

# 桑树肌动蛋白 *actin* 基因全长序列的克隆与分析

孔卫青, 杨金宏

(安康学院 陕西省蚕桑重点实验室, 陕西 安康 725000)

**摘要:** 肌动蛋白在植物的各种生理活动中起着重要作用,是研究基因表达与调控模式的内标参考。通过染色体步移方法获得了桑树肌动蛋白 *actin* 基因 1 612 bp 的序列,该基因 CDS 长 1 312 bp(GenBank 登录号: HM623866),编码 377 个氨基酸残基,与水稻、葡萄等的同源基因的序列一致性在 90% 以上。基因内含子的数目及其在基因组上的位置也与水稻、葡萄等的相似。对来自不同物种的 24 个肌动蛋白基因进行聚类分析的结果显示,基因被分为 Class I 和 Class II 两个明显的亚群。

**关键词:** 桑树; 肌动蛋白; 染色体步移; 克隆; 序列分析

中图分类号: Q943 文献标识码: A 文章编号: 1000-3142(2012)03-0362-05

## Cloning and sequence analysis of full length of mulberry (*Morus alba*) *actin* gene

KONG Wei-Qing, YANG Jin-Hong

(Key Sericultural Laboratory of Shaanxi, Ankang University, Ankang 725000, China)

**Abstract:** As a standard reference model of gene expression and regulation, actins play important roles in plants. In this study, the full length 1 612 bp of *actin* gene from mulberry was cloned by means of chromosome walking. Its CDS was 1 312 bp in length, encoding 377 amino acid residues. Homology analysis of the deduced amino acids showed above 90% identity with the actins from rice and grape, *et al.* The number and site of introns was conservative with rice and grape, too. Cluster analysis of actins from 24 species indicated that these genes were divided into two subgroups of Class I and Class II.

**Key words:** *Morus alba*; *actin*; chromosome walking; cloning; sequence analysis

植物肌动蛋白(Actin)广泛存在于植物界,是一种高度保守的蛋白质。由肌动蛋白单体聚合形成的微丝,参与了真核细胞骨架网状蛋白体系的形成,并赋予细胞强度以及对各细胞器及某些大分子的空间组织安排(Collings 等,2002; Vidali 等,2002)。研究发现,肌动蛋白在高等植物的花粉、茎韧皮部、叶表皮细胞、叶鞘细胞、根毛、卷须、内果皮和茎形成层等组织中均有表达(李园莉等,2002),肌动蛋白骨架体系可能与许多病毒的侵染、运输、复制以及子代病毒粒子的装配等过程有关(Phillip 等,2009)。阎隆

飞等(1963)首次提出高等植物中存在肌动蛋白的表达。目前已有多个植物如水稻(*Oryza sativa*)、拟南芥(*Arabidopsis thaliana*)、棉花(*Gossypium hirsutum*)、烟草(*Nicotiana tabacum*)等的 *actin* 基因被克隆和测序分析(McElroy 等,1990; McDowell 等,1996; Li 等,2005; Thangavelu 等,1993)。这些基因高度相似,可能由共同的祖先不断复制和分化遗传下来,且基因的表达也比较稳定,常被用作植物基因表达调控研究的目的基因在 mRNA 水平表达的内标基因(Thellin 等,1999)。

\* 收稿日期: 2011-10-31 修回日期: 2012-01-09

基金项目: 陕西省教育厅科技计划项目(09JS011,2010JK391)[Supported by Science and Technology Planning Project of Educational Commission of Shaanxi Province(09JS011,2010JK391)]

作者简介: 孔卫青(1980-),女,山东成武县人,博士,副教授,主要从事生物技术研究,(E-mail)weiqing\_kongwq@126.com。

桑树 (*Morus alba*) 不仅是蚕桑生产的主要物质基础, 还是一种重要的降血糖药用植物资源 (徐建飞等, 2010)。近年来, 已从分子生物学的基础上对桑树进行研究。肌动蛋白作为一个重要的基因, Gang 等 (2008) 报道了第一条桑树肌动蛋白序列的片段 (登录号: DQ785808.1), 本文在该序列片段的基础上, 采用热不对称性 PCR (thermal asymmetric interlaced PCR, TAIL-PCR) 方法获得桑树肌动蛋白 *actin* 基因的全长序列, 并分析了基因的结构, 为研究桑树基因表达与调控模式提供内标参考。

## 1 材料与方法

### 1.1 材料与主要试剂

选取桑树“沙 2×伦 109”十天生叶片。基因组提取试剂盒、Genome Walking Kit、EX taq 酶、胶回收试剂盒、pMD19-T simple vector 购自大连宝生物

工程有限公司。大肠杆菌 DH5 $\alpha$  购自上海生物工程有限公司, 质粒提取试剂盒购自北京天根。

### 1.2 方法

1.2.1 桑树基因组 DNA 的提取与内含子序列的 PCR 扩增 以基因组提取试剂盒提取的“沙 2×伦 109”桑树十天生叶片基因组为模板, 以 Primer premier 5.0 软件根据已有 *actin* 基因的序列设计的引物 AcF 和 5sp1 (表 1), 进行 PCR 扩增桑树 *actin* 基因位于该区域的第一、二内含子。扩增产物经 1.0% 琼脂糖凝胶电泳分析和胶回收纯化, 连接至 pMD19-T simple 载体, 并转化大肠杆菌 DH5 $\alpha$  感受态细胞, 阳性重组质粒送 invitrogen 公司测序。

1.2.2 基因组步移法获取桑树 *actin* 基因 3' 端和 5' 端序列 根据基因组步移试剂盒使用说明和本研究已获得的桑树 *actin* 基因的 DNA 序列, 设计引物 (表 1) 扩增基因的编码序列以及 5' 端序列和 3' 端序列。具体步骤为: 以桑树基因组 DNA 为模板, AP1

表 1 本研究使用的引物  
Table 1 Primers used in this study

引物 Primers	核酸序列 (5'-3') Nucleotide sequence (5' to 3')	引物 Primers	核酸序列 (5'-3') Nucleotide sequence (5' to 3')
AcF	GACAATGGAAGCTGGAATGGTG	3sp1	GCCAAGAGCGGTTCTCCGGTTGA
5sp1	GACTTGTGGGCATCGGTATCTCT	3sp2	GAGATACCGATGCCACAAGTCTT
5sp2	ACGACCACTGGCGTAAAGGGAGA	3sp3	TGTGCTCAGTGGTGGTTCAACTATG
5sp3	CATCCTTCTGACCATCCCAAC	ActinF	GCGATTGCTTTCTTTCATTCTC
5sp4	CGACCAACAATACTGGGAAAAGCTG	ActinR	ACTCGCCCTTGAAATCCACA

和 5sp2 进行第一次热不对称 PCR 反应, PCR 程序: 94 °C 1 m; 98 °C 1 m; 94 °C 30 s, 60 °C 1 m, 72 °C 2 m, 5 cycles; 94 °C 30 s, 25 °C 3 m, 72 °C 2 m 后, 进行 15 个循环的 94 °C 30 s, 60 °C 1 m, 72 °C 2 m, 2 cycles 和 94 °C 30 s, 44 °C 1 m, 72 °C 2 m 反应, 最后 72 °C 10 m。后 AP1、5sp3 引物与 AP1、5sp4 引物分别进行第二次和第三次 PCR 反应, 模板分别使用为稀释 100 倍后的第一次 PCR 反应液和第二次 PCR 反应液 1  $\mu$ L, 两次 PCR 反应程序与第一次 PCR 反应的后 15 个循环及其后的延伸程序相同。

同样分别以 3sp1、3sp2 和 3sp3 与 AP1 组合进行 3 次 PCR 反应, 扩增基因的 3' 端序列, 3 次 PCR 反应程序同前。分别切胶回收第三次 PCR 产物, 与 pMD19-T simple 载体连接, 转化大肠杆菌 DH5 $\alpha$  感受态细胞, 阳性重组质粒送 invitrogen 公司使用 M13 primers 进行测序。

1.2.3 桑树 *actin* 基因全长序列的获取与序列分析 根据获得的 3' 和 5' 端序列与已知核心序列推测

的编码框设计引物 ActinF 和 ActinR, 并以桑树基因组的 DNA 和 cDNA 为模板进行 PCR 扩增基因序列。Sim4 程序分析基因结构, Clustal X1.83 软件进行基因的多序列比对并对基因进化特征进行分析。

## 2 结果与分析

### 2.1 桑树 *actin* 基因序列的获得

琼脂糖凝胶电泳 AcF 和 5sp1 引物的 PCR 扩增产物, 结果在 1 000 bp 处出现一扩增条带, 回收纯化后进行克隆、测序 (图 1)。测序结果与原序列进行比对, 获得了 *actin* 基因的第一、二内含子。根据染色体步移试剂盒使用说明进行 TAIL-PCR 反应扩增基因的 5' 端序列, 扩增产物的 1.0% 琼脂糖凝胶电泳在约 360 bp 处有一扩增条带 (图 2), 回收该条带并进行克隆测序和序列拼接, 获得基因的 5' 端序列。同样分别以 3sp1、3sp2 和 3sp3 与 AP1 组合进行 PCR 反应, 获得了基因的第三个内含子和 3'

末端序列。

所有序列进行拼接,并根据拼接序列设计引物 ActinF 和 AcitnR,分别以桑树基因组 DNA 和桑树 RNA 反转录的 cDNA 为模板进行 PCR 扩增基因的 DNA 和 cDNA 序列,结果分别在约 1 400 bp 和 1 100 bp 处获得单一条带(图 3),测序该条带验证其确实为所得序列。

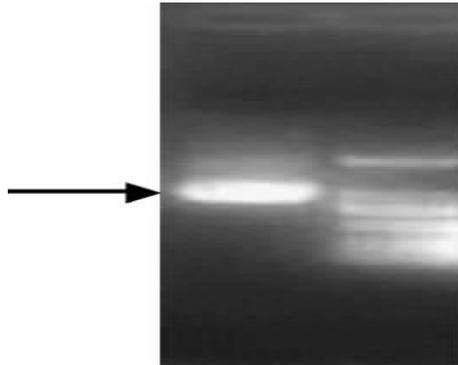


图 1 桑树 *actin* 基因第一、二内含子的 PCR 扩增(箭头所指为扩增条带)

Fig. 1 PCR amplification of 1<sup>st</sup> and 2<sup>nd</sup> intron of mulberry *actin* gene (The band is pointed by arrow)

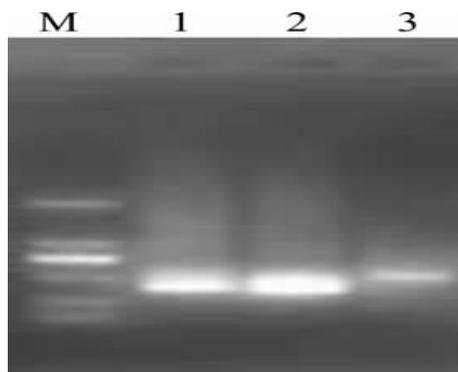


图 2 *actin* 基因 5' 末端的染色体步移扩增

Fig. 2 TAIL-PCR of the end of *actin* gene 5'

M. DL2000 分子标记; 1, 2, 3 分别为第三次、第二次和第一次扩增。

M. DL2000 marker; 1, 2, 3 are the 3rd, 2nd, and 1st PCR amplification, respectively.

## 2.2 桑树 *actin* 基因序列分析

桑树 *actin* 基因 cds 长 1 312 bp(图 4, GenBank 登录号: HM623866), 编码 377 个氨基酸, 预测蛋白分子质量为 41.6 kDa, 氨基酸序列中含有肌动蛋白家族保守的 ATP、Gelsolin 和 Profilin 结合位点。sim4 程序分析该基因有 4 个外显子, 外显子/内含子边界符合 GT-AG 规则。内含子分别位于距起始密码子 60、454 和 1 068 bp 处, 大小分别为 119、81 和 92 bp(图 4), 与其它物种水稻、葡萄(*Vitis vinif-*

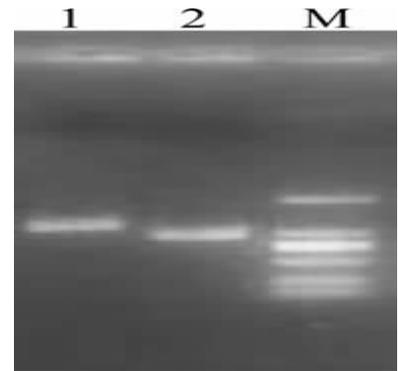


图 3 桑树 *actin* 基因的 PCR 扩增

Fig. 3 PCR amplification of mulberry *actin*

1. DNA 模板; 2. cDNA 模板; M. DL2000 分子标记。

1. DNA template; 2. cDNA template; M. DL2000 marker

*era*)、毛白杨(*Populus trichocarpa*)等所含内含子的数目、位置较为相似(图 5), 但序列差异较大。而基因的核苷酸、氨基酸序列与同源基因的多序列比对结果表明: 外显子部分核苷酸和氨基酸序列均有较高的相似性, 尤其氨基酸序列一致性在 90% 以上。

以酵母(*Saccharomyces cerevisiae*)肌动蛋白基因为外群, 对来自葡萄、毛白杨、玉米(*Zea mays*)、棉花和拟南芥等共 24 个肌动蛋白基因进行聚类分析(图 6), 结果所有肌动蛋白被分为两个亚群 Class I 和 Class II, 这与 Zhang 等(2010)所构建的结果相似, 说明肌动蛋白基因的分化可能发生在物种分化之前, 本研究获得的桑树 *actin* 基因归入 Class I, 与毛白杨的 ptact9(XP\_002316289)同源性最近, 这也与其形态学分类一致。

## 3 结论与讨论

肌动蛋白是生物体内重要的结构蛋白, 广泛存在于高等植物的各种组织细胞中, 参与许多重要的生命活动。高等植物的肌动蛋白由 375~377 个氨基酸组成, 具有高度的保守性和同源性, 这与肌动蛋白作为细胞骨架的保守功能是一致的(Liang 等, 2004)。与该结果一致, 本研究克隆的桑树肌动蛋白也由 377 个氨基酸组成, 序列中含有肌动蛋白家族保守的 ATP、Gelsolin 和 Profilin 结合位点, 与水稻、葡萄、毛白杨等的同源序列的一致性在 90% 以上。

在基因结构上, 本研究克隆的桑树 *actin* 基因由 4 个外显子和 3 个内含子组成, 内含子的两端序

```

1   TGGTTCGCTGATTATTGTCCTCCATTTTCATCTCCGCCGCTCCAGACAAAGCAAGAGTGCAGAGAACCAGGATTCTCTTGCTCT
91  TCTTCTTCCGATC GCGATTGCGTTTCTTCAATTC GAGGCTATAAAAGATGGCCGATGCAGAGGATATTCAGCCCTCGTCTGC GAC
      ActinF      M A D A E D I Q P L V C D
181 AATGGAAGCTGGAATGGTG AAGgtagtccctttcttctctgctccagtgaataagctcaatgttctattattgttaacgagtttgt
      N G T G M V K      I
      AcF
271 caattttagatgcaattttctaataaggaattttattatttctatatgcag GCTGGGTTTGTGGTATGATGCACCCAGGG CAGTTFITCC
      A G F A G D D A P R A V F
361 CAAGTATGTTGGTTCG TCCAGACACACTGGTGTATG GTTGGGATGGTTCAGAAGGATG CTTACGTAGGTGATGAAGCCCAATCCAAGA
      P S I V G R P R H T G V M V G M G Q K D A Y V G D E A Q S K
      5 sp4      5 sp3
451 GAGGTATCCTTACCTTGAAGTATCCCATTTGAACATGGTATTGTGCAGCAACTGGGATGACATGGAGAAGATTGGCCATCACACTTCTTACA
      R G I L T L K Y P I E H G I V S N W D D M E K I W H H T F Y
541 ACGAGCTTCGTGTGCTCCAGAAGAGCACCCAGTGTCTTCTACTGAGGCTCCTTTGAACCCCAAGGCTAACAGAGAGAAGATGACCCAGA
      N E L R V A P E E H P V L L T E A P L N P K A N R E K M T Q
631 TTATGTTTGGACATTCAATGTGCCTGCCATGTATGTGCCATCCAGGCCGTTT TCTCCCTTTACGCCAGTGGTCTGT ACAACTG gtttag
      I M F E T F N V P A M Y V A I Q A V L S L Y A S G R T T
      5 sp2
721 atccacaaccatttgaaattgtatttttggctacagctattgtttatcatttttgcataattggaaactttgaacag GTATTGTGTGGATT
      II
811 CTGGTGTGTTGTGAGTCACACCGTACCAATCTACGAAGTTATGCCCCTCCCATGCTATTCTCCGCTCCGACCTCGCTGGACGTGACC
      S G D G V S H T V P I Y E G Y A L P H A I L R L D L A G R D
901 TCACAGATGCTTTGATGAAGATTCTTACTGAGAGGTTACATGTTACCACCCTGCGGAACGGGAAATTTCCGTGACATGAAGGGGA
      L T D A L M K I L T E R G Y M F T T T A E R E I V R D M K G
991 AGCTGCTTATGTTGCCTTGGACTACGAGCAGGAACCTCGAGACT GCCAAGAGCGGTTCTCGGTTGA GAAGAACTATGAGCTGCCCGATG
      K L A Y V A L D Y E Q E L E T A K S G S S V E K N Y E L P D
      3 sp1
1081 GCCAAGTCATCACAATCGGAGCTG AGAGATACCCGATGCCCAAGTCT CTTCCAACCTCTCTCAATGGAATGGAAGCTGCAGGAATCC
      G Q V I T I G A E R Y R C P Q V L F Q P S L I G M E A A G I
      5 sp1 3sp2
1171 ACGAGACTACTTACAACCTATCATGAAGTGTGACGTGGATACAGAAAAGATCTCTATGGAAATAT TGTGCTCAGTGGTGGTTCACACTA
      H E T T Y N S I M K C D V D I R K D L Y G N I V L S G G S T
      3 sp3
1261 TGTTCCCTGGTATTGCAGACAGGATGAGCAAGGATACAGCCCTTGCTCCTAGCAGCATGAAGATTAAGGTTGTTGCTCCACCAGAGA
      M F P G I A D R M S K E I T A L A P S S M K I K V A P P E
1351 GAAAGTACAGTGTCTGGATTGGAGGGTCAATTCTTGCTTCCCTCAGTACCTTCCAGCAG gtaaaaaattcttctagataaacccaatta
      R K Y S V W I G G S I L A S L S T F Q Q
1441 tttattgaaccacttggcgaattgtaattctaataatgcttgtttctgttatgattcag ATGTTGATTCCAAGGCCGAGT ACGACGA
      III M W I S K G E Y D E
      ActinR
1531 GTCTGGTCCATCAATTGTCCACAGGAAGTCTTCTAAGTFTTATAAGTCTTGTATGGTGAATTTATTTTACAT
      S G P S I V H R K C F

```

图 4 桑树肌动蛋白基因及推导的氨基酸序列

Fig. 4 Gene and deduced amino acid sequence of mulberry *actin*

斜体部分 (I、II、III 核苷酸) 为序列内含子; 加粗部分为引物序列。

Italic nucleotides of I, II, and III indicate the introns; bold-faced nucleotides are the primer sequences.

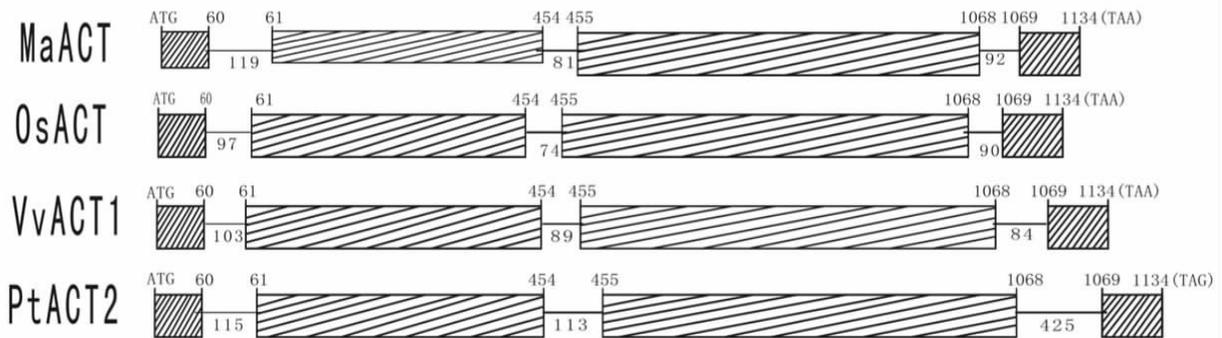


图 5 桑树与水稻、葡萄、毛白杨肌动蛋白基因所含内含子的数目、位置结构图

Fig. 5 Number and location map of *actin* gene from mulberry, rice, grape, and poplar

列符合典型的 GT-AG 法则,属于真核生物拼接位点很保守的第二类内含子。与水稻等的同源基因的内含子数量及其在基因的位置也极其相似,尤其第一个内含子通常都位于第 20 个密码子处,由此可见,肌动蛋白是生物进化中的一类高度保守的基因。

TAIL-PCR 是在 PCR 技术的基础上发展起来的染色体步移技术,用于扩增已知序列旁侧的未知 DNA 片段。其原理是通过高退火温度的长特异引

物和短的低退火温度简并引物的热不对称(高严谨性 PCR 和低严谨性 PCR 交替)循环程序,有效扩增特异产物,具有简单快速、特异性高等特点,近年来,成为分子生物学研究中非常实用的基因侧翼序列克隆技术(Qin 等,2003; Antal 等,2004)。但是该技术所获得的侧翼序列的长度存在很大的随机性,本研究利用该技术分别获取了桑树 *actin* 基因 5' 端和 3' 端均约 360 bp 左右的序列。本研究在已获得序列

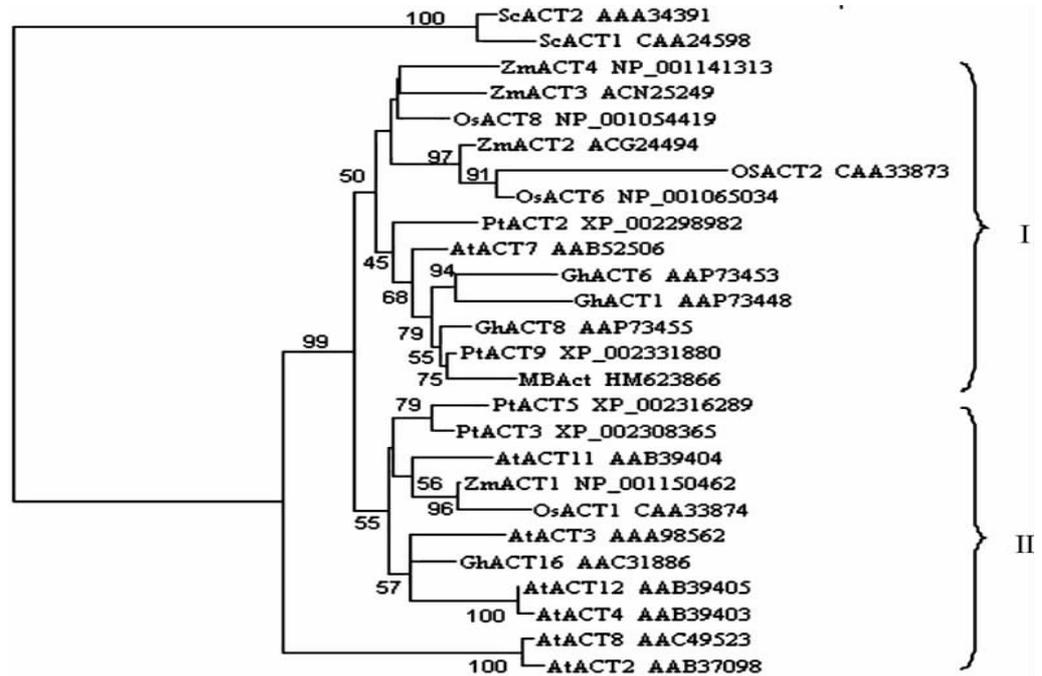


图 6 不同物种肌动蛋白基因聚类分析

Fig. 6 Clustering analysis of actin from different species

的基础上再次使用该技术以期进一步获得基因的启动子序列,未获得成功,说明在一个位置可能不能重复使用该技术进行进一步延伸。

#### 参考文献:

- Antal Z, Rasclé C, Fevre M, *et al.* 2004. Single oligo nucleotide nested PCR; a rapid method for the isolation of genes and their flanking regions from expressed sequence tags[J]. *Curt Genet*, **46**:240-246
- Collings DA, Harper JDI, Marc J, *et al.* 2002. Life in the fast lane; actin-based motility of plant peroxisomes[J]. *Can J Bot*, **80**:430-441
- Gang P, Cheng FL. 2008. Isolation of an L-aminocyclopropane-L-carboxylate oxidase gene from mulberry (*Morus alba*) and analysis of the function of this gene in plant development and stresses response[J]. *J Plant Physiol*, **165**:1 204-1 213
- Liang WH, Tang CR, Wu NH. 2004. Cloning and characterization of a new actin gene from *Oryza sativa* [J]. *Prog Nat Sci*, **14**:867-874
- Li XB, Fan XP, Wang XL, *et al.* 2005. The cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation[J]. *Plant Cell*, **17**:859-875
- Li YL(李园莉), Jiang YQ(江元清), Zhao WL(赵武玲), *et al.* 2002. Cloning and evolutionary analysis of actin gene from sunflower (*Helianthus annuus*) (向日葵肌动蛋白基因的 cDNA 克隆及进化分析)[J]. *J Agric Biotechnol* (农业生物技术学报), **10**(1):67-71
- McDowell JM, Huang SJ, McKinney EC, *et al.* 1996. Structure and evolution of the actin gene family in *Arabidopsis thaliana* [J]. *Genetics*, **142**:587-602
- McElroy D, Rothenberg M, Reece KS, *et al.* 1990. Characterization of the rice actin gene family[J]. *Plant Mol Biol*, **15**:257-268
- Phillip A, Harriesa, Park JW, *et al.* 2009. Differing requirements for actin and myosin by plant viruses for sustained intercellular movement[J]. *PANS*, **106**:17 594-17 599
- Qin GJ, Kang DM, Dong YY, *et al.* 2003. Obtaining and analysis of flanking sequences from T-DNA transformants of *Arabidopsis* [J]. *Plant Sci*, **165**:941-949
- Thangavelu M, Belostotsky D, Bevan MW, *et al.* 1993. Partial characterization of the *Nicotiana tabacum* actin gene family; evidence for pollen specific expression of one of the gene family members[J]. *Mol Gen Genet*, **240**:290-295
- Thellin O, Zorzi W, Lakaye B, *et al.* 1999. Housekeeping genes as internal standards; use and limits by [J]. *J Biotech*, **75**:291-295
- Vidali L, Hepler PK. 2002. Actin and pollen tube growth [J]. *Protoplasma*, **215**:64-76
- Xu JF(徐建飞), Yi XH(义祥辉), Chen QB(陈全斌). 2010. Effect of different chemical compositions from mulberry leaf on decreasing glucose in blood(桑叶不同化学成分的降血糖作用) [J]. *Guihaia* (广西植物), **30**(2):284-286
- Yang LF(阎隆飞), Shi DQ(石德权). 1963. contractile proteins in higher plant(高等植物中的收缩蛋白)[J]. *Acta Biochim Biophys Sin* (生物化学与生物物理学报), **3**:90-96
- Zhang DQ, Du QZ, Xu BH, *et al.* 2010. The actin multi-gene family in populus; organization, expression and phylogenetic analysis[J]. *Mol Genet Genom*, **284**:105-119