



## 陆地棉水通道蛋白 *GhNIP6.1* 基因的克隆及表达分析

巩元勇, 郭书巧, 束红梅, 倪万潮\*

(江苏省农业科学院经济作物研究所 / 农业部长江下游棉花和油菜重点实验室, 江苏 南京 210014)

**摘要:** 根据 EST 拼接的序列设计引物, 利用 RT-PCR 技术, 从陆地棉品种苏棉 18 中克隆获得了一个水通道蛋白基因 *GhNIP6.1*。该基因开放阅读框包含 903 个核苷酸, 编码 300 个氨基酸, 分子量为 30.97 kD, 理论等电点为 8.99。*GhNIP6.1* 蛋白的生物信息学分析表明:*GhNIP6.1* 含有 6 个跨膜区, 由 5 个环相连, 其中环 A、C 和 E 在细胞膜外, 环 B、D 和蛋白的 N、C 末端都位于细胞膜内, N 末端 79 个氨基酸, C 末端 18 个氨基酸, 具有 MIP 家族典型的保守氨基酸序列; 该蛋白的三级结构同拟南芥的 *AtNIP6.1* 非常相似, 亚细胞定位也同 *AtNIP6.1* 一致, 可能位于质膜中。进化分析发现, *GhNIP6.1* 蛋白同拟南芥的 *AtNIP6.1* 蛋白相似性最高。进一步扩增棉花核基因组获得了 2021 bp 的 DNA 序列, 它包含 5 个外显子和 4 个内含子, 所有外显子/内含子交接点都遵从 gt/ag 剪接规则。半定量分析表明, 该基因在根、茎和叶中都有表达, 其中真叶含量较高, 子叶中没有表达。

**关键词:** 陆地棉; 水通道蛋白; *GhNIP6.1*; 表达序列标签

**中图分类号:** S562.035.3 **文献标志码:** A

**文章编号:** 1002-7807(2013)01-0001-08

## Cloning of Aquaporin Gene *GhNIP6.1* from the Upland Cotton and Its Expression in Different Cotton Tissues

GONG Yuan-yong, GUO Shu-qiao, SHU Hong-mei, NI Wan-chao\*

(Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences / Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture, Nanjing, Jiangsu 210014, China)

**Abstract:** Primers were designed according to the assembled ESTs of *Gossypium hirsutum*, and an aquaporin gene *GhNIP6.1* was obtained from upland cotton Sumian 18 in this experiment by RT-PCR. The gene contains an ORF comprising of 903 nucleotides which encodes 300 amino acids with a predicted molecular mass of 30.97 kD and theoretical isoelectric point of 8.99. Bioinformatics analysis indicates the *GhNIP6.1* exhibited a typical structure with six trans-membrane domains which linked by five loops, loop A, C and E were at outside of membrane, loop B, D, N terminal and C terminal were at inside of membrane. N terminal has 79 amino acids and C terminal has 18 amino acids. The amino acid sequence of *GhNIP6.1* has the conserved peptides of MIP family. The result of predicted protein 3D structure of *GhNIP6.1* showed that it was highly similitude with that of *AtNIP6.1*. The subcellular localization of *GhNIP6.1* maybe localize at plasma membrane, which is consistent with *AtNIP6.1*. Phylogenic analysis displays *GhNIP6.1* has the highest similarity with *Arabidopsis thaliana* *AtNIP6.1* protein. A genomic DNA sequence with 2021 bp clone containing an aquaporin was isolated from the upland cotton in this experiment, and the *GhNIP6.1* gene contains eight introns and seven exons. All exon/intron junctions of the gene contain the gt/ag consensus splicing site. Semi-quantitative analysis of *GhNIP6.1* expression showed that the gene expressed extensively in cotton roots, stems, leaves, except cotyledon, and leaf is the most highly expressed tissue.

**Key words:** upland cotton; aquaporin; *GhNIP6.1*; EST

水通道蛋白 (Aquaporin, AQP), 又称水孔蛋白, 是一类广泛存在于动物、植物和微生物中的

转运水分子和某些小分子物质的跨膜蛋白。第一个水通道蛋白是美国 John Hopkins 大学的 Peter

收稿日期: 2012-06-08

作者简介: 巩元勇 (1982-), 男, 博士, 助理研究员, [gongyuanrong1982@163.com](mailto:gongyuanrong1982@163.com); \* 通讯作者, [niwanchao2002@yahoo.com.cn](mailto:niwanchao2002@yahoo.com.cn)

基金项目: 国家转基因重大专项 (2011ZX08005-001); 国家自然科学基金 (31101452); 江苏省农业科学院博士后科研资助计划 (012066511108); 江苏省自然科学基金 (BK2010465)

agre 研究小组于 1988 年从血红细胞和肾小管中分离纯化获得,根据该蛋白的分子量大小将其命名为 CHIP28(Channel-forming integral membrane protein 28 kD)<sup>[1-2]</sup>。将 CHIP28 蛋白在爪蟾卵母细胞中表达,鉴定出该蛋白是一种运输水分子的通道蛋白,CHIP28 蛋白因此被重新命名为 1 号水通道蛋白(Aquaporin1, AQP1)<sup>[3]</sup>。1992 年,第一个植物水通道蛋白是从拟南芥液泡膜上分离获得,被命名为 TIP(Tonoplast intrinsic protein)<sup>[4]</sup>。到目前为止,已经在拟南芥、玉米、烟草、水稻、棉花等多种植物中发现了 AQP。AQP 是古老的通道蛋白 MIP(Membrane intrinsic protein)家族成员之一,广泛存在于植物不同组织和器官以及不同发育时期<sup>[5]</sup>。根据结构和亚细胞定位,AQP 主要分为 4 种类型,即质膜内在蛋白(Plasma membrane intrinsic protein, PIP)、液泡膜内在蛋白(Tonoplast membrane intrinsic protein, TIP)、与大豆(*Glycine max*)根瘤菌周膜上水通道蛋白类 NOD26 膜内在蛋白(Nodulin 26-like intrinsic protein, NIP)和定位于内质网和质膜上的小分子碱性膜内在蛋白(Small and basic intrinsic protein, SIP)<sup>[6-8]</sup>。AQP 除具有运输水分子的功能外,还具有运输硅、硼等营养元素<sup>[7-9]</sup>和甘油、尿素、甲酰胺等中性分子<sup>[10]</sup>的功能,在维持植物细胞水分平衡和植物生长发育等生理过程中具有重要作用<sup>[10]</sup>。

通过对基因组和转录组的分析发现,拟南芥共有 35 条水通道蛋白,其中 PIP 类型最多,有 13 个成员,TIP 和 NIP 类型分别有 10 个和 9 个成员,SIP 类型最少,只包含 3 个成员<sup>[7]</sup>。运用生物信息学同源性搜索和分子克隆,在陆地棉中发现有 71 个水通道蛋白基因,PIP、TIP、NIP 和 SIP 类型的成员分别有 28 个、23 个、12 个和 7 个,此外还发现有 XIP(X intrinsic protein)型成员 1 个<sup>[11]</sup>。在棉花中存在如此众多的 AQP 基因,由此可见,AQP 必定在棉花的生长发育过程中担负着非常重要的生理功能。这些 AQP 基因,很多还没有获得全长序列,有些只有部分序列被克隆,如 *GhNIP6.1* 在 NCBI 中只提交了能够编码 280 个氨基酸的序列片段<sup>[11]</sup>。本研究在此基础上,通过对 *GhNIP6.1* EST 序列的搜索,克隆获得了该基因的 mRNA 全长序列,并初步研究了该基因的表达情

况,旨在为深入研究 *GhNIP6.1* 基因的功能提供研究基础。

## 1 材料和方法

### 1.1 试验材料

**1.1.1 供试植物品种及菌株。**陆地棉(*Gossypium hirsutum* L.)品种苏棉 18(Sumian18)由本实验室保存,*E.coli* Trans1-T1 Phage resistant 菌株购自北京全式金生物技术有限公司。

**1.1.2 试剂、酶和试剂盒。**一般化学试剂均购自南京化学试剂有限公司,为分析纯级别;*pEASY-T3* cloning kit 和 *TransTaq-T* DNA polymerase 是北京全式金生物技术有限公司产品;*TaKaRa* LA Taq<sup>®</sup>和 DNase I 是 *TaKaRa* 产品;BU-SuperScript RT kit(Biouniquer)和 DNA marker 购自南京天为生物科技有限公司;AxyPrep<sup>™</sup> DNA gel extraction kit 是爱思进生物技术(杭州)有限公司产品;引物合成和基因测序都在英潍捷基(上海)贸易有限公司(Invitrogen)完成。

### 1.2 试验方法

**1.2.1 植物培养。**经浓硫酸脱绒的光子在 9% 的双氧水浸泡 10~15 min 催芽,在直径 12 cm 的玻璃培养皿底部铺 3 层普通吸水纸,用水润湿,将种子均匀铺在吸水纸上(每皿约 20~30 粒种子),种子上面盖 2 层湿润的吸水纸,盖好培养皿盖,在培养箱(温度:27~29 °C;湿度:65%;光/暗:14 h/10 h)中发芽;待到种子胚芽露出时,将种子转入培养基质(营养土:蛭石=1:1)继续生长,子叶长到距离培养基质表面约 3 cm 时,将小苗移到 Hoagland 营养液中培养。每 3 d 换一次营养液,生长 15 d 左右,分别收取棉花的根、茎、子叶和真叶用于后续试验。

**1.2.2 总 RNA 提取及反转录合成 cDNA 第一链。**RNA 提取按照胡根海等<sup>[12]</sup>的方法,分别提取棉花的根、茎、子叶和真叶的 RNA,紫外分光光度计测定 OD<sub>260</sub> 及 OD<sub>280</sub> 值,分析 RNA 浓度与纯度;用 1.2% 琼脂糖电泳检测 RNA 的完整性,-80 °C 贮存备用。

cDNA 合成按照 BU-SuperScript RT KIT 说明书进行,在 PCR 反应管中依次加入 1 μg 总 RNA、5×BU-Script buffer 4 μL、10 mmol·L<sup>-1</sup> dNTP

mix 1  $\mu\text{L}$ 、10  $\text{mmol}\cdot\text{L}^{-1}$  Oligo (dt)18 Primer 1  $\mu\text{L}$ 、10  $\text{U}\cdot\mu\text{L}^{-1}$  RNase Inhibitor 1  $\mu\text{L}$ 、200  $\text{U}\cdot\mu\text{L}^{-1}$  BU-Script Reverse Transcriptase 1  $\mu\text{L}$  和 RNase-free  $\text{H}_2\text{O}$  至总体积 20  $\mu\text{L}$ , 在 PCR 仪中反应程序为:30  $^{\circ}\text{C}$  10 min、42  $^{\circ}\text{C}$  60 min、99  $^{\circ}\text{C}$  5 min 和 4  $^{\circ}\text{C}$  5 min。反应结束后稀释 5 倍用于后续试验。

**1.2.3 DNA 提取。** DNA 提取采用孙志栋等<sup>[13]</sup>改良的 CTAB 法,从新鲜的幼嫩叶片中提取基因组 DNA。

**1.2.4 电子克隆。** 根据同源基因的保守序列特性,以拟南芥 *AtNIP6.1* 基因 (NCBI 登录号:NM\_106724) 序列为模板,用 BLAST 检索程序检索 PlantGDB 数据库中 ([www.plantgdb.org](http://www.plantgdb.org)) 双子叶植物中已拼接的 (PUT) EST 序列。

**1.2.5 基因克隆。** *GhNIP6.1* 基因 cDNA 克隆:在检索获得的拼接的 (PUT) EST 序列的两端用 Primer premier 5.0 设计特异性引物,引物序列  $P_{\text{cDNA}}\text{-F}$ :CAAAAATGAAATAAACAACACTGA,  $P_{\text{cDNA}}\text{-R}$ :TTTTTCTCATCTTAATTTCAA。以 RNA 反转录产物为模板,用 LA Taq 进行 PCR 扩展,PCR 扩增程序,94  $^{\circ}\text{C}$  预变性 4 min 后,按以下程序进行:94  $^{\circ}\text{C}$  30 s,50  $^{\circ}\text{C}$  30 s,72  $^{\circ}\text{C}$  2 min,共 32 个循环。用 1% 琼脂糖凝胶电泳检测,切胶回收 PCR 产物,将回收产物通过“T-A”连接插入 *pEASY-T3* 载体,转化 *E.coli* Trans1-T1 phage resistant 菌株,在含有氨苄抗性的 LB 培养皿上筛选阳性单克隆菌落,送到 Invitrogen 公司测序。

*GhNIP6.1* 基因 DNA 克隆:根据获得的 *GhNIP6.1* 基因的 cDNA 序列,在开放阅读框两端用 Primer premier 5.0 设计特异性设计引物,引物序列为: $P_{\text{gDNA}}\text{-F}$ :ATGGAGAATGAGGATGTTCCATC,  $P_{\text{gDNA}}\text{-R}$ :TTACCTCCTGAAGCTCCTCACA。以提取的基因组 DNA 为模板,用 LA Taq 进行 PCR 扩展,PCR 扩增程序,94  $^{\circ}\text{C}$  预变性 4 min 后,按以下程序进行:94  $^{\circ}\text{C}$  30 s,60  $^{\circ}\text{C}$  30 s,72  $^{\circ}\text{C}$  2 min,共 32 个循环。后面的过程同 *GhNIP6.1* 基因的 cDNA 克隆。

**1.2.6 序列分析与比较。**用 ProtParam (<http://web.expasy.org/protparam/>) 对蛋白质进行理化性质分析;用 TMHMM-2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) 预测蛋白质的跨膜结构域;用

NetPhos 2.0 Server (<http://www.cbs.dtu.dk/services/NetPhos/>) 预测蛋白质翻译后的磷酸化修饰位点;用 Scratch Protein Predictor (<http://scratch.proteomics.ics.uci.edu/>) 预测蛋白质的二级结构;用 Swiss-Model (<http://swissmodel.expasy.org/>) 预测模拟蛋白质的三级结构,用 Molegro Molecular Viewer 查看蛋白质的 PDB 文件;用 NCBI 的 Blast P 对氨基酸序列进行蛋白保守区分析;用 PSORT (<http://psort.hgc.jp/>) 进行信号肽及亚细胞定位分析;用 MEGA4 进行蛋白质的进化树分析,用 DNAMAN 进行蛋白质的同源树分析。

**1.2.7 *GhNIP6.1* 基因在不同组织中表达分析。**在光照培养箱中营养液培养 15 d 的棉花幼苗,分别取新鲜的根、茎、子叶和真叶提取总 RNA,重复 3 次,每重复 3 株样本。根据克隆的 cDNA 序列设计半定量 PCR 引物; $P_{\text{RT-PCR}}\text{-F}$ :TACTCTCCACTGGCCACATC,  $P_{\text{RT-PCR}}\text{-R}$ :TGGCTTTTCACCATCTTCCTCA,扩增产物长度 515 bp;PCR 扩增程序,94  $^{\circ}\text{C}$  预变性 4 min 后,按以下程序进行:94  $^{\circ}\text{C}$  30 s,57  $^{\circ}\text{C}$  30 s,72  $^{\circ}\text{C}$  30 s,共 32 个循环。本研究采用的内参基因是棉花肌动蛋白 *Actin2*,用的是已经发表的引物<sup>[14]</sup>, $P_{\text{Actin}}\text{-F}$ :GGCATCACTCAGCACCTTC,  $P_{\text{Actin}}\text{-R}$ :TCCCTAACTTCCCGCACAT,扩增产物长度 280 bp;PCR 扩增程序,94  $^{\circ}\text{C}$  预变性 4 min 后,按以下程序进行:94  $^{\circ}\text{C}$  30 s,55  $^{\circ}\text{C}$  30 s,72  $^{\circ}\text{C}$  30 s,分别在 25、28、32 个循环后取 PCR 产物 1.5% 琼脂糖凝胶电泳检测。

## 2 结果与分析

### 2.1 棉花 *GhNIP6.1* 基因 cDNA 克隆及序列分析

使用拟南芥 *AtNIP6.1* (NCBI 登录号:NM\_106724) 基因序列,在 PlantGDB 双子叶 PUT 数据库中进行同源检索,得到 1 条来自陆地棉的 EST 拼接序列 (序列序号:PUT-165a-Gossypium\_hirsutum-83990)。该 ESTs 有 2 个 EST 序列拼接而成 (表 1),全长 1189 bp,与拟南芥 *AtNIP6.1* 基因开放阅读框序列同源性达 60%,翻译后的氨基酸序列一致性达到 80%。在检索获得的序列两端设计特异引物,用 RT-PCR 从苏棉 18 叶中扩增获得长度略大于 1000 bp 的片段 (图 1),切胶回收后,连接到 *pEasy-T3* 载体中,测序得

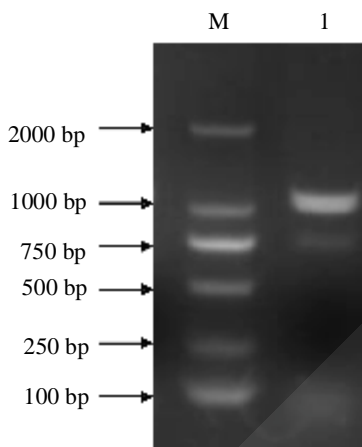
到 1180 bp DNA 序列(NCBI 登录号:JQ934895)。该序列同检索获得的序列相似度为 100%，表明 PCR 扩增获得的基因片段是本研究的的目的基因；只是由于引物设计的原因,试验获得的序列前后两端比 PlantGDB 检索到的序列少 4 bp 和 5 bp 的碱基。该序列包含长为 903 bp 的全部开放阅读

框,共编码 300 个氨基酸残基和 1 个终止密码子 TAA,起始密码子 ATG 后的核苷酸是 G,符合 kozak 规则。克隆得到的序列包含长度为 124 bp 和 154 bp 的 5' 非编码区和 3' 非编码区,并且 3' 非编码区含有 PolyA 尾,说明本研究克隆得到的基本是 *GhNIP6.1* 基因对应的 mRNA 全长序列。

表 1 *GhNIP6.1* 基因 EST 拼接序列分布

Table 1 Distribution of member ESTs among libraries of *GhNIP6.1* assembled unique transcript fragment

基因文库 Gene library	ESTs 编号 ESTs number
播种后 7~10 周根的文库 Root (free) - 7-10 weeks after planting	109836176
花后 0~10 d 的棉纤维文库 Cotton fiber 0-10 day post anthesis	164309692



M: Marker, D2000 DNA Ladder; 1: *GhNIP6.1* gene cDNA fragment.

图 1 棉花 *GhNIP6.1* 基因 RT-PCR 扩增片段

Fig. 1 Cotton *GhNIP6.1* gene by RT-PCR

## 2.2 棉花 *GhNIP6.1* 蛋白的生物信息学分析

从 *GhNIP6.1* 基因 cDNA 序列可推测 *GhNIP6.1* 蛋白有 300 个氨基酸残基构成,通过 Prot Param 预测该蛋白的基本理化性质表明,*GhNIP6.1* 的分子量是 30.97 kD,理论等电点为 8.99,负电荷的氨基酸残基数(Asp + Glu)是 17,正电荷的氨基酸残基数(Arg + Lys)是 22;脂肪系数为 93.03,总平均疏水性为 0.424,不稳定系数是 32.16,是稳定的蛋白。用 TMHMM-2.0 对 *GhNIP6.1* 蛋白跨膜结构域预测分析表明,该水通道蛋白具有 6 个跨膜螺旋(TM1-TM6),由 5 个环相连,其中蛋白的 N、C 末端和环 B、D 都在细胞膜内,环 A、C 和 E 位于细胞膜外,N 末端 79 个氨基酸,C 末端 18 个氨基酸(图 2)。在 NCBI 的蛋白保守区数据库(Conserved domain database, CDD)用 Blast P 对 *GhNIP6.1* 氨基酸序列进行蛋

白保守区分析,结果与该氨基酸序列匹配的是 MIP(Major intrinsic protein)超家族,具有 MIP 家族典型的 NPA(Asp-Pro-Ala 基序)保守肽段和两亲性通道(Amphiphilic channel),由此推测该基因应该是水通道蛋白基因家族的一员。

在 Scratch protein predictor 对 *GhNIP6.1* 蛋白的二级结构预测,组成该蛋白的 300 个氨基酸中,117 个氨基酸可能形成  $\alpha$  螺旋结构,150 个氨基酸可能形成无规则卷曲,33 个氨基酸可能形成  $\beta$  折叠片。用 NetPhos 2.0 Server 预测 *GhNIP6.1* 蛋白翻译后修饰的磷酸化位点,结果表明,有 11、43、48、52、71、293 和 297 位的 7 个 Ser(丝氨酸)和 12、18、73、248、294 位的 5 个 Thr(苏氨酸)可能发生翻译后的磷酸化修饰。PSORT 进行信号肽及亚细胞定位结果表明,*GhNIP6.1* 的氨基酸序列不存在信号肽,该蛋白最有可能定位在质膜上。用 Swiss-Model 中的自动搜索模式对 *GhNIP6.1* 和 *AtNIP6.1* 氨基酸三级结构进行预测,结果显示,*GhNIP6.1* 的模型残基位于 69~284 位,*AtNIP6.1* 的模型残基位于 75~290 位,三级结构预测的模板都是 3ne2C;用 Molegro molecular viewer 查看 *GhNIP6.1* 和 *AtNIP6.1* 的 PDB 文件,发现两者的结构非常相似(图 3)。

## 2.3 棉花 *GhNIP6.1* 蛋白的系统进化分析

本研究根据 *GhNIP6.1* 氨基酸序列在 NCBI 中 Blast P 的比对结果,选取已经公布 NIP6.1 序列的拟南芥(*Arabidopsis thaliana*, NP\_1781915)、大豆(*Glycin max*, XP\_003547292)、莲花(*Lotus japonicus*, ABY19374)、蒺藜苜蓿(*Medicago truncatula*, XP\_003604209)和葡萄(*Vitis vinifera*,

```

1   caaaaatgaataaacaactgatcaagtacagaaacttaagcctctctctcttcttactacataatttattactttaaaacttcctttcat
91  agtaattgttctgaattatacaagtgatatatggagaatgaggatgttccatcagctccttccaccctgtaacaccggggactcct
      M E N E D V P S A P S T P V T P G T P   19
181 ggtgctccactctttgggtgggttaaggagaccacagaggtgggtttaataaaaaaacctcttcaaaagctgcaaatgcttcagtggt
      G A P L F G G F K G D H R G C F N K K S L L K S C K C F S V   49
      N端 (1-79aa)
271 gaagattccatggaagaaggagattaccctctgtttctgtcattgccacctcctctgtctcactcacaagaagggtgggagctgag
      E D S M E E G R L P P V S C S L P P P P V S L T R K V G A E   79
361 ttcataggcacttttataactaactcttctgctggtagcaccactgctattgtgaaccaaaagaccgaaggctctgaaactctgattggcctg
      F I G T F I L I F A G T A T A I V N Q K T Q G S E T L I C L   109
      TM1 (80-97aa)           A环 (98-106aa)
451 gccgcctcaactggccttctgtcattgattgtcactctccactggccacatctcgggagcacatctcaatccagccctcaccattgcc
      A A S T G L A V M I V I L S T G H I S G A H L N P A V T I A   139
      TM2 (107-129aa)           B环 (130-149aa)
541 tttgctgactcaaacactttcttagaaacatgtaccggctctacatcgggacacaagtaatggcgtcattgtgtgctcatttgggcta
      F A A L K H F P R K H V P V Y I G A Q V M A S L C A A F G L   169
      TM3 (150-172aa)
631 aaaggagtgtttcatccgatgatgggaggggagtgacagtctctcagggggatttgggtcaagcttttctctggaattatcattagc
      K G V F H P M M G G G V I V P S G G F G Q A F A L E F I I S   199
      C环 (173-186aa)           TM4 (187-209aa)
721 ttcaatctcatgttctgctcactgctgggttccaccgacaccggagcagtggtgagttggcaggaatcgcggttggagctactgtcatg
      F N L M F V V T A V A I D I R A V G E L A G I A V G A T V M   229
      D环 (210-215aa)           TM5 (216-238aa)
811 ctcaatattctcatgcccggaccaattacaggtgcttctgatgaatccagtaagaacattggggcccgcattagcagcaataactacaaa
      L N I L I A G P I I G A S M N P V R T L G P A I A A N N Y K   259
      E环 (239-259aa)
901 gctatatgggtgtacttcacagccccatctcgggagcccttggcgggtgcaggaacctatactgcagtgagctgcctgaggaagatggt
      A I W V Y F T A P I L G A L C G A G T Y T A V K L P E E D G   289
      TM6 (260-282aa)
991 gaaaagccatcgactgtgaggagcttcaggaggtaatttagcatccattcaccgacctctattatcacaacttcttggaggagacattggct
      E K P S T V R S F R R *   300
      C端 (283-300aa)
1081 tctggatataaataagtgtactctatgtatgtatcaatcacggttacagaatagtcagtgatattgggggaaacaattgaaattaaga
1171 tgagaaaaaa

```

下划线:环;TM:跨膜螺旋;灰色:跨膜结构域。

Underline: Loop; TM: Trans-membrane helix; Gray: Trans-membrane domain.

图 2 棉花 *GhNIP6.1* 基因 cDNA 全长序列和推导出的氨基酸序列

Fig. 2 The full-length cDNA sequence and the deduced amino acid sequence of the *GhNIP6.1* for *Gossypium hirsutum*

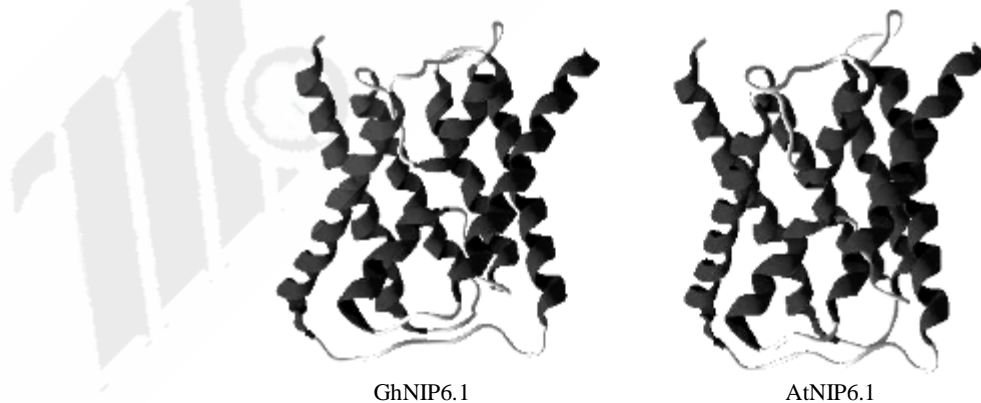


图 3 *GhNIP6.1* 和 *AtNIP6.1* 预测的三级结构的比较

Fig. 3 The comparison of the predicted 3-D structure of *GhNIP6.1* and *AtNIP6.1*

XP\_002272988) 等 5 种植物与棉花 *GhNIP6.1* 氨基酸序列构建系统发育树,进化树分析表明,*GhNIP6.1* 同拟南芥的 *NIP6.1* 蛋白在同一个进化分支,进化距离最近(图 4)。*GhNIP6.1* 氨基酸序列

同拟南芥 *NIP* 家族所有成员(共 9 个成员)的氨基酸序列同源性分析同样表明,*GhNIP6.1* 和 *AtNIP6.1* 的同源关系最近(图 5),所以我们将本研究克隆到的基因命名为 *GhNIP6.1*。

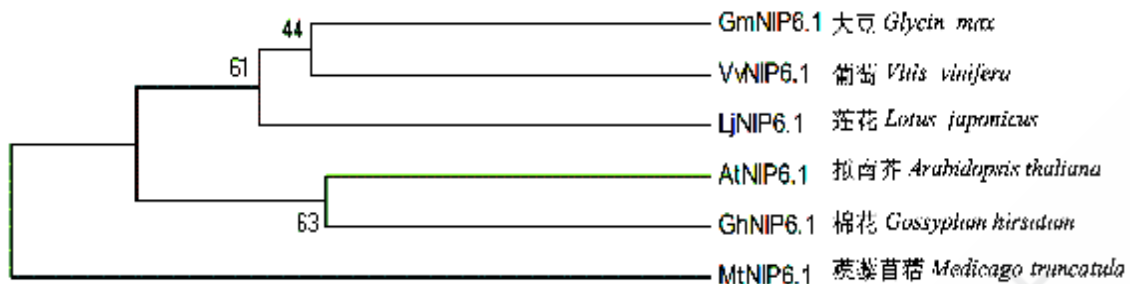


图 4 *GhNIP6.1* 基因与其它植物 *NIP6.1* 基因氨基酸进化树分析

Fig. 4 Phylogenetic tree analysis based on the amino acid of *GhNIP6.1* and other plant *NIP6.1* genes alignment

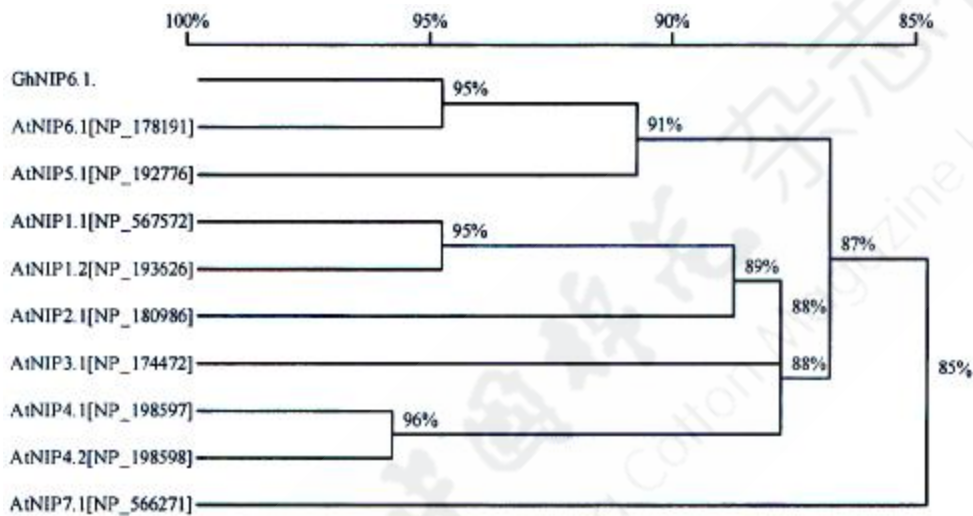


图 5 *GhNIP6.1* 基因与拟南芥 *NIP* 基因氨基酸同源性比对树状图

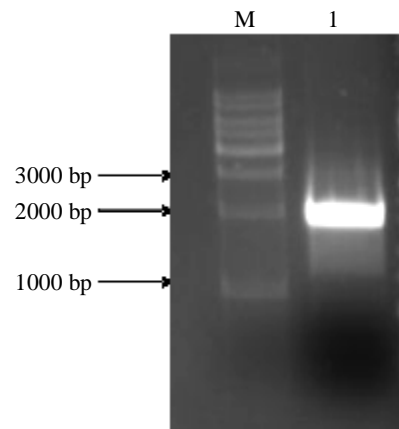
Fig. 5 Homology tree based on the amino acid of *GhNIP6.1* and *Arabidopsis thaliana* *NIP* genes alignment

#### 2.4 棉花 *GhNIP6.1* 基因 DNA 序列

根据获得的 *GhNIP6.1* 基因的 cDNA 序列, 在开放阅读框两端设计引物, 对苏棉 18 基因组进行 PCR 扩增, 得到 1 条约 2 kb 的序列片段(图 6)。切胶回收后, 连接到 *pEasy-T3* 载体中, 测序得到 2021 bp DNA 序列(NCBI 登录号: JQ934896)。对获得的 DNA 序列与 cDNA 序列对比发现, *GhNIP6.1* 基因编码区序列包含 5 个外显子和 4 个内含子(图 7), 所有内含子的左右边界均为 *gt/ag* 结构, 保证在 RNA 转录后加工过程中内含子能够被正确识别并切除(表 2)。

通过比较发现, 棉花 *GhNIP6.1* 基因和拟南芥 *AtNIP6.1* 基因包含相同数量的内含子和外显子(图 7)。*GhNIP6.1* 基因和 *AtNIP6.1* 基因除第一个和第五个外显子长度存在微小差异外, 其它 3 个外显子长度都相等, 造成基因组差异的原因是内含子的长短不同造成的。如 7 所示, *GhNIP6.1* 基因的 4 个内含子长度分别为 98 bp、82 bp、476 bp 和 462 bp, 而 *AtNIP6.1* 基因的 4 个内含子长

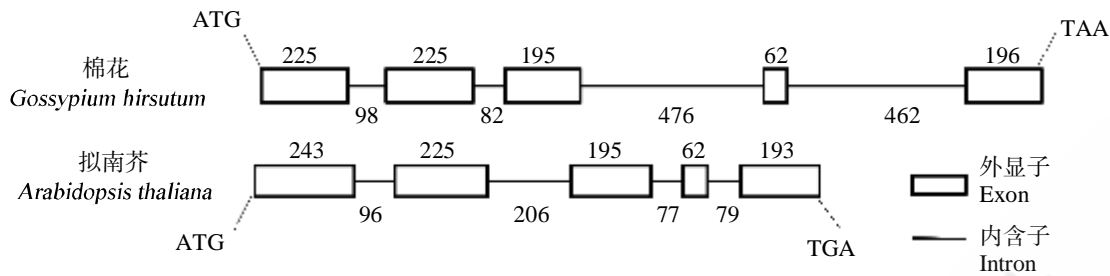
度则分别为 96 bp、206 bp、77 bp 和 79 bp, 这 2 个基因除了第一个内含子只相差 2 bp 以外, 其它 3 个内含子的长度差异都在 100 bp 以上, 差异最大的第三个内含子达到 399 bp。从外显子来看, 该基因的保守性非常高, 由此也表明棉花与拟南芥在进化上存在亲缘关系。



M: Marker, 1 Kb DNA Ladder; 1: *GhNIP6.1* gene DNA fragment.

图 6 棉花 *GhNIP6.1* 基因组 DNA PCR 扩增片段

Fig. 6 Cotton *GhNIP6.1* gene by PCR in genomic DNA

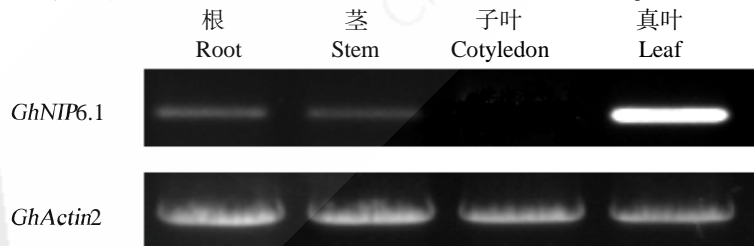
图 7 棉花 *GhNIP6.1* 基因与拟南芥 *AtNIP6.1* 基因结构Fig. 7 Genomic organization of the *NIP6.1* gene in cotton and *Arabidopsis*表 2 棉花 *GhNIP6.1* 基因外显子和内含子的拼接点Table 2 The intron/exon junctions of the *GhNIP6.1* gene

外显子及其长度 Exon and length of exon/bp	5'- 剪接供体端 5'- splice donor	内含子及其长度 Intron and length of intron/bp	3'- 剪接受体端 3'- splice acceptor	外显子 Exon
E1(225)	AGAAAG/gttaga	1(98)	ttgcag/GTGGGA	E2
E2(225)	AAACAT/gtaggg	2(82)	gggcag/GTACCG	E3
E3(195)	CGAGCA/gtaagt	3(476)	gtgcag/GTGGGT	E4
E4(62)	CGCCGG/gtatgt	4(462)	aaacag/ACCAAT	E5
E5(196)				

### 2.5 棉花发育过程 *GhNIP6.1* 基因的表达

用 RT-PCR 半定量来分析用陆地棉 *GhNIP6.1* 基因在棉花各组织器官中的表达情况。棉花长至 2 叶期时,分别从根、茎、子叶和真叶中提取总 RNA,用基因特异性引物( $P_{RT-PCR-F}$  和  $P_{RT-PCR-R}$ )所对应的 *GhNIP6.1* 基因序列来检测基因的特异性

表达,*GhActin2* 基因作为内参基因。从图 8 可以看出,*GhNIP6.1* 基因在棉花的根、茎和真叶中都有表达,在真叶中表达量最高,而在子叶中没有表达。该结果同 *GhNIP6.2* 基因的表达类似,*GhNIP6.2* 在幼根、成熟叶和茎中有表达,而在子叶中几乎没有表达<sup>[1]</sup>。

图 8 RT-PCR 半定量分析 *GhNIP6.1* 基因在苏棉 18 不同组织中的表达情况Fig. 8 Semi-quantitative analysis of *GhNIP6.1* gene expression in different tissues of Sumian 18

### 3 讨论

通过 EST 拼接序列及 RT-PCR,获得了全长 1180 bp 棉花 *GhNIP6.1* 基因的 cDNA 序列,该序列开放阅读框长 903 bp,编码 300 个氨基酸残基。该蛋白具有 6 个跨膜螺旋(TM1~TM6),由 5 个环相连,其中蛋白的 N、C 末端和环 B、D 都在细胞膜内,环 A、C 和 E 位于细胞膜外,B 环有一段 MIP 家族高度保守的 Asp-Pro-Ala 基序(NPA),这些都与 AQP 的基本结构相吻合<sup>[15]</sup>。

植物 AQP 的磷酸化是细胞信号转导与水分运输调节的媒介,通过磷酸化和去磷酸化作用,

可以快速、直接、可逆地实现植物 AQP 的门控调节<sup>[16]</sup>。*GhNIP6.1* 蛋白序列存在多个可能发生磷酸化修饰的位点,这些位点的存在都有可能直接或间接地参与该水通道蛋白对水分或其它物质运输的调节。

研究证实,拟南芥 *AtNIP5.1* 和 *AtNIP6.1* 都是硼酸通道蛋白。*AtNIP5.1* 定位于质膜,在根的表皮细胞、皮层细胞和内皮层细胞中表达<sup>[17]</sup>,拟南芥根伸长区和根毛区的 *AtNIP5.1* 基因在缺硼条件的表达受到强烈的上调;*AtNIP5.1* 基因缺失的突变体表现为根系对硼的吸收减少,生物量降低,根系和地上部枝条对缺硼的敏感性增加,由

此表明, *AtNIP5.1* 是拟南芥根部主要的硼酸通道蛋白, 对拟南芥根系在缺硼条件下对硼的有效吸收以及植株正常的生长发育至关重要<sup>[9]</sup>。*AtNIP6.1* 也定位于质膜, 主要在植株地上部结节部位, 尤其是维管组织的韧皮部表达, 主要负责硼酸在结节部位木质部到韧皮部的转移; 缺硼胁迫下 *AtNIP6.1* 基因在地上部表达增加, 在地上部幼嫩组织硼的合理分配过程中担负重要角色<sup>[18]</sup>。*GhNIP6.1* 和 *AtNIP6.1* 的氨基酸序列的同源性达到 95%, 两者蛋白质的三维结构也非常类似, 那么 *GhNIP6.1* 是否同 *AtNIP6.1* 一样也是硼酸通道蛋白呢? 这还需要深入的研究。根据推测的 *GhNIP6.1* 基因可能具有水分子和硼酸运输的生理功能分析, 该基因在子叶中没有或很少表达的原因可能是, 子叶只是为种子的萌发提供营养物质以及在植株发育苗期承担重要营养功能的过渡器官, 而不是水分子转运和硼酸运输活跃的部位, 尤其是在真叶长成之后。*GhNIP6.1* 基因在真叶中表达很高, 似乎也预示着该基因在地上部(尤其是真叶)可能发挥重要功能。

#### 参考文献:

- [1] DENKER B M, Smith B L, Kuhaida F P, et al. Identification, purification, and partial characterization of a novel Mr 28 000 integral membrane protein from erythrocytes and renal tubules[J]. *Journal of Biological Chemistry*, 1988, 263(30): 15634-15642.
- [2] PRESTON G M, Agre P. Isolation of the cDNA for erythrocyte integral membrane protein of 28 kilodaltons: member of an ancient channel family[J]. *Proceedings of the National Academy of Sciences*, 1991, 88(24): 11110-11114.
- [3] PRESTON G M, Carroll T P, Guggino W B, et al. Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein [J]. *Science*, 1992, 17(256): 385-387.
- [4] HÖFTE H, Hubbard L, Reizer J, et al. Vegetative and seed-specific forms of tonoplast intrinsic protein in the vacuolar membrane of *Arabidopsis thaliana* [J]. *Plant Physiology*, 1992, 99(2): 561-570.
- [5] BAIGE I, Schaffner A R, Affenzeller M J, et al. Plant aquaporins [J]. *Physiologia Plantarum*, 2002, 115 (2): 175-182.
- [6] CHAUMONT F, Barrieu F, Wojcik E, et al. Aquaporins constitute a large and highly divergent protein family in maize[J]. *Plant Physiology*, 2001, 125 (3): 1206-1215.
- [7] QUIGLEY F, Rosenberg J M, Shachar-Hill Y, et al. From genome to function: the *Arabidopsis* aquaporins[J]. *Genome Biology*, 2001, 3(1): 1-17.
- [8] CHANMONT F, Moshelion M, Daniels M J. Regulation of plant aquaporin activity[J]. *Biology Cell*, 2005, 97(10): 749-764.
- [9] TAKANO J, Wada M, Ludewig U, et al. The *Arabidopsis* major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation [J]. *Plant Cell*, 2006, 18(6): 1498-1509.
- [10] MAUREL C, Verdoucq L, Luu D T, et al. Plant aquaporins: membrane channels with multiple integrated functions [J]. *Annual Review of Plant Biology*, 2008(59): 595-624.
- [11] PARK W, Scheffler B E, Bauer P J, et al. Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.) [J]. *BMC Plant Biology*, 2010 (10): 142-158.
- [12] 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.
- [13] 孙志栋, 王学德, 倪西源, 等. 棉花 DNA 提取方法的探讨 [J]. *浙江农业学报*, 2004, 16(4): 177-181.  
SUN Zhi-dong, Wang Xue-de, Ni Xi-yuan, et al. Primarily study on isolation of genomic DNA from cotton [J]. *Acta Agricuturae Zhejiangensis*, 2004, 16(4): 177-181.
- [14] CHEN Jie-yin, Dai Xiao-feng. Cloning and characterization of the *Gossypium hirsutum* major latex protein gene and functional analysis in *Arabidopsis thaliana* [J]. *Planta*, 2010, 231 (4): 861-873.
- [15] FUIYOSHI Y, Mitsuoka K, de Groot B L, et al. Structure and function of water channels[J]. *Current Opinion in Structural Biology*, 2002, 12(4): 509-515.
- [16] CHANMONT F, Moshelion, Daniels M J. Regulation of plant aquaporin activity[J]. *Biology Cell*, 2005, 97(10): 749-764.
- [17] TAKANO J, Tanaka M, Toyoda A, et al. Polar localization and degradation of *Arabidopsis* boron transporters through distinct trafficking pathways [J]. *Proceedings of the National Academy of Sciences*, 2010, 107(11): 5220-5225.
- [18] TANAKA M, Wallace I S, Takano J, et al. NIP6;1 is a boric acid channel for preferential transport of boron to growing shoot tissues in *Arabidopsis* [J]. *Plant Cell*, 2008, 20 (10): 2860-2875. ●