

甘蔗 *MOC1* 基因 (*ScMOC1*) 的克隆与表达分析

李旭娟, 李纯佳, 徐超华, 刘洪博, 吴转娣, 林秀琴, 陆鑫, 毛钧, 字秋艳, 刘新龙

(云南省农业科学院甘蔗研究所/云南省甘蔗遗传改良重点实验室, 开远 661699)

摘要: *MONOCULM 1 (MOC1)* 基因在植物腋分生组织和腋芽的形成中发挥重要作用, 是植物分蘖关键调控基因。本研究利用同源基因克隆法结合 RT-PCR、RACE 技术从甘蔗品种 ROC22 中克隆获得 *MOC1* 的同源基因, 命名为 *ScMOC1*。生物信息学分析发现该基因的 cDNA 序列包含一个长度为 1299 bp 的开放阅读框, 编码 432 个氨基酸残基组成的含有 GRAS 保守结构域的非分泌蛋白, 其分子量为 45.43 kD, 理化等电点 pI 为 6.98。序列比对分析显示 *ScMOC1* 与高粱 (*Sorghum bicolor* (L.) Moench)、赖草 (*Leymus secalinus* (Georgi) Tzvel.)、水稻 (*Oryza sativa* L.) 等禾本科植物同源蛋白氨基酸序列一致性较高; 系统进化树分析显示其与高粱、小米草 (*Setaria italica* (L.) P. Beauv.)、玉米 (*Zea mays* L.) 等禾本科植物 *MOC1* 同源蛋白亲缘关系最近; *ScMOC1* 在甘蔗品种 ROC22 中的序列变异分析发现, 30 个克隆得到的序列中共有 46 个 SNP 位点和 29 处 InDel 位点, 其中 1 个位点发生的单个碱基缺失和另一个位点的 4 个碱基插入是造成基因编码蛋白序列变化的主要原因。中性检测表明, *ScMOC1* 在 ROC22 中遵循中性进化模型。采用实时荧光定量 PCR 分析 *ScMOC1* 在甘蔗品种 ROC22 分蘖期不同组织部位 (根、茎、叶、分蘖芽、叶鞘、生长点)、茎尖生长点和不同发育阶段腋芽 (幼嫩腋芽、半大腋芽、较大腋芽、成熟休眠腋芽) 的表达特征, 结果显示 *ScMOC1* 在分蘖期的 ROC22 中的表达具有组织特异性, 在生命活跃的茎尖生长点处表达量最高; 在腋芽形成发育过程中该基因表达总体呈现出“升-降-升-降”的趋势, 说明 *ScMOC1* 基因可能在甘蔗腋芽形成发育阶段中发挥作用。以上研究可推测 *ScMOC1* 在甘蔗的分蘖性状调控上扮演重要的角色。本研究为 *ScMOC1* 的功能研究及其在甘蔗产量分子辅助育种中的利用奠定基础。

关键词: 甘蔗; 分蘖; *MOC1* 基因; 序列变异分析; 表达分析

Cloning and Expression Analysis of the *MOC1* Gene (*ScMOC1*) in Sugarcane (*Saccharum officinarum*)

LI Xu-juan, LI Chun-jia, XU Chao-hua, LIU Hong-bo, WU Zhuan-di, LIN Xiu-qin,

LU Xin, MAO Jun, ZI Qiu-yan, LIU Xin-long

(Sugarcane Research Institute, Yunnan Academy of Agricultural Sciences/Yunnan

Key Laboratory of Sugarcane Genetic Improvement, Kaiyuan 661699)

Abstract: *MONOCULM 1 (MOC1)*, a gene that is important in the regulation of plant tillering, which plays a significant role in the formation of plant axillary meristem and axillary bud. In this study, the homolog of *MOC1* which named *ScMOC1* was cloned from sugarcane variety ROC22 by a series of technologies including homologous gene cloning method, reverse transcription PCR (RT-PCR) and rapid amplification of cDNA ends (RACE). Bioinformatics analysis showed that the cDNA of *ScMOC1* contained a complete open reading frame with a length of 1299 bp and encoded a secreted protein containing a GRAS conservative structure domain. The protein was predicted to encode 432 amino acid residues with a molecular mass of 45.43 kD and bioelectric point value of 6.98. Sequence alignment analysis showed *ScMOC1* keep a high sequence identity with some poaceae family plants homologous proteins such as *Sorghum bicolor* (L.) Moench, *Leymus secalinus* (Georgi) Tzvel. and *Oryza sativa* L.. The genetic

收稿日期: 2016-11-30 修回日期: 2017-01-06 网络出版日期: 2017-06-13

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20170613.0846.016.html>

基金项目: 国家自然科学基金 (31360359, 31601362); 云南省基础研究计划青年项目 (2015FD063); 云南省中青年学术技术带头人后备人才 (2014HB038)

第一作者主要从事甘蔗分子遗传学研究工作。E-mail: lixujuan2011@163.com

通信作者: 刘新龙, 主要从事甘蔗分子遗传学研究。E-mail: lxlgood868@163.com

relationship of *ScMOC1* and its homologous proteins was analyzed and showed that *ScMOC1* had a closest evolutionary relationship with its homologous proteins from *Sorghum bicolor*(L.) Moench, *Setaria italica*(L.) P. Beauv. and *Zea mays* L. in the NJ phylogenetic tree. In order to understand the sequence variation of *ScMOC1* in ROC22, we cloned *ScMOC1* from sugarcane variety ROC22 and genetic transformed to *Escherichia coli*, then 30 clones were randomly selected to sequence and analysis, totally 46 SNPs and 29 InDels were identified in these clones. Among these variation sites, one deletion and the other 4 insertion of bases were the main causes of the changes for the gene encoding protein sequences. Neutrality tests showed that no purifying selection occurred of *ScMOC1* in ROC22. Real time fluorescent quantitative-PCR (qPCR) was used to analyze the expression pattern of the gene in six different tissues (root, stem, leaf, tiller bud, leaf sheath and stem apex) of tillering stage sugarcane and stem apex, axillary buds with different size (tender axillary bud, medium axillary bud, largish axillary bud, mature dormant axillary bud), the result showed that *ScMOC1* was constitutively expressed in ROC22 and highest expressed in stem apex at the tillering stage, while presented an expression tendency of “up-down-up-down” at the axillary buds formation and development period, so we predicted that *ScMOC1* played a role in the formation and development of sugarcane axillary buds. According to these researches, we speculated that *ScMOC1* might play a key role in the regulation of tillering trait of sugarcane. These results obtained above will provide a theoretical foundation for function analysis of *ScMOC1* and molecular assisted selection of sugarcane yield in the future.

Key words: sugarcane; tillering; *MOC1* gene; sequence variation analysis; expression analysis

分蘖(分枝)是禾谷类作物和其他许多单子叶植物生长发育过程中最重要的农艺性状之一,是影响水稻、小麦(*Triticum aestivum* L.)、大麦(*Hordeum vulgare* L.)等主要农作物穗数或茎数以及形态建成进而影响其产量的重要性状因子^[1-3]。对于无性繁殖经济作物-甘蔗(*Saccharum officinarum* L.)来说,分蘖性状显得尤为重要^[4],主要因为甘蔗的收获产品主要是地上蔗茎部分,而蔗茎由主茎和分蘖茎组成,可见分蘖茎的多少和大小直接决定了甘蔗产量的高低。因此,挖掘甘蔗分蘖调控关键基因并加以开发利用,对进一步提高甘蔗品种产量潜力具有重要意义。

MONOCULM 1 (*MOC1*)是最先从水稻中克隆获得的调控其分蘖芽形成的关键基因,该基因控制着腋生分生组织的起始和分蘖芽的形成,同时还有促进分蘖芽伸长的功能。其克隆是近年来植物形态建成特别是侧芽形成研究领域中最重要进展之一^[5]。*MOC1*属植物特有的GRAS家族的VHHD转录因子,该家族成员参与了植物侧枝形成^[6]、根辐射方向结构分布^[7-8]、光信号传导^[9]、植物激素赤霉素的信号传导和植株高度控制等多种重要的生理和发育过程^[10-13]。番茄(*Lycopersicon esculentum* Mill.) *LS*基因^[6]、拟南芥(*Arabidopsis thaliana*(L.) Heynh.) *LAS*基因^[14]均属于该家族的直向同源基因,它们在结构和功能上都很保守^[15],都参与了植物分蘖或分枝的调控,对于植物产量的形成具有重要的调控

作用。

鉴于 *MOC1* 基因在植物分蘖或分枝调控中的重要功能,本研究通过同源基因克隆法结合 RT-PCR 和 RACE 技术从甘蔗品种 ROC22 中克隆出 *MOC1* 的同源基因 *ScMOC1*, 又借助生物信息学分析预测其结构和功能。然后提取 ROC22 基因组 DNA 并扩增获得 *ScMOC1* 全长序列,再将此序列进行连接转化后挑取 30 个阳性克隆摇菌测序,并分析 ROC22 中所得 *ScMOC1* 的 30 个不同克隆核苷酸序列变异情况,最后通过实时荧光定量 PCR 对 *ScMOC1* 进行表达分析,以期为后续该基因的功能解析和在甘蔗产量分子辅助育种中的运用奠定基础。

1 材料与方法

1.1 材料

以我国主栽甘蔗品种新台糖 22 号(ROC22)为研究材料,由国家甘蔗种质资源圃提供。

1.2 方法

1.2.1 RNA 提取及反转录 分别选取 ROC22 幼苗茎尖生长点部位,分蘖期 ROC22 的不同组织部位(根、茎、叶、分蘖芽、叶鞘、茎尖生长点)以及甘蔗成熟期 ROC22 的茎尖生长点和不同发育阶段腋芽组织(幼嫩腋芽、半大腋芽、较大腋芽、成熟休眠腋芽),使用液氮速冻组织样品,并研磨提取总 RNA,提取方法参照全式金 TransZol™ Plant RNA 试剂盒,并采用 1.0% 的琼脂糖凝胶和 Bio Drop Lite PC 超微

量可见紫外分光光度计检测质量和浓度,然后分别使用试剂盒 Trans Script One-step gDNA Removal and cDNA Synthesis Super Mix 和 SMARTER™ RACE cDNA Amplification Kit 以及宝生物 One Step SYBR Prime Script™ RT-PCR Kit(Perfect Real Time)反转录成常规 cDNA 模板和 RACE cDNA 模板后于 -20 °C 保存备用。反转录参照各试剂盒说明书方法进行。

1.2.2 ScMOC1 基因克隆 以国家水稻数据中心公布的 MOC1 基因编码区 mRNA 序列 (AY242058.1) 为探针,通过 BLAST 检索并下载其他作物 MOC1 同源基因 mRNA 序列,并使用 DNAMAN 5.0 软件对下载的序列进行比对,在保守区设计 RT-PCR 引物,引物序列信息见表 1。使用高保真 Trans Start Fast Pfu Fly DNA 聚合酶(下同)进行基因片段的扩增,扩增产物克隆测序后使用 DNAMAN

5.0 软件比对分析获得甘蔗 MOC1 基因部分编码区片段。根据得到的部分编码区片段设计 3'RACE 和 5'RACE 基因特异引物(表 1),RACE 实验参照 SMARTER™ RACE cDNA Amplification Kit 说明书进行。

使用 DNAMAN 5.0、Vector NTI 11.5 软件将获得的部分编码区序列与 RACE 序列拼接,去除重叠部分后得到 ScMOC1 的 cDNA 序列。根据拼接序列设计引物(表 1)扩增 ScMOC1 的完整编码区序列,测序并通过 DNAMAN 5.0 软件和 UniProt 在线工具比对分析其序列正确性。所有 PCR 产物使用 EasyPure Quick Gel Extraction Kit 回收,使用全式金零背景克隆载体(pEASY-Blunt Zero Cloning Kit)将回收的产物连接,转化至大肠杆菌 DH5 α 感受态细胞内,挑取至少 3 个阳性克隆交由深圳华大基因测序。

表 1 ScMOC1 基因扩增引物信息

Table 1 Primers used for ScMOC1 amplification

引物名称 Primer name	序列(5'-3') Sequences(5'-3')	功能 Function
SMOC-751-F	GACGAGACGCTGGCTGTCAACT	RT-PCR 扩增
ScMOC-1239-R	CGTCTGCCACCCGAGAAAGC	RT-PCR amplification
ScMOC1-60-3d	GCTTGCCGCTTCTTGAAGTGGGTC	3'RACE 扩增
ScMOC1-108-3d	GACCGTCGCGGAGAGGGAGACAATC	3'RACE amplification
UPM	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACCGCAGAGT	RACE 扩增
NUPM	AAGCAGTGGTATCAACGCAGAGT	RACE amplification
ScMOC1-446-5d	AGGTAGCCCTCCGACGGGTACTGC	5'RACE 扩增
ScMOC1-803-5d	TGCCCCGCTAGCTTGTGCAAGAACA	5'RACE amplification
ScMOC1-123F	CGTCGAAGCTATGCTCA	编码区全长扩增
ScMOC1-1538R	TAGTGAGCGAGCCAAAC	CDS amplification of gene
GAPDH-F	CACGGCCACTGGAAGCA	荧光定量 PCR
GAPDH-R	TCCTCAGGGTTCTCTGATGCC	Real-time PCR
QMOC1-782F	AGACGCTGGCCGTCAACT	
QMOC1-866R	GCCTTGACCCACTTCAAGA	

1.2.3 ScMOC1 的生物信息学分析 首先利用在线工具 ORF Finder 查找所得序列的开放阅读框(ORF),然后使用 ExPASy 服务器中的 ProtScale 和 ProtParam 软件、SIGNA-IP 软件、ProtComp Version 9.0、cNLS mpper、SOP-MA 在线软件、SWISS-MODEL、ProtFun 软件分别预测 ScMOC1 的理化性质、信号肽、亚细胞定位、核定位信号、二级结构、三级结构及功能;再通过 DNAMAN 5.0 和 MEGA 6.0 等生物学软件对 MOC1 同源蛋白进行序列比对和 NJ 系统进化树构建,进化树的 bootstrap 设置为 1000。

1.2.4 ScMOC1 全长在甘蔗主栽品种 ROC22 中的克隆及序列变异分析 使用全式金 PlantZol 试剂盒提取 ROC22 基因组 DNA,所提取 DNA 质量和浓度通过 0.8% 的琼脂糖凝胶电泳和 Bio Drop Lite PC 超微量可见紫外分光光度计进行检测,然后稀释成 20 ng/ μ L 于 -20 °C 保存备用。以 ROC22 基因组 DNA 为模板,编码区全长引物 ScMOC1-123F 和 ScMOC1-1538R 进行 PCR 扩增 ScMOC1 全长,然后连接转化后挑取 30 个阳性克隆摇菌后送菌液测序(具体操作同上基因克隆部分)。测序结果使用软件 DNAMAN 5.0、MEGA 6.0 和 DnaSP 5.10 进行序

列分析。

1.2.5 *ScMOCI* 基因的实时荧光定量 PCR 表达分析 根据已获得的 *ScMOCI* 序列,使用 Primer Express 3.0.1 设计荧光定量 PCR 引物,以 H. Ling 等^[16]发表的 *GAPDH* 基因引物序列作为内参基因引物,引物序列见上表 1。选用 TaKaRa 公司的 PrimeScript™ RT reagent Kit with gDNA Eraser(Perfect Real Time)和 SYBR® Premix Ex Taq™ II (Tli RNaseH Plus),ROX plus 试剂盒,按试剂盒说明书进行操作,定量 PCR 体系为 20 μ L,包括 cDNA 模板(50 ng/ μ L) 2 μ L、SYBR®Premix Ex Taq™ II(Tli RNaseH Plus)(2 \times) 10.0 μ L、上下游引物(10 μ mol/L) 0.8 μ L、ROX Reference Dye II(50 \times)0.4 μ L 和 H₂O 6 μ L。采用两步法 PCR 扩增:95 $^{\circ}$ C 30 s;95 $^{\circ}$ C 5 s,60 $^{\circ}$ C 34 s,40 个循环;95 $^{\circ}$ C 15 s;60 $^{\circ}$ C 1 min;95 $^{\circ}$ C 15 s。每个样品设 3 次生物学重复和 3 次技术重复,在 ABI Vii7 Real time PCR System(Applied Biosystems, USA)上进行试验。最后利用 2^{- Δ ACT}法计算 *ScMOCI* 的相对表达量。

2 结果与分析

2.1 总 RNA 提取及 *ScMOCI* 部分编码区扩增

ROC22 生长点 RNA 提取后电泳显示 3 条完整的条带,OD 260/280 在 1.98 ~ 2.00 之间,OD 260/230 在 2.00 ~ 2.20 间,表明所提取 RNA 质量完整,盐离子去除干净,符合实验要求(图 1A)。保守区片段 RT-PCR 扩增测序分析后得到一条 489 bp 的

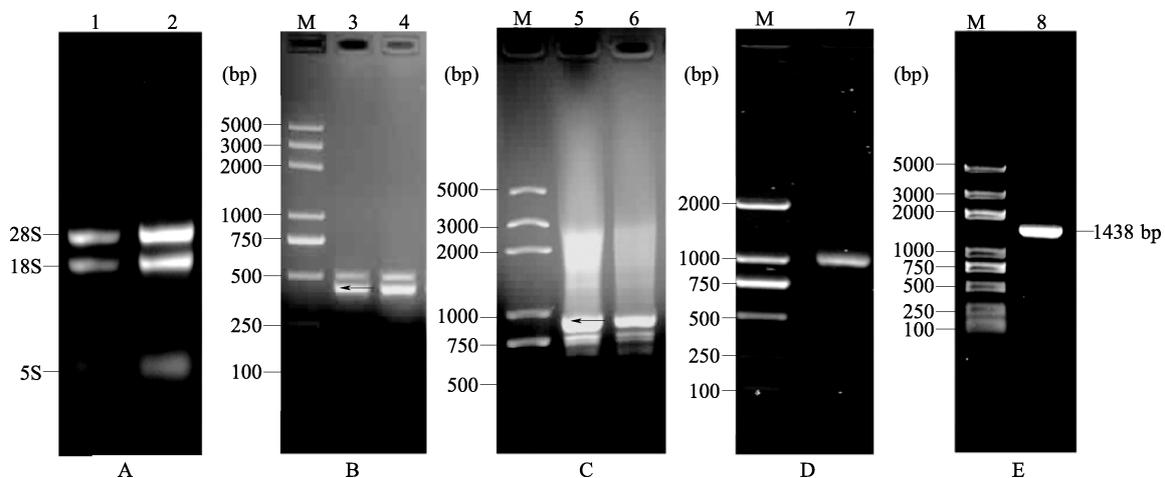
片段(图 1B),分析比对显示该序列与高粱、小麦、水稻等作物的 *MOCI* 同源基因相似性很高,其中与高粱 *MOC1A* 基因(JQ951675.1)相似性高达 99.4%,初步表明所得序列即为 *ScMOCI* 部分编码区序列。

2.2 *ScMOCI* 基因 3'和 5'RACE 结果

3'RACE PCR 扩增获得 800 bp 左右的产物(图 1C),测序分析表明该序列有效部分长 658 bp(去除通用引物),带有 poly A 结构,与高粱、双花草(*Dichanthium annulatum*(Forssk.) Stapf) *MOC1A* 基因对应序列相似性分别高达 99.2% 和 97.6%,验证了所得 *ScMOCI* 3'RACE 序列的正确性。5'RACE PCR 扩增得到 1000 bp 左右条带(图 1D),序列分析其长度为 967 bp(去除通用引物),与高秆野生稻(*Oryza alta* L.)、小麦 *MOCI* 同源基因对应序列相似性达 72.2% 和 71%,证明所获得序列是正确的 *ScMOCI* 5'RACE 序列。

2.3 *ScMOCI* 基因完整编码区扩增

上述 3 段序列拼接得到一条 1674 bp 的拼接序列,其包含的最大开放阅读框有 1410 bp,与高粱 *MOC1A* 基因对应序列相似性高达 98.3%,初步表明拼接得到的序列为正确的 *ScMOCI* 编码区序列。完整编码区 PCR 扩增测序分析得到一条 1438 bp 的条带(图 1E),该序列与 *ScMOCI* 拼接编码区序列和高粱 *MOC1A* 基因相应序列相似性分别高达 97.6%,验证扩增到的 *ScMOCI* 编码区的正确性,所得序列上传至 NCBI(Gen-Bank ID:KP876558.1)。



A:1,2 为 ROC22 茎尖生长点总 RNA 电泳结果;B:3,4 为 *ScMOCI* 部分编码区扩增结果(箭头所指条带);
C:5,6 为 *ScMOCI* 3'RACE 结果(箭头所指条带);D:7 为 *ScMOCI* 5'RACE 结果;E:8 为 *ScMOCI* 完整编码区扩增结果
A:1,2 are agarose gel electrophoresis results for total RNA from stem apex of ROC22,B:3,4 are amplification results of *ScMOCI*'s
parts of coding region(the band marked by arrow),C:5,6 are the 3'RACE results of *ScMOCI*(the band marked by arrow),
D:7 is the 5'RACE result of *ScMOCI*,E:8 is the result of complete coding region of *ScMOCI*

图 1 ROC22 茎尖生长点总 RNA 提取及 *ScMOCI* 反转录扩增产物电泳结果

Fig. 1 Agarose gel electrophoresis for total RNA from stem apex of ROC22 and RT-PCR amplification products of *ScMOCI*

2.4 ScMOC1 基因生物信息学分析

2.4.1 ScMOC1 序列结构分析 通过在线工具 ORF finder 分析发现 ScMOC1 的 cDNA 包含一个 1299 bp 的完整开放阅读框, 编码 432 个氨基酸

```

11 atgctcagctcactccaccaccaccactcctcctcctcctccgac 641 cgcgcttcgcccgtccatccagctcccttccacttcacgcc
   M L S S L H H H S S S S S D   R A F A R S I Q L P F H F T P
56 accgacaacaacaatggcaacaagaacggcagcagcggcaacagc 686 ctctgtctcctgcgccctaccatcatcatcagcaggtcgcc
   T D N N N G N K N G S S G N S   L L L S C A A T H H H Q Q V A
101 ggccggctcctcgcggcgccgggtccgtccgcgggacactcgtg 731 agcacgagcaccacgaccaccacttcgagtgccggccacgagcctg
   G G V L A A A V P S A R D L V   S T S T T T T T S S A A T S L
146 ctggcctgcccagcctgctgcagcggggaccccgagcccgccg 776 gagctgccccggagcagagcgtggcctcaactcgtcatgttc
   L A C A D L L Q R G D L Q A A   E L H P D E T L A V N C V M F
191 cggcggccgggggtcctcctgtccggcgctccccgcggcc 821 ttgcacaagctaggcggggcaggacgagcttgcgcttcttgaag
   R R A A G V L L S A A S P R A   L H K L G G Q D E L A A F L K
236 gacgcccggaccggctcgcctaccacttcgcccggcgctcgcg 866 tgggtcaaggccatggcccccccggtgacogtgcgggaggg
   D A A D R L A Y H F A R A L A   W V K A M A P A V V T V A E R
281 ctccaggcggatgtcaggcccgccggcgggcgctcgtcgcc 911 gagacaatcggcggggtttcgaccgcatcgacgacctgccccag
   L Q A D V R A A A A G R V V A   E T I G G G F D R I D D L P Q
326 ccggccttgctgactcccggctacttctccggccggcgccgg 956 cgggcccagtttccatggatcactactcggccogtgttcaggcgg
   P G L V A A A A T S A A A R   R A A V S M D H Y S A F F E A
371 ccggcgtcgtcggcgcctaccctcggcttcaaccagatcgcggc 1001 ctcgaggcgcactgtgcccggggagccgggagcggtacggtg
   P A S S G A Y L A F N Q I A P   L E A T V P P G S R E R L A V
416 ttctgcttgcggcacctaacggcgaaaccagccatcctcgac 1046 gagcaggaggtgctcggcaggagatcgaccggcgctgggtgccc
   F L R F A H L T A N Q A I L D   E Q E V L G R E I D A A L G A
461 gccgtcggggcgccggcgaatccacatcctcgacctggacgcc 1091 tccggggggaggtggggcgggactcgaacggggggggccggcg
   A V E G A R R R I H I L D L D A   S G G R W W R G L E R W G A A
506 gcgcacggcgtgagtgccggcctcctcctcagggcagatcgcc 1136 gcgcggccggcggattcggcgccggcctcagcgcgttcggcc
   A H G V Q W P P L L Q A I A E   A R A A G F A A R P L S A F A
551 cggcagcaccggcggcgggaccggcggaggtgcatcaccggc 1181 gtgtcgaggcagcggctgctgctgctgctgctgctgctgctgctg
   R D D P A A G P P E V R I T G   V S Q A R L L L R L H Y P S E
596 gccggcggcggcggcggcggcggcggcggcggcggcggcggc 1226 ggctacctcgtgaggagggcgccggcgctgcttctcgggtgg
   A G A D R D T L R T G S R L   G Y L V Q E A R G A C F L G W
641 cgcgcttcgcccgtccatccagctcccttccacttcacggc 1271 cagacggcggcgtgctgctgctgctgctgctgctgctgctgctgctg
   R A F A R S I Q L P F H F T P   Q T R P L L S V S S W Q *

```

图2 ScMOC1 的 ORF 及其推导的编码氨基酸序列

Fig. 2 The complete open reading frame of ScMOC1 and deduced amino acid sequence

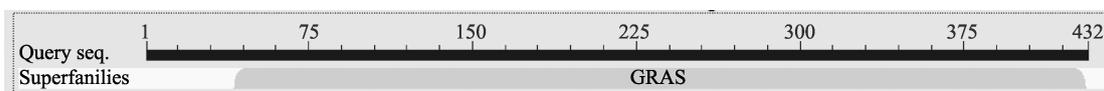


图3 ScMOC1 保守结构域预测

Fig. 3 Prediction of conserved domain of ScMOC1 protein

2.4.2 ScMOC1 理化性质分析 使用 ProtParam 和 ProtScale 工具对 ScMOC1 的理化性质预测结果表明, 该蛋白分子质量约为 45.43 kD、理化等电点 pI 为 6.98、共包含 6374 个原子, 分子式为 $C_{1999}H_{3169}N_{603}O_{595}S_8$; 在组成该蛋白的 20 种氨基酸中丙氨酸 (Ala) 所占比例最高, 达到 19.7%, 而半胱氨酸 (Cys)、赖氨酸 (Lys) 以及蛋氨酸 (Met) 所占比例最低, 为 0.9%; 该蛋白不稳定指数 (II) 为 42.18, 大于 40, 表明该蛋白不太稳定, 脂肪指数 (AI) 为 92.87, 总平均亲水性 0.038, 介于 $-0.5 \sim 0.5$ 之间, 推测其可能是两性蛋白, 其亲/疏水信号如图 4, 其中第 48 位具有最高分值, 为 2.444, 疏水性最强, 第 20 位具有最低分值 -2.889, 亲水性较强。

(Gen-Bank ID: AMB19792.1)。Conserved domain search 在线分析显示其包含一个 GRAS 保守结构域, 属于 GRAS 家族蛋白(图 2, 图 3)。

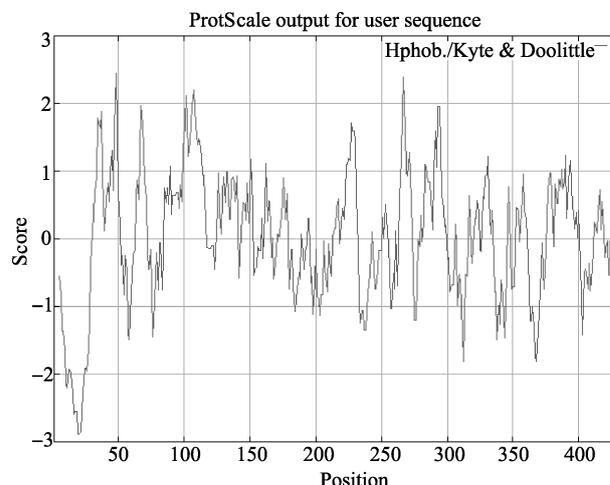


图4 ScMOC1 亲疏水性分析

Fig. 4 Hydrophilicity/hydrophobicity prediction of ScMOC1

2.4.3 ScMOC1 信号肽、核定位信号和蛋白亚细胞定位预测 使用 SignalP 4.1 工具和 cNLS mpper 软件预测 ScMOC1 蛋白的信号肽和核定位信号,结果表明 ScMOC1 第 48 位的半胱氨酸(Cys)残基具有最高的剪切位点分值 0.122(图 5),第 48 位的酪氨酸(Tyr)残基具有最高的结合剪切位点分值 0.115,第 47 位的丝氨酸(Ser)残基具有最高的信号肽分值 0.171;由于最后算得信号肽平均值为 0.107,小于 0.5,推测 *ScMOC1* 基因所编码的蛋白不存在信号肽,为非分泌蛋白,且没有检测到核定位信号。ProtComp Version 9.0 在线软件预测表明 ScMOC1 定位于细胞核的积分值为 9.87,其余部位除了高尔基体的积分值为 0.12,液泡处的积分值为 0.01 外,均为 0.00,从而推断 ScMOC1 蛋白定位于细胞核内的可能性很大,可能是一个在细胞核内发挥调控作用的转录因子。对于该蛋白没有核定位信号,却能够定位到核上这一特点与水稻的 MOC1 蛋白相似^[5]。

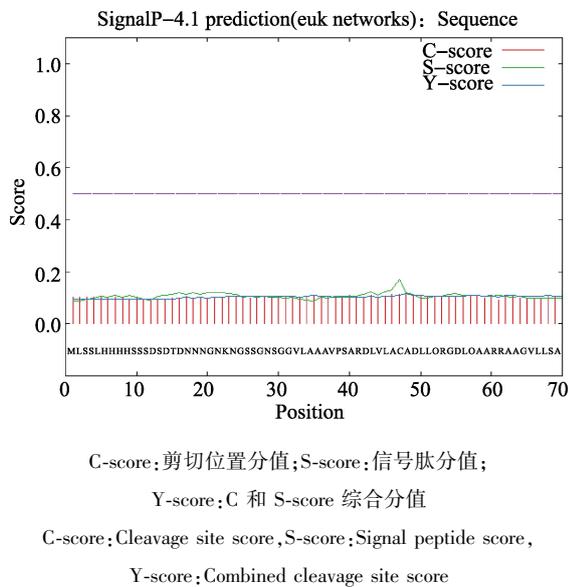


图 5 ScMOC1 信号肽预测

2.4.4 ScMOC1 的二级结构和功能预测 使用 SOPMA 软件预测 ScMOC1 蛋白的二级结构,结果如图 6 所示,该蛋白主要由 α -螺旋(h)、无规则卷曲(c)、 β -折叠(e)和 β -转角(t)构成,其中 α 螺旋最多,比例为 48.84%,其次为无规则卷曲(29.63%) 和 β -折叠(12.27%), β -转角最少(9.26%)。使用在线软件 ProtFun 2.2 预测 ScMOC1 蛋白功能,预测结果如下表 2 所示,分为三部分内容,第一部分显示的是功能分类(Functional category),第二部分是酶

类或非酶类(Enzyme/nonenzyme),第三部分是基因本体分类(Gene ontology category),结果表明该蛋白在嘌呤和嘧啶、能量代谢、转运和结合、辅因子和氨基酸生物合成过程中发挥功能的可能性分别为 0.549、0.309、0.195、0.181 和 0.162;为酶类和非酶类的可能性分别为 0.754 和 0.246;作为信号转导分子、转录因子、转运因子、转录调控因子的可能性分别为 0.122、0.051、0.025 和 0.021,明显高于其他功能。综合考虑所有的分析结果,表明 ScMOC1 蛋白可能在嘌呤和嘧啶以及能量代谢等过程中发挥信号转导、转录调控等重要作用。



图 6 ScMOC1 二级结构预测

Fig. 6 Secondary structure analysis of ScMOC1

表 2 ScMOC1 功能预测结果

Table 2 Function prediction results of ScMOC1

功能分类 Functional category	概率 Probability	几率 Odds	基因本体分类 Gene ontology category	概率 Probability	几率 Odds
氨基酸生物合成	0.162	7.360	信号转导分子	0.122	0.571
辅因子生物合成	0.181	2.513	受体	0.003	0.018
细胞被膜	0.060	0.992	激素	0.001	0.206
细胞加工	0.027	0.373	生长因子	0.018	1.314
中央中间代谢	0.052	0.833	转运因子	0.025	0.230
能量代谢	0.309	3.431	转录调控因子	0.021	0.167
脂肪酸代谢	0.029	2.194	电压激活离子通道	0.005	0.249
嘌呤和嘧啶	0.549	2.259	结构蛋白	0.003	0.116
调控功能	0.022	0.135	阳离子通道	0.010	0.169
复制和转录	0.037	0.138	应激反应	0.009	0.102
翻译	0.045	1.013	免疫应答	0.016	0.186
转运和结合	0.195	0.477	金属离子转运	0.009	0.020
酶类	0.754	2.632	离子通道	0.010	0.220
非酶类	0.246	0.345	转录因子	0.051	0.397

2.4.5 ScMOC1 三级结构预测 使用 phyre 2 软件通过线串法构建 ScMOC1 与其同源蛋白的三级结构,根据预测结果可知 ScMOC1 的空间结构以 α 螺旋和无规则卷曲为主(图 7),这和其二级结构预测

结果相吻合。将该蛋白与高粱、小麦和水稻的 MOC1 同源蛋白的三级空间结构图进行比较,结果显示它们的三级空间结构高度相似。

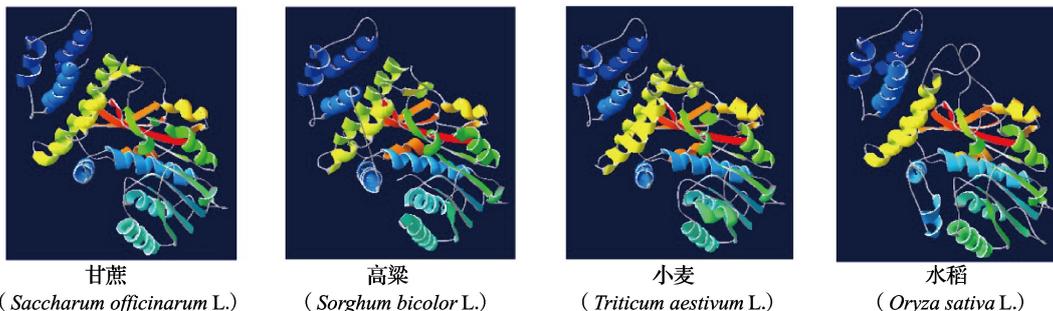


图 7 MOC1 及其同源蛋白三级结构预测图

Fig. 7 Prediction of the tertiary structure of ScMOC1 and it's homologous protein

2.4.6 MOC1 同源性分析 为了明确 ScMOC1 与其他植物同源蛋白的系统进化关系,使用 DNAMAN 软件对甘蔗、高粱、大麦、赖草 (*Leymus secalinus* (Georgi) Tzvelev)、粳稻 (*Oryza sativa* L. subsp. *japonica* S. Kato)、稻属 (*Oryzactata* L.)、小麦 7 种植物的 MOC1 同源蛋白进行序列比对分析,结果见图 8,发现这几种

植物氨基酸序列中存在丝氨酸同聚物链, VHID 基序以及酪氨酸等多处保守结构域,总体呈现为 N 端变异较大,但 C 端更为保守,总体一致性达到 87.37%。其中 ScMOC1 与高粱、大麦以及粳稻同源蛋白相似性分别高达 94%、78.3% 和 77.6%,推测它们具有相似或相近的功能。

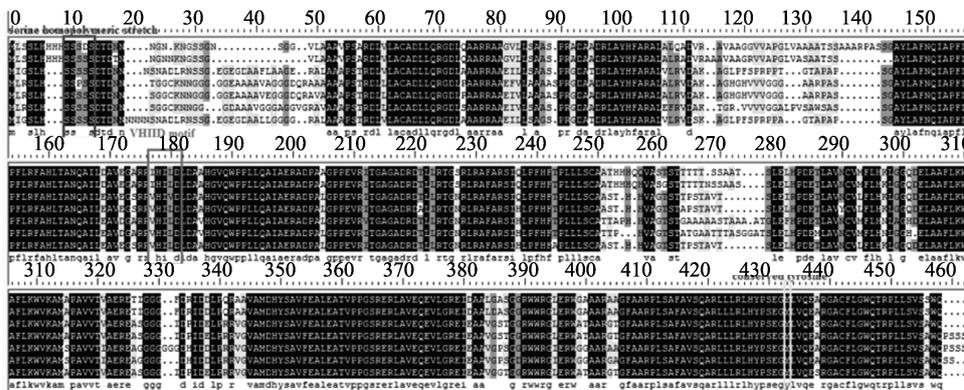


图 8 ScMOC1 及其同源蛋白氨基酸序列一致性比对结果

Fig. 8 Sequence alignment of amino acid residues of ScMOC1 and it's homologous protein

2.4.7 系统进化树分析 为了进一步明确 ScMOC1 的功能和进化关系,将其与双花草、高粱、曲序香茅草 (*Cymbopogon flexuosus* (Nees ex Steud.) Will. Wats.)、小米草、玉米、马来野生稻 (*Oryza ridleyi* Hook. f.)、短花药野生稻 (*Oryza brachyantha* A. Chev. & Roeh.)、药用稻 (*Oryza officinalis* Wall. ex. G. Watt)、赖草、大麦、小麦、二穗短柄草 (*Brachypodium distachyum* (L.) P. Beauv.)、拟南芥、烟草 (*Nicotiana tabacum* L.)、番茄等 15 种植物 MOC1 同源蛋白的氨基酸序列使用 MEGA 6.0 软件进行系统进化树构建,Bootstrap 值设为 1000,结果显示,可将 16 种植物蛋白分为 A、B 两个组。其中

A 组是单子叶植物, B 组是双子叶植物,而且 ScMOC1 与高粱、小米草等禾本科植物 MOC1 同源蛋白亲缘关系更近,推测它们功能也相似,具体如图 9 所示。

2.5 ScMOC1 全长在甘蔗主栽品种 ROC22 中的克隆与序列变异分析

ROC22 基因组 DNA 提取后条带单一, OD260/OD280 值为 1.82,表明提取基因组 DNA 质量好,符合实验要求。以其为模板扩增得到 1.4 kb 左右的条带,如图 10A 所示,将此条带回收克隆测序分析正确后连接转化克隆,部分克隆 PCR 检测得到 1.5 kb 左右的条带如图 10B 所示。

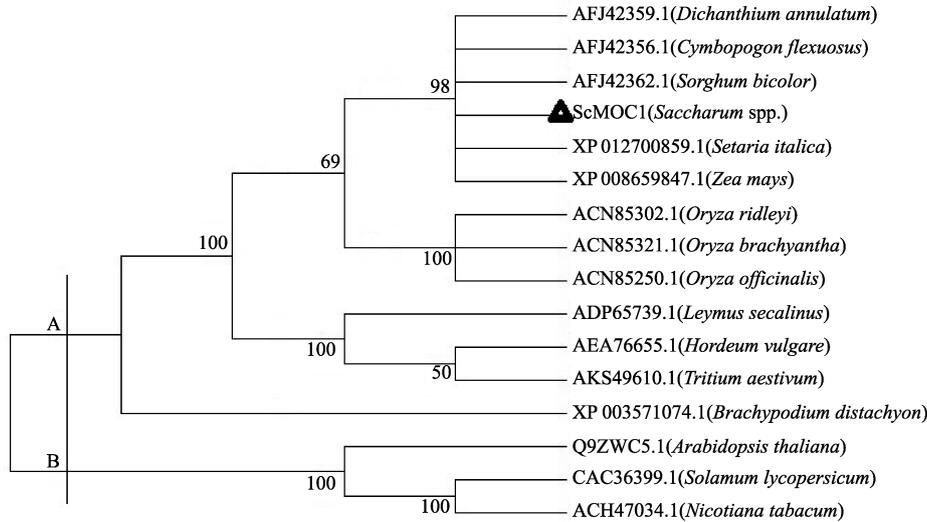


图 9 16 种植物 MOC1 同源蛋白的 NJ 分子进化树

Fig. 9 NJ phylogenetic tree of *ScMOC1* and it's homologous proteins of other plants

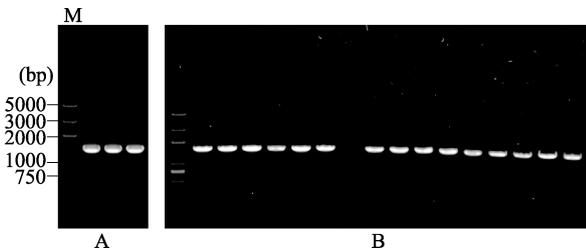


图 10 *ScMOC1* 全长 PCR 扩增及阳性克隆电泳结果

Fig. 10 Agarose gel electrophoresis for the PCR products of *ScMOC1* and it's positive clones

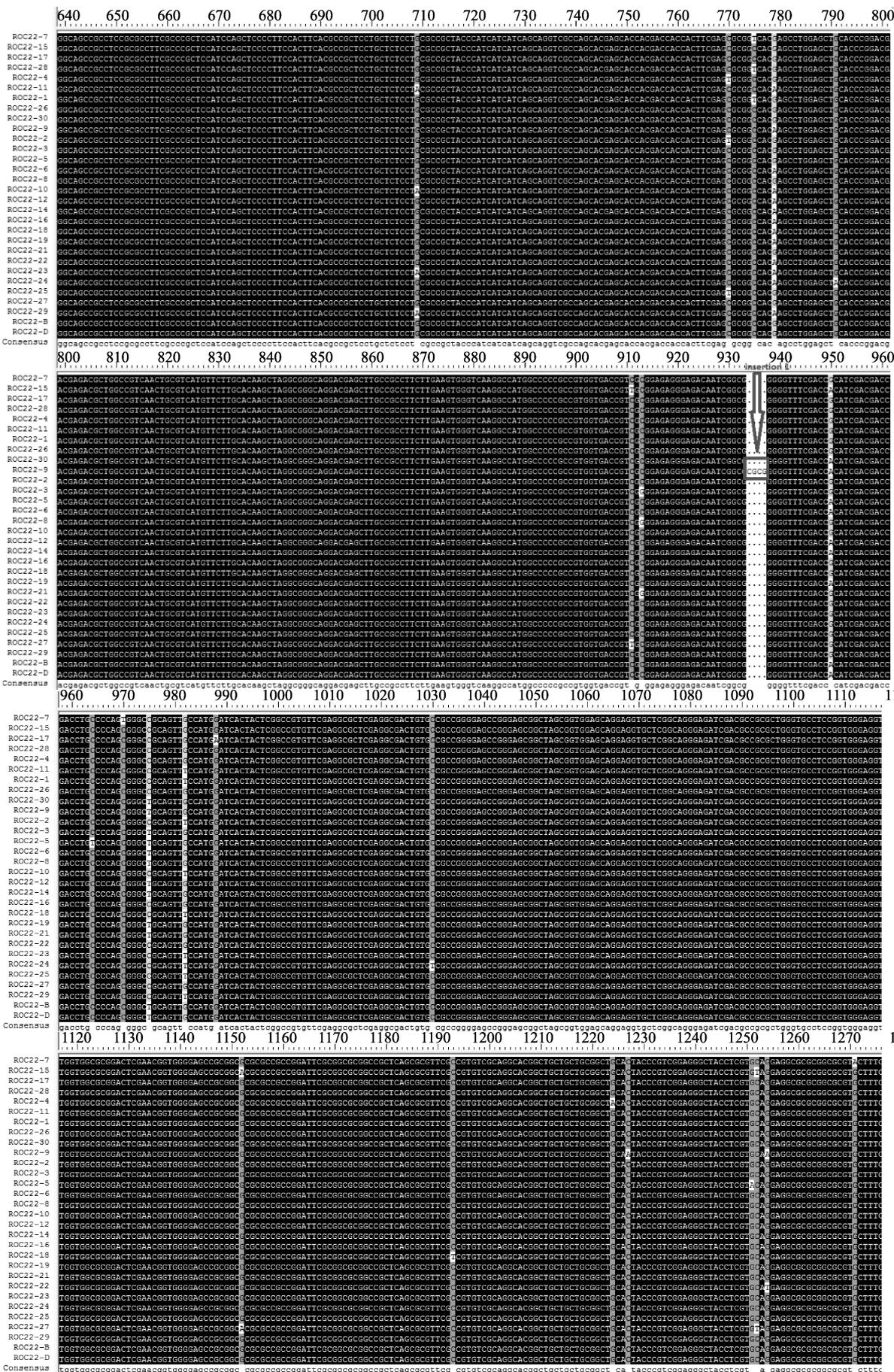
阳性克隆测序结果分析表明 ROC22 中扩增序列转化得到的 30 个克隆和水稻 *MOC1* 及小麦 *TaMOC1* 一样没有内含子,仅包含 1 个外显子,长 1324 bp。统计分析发现目标序列中共有 46 个 SNP 位点,其中碱基转换 (T→C, A→G) 发生了 35 处,碱基颠换 (A→C, A→T, C→G, G→T) 发生了 11 处,转换与颠换的比例为 3.2:1,说明 ROC22 中 *ScMOC1* 发生的碱基突变主要类型是转换;InDel 位点共有 29 处。转换成氨基酸序列后,30 条核苷酸序列中同义突变有 19 处,非同义突变有 27 处;另外,单倍型分析显示 ROC22 中 *ScMOC1* 的平均核苷酸差异 θ 为 0.00936,核苷酸多样性 P_i 为 0.00668,单倍型数目为 25 个,单倍型多样性为 0.984。Tajima's D 为 -1.06264,且 $P > 0.10$,未达到显著性水平,说明 *ScMOC1* 在 ROC22 品种中的进化受到了选择压力,但遵循中性进化模型,没有发生纯化选择。序列比对发现 30 个克隆的 DNA 序列相似性很高,为 98.04%,而翻译成氨基酸后的序列相似性仅 89.27%。30 条序列间存在随机分布的碱基突变,

在 30~40 bp、50~60 bp、90~100 bp、310~340 bp 以及 930~940 bp 处发生了碱基插入或缺失 (图 11 中方框标注处),进一步分析发现,B、D 及 9 号克隆氨基酸序列与其他克隆差异较大 (图 12),是造成氨基酸序列差异的主要原因,移除这 3 个序列后氨基酸相似性高达 98.12%,其中前两个克隆发生明显差异从第 18 位氨基酸开始,主要是由于 50~60 bp 间缺失了一个碱基 A 导致的,9 号克隆氨基酸序列差异从 310 个氨基酸后开始,主要由 930~940 bp 的 CGCG 碱基插入引起的 (图中已用方框及箭头标注)。

2.6 *ScMOC1* 的表达分析

2.6.1 *ScMOC1* 的组织特异性分析 qPCR 实验显示 *ScMOC1* 基因在分蘖期的 ROC22 不同组织部位均有表达,但表达量有所差异。在生长点中的表达量最高,且显著高于其他组织部位,其次为在根和茎中,紧接着是在分蘖芽中,在叶和叶鞘中表达量较低 (图 13)。由此表明,*ScMOC1* 基因在分蘖期甘蔗不同组织部位的表达具有明显的组织特异性,在生命活跃的甘蔗茎尖生长点处表达量最高。

2.6.2 *ScMOC1* 基因在不同发育阶段腋芽中的表达分析 qPCR 分析表明,*ScMOC1* 在不同发育阶段腋芽组织中的表达量有所差异,所有阶段的腋芽组织中的表达都显著高于茎尖生长点部位。在腋芽形成发育过程中总体呈现出“升-降-升-降”的趋势,其中在较大腋芽中表达量显著高于其他腋芽发育阶段,其次为幼嫩腋芽,然后随着腋芽的成熟,表达有所下降 (图 14)。以上说明 *ScMOC1* 基因可能在甘蔗腋芽的形成发育时期发挥作用。



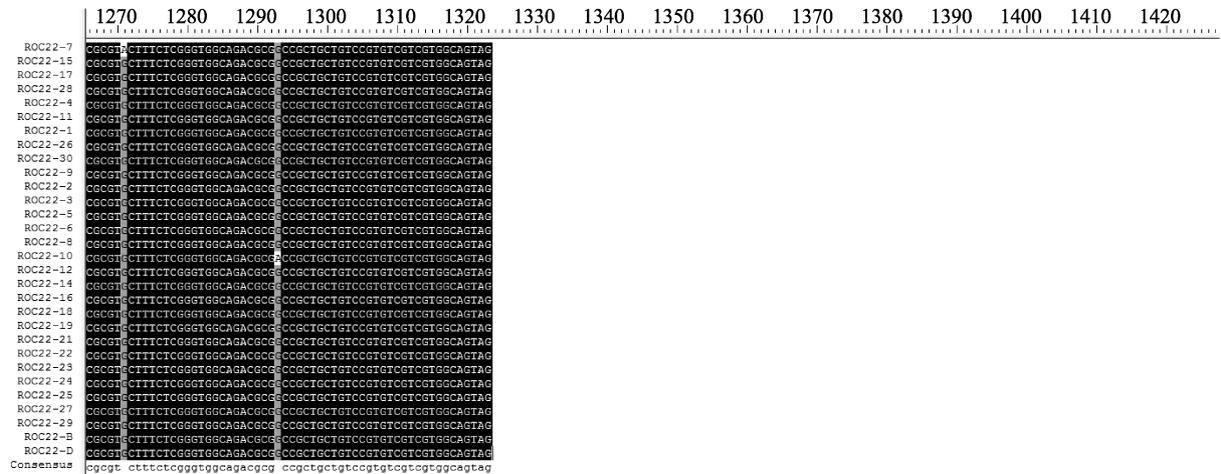
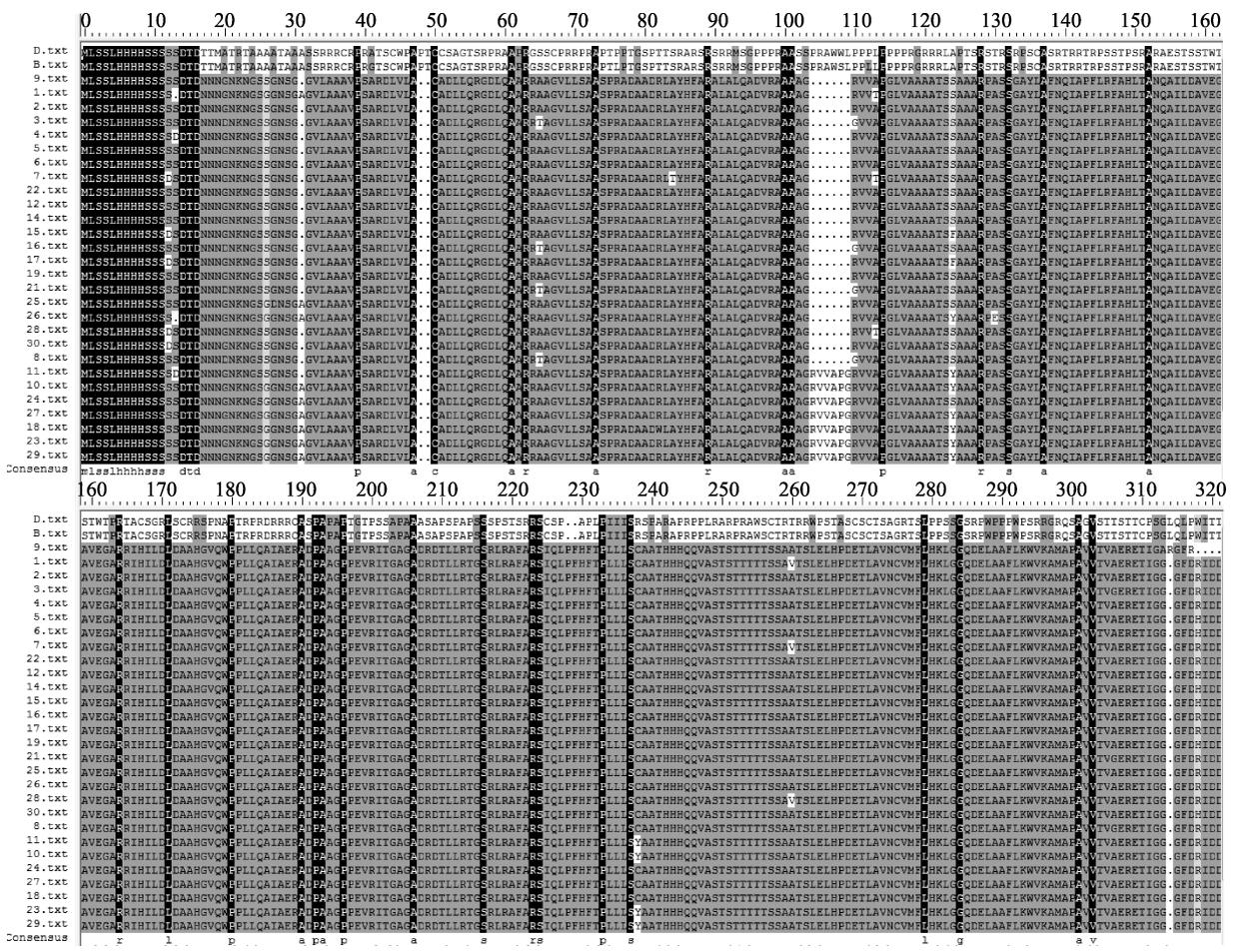


图 11 ROC22 中 *ScMOCI* 的 DNA 序列同源性分析
 Fig. 11 DNA sequence similarity of *ScMOCI* in ROC22



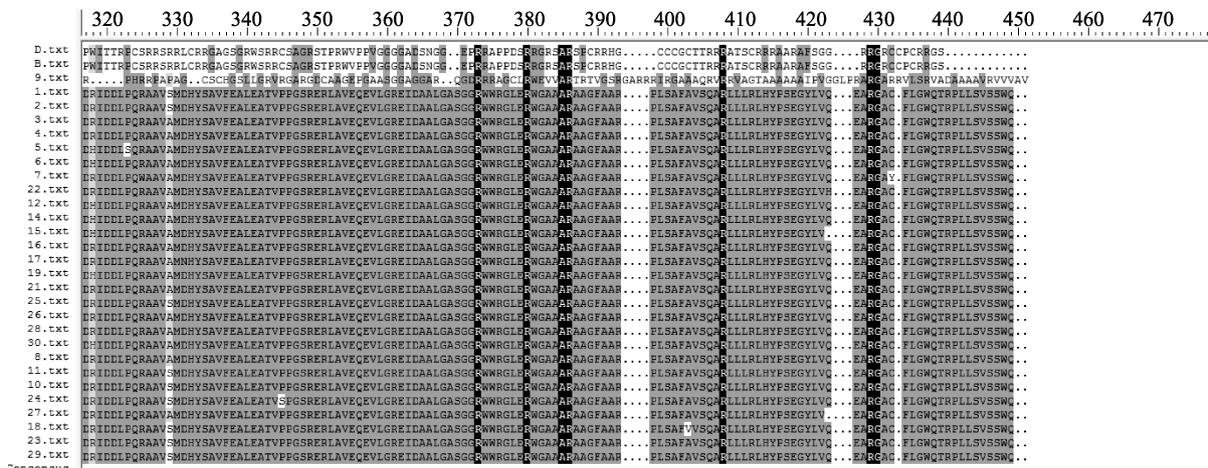
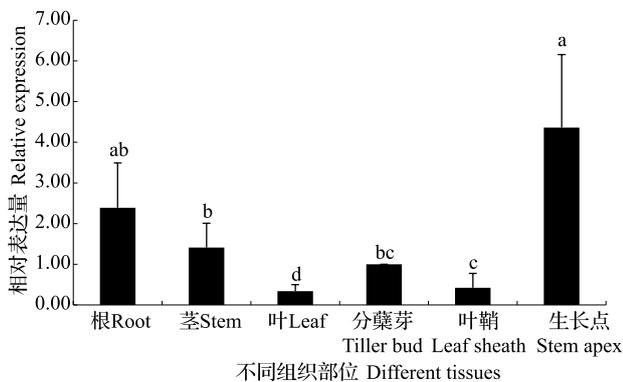


图 12 ROC22 中 *ScMOC1* 氨基酸序列同源性分析

Fig. 12 Amino acid sequence similarity of *ScMOC1* in ROC22



不同字母表示差异显著 ($P < 0.05$), $n = 3$ (下同)
Bars with different superscripts differ significantly ($P < 0.05$), $n = 3$ (the same as below)

图 13 *ScMOC1* 在甘蔗不同组织部位的表达分析

Fig. 13 Expression analysis of *ScMOC1* in different sugarcane tissues

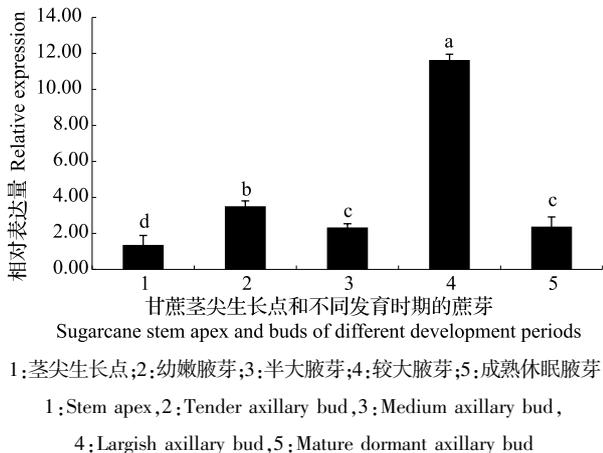


图 14 甘蔗茎尖生长点及不同发育时期蔗芽中 *ScMOC1* 的表达分析

Fig. 14 Expression analysis of *ScMOC1* in stem apex and sugarcane buds of different development periods

3 讨论

GRAS 转录因子是植物特有的重要转录因子家族之一,广泛参与植株侧枝形成、分生组织维持、胁迫应答、信号转导等多种生理过程^[17]。该家族名字起源于 3 个最初从拟南芥中分离出来的成员 GAI[GA (GIBBERELIC ACID)-INSENSITIVE]^[10], RGA (REPRESSOR OF GAI)^[11] 和 SCR (SCARECROW)^[7]。*LS/LAS/MOC1* 基因均编码 GRAS 转录因子^[18]。番茄 *LS* 基因的功能缺失突变体表现为无分枝,无花被,雌雄育性降低,这些表型均是由于腋芽分生组织的起始被阻断所致^[6]。*MOC1* 过表达能明显增加水稻分蘖数,而 *moc1* 突变体则不能形成腋芽和分蘖,仅有一个主干^[5]。研究发现^[14],将 *LAS* 基因转到 *ls* 突变体后能够恢复番茄的表型,上述研究均表明 *MOC1* 及其同源基因在植物腋分生组织的形成和分蘖(枝)数目调控方面扮演着十分重要的角色,且功能保守。

本研究克隆的 *ScMOC1* 包含长度为 1299 bp 的 ORF,编码 432 个氨基酸残基,该基因没有内含子,和水稻、小麦、玉米等作物 N 端存在变异区而 C 端较为保守,包含丝氨酸同聚物链,VIHID 基序以及酪氨酸等结构域,与 GRAS 家族蛋白结构完全吻合^[19]。信号肽和核定位信号预测表明 *ScMOC1* 不存在信号肽,为非分泌蛋白,也不存在核定位信号;但亚细胞定位预测 *ScMOC1* 主要定位于细胞核,这似乎与不存在核定位信号相矛盾,但与水稻^[5]、小麦 *MOC1*^[15] 预测结果一致,而且 X. Li 等^[5] 研究中也指出,水稻 *MOC1* 中没有明确的核定位信号,可能

是通过某种未经确认的核定位信号或其他不依赖于核定位信号的机制进入到细胞核, *ScMOC1* 也可能与水稻 *MOC1* 这一情况一致。二三级结构预测显示 *ScMOC1* 空间结构以 α 螺旋和无规则卷曲为主, 与高粱、小麦、水稻等同源蛋白结构高度相似。系统进化树结果也显示 *ScMOC1* 与高粱、小米草、玉米等 *MOC1* 蛋白亲缘关系近。

功能基因核苷酸的变异是物种遗传多样性的前提, 丰富的遗传多样性是作物遗传改良的基础。对目的基因进行核苷酸多态性分析对掌握其遗传多样性, 并进而通过遗传转化等途径在分子水平上有目的地进行作物的遗传改良创造条件^[20]。SNP 和 InDel 变异是基因组中最常见的变异, 其中 SNP 是基因组中发生频率最高、数量最多而且密度最大的变异。不论 SNP 还是 InDel 均可引起基因的突变并对表达产物产生影响, 而且部分变异能够引起作物表型产生巨大的变化^[21-22]。本研究对 *ScMOC1* 在甘蔗品种 ROC22 中进行序列变异分析发现, 序列中存在较多处 SNP 和 InDel, 而引起编码产物发生变化较大的主要是 50 ~ 60 bp 间的 1 个碱基缺失及 930 ~ 940 bp 的 4 个碱基插入, 此变异引起的基因功能改变还有待进一步分析。

组织特异性分析显示 *ScMOC1* 在分蘖期甘蔗的茎尖生长点处表达量最高, 其次是根、茎、分蘖芽, 在叶和叶鞘中表达量较低, 这与孕穗期小麦和初花期赖草 *MOC1* 同源基因表达情况类似^[15, 23]。另外, 在甘蔗腋芽形成发育不同阶段, *ScMOC1* 的表达量均有明显增加。以上研究初步表明本研究成功克隆到了甘蔗 *MOC1* 基因, 且该基因在甘蔗腋芽形成发育中有重要调控作用。虽然同源基因在结构和功能方面通常都是保守的, 但不同物种间的功能特异性可能有所差异^[15], 因此 *ScMOC1* 功能还有待进一步实验验证。本研究结果可为今后 *ScMOC1* 的功能、作用机制研究及其在甘蔗产量分子辅助育种中的利用奠定重要基础。

参考文献

- [1] Wang Y, Li J. Branching in rice[J]. *Curr Opin Plant Biol*, 2011, 14(1): 94-99
- [2] Xing Y, Zhang Q. Genetic and molecular bases of rice yield[J]. *Plant Biol*, 2010, 61(11): 1-22
- [3] Sreenivasulu N, Schnurbusch T. A genetic playground for enhancing grain number in cereals[J]. *Trends Plant Sci*, 2012, 17(2): 91-101
- [4] 陈如凯. 现代甘蔗遗传育种[M]. 北京: 中国农业出版社, 2011: 1-19
- [5] Li X, Qian Q, Fu Z, et al. Control of tillering in rice[J]. *Nature*, 2003, 422(6932): 618-621
- [6] Schumacher K, Schmitt T, Rossberg M, et al. The *Lateral suppressor (Ls)* gene of tomato encodes a new member of the VHIID protein family[J]. *Proc Natl Acad Sci USA*, 1999, 96(1): 290-295
- [7] Di L L, Wysocka-Diller J, Malamy J E, et al. The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root[J]. *Cell*, 1996, 86(3): 423-433
- [8] Helariutta Y, Fukaki H, Wysockadiller J, et al. The *SHORT-ROOT* gene controls radial patterning of the *Arabidopsis* root through radial signaling[J]. *Cell*, 2000, 101(5): 555-567
- [9] Chance B, Nishimura M. PAT1, a new member of the GRAS family, is involved in phytochrome A signal transduction[J]. *Genes Dev*, 2000, 14(10): 1269-1278
- [10] Peng J, Carol P, Richards D E, et al. The *Arabidopsis GAI* gene defines a signaling pathway that negatively regulates gibberellin responses[J]. *Genes Dev*, 1998, 11(23): 3194-3205
- [11] Silverstone A L, Ciampaglio C N, Sun T. The *Arabidopsis RGA* gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway[J]. *Plant Cell*, 1998, 10(2): 155-169
- [12] Itoh H, Ueguchi-Tanaka M, Sato Y, et al. The gibberellin signaling pathway is regulated by the appearance and disappearance of SLENDER RICE1 in nuclei[J]. *Plant Cell*, 2002, 14(1): 57-70
- [13] Peng J, Richards D E, Hartley N M, et al. Green revolution' genes encode mutant gibberellin response modulators[J]. *Nature*, 1999, 400(6741): 256-261
- [14] Greb T, Clarenz O, Schafer E, et al. Molecular analysis of the *LATERAL SUPPRESSOR* gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation[J]. *Genes Dev*, 2003, 17(9): 1175-1187
- [15] Zhang B, Liu X, Xu W, et al. Novel function of a putative *MOC1* ortholog associated with spikelet number per spike in common wheat[J]. *Sci Rep*, 2015, 5: 12211
- [16] Ling H, Wu Q, Guo J, et al. Comprehensive selection of reference genes for gene expression normalization in sugarcane by real time quantitative RT-PCR[J]. *PLoS One*, 2014, 9(5): e97469
- [17] 张文霞, 刁志娟, 吴为人. 植物 GRAS 蛋白研究的新进展[J]. *分子植物育种*, 2016, 15(5): 1159-1165
- [18] Bolle C. The role of GRAS proteins in plant signal transduction and development[J]. *Planta*, 2004, 218(5): 683-692
- [19] Sun X, Jones W T, Rikkerink E H. GRAS proteins: the versatile roles of intrinsically disordered proteins in plant signalling[J]. *Biochem J*, 2012, 442(1): 1-12
- [20] 徐丽, 陈新, 张力思, 等. 核桃 *JrCBF* 基因的克隆与表达和单核苷酸多态性分析[J]. *植物遗传资源学报*, 2014, 15(2): 320-326
- [21] 焦健, 刘崇怀, 樊秀彩, 等. 中国野生种葡萄 mybA 转录因子 SNP 特征分析[J]. *植物遗传资源学报*, 2013, 14(5): 885-891
- [22] Zhang E, Yang Z, Wang Y, et al. Nucleotide polymorphisms and haplotype diversity of *RTCS* gene in China elite maize inbred lines[J]. *PLoS One*, 2013, 8(2): e56495
- [23] 何文兴, 李洪梅, 叶春江, 等. 赖草根茎分蘖相关基因 *LRC1* 的克隆及表达[J]. *中国草地学报*, 2011, 33(5): 7-14