: DNA 分子量标记(100 bp DNA ladder); 1: *GhCO*。

图 2 棉花 *GhCO* PCR 扩增结果

Fig. 2 PCR product of *G. hirsutum GhCO*

2.3 *GhCO* 蛋白结构比对和进化树分析

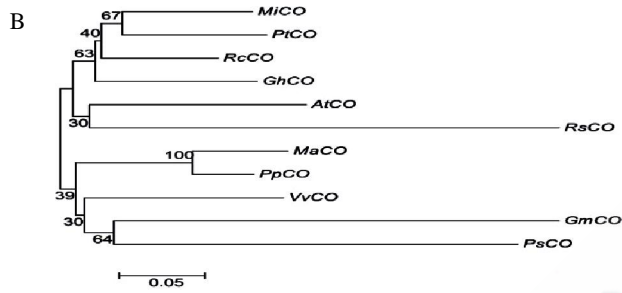
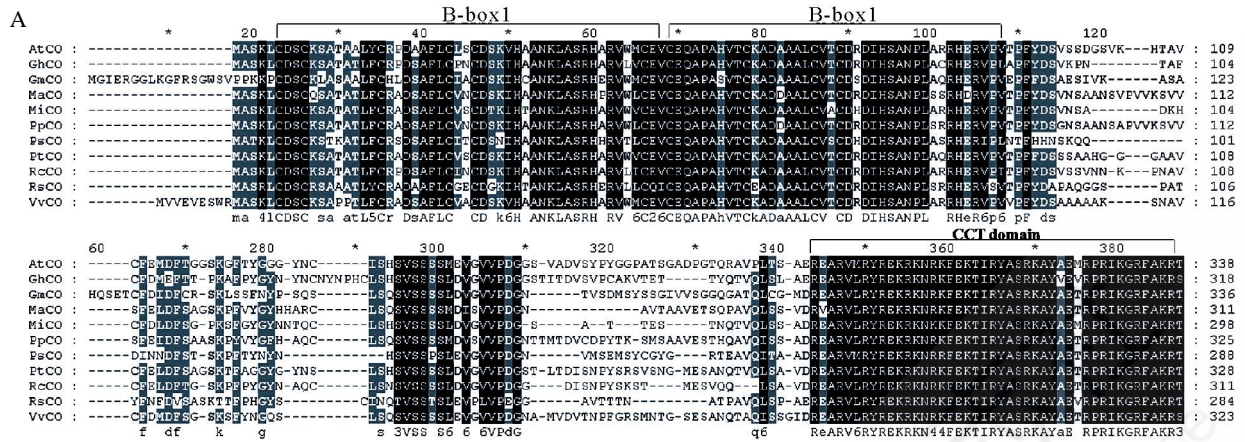
对 *GhCO* 基因编码的氨基酸序列进行 BLASTP 分析表明,与蓖麻(*Ricinus communis*)、芒果 (*Mangifera indica*)、毛果杨 (*Populus trichocarpa*)等物种中 CO 基因或 CO 基因类似物编码氨基酸序列同源性达到 70% 以上,均含有两类保守结构域,N 端的 B-box 型锌指结构域和靠近 C 端的 CCT 结构域 (CO, CO-Like, TOC1 蛋白) (图 3A)。为明确棉花 *GhCO* 蛋白与其它植物 CO 蛋白的进化关系,选取拟南芥、大豆、苹果、芒果、桃等 10 个 CO 蛋白构建了进化树。如图 3B 所示,棉花 *GhCO* 蛋白与芒果、杨树、蓖麻、拟南芥、萝卜 CO 蛋白在进化关系上比较接近,与其它植物中 CO 蛋白进化关系较远。由于序列的同源性与其功能的相似性往往存在一致,已经证明 AtCO 在花发育过程中对开花时间起正向调节因子的作用,因此,我们推测 *GhCO* 蛋白在棉花开花途径中起重要作用。

2.4 棉花 *GhCO* 基因的表达分析

为研究 *GhCO* 在棉花中表达的时空特性,提取不同组织和主茎生长点不同发育时期的总 RNA,以棉花 18S 基因作内标,进行 QRT-PCR 分析。结果表明,*GhCO* 基因在棉花的花、胚珠、茎、叶、主茎生长点、苞叶、根、花瓣、雄蕊、雌蕊、萼片中均有表达,在蕾和花中具有较高的表达量,说明这个基因可能与蕾和花的发育有关(图 4A)。为了进一步研究 *GhCO* 在蕾和花中的表达趋势,分别分析了 *GhCO* 在长度 3 mm、4~5 mm、1 cm、2 cm 蕾中和 -1 d、0 d、+1 d、+2 d 花中的表达量。从图 4B 和图 4C 中分别可以看出,*GhCO* 在长度为 1 cm 的蕾中表达量最高,在开花 -1 d、0 d、+1 d、+2 d 的花中都有表达,其中,0 d 时花中 *GhCO* 优势表达。*GhCO* 在苗期 10 SDs (short days, 短日照处理出苗后 10 d, 1 片真叶展平时) 时表达量很高,在 20 SDs (出苗后 20 d, 2 片真叶展平时) 时表达量下降。而在 TM-1 中,10 SDs 时表达量很低,在 20 SDs 时表达量升高(图 4D)。

2.5 棉花 *GhCO* 蛋白基因表达载体的构建

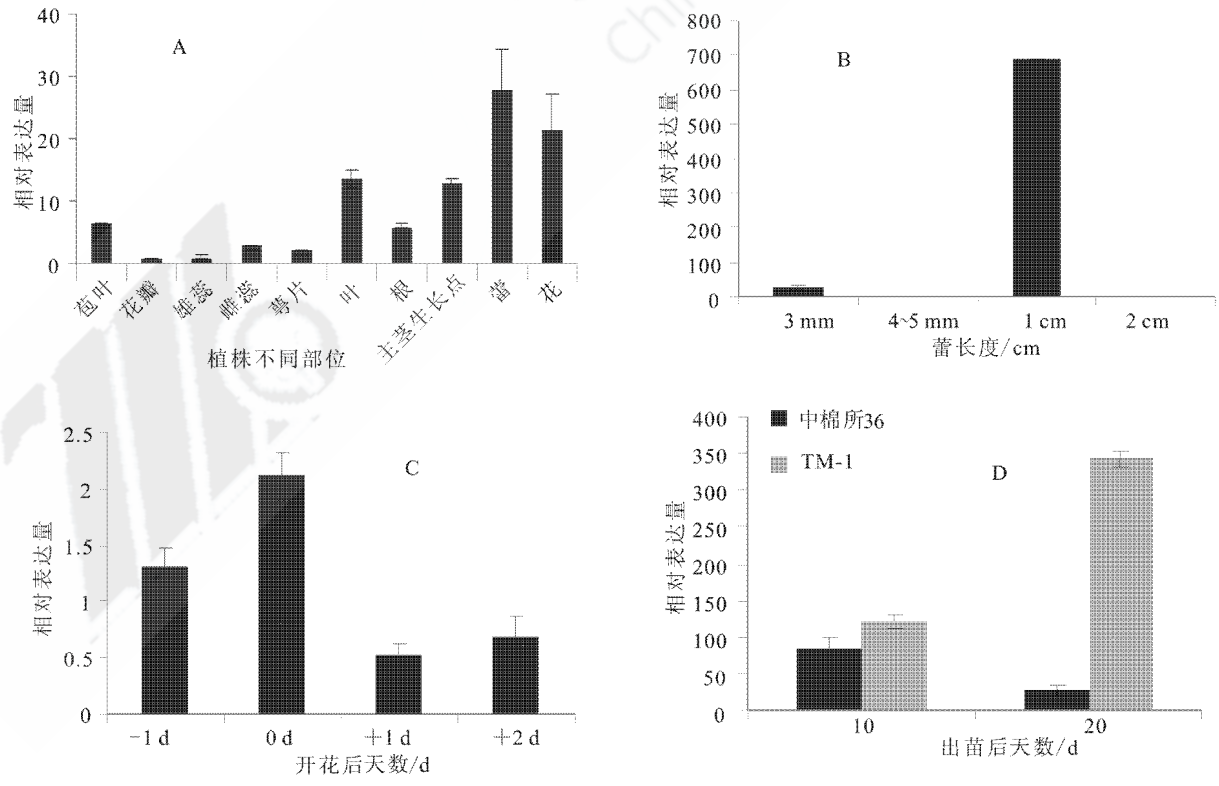
根据 pBI121 载体的 GUS 基因两边的酶切位点,重新设计合适的引物(分别在正反引物的 5' 端加上酶切位点 *Sma*I 和 *Xba*I),用这 2 组引



AtCO: 拟南芥编码 CO 蛋白基因(Q940T9.2); GmCO: 大豆编码 CO 蛋白基因(ACX42572.1); MaCO: 苹果 CO 蛋白(AAC99310.1), MiCO: 芒果 CO 蛋白(ACN62415.1), PpCO: 桃 CO 蛋白 (ACH73166.1), PsCO: 豌豆 CO 蛋白 (AAX47173.1), PtCO: 杨树 CO 蛋白(XP_002309695.1), RcCO: 蓖麻 CO 蛋白(XP_002515382.1), RsCO: 萝卜 CO 蛋白 (AAC35496.1), VvCO: 葡萄 CO 蛋白(XP_002263458)。

图 3 棉花 GhCO 的推定氨基酸序列及同源序列比较(A)及进化树分析(B)

Fig. 3 Sequence alignment(A) and phylogenetic analysis(B) of the predicted amino acid sequence of *G. hirsutum* GhCO



A. 植株不同部位; B. 蕾不同时期; C. 花不同时期; D. 主茎生长点不同时期。

图 4 GhCO 基因的 QRT-PCR 分析

Fig. 4 The QRT-PCR analysis of the expressions of GhCO

物再次进行 RT-PCR。将 PCR 产物和 pBI121 载体分别进行 *Sma*I 和 *Xba*I 双酶切,回收并使用 T4 连接酶连接带有相同粘性末端的 PCR 产物和 pBI121 载体大片段,得到表达载体 pBIGhCO (图5)。

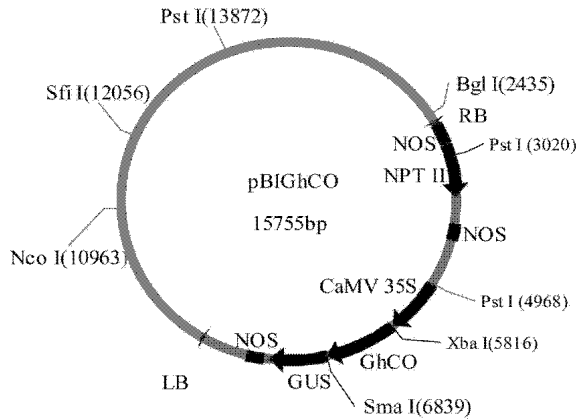


图 5 植物表达载体 pBIGhCO 的构建

Fig. 5 Expression vectors of *GhCO* is constructed by using pBI121

2.6 pGET-easy-*GhCO* 重组质粒菌落 PCR 鉴定

将 RT-PCR 扩增所获得的 *GhCO* 基因产物经过电泳回收纯化步骤直接与质粒 pGET-easy 进行连接,转化大肠杆菌 DH5 α 感受态细胞,再涂布于含氨苄青霉素 ($50 \mu\text{g}\cdot\text{mL}^{-1}$) 抗性平板上培养,待长出菌落后挑选单菌落,采用扩增 *GhCO* 全长的引物进行菌落 PCR。结果表明,pGET-easy-*GhCO* 重组质粒的菌落 PCR 结果条带大小一致。将含有 pGET-easy-*GhCO* 重组质粒的单菌落送上海北京华大基因研究中心进行 DNA 测



M: DNA 分子量标记 (100 bp DNA ladder); 1-7: *GhCO* 重组质粒为模板菌落 PCR 结果。

图 6 pGET-easy-*GhCO* 重组质粒菌落 PCR 琼脂糖凝胶电泳图

Fig. 6 PCR products in gel electrophoresis of pGET-easy-*GhCO* recombinant colonies

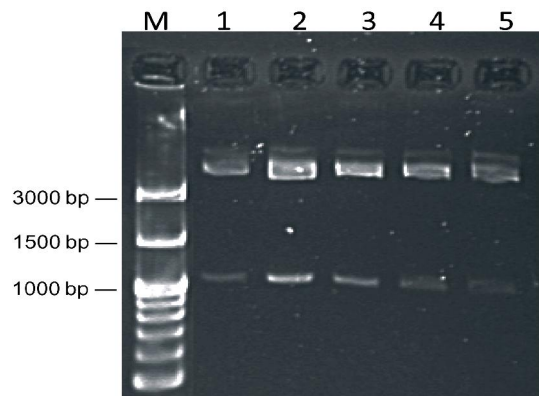
序。pGET-easy-*GhCO* 重组质粒中 *GhCO* 与本实验室中棉所 36 均一化全长 cDNA 全长文库中测序所得到的 mRNA 序列完全一致,能够编码全长氨基酸序列(图 6)。

2.7 pBI121-*GhCO* 重组质粒鉴定

挑选含有 pBI121-*GhCO* 重组质粒的单个菌落,于含卡那霉素的抗性 LB 培养基中培养,提取 pBI121-*GhCO* 重组质粒,并用限制性内切酶 *Sma*I 和 *Xba*I 双酶切,结果如图 7 所示。酶切后产生 2 条片段,其中小片段均与预期大小相等,说明植物过量表达载体 pBI121-*GhCO* 已成功构建。含此转录本的植物表达载体命名为 pBIGhCO。然后,pBIGhCO 经遗传转化到农杆菌中,用于下一步转基因烟草和棉花。

3 讨论

CO 是植物开花光周期反应过程中控制植物开花时间的一个关键基因^[6,9]。本实验通过筛选中棉所 36 均一化全长 cDNA 文库^[10],然后利用 RT-PCR 技术克隆到了 *GhCO* 的 cDNA 片段,这个基因编码的氨基酸序列分别与芒果、杨树、拟南芥 *CO* 蛋白序列同源性达到 70% 以上。这些基因均含有 2 类保守结构域,N 端的 B-box 型锌指结构域和靠近 C 端的 CCT 结构域(*CO*, *CO*-Like, *TOC1* 蛋白)。这也说明了植物 *CO* 类似基因在开花时间控制途径上的保守性^[6-7]。QRT-PCR 分析结果表明,*GhCO* 在长度为 1 cm 的蕾中表达量最



M: DNA 分子量标记 (100 bp DNA ladder); 1-5: pBI-*GhCO* 重组质粒 *Sma*I 和 *Xba*I 双酶切结果。

图 7 pBI-*GhCO* 重组质粒酶切琼脂糖凝胶电泳图

Fig. 7 PCR products in gel electrophoresis of pBI-*GhCO* recombinant colonies

高,在开花当天时 *GhCO* 表达量最高,这与前人的研究结果一致。说明 *GhCO* 的确参与了棉花蕾和花的发育,可能在开花途径中也处于非常重要的位置。根据本实验室研究发现,中棉所 36 在 2 片真叶展平时开始花芽分化,而 TM-1 在 3 片真叶展平时开始花芽分化(文章尚未发表)。*GhCO* 在苗期 10 SDs 表达量很高,在 20 SDs 表达量下降。而在 TM-1 中,10 SDs 时表达量很低,在 20 SDs 表达量升高,说明 *GhCO* 在早熟品种中棉所 36 中花芽分化前开花决定期(1 片真叶展平时)就已经开始高调表达,促进 2 片真叶展平时开始花芽分化。而在晚熟品种 TM-1 中 2 片真叶时高调表达,促进 3 片真叶展平时的花芽分化。说明 *GhCO* 在花芽分化形态出现以前就已经高调表达,而在花芽分化从形态上出现分化时表达量反而降低。由此推测,*GhCO* 可能在花芽分化前的开花决定期就已经高调表达,从而来促进花芽的分化。

CO 基因在不同物种中具有保守的锌指结构和核定位区域,但是不同植物中的作用机理并不完全相同。序列分析表明,该基因在被子植物与裸子植物之间、双子叶植物与单子叶植物之间以及不同科、属的植物之间均有明显分化^[11]。短日照植物成花机理有可能与长日照植物不同^[12-13]。总之,目前对 *CO* 基因的研究还多集中于对模式植物拟南芥和水稻的研究,而对其它经济作物还很少。虽然这些基因具有相当的保守性,但基因在一定程度上仍具有种属特异性。因此,构建了 pBIGhCO 过量表达载体,为进一步探讨 *GhCO* 的功能奠定了重要的基础。

参考文献:

- [1] GAMER W W, Allard H A. Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants[J]. Journal of Agricultural Research, 1920, 18: 553-606.
- [2] HAYAMA R, Coupland G. The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice[J]. Plant Physiology, 2004, 135: 677-684.
- [3] 胡根海,喻树迅. 利用改良的 CTAB 法提取棉花叶片总 RNA[J]. 棉花学报,2007,19(1):69-70.
- HU Gen-hai, Yu Shu-xun. Extraction of high-quality total RNA in cotton leaf with improved CTAB method[J]. Cotton Science, 2007, 19(1): 69-70.
- [4] 刘天明,胡银岗,张宏,等. 条锈菌诱导的抗锈小麦种质的基因表达分析[J]. 西北植物学报,2006,26(3):521-526.
- LIU Tian-ming, Hu Yin-gang, Zhang Hong, et al. *Puccinia striiformis* west induced gene expression of wheat germplasm with stripe rust resistance[J]. Acta Bot Boreal-Occident Sin, 2006, 26(3): 521-526.
- [5] 宋国琦,胡银岗,林凡云,等. YS 型小麦温敏雄性不育系 A3017 控温条件下的花粉育性比较[J]. 麦类作物学报,2006,26(1):17-20.
- SONG Guo-qi, Hu Yin-gang, Lin Fan-yun, et al. Comparison on the pollen fertility of YS type thermo sensitive male sterile wheat line A3017 under artificial temperature conditions [J]. Journal of Triticeae Crops, 2006, 26(1): 17-20.
- [6] ROBERT L S, Robson F, Sharpe A, et al. Conserved structure and function of the *Arabidopsis* flowering time gene *CONSTANS* in *Brassica napus*[J]. Plant Mol Biol, 1998, 37(5): 763-772.
- [7] LIU J, Yu J, Mcintosh L, et al. Isolation of a *CONSTANS* ortholog from *Pharbitis nil* and its role in flowering[J]. Plant Physiology, 2001, 125(4): 1821-1830.
- [8] YASUE N, Mayumi K, Takuichi F, et al. Characterization and functional analysis of three wheat genes with homology to the *CONSTANS* flowering time gene in transgenic rice[J]. Plant J, 2003, 36(1): 82-93.
- [9] HECHT V, Foucher F, Ferrandiz C, et al. Conservation of *Arabidopsis* flowering genes in model legumes[J]. Plant Physiology, 2005, 137(4): 1420-1431.
- [10] 吴东,刘俊杰,喻树迅,等. 中棉所 36 均一化全长 cDNA 文库的构建与鉴定[J]. 作物学报,2009,35(4):602-607.
- WU Dong, Liu Jun-jie, Yu Shu-xun, et al. Establishment and identification of a normalized full length cDNA library of CCRI 36[J]. Acta Agronomica Sinica, 2009, 35(4): 602-607.
- [11] 樊丽娜,邓海华,齐永文. 植物 *CO* 基因研究进展[J]. 西北植物学报,2008,28(6):1281-1287.
- FAN Li-na, Deng Hai-hua, Qi Yong-wen. Research advances in *CO* genes of plants[J]. Acta Bot Boreal-Occident Sin, 2008, 28(6): 1281-1287.
- [12] HAYAMA R, Yokoi S, Tamaki S, et al. Adaptation of photoperiodic control pathways short-day flowering in rice[J]. Nature, 2003, 422: 719-722.
- [13] RYOSUKE H, Bhavna A, Elisabeth L, et al. A circadian rhythm set by dusk determines the expression of FT homologs and the short-day photoperiodic flowering response in *Pharbitis* [J]. Plant Cell, 2007, 19: 2988-3000. ●