

DOI:10.13275/j.cnki.lykxyj.2022.02.015

油茶转录因子基因 *CoSOC1-like* 的克隆和表达分析

黄国文*, 管天球, 赵雨云, 陈莫林, 刘宏辉

(湖南科技学院化生院, 湖南 永州 425199)

摘要: [目的] 克隆油茶的 *SOC1* 同源基因 (*CoSOC1-like*), 分析其序列特征和表达模式及其蛋白进化。 [方法] 以3年生油茶嫩叶为材料提取 RNA, 利用 RT-PCR 和 RACE 方法克隆油茶的 *CoSOC1-like* 基因, 用生物信息学工具分析其序列特征, 用荧光定量 PCR 分析其表达模式, 用基因瞬时表达法分析其蛋白的亚细胞定位。 [结果] 油茶 *CoSOC1-like* 基因的 CDS 全长为 654 bp, 推测蛋白质由 127 个氨基酸组成, 蛋白分子量为 24.958 kD, 等电点 (pI) 为 6.8, 基因库的登录号为 MT036382。 *CoSOC1-like* 具有植物 II 型 MADS-box 基因的蛋白结构, 且 C 末端含有 MOTIF 结构域, 是一个 MADS-box 家族转录因子。 *CoSOC1-like* 蛋白有 31 个磷酸化位点, 建立在二级结构基础上的 *CoSOC1-like* 蛋白的三级结构具有明显活性部位; *CoSOC1-like* 基因瞬时表达分析表明, 油茶 *CoSOC1-like* 蛋白定位于细胞核, 符合转录因子的细胞核定位特征。系统进化分析表明: 油茶 *CoSOC1-like* 与茶树 *SOC1-like* 蛋白聚类在同一个进化支上。荧光定量 PCR 分析表明: *CoSOC1-like* 基因存在于油茶所有的器官中, 尤其在花芽中的相对表达量最多。 [结论] *CoSOC1-like* 基因在油茶的花芽分化中起重要作用, 也可能参与根、茎、叶和种子等器官的生长发育。本结果为进一步研究油茶成花的分子机制奠定了基础。

关键词: 油茶; *SOC1-like*; 克隆; 序列分析; 表达分析

中图分类号: S794.4

文献标志码: A

文章编号: 1001-1498(2022)02-0129-11

油茶 (*Camellia oleifera* Abel.) 为山茶科山茶属常绿灌木或小乔木, 是一种重要的木本油料树种^[1], 主要分布在我国长江流域及以南地区的湖南、安徽、广东、广西、福建等省区。茶油含有 90% 以上不饱和脂肪酸, 具有降血脂、预防心脑血管疾病等功效, 是有利于人类健康的食用油。油茶是虫媒、两性花的异花授粉植物。油茶的成花时间长。“湘林 1 号”油茶 (*Camellia oleifera* ‘Xianglin1’) 的花芽分化从 5 月下旬开始到 9 月上旬结束, 经过生理分化期、萼片形成期、花瓣形成期、雌雄蕊形成期、子房和花药形成期、雌雄蕊成熟期等阶段^[2], 然而, “长林 4 号”油茶 (*Camellia oleifera*

‘Changlin4’) 花芽分化的开始时间是 6 月上旬^[3], 10 月中下旬开花结果。植物成花受光周期途经、自主春化途径、赤霉素途径和糖类途经等的基因调节, 也受开花整合子 *SOC1*、*FT*、*LFY* 等基因的调节, 最后受同源异形基因 *AP1*、*SEP*、*AG* 等的调控完成花器官发育^[4-5]。*SOC1* 基因是 MADS-box 家族基因之一。MADS-box 基因家族在调节植物的花器官发育、开花时间、胚珠发育、根瘤形成和对环境信号反应等方面起作用^[6-7]。在真核生物中, MADS-box 基因家族是一类编码转录因子的基因组成, 其蛋白质具有高度保守的 MADS 结构域^[8], 能够使本身蛋白质以二聚体的形式结合其它

收稿日期: 2021-08-18 修回日期: 2021-11-28

基金项目: 永州市科技创新指导性计划项目 (2018ZD02); 湖南省科技计划重点项目 (2014NK2021); 湘南优势植物资源综合利用湖南省重点实验室资助项目 (XNZW15C16)

* 通讯作者: 黄国文。Email: huanggdax@163.com

辅助因子和 DNA, 调节基因的时空表达。在被子植物中, 大多数 *SOC1* 基因含有 7 个外显子和 6 个内含子, 其编码的蛋白除了含有保守的 MADS-box 以外, 还含有 K-box 和非保守的 I 区、C 区, 因此, 属于 MIKC 型^[9]。K-box 是一个半保守片段, 主要由有 3 个 α 螺旋组成, 其中, 前面 2 个螺旋和第 3 个螺旋的作用分别是决定本身蛋白二聚化的特异性和形成高级复合物, 是蛋白质相互作用部位^[10]; I 区的序列变化较大, 主要促进蛋白二聚体本身与 DNA 的结合; C 区由疏水性氨基酸组成的最不保守区域, 是基因转录的激活区, 但有些 *SOC1* 蛋白也具有一个保守的 MOTIF 结构域, 因此, 被划入 SOC1/TM3 亚家族^[11], *SOC1* 的不同功能主要由不同的 C 区基序完成^[12]。目前, 在大豆 (*Glycine max* (Linn.) Merr.)^[13-14]、玉米 (*Zea mays* L.)^[15]、烟草 (*Nicotiana tabacum* L.)^[16]、芒果 (*Mangifera indica* L.)^[17]、油菜 (*Brassica campestris* L.)^[18]、拟南芥 (*Arabidopsis thaliana* (L.) Heynh)^[19-20]、茶树 (*Camellia sinensis* (L.) O. Ktze)^[21-22]、橘子 (*Citrus reticulata* Blanco)^[23]、火龙果 (*Hylocereus polyrhizus* (Haw) Britt & Ros)^[24]、牡丹 (*Paeonia suffruticosa* Andr.)^[25] 等植物中的 *SOC1* 基因序列和作用都有报道, 但关于油茶中 *SOC1* 同源基因的序列和作用的报道很少。本研究采用 RT-PCR 技术和 RACE (Rapid Amplification of cDNA Ends) 技术克隆油茶的 *CoSOC1-like* 基因, 用生物信息学方法研究其序列特征, 用荧光定量 PCR 技术研究 *CoSOC1-like* 基因的表达模式, 用基因瞬时表达法研究蛋白的亚细胞定位, 为进一步研究油茶 *CoSOC1-like* 基因的功能打下基础, 为揭示油茶成花的分子机制和花期调控提供理论依据。

1 材料与方 法

1.1 材 料

油茶材料采摘于湖南科技学院油茶示范基地的 3 年生油茶‘湘林 210’ (*Camellia oleifera* ‘Xianglin 210’) 品种。采摘后的叶片立即放于干冰盒中冷冻保存运输, 后转存入 $-70\text{ }^{\circ}\text{C}$ 超低温冰箱, 在 48 h 内提取 RNA。

1.2 试 验 方 法

1.2.1 油茶 *CoSOC1-like* 基因的克隆 采用 RACE 方法克隆油茶 *CoSOC1-like* 基因^[26]。以油茶幼嫩

叶片为材料, 用 RNA 提取试剂盒 (北京索莱宝科技有限公司) 提取叶片总 RNA。使用 1.0% 琼脂糖凝胶电泳测定 RNA 的质量。以 RNA 完整且无蛋白质污染的 RNA 作为底物, 按照反转录试剂盒 (北京全式金生物技术有限公司) 合成 cDNA 第一链。扩增 *CoSOC1-like* 基因的引物序列见表 1。以 cDNA 第一链为模板, F1 和 R1、F2 和 R2 为引物和高保真的 PFU 酶 (北京全式金生物技术有限公司) 构建反应体系, 进行巢式 PCR 扩增, 获得基因的核心片段。使用 3' 和 5'RACE cDNA 扩增试剂盒 (北京百泰克生物技术有限公司) 提供的引物、F3 和 R3、F4 和 R4 引物进行 PCR 扩增, 获得基因的 3' 端序列片段和 5' 端序列片段。每个片段经过琼脂糖凝胶回收试剂盒 (上海生物工程股份有限公司) 回收, 与 pGM-T 载体 (北京索莱宝公司) 连接并转化大肠杆菌 DH5 α 细胞, 并涂布在含有 $100\text{ }\mu\text{g}\cdot\text{mL}^{-1}$ 氨苄青霉素的固体 LB 培养基上进行蓝白斑筛选, $37\text{ }^{\circ}\text{C}$ 培养 16 h。用 PCR 法鉴定白色菌落的阳性克隆。用质粒提取试剂盒 (上海生物工程股份有限公司) 提取 5~10 个阳性克隆的质粒, 送至北京擎科生物科技有限公司湖南分公司测序。将每个序列在 DNAMAN 软件中拼接, 获得含有全长 *CoSOC1-like* 基因序列的片段。

表 1 *CoSOC1-like* 基因的引物序列
Table 1 Primer sequences of *CoSOC1-like*

引物名称 Primer name	寡核苷酸序列 5'-3' Oligonucleotide sequence 5'-3'
F1	ATGGTGAGNGGNAARACNCA
R1	AGCTGCTCDATYTYGYTCYTT
F2	ACAAGCAGNCARGTNACNTT
R2	TTCTCAAGYTYGYTCDAT
F3	AGGTTGCTCTCATCATCTTCTCTC
R3	TTTGAACATCTTTTGATGCCTCT
F4	CAGAGGCATACAAAAGATGTTCAA
R4	TAGGAGAGAAGATGATGAGAGCAA
F5	GCTCTAGAGATGGTGAGAGGGAAGACTCAGATGA
R5	TCCCCCGGGGTAACCTTCTGAGGCGAAGCACGCT

1.2.2 油茶 *CoSOC1-like* 基因序列的生物信息学分析 用 DNASTar 5.0 软件分析油茶 *CoSOC1-like* 基因开放阅读框, 并推导氨基酸序列。用 DNAMAN 软件预测油茶 *CoSOC1-like* 蛋白质序列的疏水性、亲水性和跨膜结构域。采用 SOPMA

在线工具 (<http://npsa-pbil.ibcp.fr/cgi-bin/>) 分析蛋白质二级结构; 用 Phyre2 在线工具分析蛋白质三级结构; 在 PSORT 在线工具 (<http://psort.hgc.jp/>) 上进行亚细胞定位预测; 运用在线 KinasePhos 2.0 软件对 *CoSOC1-like* 蛋白作用方式预测; 利用 MEGA5.0 软件的 Clustal W 法进行比对, 用邻接法构建同源蛋白系统发育树 (自展值 Bootstrap 设定为 1 000, 其它参数的缺省值不变)。

1.2.3 油茶 *CoSOC1-like* 基因表达模式分析 在 2020 年 12 月 20 日取处于开花后期油茶的根、幼茎、叶片、营养芽、花瓣、雄蕊、柄托雌 (花柄、花托和雌蕊部分的总称, 因为花柄短并且花托和子房紧密结合, 不易区分且很难分离) 等部分提取总 RNA。在 2021 年 5 月 30 日取处于花芽生理分化期油茶的根、幼茎、叶片、顶芽 (未出现分化的芽)、营养芽、花芽和幼果等部分提取 RNA。在 2021 年 6 月 30 日取处于雌蕊和雄蕊形成期油茶的根、茎、叶片、营养芽 (0.4 cm 左右)、花芽 (0.8 cm 左右)、果皮和种子等部分提取 RNA。在 2021 年 7 月 30 日取处于花芽子房和花药形成期 (去掉苞片的花芽中有肉眼可见的约 2 mm 左右分叉柱头、1 mm 左右的雄蕊) 油茶的根、茎、叶片、营养芽 (0.5 cm 左右)、花芽 (1.3 cm 左右)、果皮和种子等部分提取 RNA。所有的 RNA 用反转录试剂盒合成 cDNA 后用于荧光定量 PCR。测定基因表达量的引物为, 油茶 *CoSOC1-like* 基因引物为 F: 5'TCTCTGCGATGCTGAGGTTG 3', R: 5' TCTATCTGCTTTGCCATGTCTG 3', 片段长度 195 bp。ACTIN 基因引物为 F: 5' TAGACTTGC GGCATCAGTTAGA 3', R: 5' TTCACGGTTTTT GGACGGATT 3', 片段长度 176 bp。所用仪器为 BIO-RAD 的 CFX Connect Real-Time System, 试剂是赛默飞世尔科技有限公司的 Power SYBR™ Green PCR 预混液 (货号: 4367659)。反应体积为 20 μ L, 反应程序为: 95.0 $^{\circ}$ C 3 min, 95.0 $^{\circ}$ C 15 s, 55.0 $^{\circ}$ C 20 s, 72.0 $^{\circ}$ C 30 s, 35 个循环, 72.0 $^{\circ}$ C 5 min, 4 $^{\circ}$ C 保存。每个样品设置 3 次重复, 采用 $2^{-\Delta\Delta CT}$ 方法计算基因的相对表达量, 利用 SPSS19 软件的单变量方差分析的 Duncan 多重比较法进行显著性差异和标准误分析。

1.2.4 油茶 *CoSOC1-like* 蛋白的亚细胞定位分析

以扩增油茶 *CoSOC1-like* 基因全长的加入 *Xba* I 酶切位点的正向引物 (F: 5' GCTCTAGAG ATGGTG

AGAGGGAAGACTCAGATGA 3')、加入了 *Xma* I 酶切位点的反向引物 (R: 5' TCCCCCGGGGTA ACT TCTGAGGCGAAGCACGCT 3') 和高保真的 PFU 酶进行 PCR 扩增出基因全长。对 *CoSOC1-like* 基因全长和 pBI121-EGFP 质粒 (湖南丰辉生物科技有限公司) 分别进行 *Xba* I 和 *Xma* I (北京 NEB 公司) 双酶切, 用琼脂糖凝胶 DNA 片段回收试剂盒对酶切片进行回收, 然后用 T4 DNA 连接酶 (美国 clontech 公司) 连接这 2 个片段, 转化大肠杆菌 DH5 α 感受态细胞, 用含有抗生素 Amp 的 LB 培养基培养, 对菌斑进行 PCR 鉴定和测序鉴定获得阳性克隆。将此阳性克隆用热激法转化到农杆菌 EHA105 菌株中。用含有 0.1 mol·L⁻¹ 乙酰丁香酮和 25 μ g·mL⁻¹ 利福平的 YEB 培养基培养农杆菌 EHA105 到 OD₆₀₀ 为 0.5~0.6 时, 转速 5 000 rpm 离心 10 min 收集农杆菌, 用重悬液 (10 mmol·L⁻¹ MES, 0.1 mmol·L⁻¹ 乙酰丁香酮和 10 mmol·L⁻¹ 氯化镁, 过滤除菌) 悬浮农杆菌, 用注射器将农杆菌注射到洋葱 (*Allium cepa* L.) 鳞片内表皮中, 1 d 后取洋葱鳞片内表皮在荧光显微镜 (日本 Nikon ECLIPSE Ni-U) 下观察 EGFP 蛋白的绿色荧光。

2 结果与分析

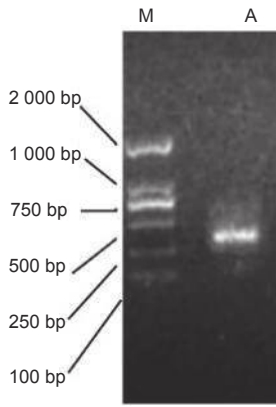
2.1 油茶 *CoSOC1-like* 基因的克隆和序列分析

CoSOC1-like 基因的核心片段用 PCR 反应扩增, 其产物经凝胶电泳检测, 表明扩增产物约在 350 bp 左右处有一条亮带, 无其它杂带 (图 1)。用 5'RACE 反应和 3'RACE 反应, 扩增基因的 5'端序列和 3'末端序列; 经过测序和拼接, 获得序列 1 025 bp 片段。经过开放阅读框分析, 此 DNA 片段包含 1 个完整的 654 bp CDS 和 150 bp 5'UTR、221 bp 3'UTR (图 2)。

2.2 油茶 *CoSOC1-like* 蛋白质的氨基酸特征和一级结构分析

油茶 *CoSOC1-like* 基因的 CDS 序列编码 217 个氨基酸组成的蛋白质 (图 2)。蛋白质分子量为 24.958 kDa, 等电点为 6.8。带负电荷的氨基酸 (Asp + Glu) 数量为 37 个, 带正电荷的氨基酸 (Arg + Lys) 数量为 36 个, 平均疏水性值为 -0.77, 平均亲水性值为 0.361, 属于亲水性蛋白。该蛋白无信号肽位点, 是非分泌蛋白, 也无跨膜螺旋区。

油茶 *CoSOC1-like* 蛋白质与其它物种 SOC1 蛋白的同源性比对结果 (图 3) 表明: 它与茶树



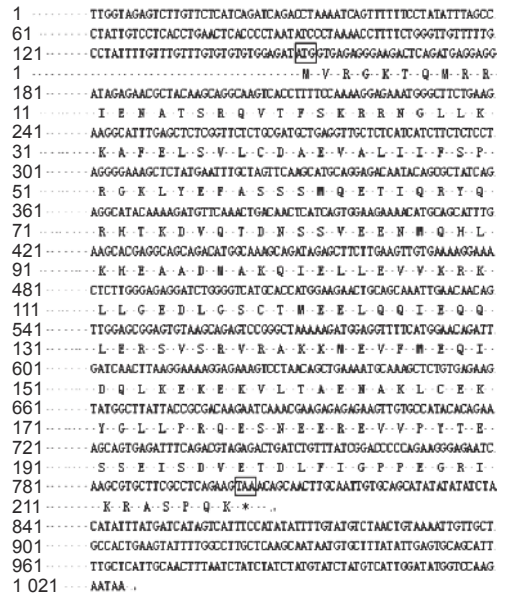
注: M, DL2000 DNA Marker; A, 核心片段的产物

Notes: M, DL2000 DNA Marker; A, Core fragment products

图 1 油茶 *CoSOC1-like* 基因片段的扩增产物电泳图

Fig. 1 Electrophoresis result of *CoSOC1-like* fragment of *Camellia oleifera* from RT-PCR

SOC1-like (XP-028068271.1) 的同源性为 97.7%、与榴莲 (XP_022769695.1) 的 SOC1-like 同源性为 78.4%、与核桃 (XP_018851690.1) 的 SOC1-like 同源性为 75.46%、与加州白栎 (XP_030930174.1) 的 SOC1 同源性为 79.1%、与猕猴桃 (AKH61958.1) SOC1e 同源性为 82.16%、与山核桃 (AHI85950.1) 的 SOC1 同源性为 75.35%、与杨梅 (KAB1200892.1) 的 SOC1 同源性为 75.94%、表明所克隆到的片段为油茶



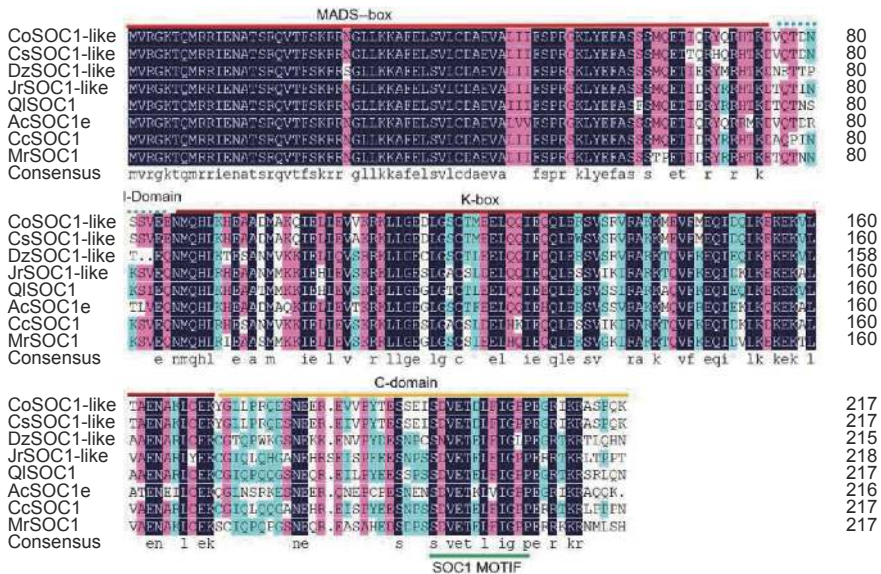
注: 带方框的 ATG 和 TAA 分别表示起始密码子和终止密码子, *表示蛋白翻译终止

Notes: ATG and TAA with boxes and * represent the starting and termination codon, and the termination of protein translation, respectively.

图 2 *CoSOC1-like* 基因 cDNA 全长序列和开放阅读框及其氨基酸序列

Fig. 2 cDNA sequence and deduced amino acid sequence of *CoSOC1-like*

SOC1 同源基因, 本文将该基因命名为 *CoSOC1-like*, 并在 GenBank 注册, 登录号为 MT036382。该



注: Co:油茶; Cs:茶树; Dz:榴莲; Jr:核桃; Ql:加州白栎; Ac:猕猴桃; Cc:山核桃; Mr:杨梅

Notes: Co: *Camellia oleifera*; Cs: *Camellia sinensis*; Dz: *Durio zibethinus*; Jr: *Juglans regia*; Ql: *Quercus lobata*; Ac: *Actinidia chinensis*; Cc: *Carya cathayensis*; Mr: *Morella rubra*

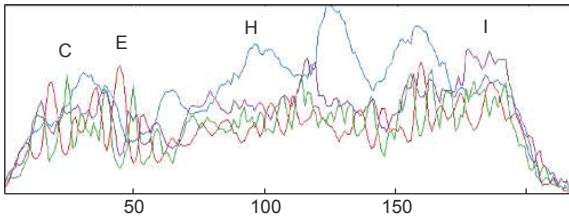
图 3 *CoSOC1-like* 基因编码的氨基酸与其它植物的 SOC1 的序列比对

Fig. 3 Alignments of the *CoSOC1-like* deduced amino acid sequences with other SOC1 proteins from plants

序列蛋白质 *CoSOC1-like* 结构包含 MADS-box (1-74 位的氨基酸)、I-domain (75-83 位的氨基酸)、K-box (84-172 位的氨基酸)、C-domain 四种结构域, 属于植物 II 型 MADS-box 基因的蛋白结构, C-domain 中含有 SOC1 MOTIF, 因此 *CoSOC1-like* 属于 SOC1/TM3 型。

2.3 油茶 *CoSOC1-like* 蛋白质的二级结构预测

蛋白质二级结构是指蛋白质多肽链中氨基酸主要依赖于氢键而建立的有规则重复的构象, 主要包括 α -螺旋、 β -折叠片、 β -转角和无规则卷曲。油茶 *CoSOC1-like* 蛋白的二级结构 (图 4) 表明: 油茶 *CoSOC1-like* 蛋白含有 α -螺旋 56.68%、 β -转角 4.15%、折叠延伸链 10.14% 和无规卷曲 29.0%。在蛋白质的 84—172 位 K-box 的氨基酸段有 3 个很明显的 α -螺旋, 这 3 个 α -螺旋是 MIKC 型 MADS-box 基因的特征序列。在蛋白质的 C-domain 的氨基酸段主要是卷曲和折叠延伸链组成。



注: C 表示转角 (绿色); E 表示折叠延伸链 (红色); H 表示 α 螺旋 (蓝色); I 表示无规卷曲 (紫色)

Notes: C: corner (green); E: folded extended strand (red); H: alpha helix (blue); I: irregular coil (purple)

图 4 *CoSOC1-like* 蛋白的二级结构预测

Fig. 4 Secondary structure prediction of *CoSOC1-like* protein

2.4 油茶 *CoSOC1-like* 蛋白质的三级结构预测

蛋白质的三级结构是指一条多肽链在二级结构甚至结构域的基础上, 依靠氨基酸侧链基团的疏水作用、范德华力及氢键和静电作用, 进一步盘旋和折叠形成特定空间结构。油茶 *CoSOC1-like* 蛋白三级结构 (图 5) 表明: *CoSOC1-like* 属于 SRF-LIKE 家族的蛋白质, 预测的油茶 *CoSOC1-like* 三级结构的肽链与二级结构的预测结果较一致。在 *CoSOC1-like* 蛋白质的氨基端有 2 个明显的 α -螺旋 (红色和黄色螺旋部分), 属于蛋白的 MADS-box 区域; 中间部分有 2 个同向平行的折叠延伸链, 还有一个较长的 α -螺旋 (青绿色 α -螺旋部分) 属于 K-box 域; 蛋白的羧基端是无规卷曲 (深蓝色部分)。整个蛋白质折叠成了一个紧密的并且有“空

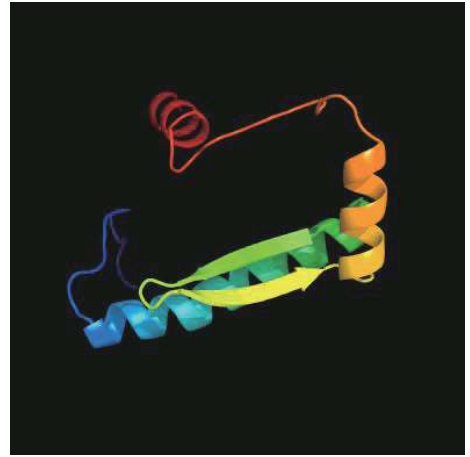


图 5 *CoSOC1-like* 蛋白质三级结构预测

Fig. 5 Tertiary structure prediction of *CoSOC1-like* protein

穴”(活性部位)的空间结构, 这种结构与它结合 DNA 和激活基因转录功能相适应。

2.5 油茶 *CoSOC1-like* 蛋白的亚细胞定位和磷酸化位点分析

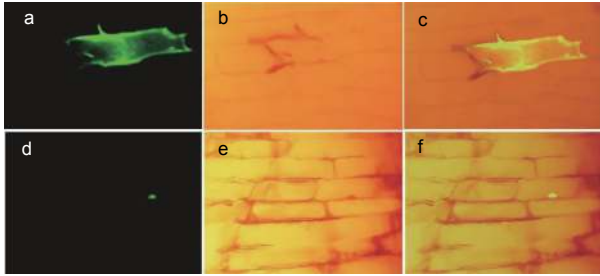
预测油茶 *CoSOC1-like* 蛋白质的亚细胞定位表明: 油茶 *CoSOC1-like* 蛋白没有信号肽存在, 不是外分泌蛋白; 在蛋白质 153 位上存在亮氨酸拉链模式 (LKEKEKVLTAENAKLCEKYGLL), 3 位上存在 SRF 型转录因子 DNA 结合和二聚化结构域 (RGKTQMRRIENATSRQVTF SKRRNGLLKKAFE LSVLCAEVALIIFSPRGKLYEF), *CoSOC1-like* 位于细胞核的可能性最大, 其次为线粒体中 (表 2)。

表 2 *CoSOC1-like* 蛋白质亚细胞定位分析结果
Table 2 Subcellular localization prediction of *CoSOC1-like* protein

位置 Position	概率 Probability/%
细胞核 Nuclear	47.8
线粒体 Mitochondrial	30.4
细胞膜 Plasma membrane	4.3
包括细胞壁的细胞外部分 Extracellular, including cell wall	4.3
高尔基体 Golgi	4.3
细胞质 Cytoplasmic	4.3
过氧化物酶体 Peroxisomal	4.3

将 *CoSOC1-like* 基因与载体 pBI121-EGFP 融合, 构建的 pBI121-*CoSOC1-like*-EGFP 融合载体转化洋葱表皮细胞, 以转化空载体的洋葱表皮细胞

作为对照,在荧光显微镜下观察发现,对照中整个洋葱表皮细胞(包括细胞质和细胞核)都发出绿色荧光(图 6 a、c),说明该空载体是有作用的;融合载体转化的洋葱表皮细胞中只有细胞核中发出绿色荧光(图 6 d、f),说明 CoSOC1-like 蛋白不是定位于线粒体,而是定位于细胞核中。



注: a、b、c: EGFP 质粒转化到洋葱细胞; d、e、f: CoSOC1-like 与 EGFP 融合质粒转化到洋葱细胞; a、d: 蓝光激发的绿色荧光; b、e: 明场下的图片; c、f: 分别为 a 和 b、d 和 e 融合的照片

Notes: a, b, c: Onion epidermal cells transformed with plasmids EGFP; d, e, f: Onion epidermal cells transformed with the CoSOC1-like-EGFP fusion plasmids; a, d: The green fluorescence in the dark field excited by blue light; b, e: The images of the same cells in the bright field; c, f: Merged images of a and b, d and e, respectively

图 6 荧光显微镜(200X)下观察 CoSOC1-like 蛋白在洋葱表皮细胞中的定位

Fig. 6 Localization of CoSOC1-like protein in onion epidermal cells observed in a fluorescence microscope (200X)

对 CoSOC1-like 蛋白序列修饰位点分析表明: CoSOC1-like 蛋白含有 17 个丝氨酸磷酸化位点、10 个苏氨酸磷酸化位点、4 个酪氨酸磷酸化位点,说明该蛋白活性受磷酸化作用。油茶 CoSOC1-like 转录因子可能是通过蛋白质磷酸化激活,与有关基因的启动子区相互作用达到调节靶基因表达的目的。

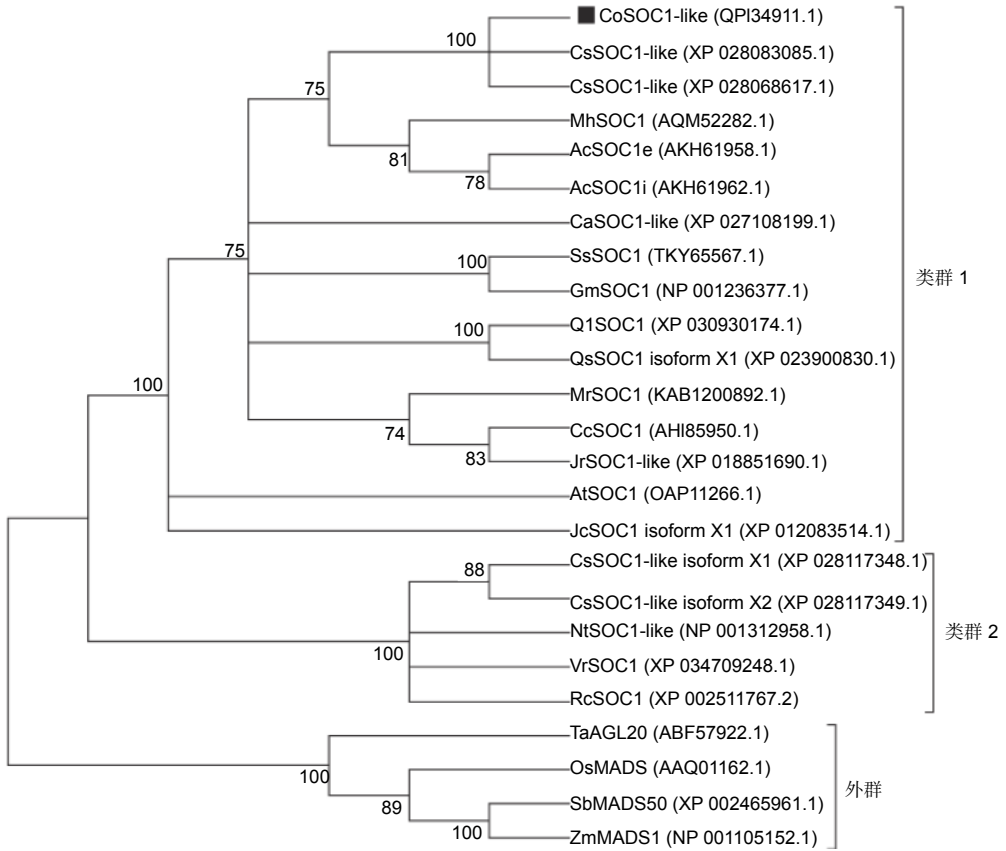
2.6 油茶 CoSOC1-like 蛋白序列系统发育分析

对油茶 CoSOC1-like 蛋白与茶树、地花(*Monotropa hypopitys* Linn.)、猕猴桃、小果咖啡(*Coffea arabica* L.)、鸡血藤(*Spatholobus suberectus* Dunn)、大豆、加州白栎、栓皮栎(*Quercus suber* L.)、杨梅、山核桃、核桃、拟南芥、麻疯树(*Jatropha curcas* L.)、烟草、河岸葡萄(*Vitis riparia* Mchx.)、蓖麻(*Ricinus communis* L.)、小麦(*Triticum aestivum* L.)、水稻(*Oryza sativa* 'Japonica Group')、高粱(*Sorghum bicolor* L.)、玉米等植物 SOC1 及其同

源蛋白,在氨基酸水平构建系统进化树,结果(图 7)表明:以小麦、水稻、高粱、玉米等单子叶植物 SOC1 及其同源蛋白作为外群,可以将比较的其他植物 SOC1 及其同源蛋白分为 2 个分化类群。第 1 个类群中包含油茶 CoSOC1-like、茶树 2 个 SOC1-like、地花 SOC1、猕猴桃的 SOC1e 和 SOC1i、小果咖啡 SOC1-like、鸡血藤 SOC1、大豆的 SOC1、加州白栎 SOC1、栓皮栎 SOC1 isoform X1、杨梅 SOC1、山核桃的 SOC1、核桃 SOC1-like、拟南芥 SOC1、麻疯树 SOC1 isoform X1,自展支持率为 100%;油茶 CoSOC1-like 与茶树 2 个 CsSOC1-like 聚类在同一个进化支上,自展支持率为 100%。第 2 个类群包括茶树 SOC1-like isoform X1 和 X2、烟草 SOC1-like、河岸葡萄 SOC1、蓖麻的 SOC1,自展支持率为 100%。说明油茶的 CoSOC1-like 与茶树 SOC1-like 的亲缘关系最近,与猕猴桃 SOC1e 和 SOC1i、地花 SOC1 的亲缘关系较近,与小果咖啡 SOC1-like、鸡血藤 SOC1、大豆的 SOC1、加州白栎 SOC1 等的亲缘关系较远,与茶树的 SOC1-like isoform X1 和 X2、烟草 SOC1-like、河岸葡萄和蓖麻的 SOC1 的亲缘关系更远。这说明油茶 CoSOC1-like 具有较高种属特性,可为今后研究油茶 CoSOC1-like 基因的功能提供相应的参考。

2.7 油茶 CoSOC1-like 基因的表达分析

为了理解油茶 CoSOC1-like 基因的表达情况,分析了不同生长发育时期的油茶各个器官的基因表达量。对处于开花后期油茶植株的根、茎、叶、营养芽、花瓣、雄蕊、柄托雌(花柄、花托和雌蕊部分)等部位进行了荧光定量 RT-PCR 分析,结果(图 8)表明: CoSOC1-like 基因在所检测的各个部位都有表达,其中,在根和茎中表达水平相对较高,在叶片、营养芽、花瓣、雄蕊和柄托雌等部位表达水平相当。对处于花芽生理分化期油茶的各个器官的荧光定量 RT-PCR 分析结果(图 9)表明: CoSOC1-like 基因在花芽中的相对表达量最多,在根、茎、叶、顶芽、营养芽、幼果中都有一定的相对表达量且表达量相当。对处于雌蕊和雄蕊形成期油茶的各个器官的荧光定量 RT-PCR 分析结果(图 10 A)表明: CoSOC1-like 基因在花芽中的相对表达量最多,在根、茎、叶、营养芽、果皮和种子中的相对表达量差别不大。对花芽处于子房和花药形成期油茶的各个器官的荧光定

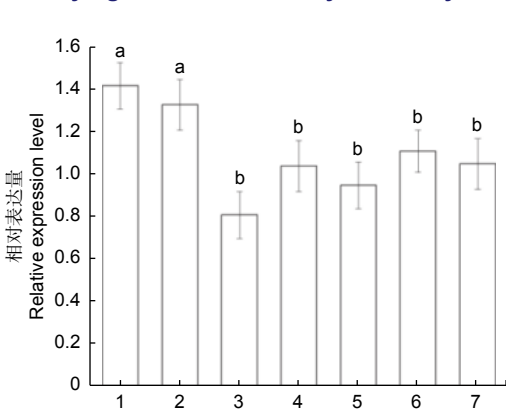


注: Co 油茶; Cs 茶树; Mh 地花; Ac 猕猴桃; Ca 小果咖啡; Ss 鸡血藤; Gm 大豆; Ql 加州白栎; Qs 栓皮栎; Mr 杨梅; Cc 山核桃; Jr 核桃; At 拟南芥; Jc 麻疯树; Nt 烟草; Vr 河岸葡萄; Rc 蓖麻; Ta 小麦; Os 水稻; Sb 高粱; Zm 玉米

Notes: Co: *Camellia oleifera*; Cs: *Camellia sinensis*; Mh: *Monotropa hypopitys*; Ac: *Actinidia chinensis*; Ca: *Coffea arabica*; Ss: *Spatholobus suberectus*; Gm: *Glycine max*; Ql: *Quercus lobata*; Qs: *Quercus suber*; Mr: *Morella rubra*; Cc: *Carya cathayensis*; Jr: *Juglans regia*; At: *Arabidopsis thaliana*; Jc: *Jatropha curcas*; Nt: *Nicotiana tabacum*; Vr: *Vitis riparia*; Rc: *Ricinus communis*; Ta: *Triticum aestivum*; Os: *Oryza sativa* Japonica Group; Sb: *Sorghum bicolor*; Zm: *Zea mays*

图 7 油茶 *CoSOC1-like* 与其他物种 *SOC1* 同源蛋白的系统进化树分析

Fig. 7 Phylogenetic evolutionary tree analysis of *CoSOC1-like* and homologous *SOC1* proteins form other plants

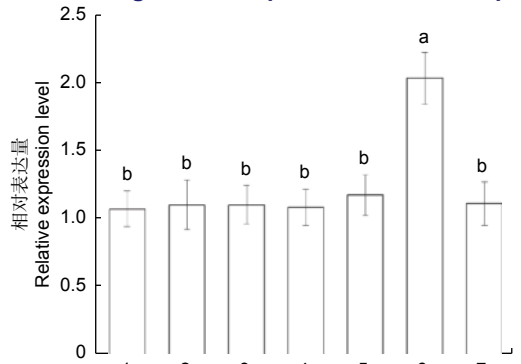


注: 1 根; 2 茎; 3 叶; 4 营养芽; 5 花瓣; 6 雄蕊; 7 柄托雌

Notes: 1 Root; 2 Stem; 3 Leaf; 4 Bud; 5 Petal; 6 Stamen; 7 Parts of Petiole, receptacle and pistil

图 8 油茶开花后期的 *CoSOC1-like* 基因在不同部位的相对表达量

Fig. 8 Relative expression level of *CoSOC1-like* gene of different parts of *Camellia oleifera* in the late flower stage



注: 1 根; 2 茎; 3 叶; 4 顶芽; 5 营养芽; 6 花芽; 7 幼果

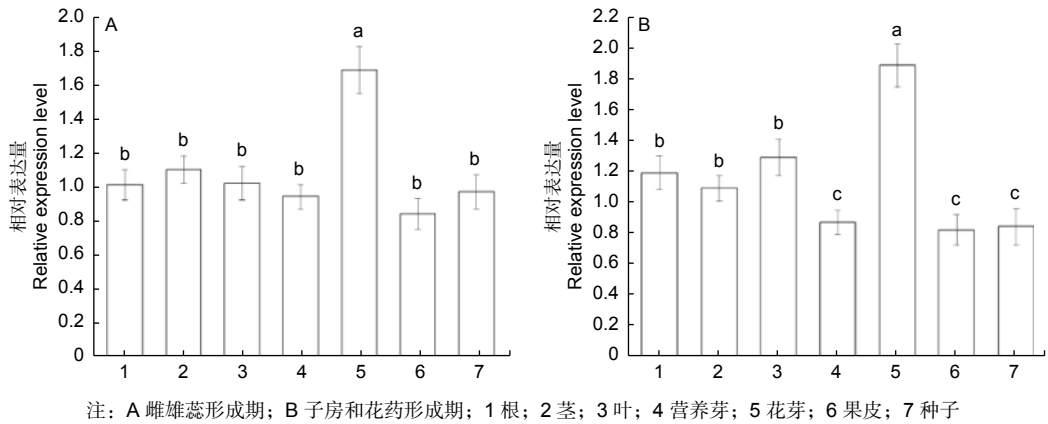
Notes: 1 Root; 2 Stem; 3 Leaf; 4 Terminal bud; 4 Vegetative bud; 5 Flower bud; 6 Young fruit

图 9 油茶花芽生理分化期的 *CoSOC1-like* 基因在不同部位的相对表达量

Fig. 9 Relative expression level of *CoSOC1-like* gene of different parts of *Camellia oleifera* in the stage of physiological differentiation of flower bud

量 RT-PCR 分析结果 (图 10 B) 表明: *CoSOC1-like* 基因在花芽中的相对表达量最多, 在根、茎和叶中的相对表达量较多, 在营养芽、果皮和种子中的相对表达量较少, 并且在营养芽、果皮和种子中有一

定的绝对表达量。这些结果说明, *CoSOC1-like* 基因不仅可以调节生理分化期、雌蕊和雄蕊形成期、子房和花药形成期的花芽生长分化, 也可以调节根、茎、叶、营养芽、果皮和种子的生长发育。



注: A 雌雄蕊形成期; B 子房和花药形成期; 1 根; 2 茎; 3 叶; 4 营养芽; 5 花芽; 6 果皮; 7 种子

Notes: A The stage of pistil and stamen formation; B The stage of ovary and anther formation; 1 Root; 2 Stem; 3 Leaf; 4 Vegetative bud; 5 Flower bud; 6 Peel; 7 Seed

图 10 油茶花芽不同分化期的 *CoSOC1-like* 基因在不同部位的相对表达量

Fig. 10 Relative expression level of *CoSOC1-like* gene of different parts of *Camellia oleifera* in the stages of different differentiation of flower bud

3 讨论

植物从幼年期生长到成熟期以后, 就进入成花阶段。这个阶段包括植物感受开花信号的刺激, 诱导植物从营养生长向生殖生长转化的成花诱导阶段, 花的分生组织分化成花原基和花器官原基的分化阶段以及花器官形成和生长等 3 个阶段^[27-28]。*SOC1* 基因是开花整合子基因, 能够感受光周期途径、自主/春化途径、赤霉素途径、糖类途径和温敏途径等开花刺激途径中的信号, 诱导花分生组织特征的表达和花器官的形成, 促进植物开花^[29-31]。

本试验根据在很多物种中存在的 *SOC1* 同源基因序列, 利用 RACE 方法, 成功地从油茶中克隆出油茶 *SOC1* 同源基因 cDNA 序列, 命名为 *CoSOC1-like*, 基因编码区 654 bp 核苷酸, 编码 217 个氨基酸, GenBank 登陆号为 MT036382。同源蛋白质比对结果表明, 油茶 *CoSOC1-like* 蛋白具有 MADS-box 家族基因的 type II 型结构, 包含 MADS-box、I-domain、K-box 和 C-domain 4 部分, 且在 C 结构域含有 *SOC1*MOTIF 结构, 属于 *SOC1*/TM3 型亚家族基因, 是一个 MADS-box 家族转录因子^[32]。蛋白序列分析表明, 油茶的 *CoSOC1-like* 蛋白没有跨膜螺旋区, 不含信号肽,

含有 SRF 型转录因子结合 DNA 的结构域; 基因瞬时表达表明 *CoSOC1-like* 蛋白定位于细胞核, 与绿竹 *SOC1-like*^[33] 和小麦 *TaSOC1-like* 蛋白^[34] 的定位是一致的, 符合转录因子的结构特点和亚细胞定位特征。油茶 *CoSOC1-like* 蛋白有 31 个磷酸化位点, 它的二级结构和三级结构的多肽链特征与它的一级结构是一致的, 该蛋白的三级结构表现出的紧密又有活性部位的空间结构, 说明 *CoSOC1-like* 蛋白通过磷酸化作用结合 DNA 并激活基因的转录作用。

系统进化树分析表明, 油茶 *CoSOC1-like* 蛋白与茶树 *SOC1-like* 蛋白聚为一枝, 亲缘关系最近, 其次与猕猴桃 *SOC1e* 和 *SOC1i*、地花的 *SOC1* 的亲缘关系较近, 与小果咖啡 *SOC1-like*、鸡血藤和大豆以及拟南芥的 *SOC1* 等的亲缘关系较远, 与茶树的 *SOC1-like isoform X1* 和 *X2*、烟草 *SOC1-like*、河岸葡萄和蓖麻的 *SOC1* 的亲缘关系更远, 说明油茶 *CoSOC1-like* 具有独特的序列特征。由于茶树 *CsSOC1-like* 具有调节茶树开花时间作用^[35], 因此, 推测油茶 *CoSOC1-like* 也能够调节油茶的开花时间。

植物的 *SOC1-like* 基因的功能具有多样性。苜蓿的 *SOC1-like* 基因 *MtSOC1a* 不仅可以促进开花,

还可以促进主茎的伸长^[36];过表达矮牵牛的 *SOC1-like* 基因能够促进烟草的花瓣和叶片的光合作用,并增强烟草对高温的耐受性^[37]。基因表达分析结果表明, *CoSOC1-like* 基因在油茶不同发育阶段的营养器官和生殖器官中都有表达,这一个特点与松树中 *SOC1-like* 基因 *MADS11* 的表达类似^[38]。*CoSOC1-like* 基因在花芽生理分化期、雌雄蕊形成期和子房花药形成期的花芽中的相对表达量最多,在根、茎、叶、顶芽、营养芽、幼果、种子中都有一定的相对表达量,这个特点与核桃 *MADS-like* 基因在花芽中的相对表达量最高相似,但是又不同于核桃的根、茎、叶中相对表达量很低的特点^[39];在子房和花药形成期的果皮和种子中 *CoSOC1-like* 的相对表达量低,这与梅花 *PmSOC1-like* 基因在果实和种子中相对表达量低一致^[40]。在开花后期, *CoSOC1-like* 基因在根和茎中相对表达量较多,在叶片、营养芽、花瓣、雄蕊和柄托雌等部位表达水平相当。说明油茶 *CoSOC1-like* 基因除了参与油茶的花芽分化以外,还可能参与油茶的根、茎、叶和种子等其他器官的生长发育过程。这些研究结果为进一步研究 *CoSOC1-like* 基因在油茶中的成花机理奠定了基础。

4 结论

本研究成功克隆了油茶的 *CoSOC1-like* 基因。*CoSOC1-like* 蛋白定位于细胞核,含有 MADS-box、K-box、I 区和 C 区,属于典型的 MIKC 型蛋白的转录因子; *CoSOC1-like* 蛋白与同属植物茶树的 *CsSOC1-like* 的亲缘关系最近。*CoSOC1-like* 基因在油茶的生理分化期、雌雄蕊形成期、子房和花药形成期的花芽中相对表达量最高,在果皮和种子中的相对表达量较低,说明 *CoSOC1-like* 基因在油茶的花芽分化中有重要的调节作用。本结果为进一步研究油茶 *CoSOC1-like* 基因的功能奠定了理论基础,为后续应用转基因技术调控 *CoSOC1-like* 基因表达水平来调节油茶花期提供了理论依据。

参考文献:

[1] 林秀艳,彭秋发,吕洪飞,等.山茶属油茶组和短柱茶组叶解剖特征及其分类学意义[J].植物分类学报,2008,46(2):183-193.
[2] 袁德义,邹锋,谭晓凤,等.油茶花芽分化及雌雄配子体发育的研究[J].中南林业科技大学学报,2011,31(3):65-70.
[3] 罗帅.施肥对油茶花芽分化及生理生化特性的影响[D].北京:

中国林业科学研究院,2018.
[4] 马腾飞,林新春.植物 *SOC1/AGL20* 基因研究进展[J].浙江农林大学学报,2013,30(6):930-937.
[5] 李贵生,孟征,孔宏智,等.ABC模型与花进化研究[J].科学通报,2003,48(23):2415-2421.
[6] 马辉,张智俊,罗淑萍.植物 MADS-box 基因研究进展[J].生物技术通报,2006,22(6):14-18.
[7] 陈翠翠,马元武,冯永君,等. MADS-box 家族蛋白在植物开花、结实及根瘤形成中的多功能调节作用[J].华北农学报,2008,23(增刊):74-77.
[8] Becker A, Theissen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants[J]. Molecular Phylogenetics and Evolution, 2003, 29(3): 464-489.
[9] Theissen G, Kim J T, Saedler H. Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes[J]. Journal of Molecular Evolution, 1996, 43(5): 484-516.
[10] Parenicova L, de Folter S, Kieffer M, et al. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world[J]. The Plant Cell Research, 2003, 15(7): 1538-1551.
[11] Nakamura T, Song I J, Fukuda T, et al. Characterization of *TrcMADS1* gene of *Trillium camtschaticense* (Trilliaceae) reveals functional evolution of the *SOC1/TM3-like* gene family[J]. Journal of Plant Research, 2005, 118(3): 229-234.
[12] Lamb K S, Irish V F. Functional divergence within the *APETALA3/PISTILLATA* floral homeotic gene lineages[J]. Proceedings of the National Academy of Sciences, 2003, 100(11): 6558-6563.
[13] Zhong X F, Dai X, Xv J H, et al. Cloning and expression analysis of *GmGAL1*, *SOC1* homolog gene in soybean[J]. Molecular Biology Reports, 2012, 39(6): 6967-6974.
[14] Na X F, Jian B, Yao W W, et al. Cloning and functional analysis of the flowering gene *GmSOC1-like*, a putative *SUPPRESSOR OF OVEREXPRESSION CO1/AGAMOUS-LIKE 20* (*SOC1/AGL20*) ortholog in soybean[J]. Plant Cell Reports, 2013, 32(8): 1219-1229.
[15] Zhao S Z, Luo Y Z, Zhang Z L, et al. *ZmSOC1*, a MADS-Box transcription factor from *Zea mays*, promotes flowering in *Arabidopsis*[J]. International Journal of Molecular Sciences, 2014, 15(11): 19987-20003.
[16] 石永春,杨永银,刘卫群.烟草 *SOC1* 基因的克隆和表达分析[J].中国烟草学报,2014,20(2):99-103.
[17] 魏军亚,唐杰,刘国银,等.芒果 *MSOC1* 基因的克隆与表达分析[J].西北植物学报,2015,35(6):1092-1097.
[18] 王祺,蒲媛媛,赵玉红,等.强冬性甘蓝型冬油菜抽薹相关基因 *SVP* 和 *SOC1* 的克隆与表达分析[J].江苏农业学报,2020,

- 36 (5): 22-31.
- [19] Borner R, Kampmann G, Chandler J, *et al.* A MADS domain in the transition gene involved to flowering in *Arabidopsis*[J]. *Plant Journal*, 2000, 24(5): 591-599.
- [20] Dorca-Fornell C, Gregis V, Grandi V, *et al.* The *arabidopsis* *SOC1-like* genes *AGL42*, *AGL71* and *AGL72* promote flowering in the shoot apical and axillary meristems[J]. *Plant Journal*, 2011, 67(6): 1006-1017.
- [21] 沈 威, 滕瑞敏, 李 辉, 等. 茶树 MADS-box 转录因子基因的克隆与非生物胁迫响应分析[J]. *茶叶科学*, 2017, 37 (6): 575-585.
- [22] Zhang Z B, Jin Y J, Wan H H, *et al.* Genome-wide identification and expression analysis of the MADS-box transcription factor family in *Camellia sinensis*[J]. *Journal of Applied Genetics*, 2021, 62(2): 249-264.
- [23] 旦帅男, 胡 颖, 何新华, 等. 金柑*FcSOC1*同源基因的克隆及表达分析[J]. *基因组学与应用生物学*, 2015, 34 (12): 2651-2659.
- [24] 严佳文, 解 璞, 袁启凤, 等. 火龙果开花调控转录因子基因*HpSOC1*的克隆与表达分析[J]. *植物生理学报*, 2018, 54 (10): 1561-1568.
- [25] 刘传娇, 王顺利, 薛憬祺, 等. 牡丹开花调控转录因子基因*PrSOC1*的克隆与表达分析[J]. *园艺学报*, 2014, 41 (11): 2259-2267.
- [26] Wang S, Peng M C, Chen X, *et al.* Molecular cloning and spatiotemporal expression of *APETALA1-like* gene in *Lonicera macranthoides*[J]. *Journal of Genetics*, 2018, 97(5): 1281-1288.
- [27] Wigge P A, Kim M C, Jaeger K E, *et al.* Integration of spatial and temporal information during floral induction in *Arabidopsis*[J]. *Science*, 2005, 309(5737): 1056-1059.
- [28] 傅永福, 孟繁静. 植物的成花决定[J]. *植物生理学通讯*, 1997, 33 (2): 81-87.
- [29] Yoo S K, Chung K S, Kim J, *et al.* *CONSTANS* activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to promote flowering in *Arabidopsis*[J]. *Plant Physiology*, 2005, 139(2): 770-778.
- [30] Lee H, Suh S S, Park E, *et al.* The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*[J]. *Genes & Development*, 2000, 14(18): 2366-2376.
- [31] 黄国文, 韩玉珍, 傅永福. 拟南芥*SUA41*基因的表达和功能分析[J]. *遗传*, 2013, 35 (1): 93-100.
- [32] Zhang Z B, Jin Y J, Wan H H, *et al.* Genome-wide identification and expression analysis of the MADS-box transcription factor family in *Camellia sinensis*[J]. *Journal of Applied Genetics*, 2021, 62(10): 249-264.
- [33] Hou D, Li L, Ma T F, *et al.* The *SOC1-like* gene *BoMADS50* is associated with the flowering of *Bambusa oldhamii*[J]. *Horticulture Research*, 2021, 8: 133-146.
- [34] 吴 琼. 小麦*TaSOC1-like*基因的分离及功能研究[D]. 泰安: 山东农业大学, 2019.
- [35] Gogoi M, Borchetta S, Bandyopadhyay T. Computational identification and analysis of MADS box genes in *Camellia sinensis*[J]. *Bioinformatics*, 2015, 11(3): 115-121.
- [36] Jaudal M, Zhang L L, Che C, *et al.* A *SOC1-like* gene *MtSOC1_a* promotes flowering and primary stem elongation in *Medicago*[J]. *Journal of Experimental Botany*, 2018, 69(20): 4867-4880.
- [37] Ning G G, Yan X, Chen H, *et al.* Genetic manipulation of *Soc1-like* genes promotes photosynthesis in flowers and leaves and enhances plant tolerance to high temperature[J]. *Plant Biotechnology Journal*, 2021, 19(1): 8-10.
- [38] Ma J, Chen X, Song Y, *et al.* MADS-box transcription factors MADS11 and DAL1 interact to mediate the vegetative-to-reproductive transition in pine[J]. *Plant Physiology*, 2021;doi: 10.1093/plphys/kiab250.
- [39] 侯传明, 郑雅文, 王正加, 等. 山核桃*MADS-like*基因的克隆与分析[J]. *浙江农林大学学报*, 2015, 32 (1): 33-39.
- [40] 李玉舒, 杨炜茹, 程堂仁, 等. 梅花*PmSOC1-like*基因的克隆与表达分析[J]. *华北农学报*, 2016, 31 (5): 78-85.

Molecular Cloning and Expression Analysis of Transcription Factor Gene *CoSOC1-like* in *Camellia oleifera*

HUANG Guo-wen, GUAN Tian-qiu, ZHAO Yu-yun, CHEN Mo-lin, LIU Hong-hui

(Department of Chemical and Biological Engineering, Hu'nan University of Science and Engineering, Yongzhou 425199, Hu'nan, China)

Abstract: [Objective] To clone the homologous gene of *SOC1* from *Camellia oleifera* (*CoSOC1-like*) and to analyze its sequence structure, expression pattern and protein evolution. [Method] Total RNA was extracted from the young leaves of three-years-old *C. oleifera*, and the *CoSOC1-like* gene was cloned by using RT-PCR technology and RACE (rapid amplification of cDNA ends) technology. The sequence and expression pattern of *CoSOC1-like* were analyzed using the bioinformatic tools and fluorescence quantitative PCR, respectively. The subcellular localization and evolution of *CoSOC1-like* protein were analyzed using the method of gene transient expression and MEGA7 software, respectively. [Result] The full length cDNA of *CoSOC1-like* contained 654 bases, encoding 217 amino acids, and the relative molecular weight was 24.958 kD and the isoelectric point was 6.8. The Genbank accession number of *CoSOC1-like* is MT036382. The *CoSOC1-like* protein had the structure of MADS-box family transcription factors of plant type II, and there was a *SOC1* MOTIF in C-domain. The *CoSOC1-like* protein had 31 phosphorylatable loci in its amino acid sequence, and its tertiary structure based on secondary structure had obvious active sites, and the transient expression of *CoSOC1-like* gene showed the *CoSOC1-like* protein located in nuclear, which was consistent with the nuclear localization characteristics of transcription factors. The phylogenetic analysis showed that the *CoSOC1-like* protein was clustered in the same evolutionary branch with *SOC1-like* protein of *Camellia sinensis*. Fluorescence quantitative PCR analysis showed that the *CoSOC1-like* gene could be detected in all organs, and there was a maximal relative expression level in flower buds of *C. oleifera*. [Conclusion] The *CoSOC1-like* gene may play an important role in the flower bud differentiation, and participate in the growth and development of other organs such as root, stem, leaf and seed of *C. oleifera*.

Keywords: *Camellia oleifera*; *SOC1-like*; cloning; sequence analysis; expression analysis

(责任编辑: 张 研)