

红麻 Ms5 基因的克隆及多转录本分析

唐向民¹ 金 刚² 李 刚³ 周 琼¹ 陈 鹏¹ 周瑞阳¹

1. 广西大学农学院, 南宁 530005; 2. 广西亚热带作物研究所, 南宁 530001; 3. 广西药用植物园, 南宁 530023

摘要 为研究 Ms5 基因在红麻(*Hibiscus canabius* L.) 花药中的转录特征, 以 L23A 不育系为试材, 基于已知的红麻 Ms5 基因 5'端 cDNA 序列, 采用 3'RACE 和 RT-PCR 技术对该基因全长 cDNA 序列进行克隆。序列比对分析结果表明, 红麻花药组织中同时存在 Ms5 基因 2 种不同的转录本, 其长度分别为 972 和 1 105 bp, 依次命名为 HcMs5- α 和 HcMs5- β 。这 2 种转录本均包含了完整且一致的编码序列, 其编码序列(CDS)全长为 858 bp, 预测编码的蛋白质含有 285 个氨基酸残基, 分子质量为 32.4 ku, 理论等电点为 7.62; 两者 5'端前 972 bp 序列完全一致, 但 3'端的非编码区(UTR)长度有所不同, HcMs5- β 在第 972 bp 以后比 HcMs5- α 多出 133 bp 的序列, 具有选择性转录终止的特征。这种转录终止信号的可选择性可能与该基因转录后水平的调控有关。

关键词 红麻; Ms5 基因; 3'RACE; 多转录本; 选择性转录终止

中图分类号 S 563.5 **文献标识码** A **文章编号** 1000-2421(2012)05-0536-05

减数分裂是有性生殖过程中配子体发生的重要事件。目前, 在拟南芥中已鉴定到多个与小孢子减数分裂有关的基因, 如 MMD1、SYN1、ASK1、MEI1 等^[1-5]。Ms5 基因也是一个在高等植物中与减数分裂密切相关的基因^[6]。Ms5 基因在拟南芥基因组中至少包含以下 3 个拷贝: Ms5、Ms5-LIKE1 和 Ms5-LIKE2^[7]。Sanders 等^[7]研究发现, 拟南芥 Ms5 基因的 T-DNA 插入突变体表现为减数分裂四分体异常, 进而导致小孢子败育, 药室内细胞退化, 但绒毡层无异常; Ms5 基因仅仅在药室内的减数分裂细胞中表达, 且只短暂地出现在减数分裂后期。因此, 推测该基因编码蛋白可能在减数分裂期对小孢子的发育起到重要的周期性调控作用。

红麻具有丰富的种质资源和遗传多样性^[8]。李刚^[9]以红麻细胞质雄性不育系(CMS)L23A 及其保持系 L23B 为材料, 采用同位素相对标记与绝对定量(iTRAQ)技术, 通过二维液相色谱分离, 结合串联质谱, 共鉴定出 93 个在不育系和保持系花药线粒体中显著差异的蛋白质, 其中包括 1 个红麻 Ms5 蛋白。但该 Ms5 基因在红麻中的表达及功能尚不明确。本研究以红麻 CMS 系 L23A 为材料, 基于已知的红麻 Ms5 基因 cDNA 5'-UTR 序列设计 3'RACE

引物进行 PCR 扩增, 克隆该基因的全长开放阅读框(ORF), 揭示了其多转录本现象, 以期为该基因的表达调控和功能研究奠定基础。

1 材料与方法

1.1 供试材料

供试材料为笔者所在课题组选育的红麻细胞质雄性不育系(CMS)L23A, 于 2010 年 4 月种植于广西大学农学院教学科研试验基地。

1.2 红麻花药总 RNA 的提取和反转录

2010 年 10 月初剥取红麻花药, 由于红麻与松树一样含多糖、多酚等化合物, 多糖的理化性质与 RNA 相似, 在提取过程中, 易与 RNA 形成胶冻状物质共沉淀^[10-11]; 而酚类物质易被氧化成多醌类物质, 与 RNA 不可逆结合, 影响 RNA 产量和活性^[12]; 因此参照 Chang 等^[13]对松树 RNA 提取的方法, 采用 CTAB 法提取总 RNA; 用 Oligo(dT)接头引物(3'RACE-W)按照北京天恩泽基因科技公司 cDNA 第一链合成试剂盒进行反转录。

1.3 目的基因的 3'RACE 克隆与鉴定

在该基因已知的 cDNA 序列 5'端 UTR 区域, 设计 2 条基因特异引物(GSP1 和 GSP2)作为上游

收稿日期: 2011-09-23

基金项目: 现代农业产业技术体系建设专项资金(CARS-19-E16)

唐向民, 博士研究生, 研究方向: 作物遗传改良与杂种优势利用。E-mail: txm906@126.com

通讯作者: 周瑞阳, 博士, 教授。研究方向: 作物雄性不育与杂种优势利用。E-mail: ry_zhou@tom.com

引物,用 1 个含有部分接头序列的通用引物 UPM 作为下游引物,进行两轮 PCR 扩增。PCR 产物经 DNA 凝胶回收试剂盒纯化后克隆,送至华大基因公司测序。

表 1 本研究所用的引物
Table 1 Primer pairs used in this study

引物 Primers	序列 Primer sequences (5'-3')
3'RACE-W	GGCCACGCGTCGACTAG-TACTTTTTTTTTTTTTTTT
UPM	GGCCACGCGTCGACTAGTAC
GSP1	GGAGAGGAGAAAGCTGAGTTC
GSP2	AGAAAGCTGAGTTCCCGAGA

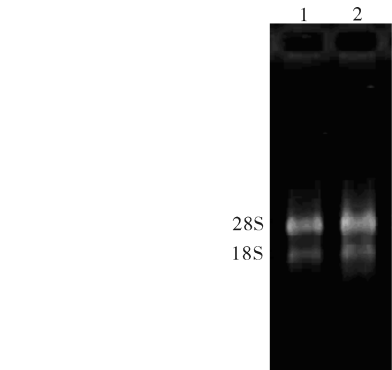
1.4 生物信息学和序列分析

利用 NCBI 的 Blast 和 ORF Finder 程序进行基因序列同源性和开放阅读框的预测;利用 <http://web.expasy.org/cgi-bin/protparam/protparam> 网站分析 *PNDREB1* 基因的氨基酸组成、理论分子质量和等电点;利用 <http://www.ncbi.nlm.nih.gov/structure/cdd/wrpsb.cgi> 进行基因编码蛋白质的功能域分析;利用 MEGA 4.0 进行该基因蛋白序列的同源比对,并采用邻接法构建进化树^[14]。

2 结果与分析

2.1 RNA 的提取与纯化

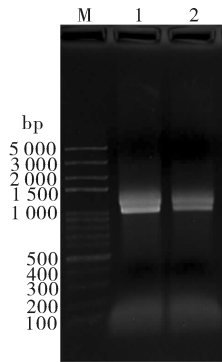
从图 1 可以看出用 CTAB 法提取的红麻花药总 RNA,28S rRNA 和 18S rRNA 条带清晰,5S rRNA 几乎看不到,且 28S rRNA 条带的亮度约为 18S rRNA 条带亮度的 2 倍,没有 DNA 与蛋白质污染,RNA 无降解,完整性好,纯度高,其质量可满足反转录为完整的 cDNA 的技术要求。



1,2:红麻花药总 RNA Total RNA extracted from kenaf anther.
图 1 CTAB 法提取红麻花药总 RNA 的琼脂糖凝胶电泳图
Fig.1 Agarose gel electrophoresis of total RNA extracted from kenaf anther with CTAB method

2.2 红麻 *Ms5* 基因的 3'RACE 克隆

以红麻花药 cDNA 第一链为模板,应用基于该基因 cDNA 5'-UTR 的 2 条特异引物与 UPM 组合进行两轮 PCR 扩增,获得了红麻 *Ms5* 2 条大小不同的条带,其大小分别为 972 和 1 105 bp(图 1)。2 种转录本均包含完整的开放阅读框,其长度为 858 bp,预测编码的蛋白质含有 285 个氨基酸残基,分子质量为 32.4 ku,理论等电点为 7.62;且所扩增的 5'-UTR 和 ORF 序列一致,预测的蛋白序列含有介导蛋白质互作的 TPR 结构域(图 2)。



M:100 bp DNA ladder plus; 1,2:3'RACE 扩增产物 Product of 3'RACE.

图 2 红麻 *Ms5* 基因 cDNA 的 3'RACE 扩增产物

Fig.2 3'RACE products of *Ms5* cDNA from *H. canabius*

2.3 红麻 *Ms5* 基因的 2 种转录本

经 *Ms5* 基因的 3'RACE 扩增,得到了 2 种不同的 *Ms5* 基因 cDNA 序列,分别命名为 *HcMs5-α* 和 *HcMs5-β*(图 3)。两者 5'端前 972 bp 序列完全一致,但 3'端的 UTR 长度有所不同,*HcMs5-β* 在第 972 bp 以后比 *HcMs5-α* 多出了 133 bp 的序列。在该基因 3'非编码区没有典型的加尾信号 AATA-AA,但具有 2 个可能的加尾信号,具有选择性转录终止的特征(图 4)。

2.4 红麻 *Ms5* 的分子进化分析

采用 MEGA4.0 软件将红麻和其他植物 *Ms5* 氨基酸序列作多重比对,构建系统进化树。结果表明:该基因在高等植物中是多拷贝的;高等植物 *Ms5* 蛋白可分为 3 个亚组。红麻 *Ms5* 所在的亚组 I 分支具有单子叶植物(如水稻 *Oryza sativa* 和玉米 *Zea mays*)和双子叶植物(如蓖麻 *Ricinus communis*、杨树 *Populus trichocarpa*、葡萄 *Vitis vinifera* 和拟南芥 *Arabidopsis thaliana*)的 *Ms5* 蛋白成员,说明该亚组 *Ms5* 蛋白的基本特征在单子叶和双子叶植物分离前就已形成(图 5)。

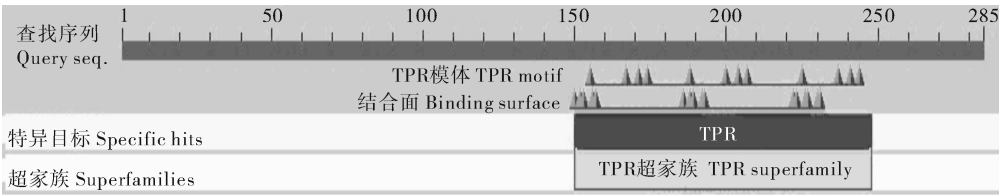


图 3 Ms 蛋白的保守域
Fig. 3 Conserved domains of Ms5

<i>HcMs5</i> - α	AGAAAGCTGAGTTC CGAGAAAAAGAAGAAAGGATGAGCAGCAAAGACCAACAGTTC CATGTTATT CATAAA	75
<i>HcMs5</i> - β	AGAAAGCTGAGTTC CAAGAAAAAGAAGAAAGGATGAGCAGCAAAGACCAACAGTTC CATGTTATT CATAAA	75
<i>HcMs5</i> - α	TCCTCAAGGTCAATTCTCCTTATGTTAAAGCAAACATGCTCAGGTGGTTGAAAGGATCCAGAAGCAGCAATTGT	150
<i>HcMs5</i> - β	TCCTCAAGGTCAATTCTCCTTATGTTAAAGCAAACATGCTCAGGTGGTTGAAAGGATCCAGAAGCAGCAATTGT	150
<i>HcMs5</i> - α	GCTGTTTTGGAAGGCTATCAATGCGGGGATAGGGTAGAGAGCGCTCTCAAAGACATGGCGGTGGTGATGAAGCA	225
<i>HcMs5</i> - β	GCTGTTTTGGAAGGCTATCAATGCGGGGATAGGGTAGAGAGCGCTCTCAAAGACATGGCGGTGGTGATGAAGCA	225
<i>HcMs5</i> - α	GCTCAACCGCTCTGAAGAAGCATTGAGGCCCTCTACTCTTTCAGGGGCTTTGCCCAAACAATCTCAGGAATC	300
<i>HcMs5</i> - β	GCTCAACCGCTCTGAAGAAGCATTGAGGCCCTCTACTCTTTCAGGGGCTTTGCCCAAACAATCTCAGGAATC	300
<i>HcMs5</i> - α	ACTTGACAATCTTCTCTTTCGACTTATACAAGAAATGTGGGAGAGTAGACGAGCAGATTGACTTGATAAAAAGAAA	375
<i>HcMs5</i> - β	ACTTGACAATCTTCTCTTTCGACTTATACAAGAAATGTGGGAGAGTAGACGAGCAGATTGACTTGATAAAAAGAAA	375
<i>HcMs5</i> - α	GCTAAGATTAAATATACAGGGAGAAATCTTCAATGGAAAACCCACCAAGACTGCACGCTCCCATGGCAAGAAGTT	450
<i>HcMs5</i> - β	GCTAAGATTAAATATACAGGGAGAAATCTTCAATGGAAAACCCACCAAGACTGCACGCTCCCATGGCAAGAAGTT	450
<i>HcMs5</i> - α	TCAAGTTTCAGTTCAGCAAGAACTTCAAGATTACTGGGAAATTTGGGGTGGGCTTACATTAAAGAAATCAAACTA	525
<i>HcMs5</i> - β	TCAAGTTTCAGTTCAGCAAGAACTTCAAGATTACTGGGAAATTTGGGGTGGGCTTACATTAAAGAAATCAAACTA	525
<i>HcMs5</i> - α	CTTGACAGCCGAGGTGGTGTACCGAAAAGCACAGATGATCGACCCGGATGCCAACAGGCCTGCAACCTGGGTTT	600
<i>HcMs5</i> - β	CTTGACAGCCGAGGTGGTGTACCGAAAAGCACAGATGATCGACCCGGATGCCAACAGGCCTGCAACCTGGGTTT	600
<i>HcMs5</i> - α	TTGCTTGATTAAACAAAGGCAGATACGATGAGGCACGTAGAGTGCTGGAAGAAACGCTGAGAGGAGAAGTCCAGG	675
<i>HcMs5</i> - β	TTGCTTGATTAAACAAAGGCAGATACGATGAGGCACGTAGAGTGCTGGAAGAAACGCTGAGAGGAGAAGTCCAGG	675
<i>HcMs5</i> - α	TTCGGAGGACATTAGAGCCAGGAACAGAGCTGCGGAATTGTTGATGGAAATCAAAGCATTTGAGCCAGCAGAGTC	750
<i>HcMs5</i> - β	TTCGGAGGACATTAGAGCCAGGAACAGAGCTGCGGAATTGTTGATGGAAATCAAAGCATTTGAGCCAGCAGAGTC	750
<i>HcMs5</i> - α	GGAAATAGCAGATATTCTGGGCCCTTGATGATGATTTCTGCTACTGGGCTTGAACTACTGATGAATGAATGGGCACC	825
<i>HcMs5</i> - β	GGAAATAGCAGATATTCTGGGCCCTTGATGATGATTTCTGCTACTGGGCTTGAACTACTGATGAATGAATGGGCACC	825
<i>HcMs5</i> - α	GCCTAGATCTAAGAGATTACCATTTTGAAGAGATCTCTCTTTTAGAGATCAATTAGCTTGCCTAATATTA	900
<i>HcMs5</i> - β	GCCTAGATCTAAGAGATTACCATTTTGAAGAGATCTCTCTTTTAGAGATCAATTAGCTTGCCTAATATTA	900
<i>HcMs5</i> - α	ATTATAGGGATAATGGCATGAGTGTCCTGTACTTTACGAAATTTTCAATGTGATGTATATATTTTCTAAA...	972
<i>HcMs5</i> - β	ATTATAGGGATAATGGCATGAGTGTCCTGTACTTTACGAAATTTTCAATGTGATGTATATATTTTCTAAACGT	975
<i>HcMs5</i> - α	972
<i>HcMs5</i> - β	TTATCGACCCCTTATACCTTTCTTCTGTTATACAGGATGTCTCCATAACACTGTTAGTTGGTAACGTGGTAGA	1050
<i>HcMs5</i> - α	972
<i>HcMs5</i> - β	TACATGTGACACACTGTTATGTGTTAATTTATTTCTAGAAAAGTCAC	1098

方框部分为起始密码子 ATG 和终止密码子 TAA Two boxes indicate ATG and TAA.

图 4 *HcMs5*- α 和 *HcMs5*- β 核苷酸序列比对
Fig. 4 DNA sequence of *HcMs5*- α and *HcMs5*- β

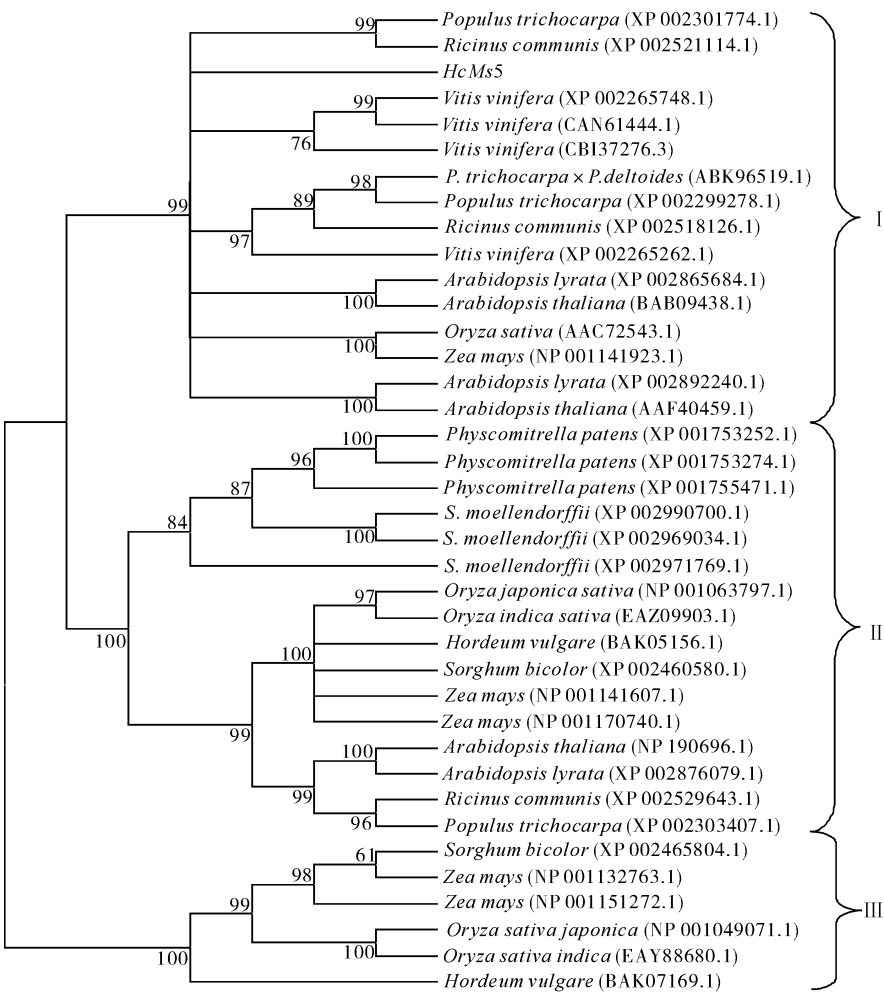


图 5 红麻 *Ms5* 与其他植物的同源蛋白的系统进化树

Fig. 5 Phylogenetic tree based on the amino acid sequence of Ms5

3 讨 论

本研究所克隆的红麻 *Ms5* 基因是一个与雄性不育相关的基因^[7,9]。本研究通过 3'RACE 克隆该基因 cDNA 序列,发现 2 种转录本 3'UTR 的长度不一样,暗示了该基因的表达具有选择性转录终止的特征。

5'UTR 和 3'UTR 通常含有涉及基因转录和转录后水平调控的序列。选择性转录起始和终止增加了转录本的多样性和复杂性^[15]。多聚腺苷酸化是真核生物成熟 mRNA 产生过程中的重要环节。3'UTR 被认为是基因表达所必需的一个调控区,其序列和长度是可变的;3'UTR 的顺式作用元件通过与反式作用因子的互作,对 mRNA 稳定性、基因表达定位以及翻译效率等过程进行调控^[16-17]。例如,戴冰冰等^[18]报道了 *PI3K γ* mRNA 3'UTR 可能存在基因表达负调控区。可变的聚腺苷酸位点还可能与基因表达的组织特异性有关^[16-17]。因此,选

择性的多聚腺苷酸位点可能也是基因表达调控的重要机制。我们推测,红麻 *Ms5* 基因的 1 种转录本选择性终止后可能通过转录后的调控机制影响到其稳定性,从而在蛋白水平上表现 *Ms5* 量上的差异,影响到花药的发育^[9]。但 *HcMs5- α* 和 *HcMs5- β* 哪一种更不稳定? 同时,*Ms5* 基因的选择性转录终止是否在高等植物中具有普遍性? 这种选择性终止会如何影响该基因的表达、翻译和组织特异性? 由于 *Ms5* 基因在高等植物中存在多个拷贝,因此 *Ms5* 基因在红麻中的拷贝数以及它们所执行的生物学功能等问题还有待于进一步的研究。

参 考 文 献

[1] YANG X H, MAKAROFF C A, MA H. The *Arabidopsis* MAE MEIOCYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis[J]. Plant Cell, 2003, 15: 1281-1295.
[2] CAI X, DONG F, EDELMANN R E, et al. The *Arabidopsis*

SYN1 cohesin protein is required for sister chromatid arm cohesion and homologous chromosome pairing [J]. Journal of Cell Science, 2003, 116(14): 2999-3007.

[3] YANG M, HU Y, LODHI M H, et al. The *Arabidopsis* SKP1-LIKE1 gene is essential for male meiosis and may control homologue separation[J]. Proceedings of the National Academy of Sciences USA, 1999, 96: 11416-11421.

[4] YANG X H, TIMOFEJEVA L, MA H, et al. The *Arabidopsis* SKP1 homolog ASK1 controls meiotic chromosome remodeling and release of chromatin from the nuclear membrane and nucleolus [J]. Journal of Cell Science, 2006, 119: 3754-3763.

[5] HE C, MASCARENHAS J P. *MEI1*, an *Arabidopsis* gene required for male meiosis: isolation and characterization [J]. Sex Plant Reprod, 1998, 10: 199-207.

[6] GLOVER J, GRELON M, CRAIG S, et al. Cloning and characterization of *Ms5* from *Arabidopsis*: a gene critical in male meiosis [J]. The Plant Journal, 1998, 15: 345-356.

[7] SANDERS P M, BUI A Q, MCINTIRE K W K, et al. Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants [J]. Sex Plant Reprod, 1999, 11: 297-322.

[8] 郑海燕, 栗建光, 戴志刚, 等. 红麻种质资源 SRAP 指纹图谱构建及遗传多样性分析[J]. 湖北农业科学, 2011, 50(2): 411-416.

[9] 李刚. 红麻细胞质雄性不育系线粒体蛋白质组差异分析与不育相关基因的发掘[D]. 南宁: 广西大学农学院, 2009.

[10] LOGEMANN J, SCHELL J, WILLMITZER L. Improved method for isolation of RNA from plant tissues[J]. Analytical Biochemistry, 1987, 163(1): 16-20.

[11] 涂国英, 漆艳香, 蒲金基, 等. 改良 CTAB 法提取高质量香蕉叶片总 RNA[J]. 广东农业科学, 2009(7): 192-195.

[12] 刘晓菊, 洪海波, 李敏, 等. 改良 CTAB 法提取核桃总 RNA 试验[J]. 山东农业科学, 2008(1): 97-99.

[13] CHANG S, PURYEAR J, CAIRNEY J. A simple and efficient method for isolating RNA from pine trees[J]. Plant Molecular Biology Reporter, 1993, 11: 113-116.

[14] TAMURA K, DUDLEY J, NEI M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0 [J]. Molecular Biology And Evolution, 2007, 24: 1596-1599.

[15] SHABALINA S A, SPIRIDONOV A N, SPIRIDONOV N A, et al. Connections between alternative transcription and alternative splicing in mammals[J]. Genome Biology And Evolution, 2010, 2: 791-799.

[16] 陈淑华, 曾卫民, 宋惠萍. 真核生物 mRNA 3'非翻译区的调控功能[J]. 国外医学: 生理、病理科学与临床分册, 2003, 23(6): 611-614.

[17] 李明, 赵巧辉, 陈其新, 等. 家兔 BMP7 基因的克隆及其生物信息学分析[J]. 遗传, 2008, 30(7): 885-892.

[18] 戴冰冰, 卢健, 王家敏, 等. P13K mRNA 3'非翻译区可能存在基因表达负调控区[J]. 中国生物化学与分子生物学报, 2003, 19(5): 630-635.

Cloning and multiple transcript analysis of *Ms5* gene in kenaf(*Hibiscus canabius* L.)

TANG Xiang-min¹ JIN Gang² LI Gang³ ZHOU Qiong¹ CHEN Peng¹ ZHOU Rui-yang¹

1. College of Agronomy, Guangxi University, Nanning 530005, China;
2. Guangxi Subtropical Crops Research Institute, Nanning 530001, China;
3. Guangxi Botanical Garden of Medicinal Plants, Nanning 530023, China

Abstract The full mRNA of kenaf *Ms5* gene was tried to clone with RT-PCR and 3'RACE methods based on the known partial 5'-end cDNA sequence. The sequencing results showed that two transcribed cDNAs of *Ms5* gene were cloned from the anther of kenaf. The length of two transcripts named as *Hc-Ms5-α* and *HcMs5-β* were 972 bp and 1 105 bp, respectively. The structure analysis showed that both of them contained an 858 bp ORF, encoding a protein of 285 amino acids with predicted molecular weight of 32.4 ku and a isoelectric point of 7.62. The sequencing results showed that 5'-end 972 bp sequence of the two transcripts was identical, but *HcMs5-β* possessed an additional 133 bp sequence at 3'-end after 972 bp compared with *HcMs5-α*. The alternative termination of transcription in the 3'UTR of kenaf *Ms5* maybe related with the post transcriptional regulation of the gene. It is the first report about the multiple transcripts of *Ms5* gene in higher plant.

Key words kenaf(*Hibiscus canabius* L.); *Ms5* gene; 3'RACE; multiple transcript; alternative termination

(责任编辑: 张志钰)