

• 药材与资源 •

灰毡毛忍冬 MADS-box 基因家族 AGL15 基因的克隆、生物信息学和表达分析

易刚强¹, 蔡嘉洛², 朱贻霖², 陈林², 樊悦², 刘峰³, 张亚丽², 欧阳琳¹, 刘湘丹², 童巧珍^{2*}

1. 湖南中医药大学第一附属医院, 湖南长沙 410000

2. 湖南中医药大学, 湖南长沙 410208

3. 湖南省农业科学院, 湖南长沙 410000

摘要: 目的 基于前期对灰毡毛忍冬 *Lonicera macranthoides* 转录组测序结果, 从灰毡毛忍冬 RNA 中克隆灰毡毛忍冬 MADS-box 基因家族蕾期延长相关基因 AGL15, 并对其进行生物信息学及表达分析。方法 对灰毡毛忍冬总 RNA 进行反转录酶链式反应 (reverse transcription-PCR, RT-PCR) 和 cDNA 末端快速扩增反应 (rapid amplification of cDNA ends, RACE), 获得完整的开放阅读框 (ORF)。运用生物信息学的方法对该序列进行同源性分析和相似性比较, 预测编码蛋白, 并对其进行各种理化性质分析。运用半定量 PCR 测得该基因在灰毡毛忍冬不同部位的表达情况。结果 获得 AGL15 基因, ORF 长 795 bp, 编码 264 个氨基酸, 与葡萄 *Vitis vinifera* 中 MADS-box 基因家族 AGL15 基因相似性最高, 且包含 MADS 和 K-box 的保守序列。AGL15 蛋白无跨膜区域, 定位于细胞质中, 在灰毡毛忍冬各部位均有表达。结论 首次从灰毡毛忍冬中克隆得到可能控制蕾期延长的基因, 分析了其在灰毡毛忍冬不同部位表达差异性, 为灰毡毛忍冬蕾期延长研究提供参考。

关键词: 灰毡毛忍冬; 蕾期延长; 基因克隆; 生物信息学分析; 基因表达

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2016)04-0640-08

DOI: 10.7501/j.issn.0253-2670.2016.04.019

Cloning of MADS-box gene AGL15 from *Lonicera macranthoides* and analysis on its bioinformatics and expression

YI Gang-qiang¹, CAI Jia-luo², ZHU Yi-lin², CHEN Lin², FAN Yue², LIU Feng³, ZHANG Ya-li², OUYANG Lin¹, LIU Xiang-dan², TONG Qiao-zhen²

1. The First Hospital of Hunan University of Chinese Medicine, Changsha 410000, China

2. Hunan University of Traditional Chinese Medicine, Changsha 410208, China

3. Hunan Academy of Agricultural Sciences, Changsha 410000, China

Abstract: Objective To clone the MADS-box gene denoted AGL15 from total RNA of *Lonicera macranthoides* based on previous transcriptome sequencing data and analyze the bioinformatics and expression of the gene. **Methods** The gene containing intact open reading frame (ORF) was cloned by reverse transcription-PCR (RT-PCR) and rapid amplification of cDNA ends (RACE). The similarity comparison and homology analysis on the sequence were carried out using bioinformatic method, the coding protein was predicted and the physicochemical properties were analyzed. The expression of the gene in different locations of *L. macranthoides* was determined by semiquantitative PCR using gene-specific primers. **Results** The Lm-AGL15 gene, containing a 795 bp ORF that encoded 264 amino acids, was cloned. The deduced protein sequence had the most similarity to the AGL15 in *Vitis vinifera* and exhibited two conserved motifs (MADS and K-BOX). Without transmembrane domain, Lm-AGL15 was located in cytoplasm and expressed only in each part. **Conclusion** For the first time from *L. macranthoides* cloning could prolong the period of bud of gene, analysis of the gene expression in different parts of *L. macranthoides* which would provide a reference for the study of prolonging *L. macranthoides* bud stage.

Key words: *Lonicera macranthoides* Hand. -Mazz.; bud stage prolonging; clone; bioinformatic analysis; gene expression

收稿日期: 2015-09-21

基金项目: 国家自然科学基金资助项目 (81203007); 湖南省自然科学基金资助 (13JJ9010); 湖南省教育厅科技创新平台 (14K071); 湖南省研究生科研创新项目 (CX2014B368); 湖南省“中药学”重点学科建设项目资助 (湘教通[2011]76号; zy201403, zy201505)

作者简介: 易刚强 (1965—), 男, 副教授, 主要从事中药资源及中药新药研究。E-mail: ygg8228@163.com

*通信作者 童巧珍 (1971—), 女, 教授, 主要从事中药资源与质量研究。E-mail: qztong88@126.com

灰毡毛忍冬 *Lonicera macranthoides* Hand. -Mazz. 为忍冬科植物。干燥花蕾或带初开的花为中药^[1], 具有清热解毒, 凉散风热的功效, 临床上广泛用于治疗痈肿疔疮、喉痹、丹毒、热毒血痢、风热感冒、温热发病等症, 也可作为提取绿原酸或挥发油的原料, 长期以来如何解决灰毡毛忍冬蕾期短、采摘不及时是灰毡毛忍冬生产和应用中一个棘手的问题。常规品种的灰毡毛忍冬花蕾期长仅 3~4 d, 开花后第 2 天即开始凋谢^[2], 采收时间仓促, 有时与农时相冲突来不及采摘, 直接影响灰毡毛忍冬药材的产量与质量。杨苗等^[3]和耿世磊等^[4]对灰毡毛忍冬不同发育阶段花结构与绿原酸量变化关系进行了研究, 结果表明花中细胞分化较早绿原酸量较高, 分化越成熟量越低。以上说明, 灰毡毛忍冬绿原酸的富集在蕾期较多, 其量随着花冠展开而下降。花期太短不仅对采摘灰毡毛忍冬药材带来麻烦, 还将直接影响灰毡毛忍冬的产量和品质。因此, 灰毡毛忍冬蕾期长短是生产中亟待解决的关键瓶颈问题。

20 世纪 90 年代末, 发现了一株灰毡毛忍冬自然变异株, 表现为花蕾整齐, 花蕾期长, 花冠不开放, 花蕾初期呈青绿色, 逐渐转为绿白色和白色, 20 多天才开始凋谢等性状。该变异性状为灰毡毛忍冬的采摘、质量及农事操作等提供了有力的保障, 于 2004 年 6 月通过省级鉴定, 命名为湘蕾灰毡毛忍冬, 被广泛引种栽培。湘蕾灰毡毛忍冬具有以上诸多优点, 为灰毡毛忍冬的开发利用开创了新的纪元, 也因此促使了诸多研究学者对变异株变异的根本原因进行探索, 期望能找到调控优良性状的原因, 为其他花类药材、食品、切花等领域指引新的方向。

目前大量不同植物种类的 MADS 家族基因已经被成功克隆, MADS-box 基因是真核生物中一类重要的转录调控因子, 在生长发育调控和信号传导中发挥着重要作用, 同时在植物中参与花器官的发育, 开花时间的调节, 但灰毡毛忍冬 MADS-box 相关基因的研究并不多^[5]。因此, 研究灰毡毛忍冬 MADS-box 家族基因对于阐明灰毡毛忍冬蕾期控制相关分子机制有重要意义^[6]。

本课题组前期对常规灰毡毛忍冬和湘蕾灰毡毛忍冬转录组进行了测序和分析, 获得多条差异明显基因, 同时有数条预测为有控制发育功能的编码序列 (coding sequence, CDS) 的 MADS-box 家族成

员, 其中 Unigene17356 基因差异明显, 根据其序列设计引物, 对湘蕾型灰毡毛忍冬总 RNA 进行逆转录 (reverse transcription) 反应和 cDNA 末端快速扩增技术 (rapid amplification of cDNA ends, RACE), 获得完整的开放阅读框 (ORF), 并命名为 *Lm-AGL15*, 首次成功从灰毡毛忍冬总 RNA 中克隆到可能参与控制蕾期延长的基因 *AGL15*。运用生物信息学的方法对该序列进行同源性分析, 理化性质分析, 并预测 *Lm-AGL15* 编码蛋白。最后通过半定量 PCR 技术对 *Lm-AGL15* 基因在灰毡毛忍冬不同部位的表达进行差异对比分析, 为后续进一步通过功能分析找出控制湘蕾灰毡毛忍冬花蕾期长度、抑制花冠展开的系列功能基因, 为将来进一步通过转基因技术将这些功能基因转入其他品种的灰毡毛忍冬、金银花或其他药用、食用、切花、香精香料原料等植株中, 培育出花蕾期长、或花冠不展开的新优良品种打下前期基础, 具有广泛的开发和应用前景, 有望带来巨大的社会效益和经济效益。

1 材料

样品均采自湖南中医药大学药用植物培植基地, 经湖南中医药大学药学院周日宝教授鉴定为湘蕾型灰毡毛忍冬与常规型灰毡毛忍冬 *Lonicera macranthoides* Hand. -Mazz.。立即用锡箔纸包裹置于液氮中保存。

Biospin 多糖多酚植物总 RNA 提取试剂盒 (杭州博日科技有限公司)、琼脂糖凝胶回收试剂盒、Trans2K DNA Marker (全式金生物科技有限公司); EASY-T1 Cloning 试剂盒、RevertAidTM First Strand cDNA Synthesis Kit (反转录试剂盒) (北京智杰方远科技有限公司); 3'-RACE 试剂盒、5'-Full RACE Kit with TAP 试剂盒、Ex TaqDNA 聚合酶、Ex TaqHS DNA 聚合酶 (Takara 公司); 其余试剂均为国产分析纯; 扩增所用引物及获得片段的序列测定委托上海生工生物工程有限公司完成 (表 1)。

2 方法

2.1 灰毡毛忍冬总 RNA 的提取与 cDNA 第一条链的合成

参照 Biospin 多糖多酚植物总 RNA 提取试剂盒说明书, 以 1% 的琼脂糖凝胶电泳检测 RNA 的完整性, 紫外分光光度计检测 RNA 纯度。以提取的 RNA 样品为模板, 使用北京智杰方远科技有限公司的反转录 cDNA 第一链合成试剂盒合成第一条链, 获得 cDNA 样品于 -20 °C 贮存备用^[6]。

表 1 引物序列

Table 1 Primer sequences

引物	序列 (5'-3')	引物用途
<i>Lm</i> -AGL15F	ATGGGTCGAGGGAAGATTGAG	核心片段上游引物
<i>Lm</i> -AGL15R	TTAACATTTCAATTTTCATTTTC	核心片段下游引物
AGL15GSP1	GATTACGCCAAGCTTCAAGGCAGAGGAACCTAAGGAGGTGGAT	3'-RACE 上游特异性引物
AGL15GSP2	GATTACGCCAAGCTTCCACCTCCTTAGGTTCTCTGCCTTGT	5'-RACE 下游特异性引物

2.2 灰毡毛忍冬 *Lm*-AGL15 核心片段的克隆

根据前期灰毡毛忍冬转录组测序获得的长度为 765 bp Unigene17356 序列, 设计特异性引物, 上游引物 *Lm*-AGL15F, 下游引物 *Lm*-AGL15R。以“2.1”项方法获得的 cDNA 为模板进行 PCR 扩增, 反应体系 25 μ L, 其中引物 *Lm*-AGL15F 和 *Lm*-AGL15R 各 2.5 μ L, 10 \times Ex Taq 缓冲液 2.5 μ L, 2.5 mmol/L dNTP 1 μ L, 模板 cDNA 3 μ L, Ex TaqDNA 聚合酶 1 μ L, 补 ddH₂O 12.5 μ L 至 25 μ L。反应条件为 95 $^{\circ}$ C 预变性 5 min, 94 $^{\circ}$ C 变性 50 s, 53 $^{\circ}$ C 退火 50 s, 72 $^{\circ}$ C 延伸 1 min, 34 个循环后 72 $^{\circ}$ C 延伸 10 min。经 1.0% 的琼脂糖凝胶电泳检测后, 回收核心片段, 连接和转化过程参考全式金 pEASY-T1 Cloning Kit 试剂盒说明书, 转化 Trans1-T1 感受态细胞。

菌液 PCR 体系: 25 μ L 体系含 1 μ L 菌液, 2 μ L dNTPs, 1 μ L Taq 酶, 2.5 μ L 10 \times Buffer, 引物各 1 μ L, 17.5 μ L ddH₂O。94 $^{\circ}$ C 预变性 7 min; 94 $^{\circ}$ C 变性 1 min, 53 $^{\circ}$ C 退火 1 min, 72 $^{\circ}$ C 延伸 1 min, 循环扩增 34 次; 72 $^{\circ}$ C 延伸 10 min。将 PCR 鉴定的阳性克隆, 按质粒提取试剂盒方法提取质粒 DNA, 送上海生工生物工程有限公司测序。

2.3 灰毡毛忍冬 *Lm*-AGL15 3' 末端序列的获得

使用 TaKaRa 3'-RACE 试剂盒将灰毡毛忍冬总 RNA 反转录为 cDNA, 以 AGL15GSP1 进行灰毡毛忍冬 3' 端 PCR 扩增。

PCR 50 μ L 反应体系: 2 μ L 反转录反应液, 8 μ L 1 \times cDNA Dilution Buffer II, 2 μ L AGL15GSP1 (10 μ mol/L), 2 μ L 3'-RACE Primer (10 μ mol/L), 5 μ L 10 \times PCR Buffer II (Mg²⁺ Free), 3 μ L MgCl₂ (25 mmol/L), 0.5 μ L Taq 酶 (2.5 U/ μ L), 28.5 μ L ddH₂O。94 $^{\circ}$ C、5 min, 94 $^{\circ}$ C、55 s, 55 $^{\circ}$ C、1 min, 72 $^{\circ}$ C、90 s; 共 32 个循环; 72 $^{\circ}$ C、10 min。

PCR 产物经凝胶回收、连接和转化, 挑取阳性克隆送上海生工生物工程有限公司测序。

2.4 灰毡毛忍冬 *Lm*-AGL15 5' 末端序列的获得

使用 TaKaRa 5'-RACE 试剂盒将灰毡毛忍冬总 RNA 反转录为 cDNA, 以 AGL15GSP2 进行灰毡毛忍冬 5' 端 PCR 扩增。

PCR 50 μ L 反应体系: 2 μ L 反转录反应液, 9 μ L 1 \times cDNA Dilution Buffer II, 3 μ L 10 \times PCR Buffer II (Mg²⁺ free), 3 μ L MgCl₂ (25 mmol/L), 0.5 μ L Taq 酶 (2.5 U/ μ L), 2 μ L AGL15 GSP2 Primer (10 μ mol/L), 2 μ L 5'-RACE Outer Primer (10 μ mol/L), 28.5 μ L ddH₂O。94 $^{\circ}$ C 预变性 4 min; 94 $^{\circ}$ C 变性 40 s、55 $^{\circ}$ C、35 s, 72 $^{\circ}$ C、90 s, 35 个循环; 72 $^{\circ}$ C 延伸 10 min^[7-8]。

取 PCR 反应液各 10 μ L, 使用 1.0% 的琼脂糖凝胶进行电泳检测, 切胶回收电泳条带, 连接、转化, 挑取阳性克隆送上海生工生物工程有限公司测序。

2.5 灰毡毛忍冬 *Lm*-AGL15 基因 cDNA 全长序列的拼接

根据已测定的核心片段、5' 和 3' 端序列及其重叠区域, 使用 Bioedit 软件进行序列分析及拼接, 获得 *Lm*-AGL15 基因 cDNA 全长序列。将获得的序列上传到 NCBI 的 ORF Finder 平台 (<http://www.ncbi.nlm.nih.gov/gorf/>), 确定是否获得全长 ORF。将确定的 ORF 片段上传到 NCBI 的 CD-search (<http://www.ncbi.nlm.nih.gov/Structure/cdd>), 对预测的氨基酸序列进行保守区域搜索, 验证该 ORF 片段是否完整^[6-10]。

2.6 灰毡毛忍冬 *Lm*-AGL15 基因的表达水平分析

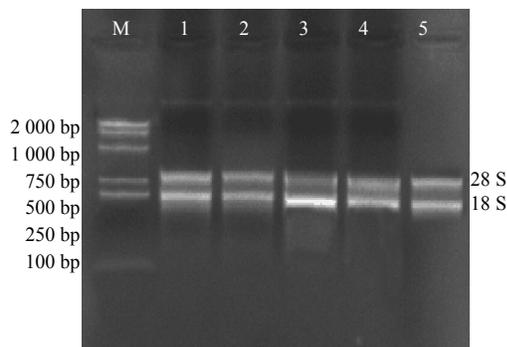
根据全长序列, 设计可扩增 *Lm*-AGL15 基因全长的特异性引物, 预计扩增长度为 760 bp。分别按“2.1”项方法提取 RNA 后反转录为 cDNA。利用扩增 *Lm*-AGL15 基因, 反应条件同“2.2”项。取 10 μ L PCR 产物在 1.0% 琼脂糖凝胶上电泳检查。

3 结果与分析

3.1 灰毡毛忍冬总 RNA 的提取

使用 Biospin 多糖多酚植物总 RNA 提取试剂盒

法提取同一批 5 份 (1~5) 灰毡毛忍冬总 RNA, 非变性琼脂糖凝胶电泳结果如图 1 所示, 总 RNA 无 DNA 残留和蛋白质杂质, 电泳结果显示 28 S、18 S 条带清晰, 无拖尾现象。说明 RNA 具有良好的完整性。紫外分光光度法测定 $A_{260\text{ nm}}/A_{230\text{ nm}}$ 值在 2.11, $A_{260\text{ nm}}/A_{280\text{ nm}}$ 值在 2.01, 说明 RNA 纯度高, 能满足后续实验要求。



M-Marker 1~5-常规灰毡毛忍冬花蕾 RNA
M-Marker 1—5-bud RNA of Xianglei *L. macranthoides* Hand.-Mazz.

图 1 灰毡毛忍冬总 RNA 电泳图

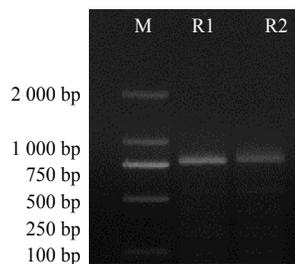
Fig. 1 Electrophoresis of total RNA in *L. macranthoides*

3.2 灰毡毛忍冬 *Lm*-AGL15 基因的获得

前期转录组测序获得长 765 bp 的 Unigene17356 序列, 经 BLAST 分析后发现其可能为 MADS-box 家族基因, 根据该序列设计引物, 扩增出目的核心片段, PCR 产物使用 1.0% 的琼脂糖凝胶电泳, 结果见图 2, 在 750 bp 左右有一条明显的亮带, 扩增产物大小与预测片段大小基本相符, 初判为目的基因片段。使用胶回收试剂盒回收目标片段, 将回收的产物连接后转化, 取阳性克隆子进行测序, 测序结果显示该序列的长度为 764 bp, 将测序结果与 NCBI 数据库进行 BLAST 对比, 该片段与数据库中葡萄 *Vitis vinifera* L. AGL15 相似性很高, 相似度为 80%, 证明该核心片段为灰毡毛忍冬 AGL15 基因的核心片段。

利用合成的 3'-RACE 特异性引物, 进行 PCR 扩增。3'-RACE PCR 的扩增结果在 800 bp 有一条亮带 (图 3)。将 PCR 产物切胶回收, 连接 pEASY-T1 载体后转化, 取阳性克隆子进行测序。

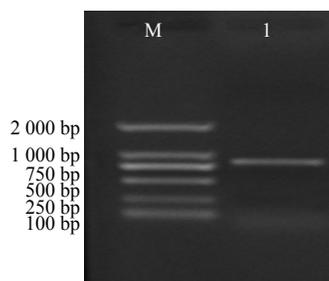
利用设计合成的 5'-RACE 特异性引物, 进行 PCR 扩增。5'-RACE 扩增结果在 500 bp 左右有一条亮带 (图 4)。将 PCR 产物回收, 连接后转化, 取阳性克隆子进行测序。



M-Marker R1~R2-AGL15 基因核心片段 PCR 产物
M-Marker R1—R2-PCR product of AGL15 gene core fragment

图 2 AGL15 基因核心片段的 PCR 扩增结果

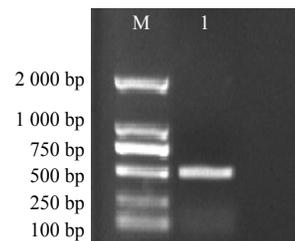
Fig. 2 PCR result of core fragment of AGL15 gene



M-Marker 1-AGL15 基因 3'-RACE PCR 产物
M-Marker 1-PCR product of AGL15 by 3'-RACE

图 3 AGL15 基因 3'-RACE PCR 结果

Fig. 3 3'-RACE PCR results of AGL15 gene



M-Marker 1-AGL15 基因 5'-RACE PCR 产物
M-Marker 1-PCR product of AGL15 by 5'-RACE

图 4 AGL15 基因 5'-RACE PCR 结果

Fig. 4 5'-RACE PCR results of AGL15 gene

使用 Bioedit 软件将上述测序所得到的核心目的序列与 3'端、5'端序列进行拼接, 得到基因全长序列, 并对其验证。设计全长引物, 进行 PCR, 回收、转化及测序, 测序结果与拼接结果一致, 则说明已经扩增得到灰毡毛忍冬 AGL15 基因的 cDNA 全长序列, 验证后将该基因命名为 *Lm*-AGL15 基因, 获得基因登录号为 KR028478。

3.3 灰毡毛忍冬 *Lm*-AGL15 基因及其编码氨基酸序列分析

Lm-AGL15 基因全长 1 399 bp, 不含内含子序列, 编码 265 个氨基酸 (GenBank 登录号为 KR028478,

结果见图 5。利用 ExPASy Proteomics Server 的在线软件 ProtParam (<http://www.cn.expasy.org/tools/protparam.html>) 对 *Lm*-AGL15 蛋白的理化性质进行预测分析, 推测其分子式为 $C_{1324}H_{2161}N_{376}O_{420}S_{12}$, 相对分子质量为 30 493, 等电点为 6.39, 为酸性蛋白, 正电残基 (Arg+Lys) 为 40, 负电残基 (Asp+Glu) 为 42。脂肪系数为 85.28, 其半衰期在体外哺乳动物网织红细胞为 100 h, 在酵母体内大于 20 h, 在大肠杆菌体内大于 10 h。不稳定系数为 52.67, 为不稳定蛋白。对该氨基酸序列进行亲疏水性分析, 结果如图 6 所示, 横坐标表示氨基酸残基的序号, 纵坐标表示残基的疏水、亲水特性, 正值为疏水, 负值为亲水。总平均疏水指数 (GRAVY) 为 -0.201, 预测其为亲水性蛋白。通过 TMHMM 软件分析得知,

Lm-AGL15 编码蛋白质无跨膜区域, 全部位于膜外。通过 Target IP 服务器的分析, 判断 *Lm*-AGL15 蛋白定位于细胞质中。通过 SMART 服务器分析, *Lm*-AGL15 蛋白第 1~60 位和第 74~162 位是 2 个高度保守的结构功能域——MADS 和 K-box, 即 MADS-box 家族成员共有的典型结构域^[10]。

3.4 灰毡毛忍冬 *Lm*-AGL15 基因编码的蛋白质结构与功能预测

运用 Predictprotein 软件预测结果显示, 软件预测 *Lm*-AGL15 蛋白的二级结构表明该蛋白含有 123 个 α 螺旋, 占 46.42%; 17 个 β 折叠, 占 6.03%; 125 个无规卷曲, 占 47.55%。SWISS-MODEL 软件三级结构预测结果显示 *Lm*-AGL15 基因编码蛋白质的三级结构见图 7。

```

CTAATCGACTCACTATAGGGCAAGCAGTGGTATCAA CGCAG AGTACATGGGGAATTTTAAAGCTATATAAAAATACACA CGTACATATA          90
TTTATATAGAGTTTGTATATAGATAGAAACCCCTCAAGACAAATTGGGGAATTTTGTAAATTTCTTTGGGTTTCTATCATATCTTTGAA          180
TTGGGTTTGTGGGTTCTGTTTATAAATGATTTTGTAGTAAATGGTGGTGGGGAAGATTGAGATAAAGAAGATCGAGAATGTGAATAGCAG          270
MGRGKIEIKKIEINVNSRQ
GCAAGTTACATTTCTCAAAGAGAAGAGCTGGGTTGCTTAAGAAAGCCCGTGAACCTGCGGTTCTTTGTGATGCTGAAGTTGGGGTTAATTGT          360
VTFSSKRRRAGLLKKARELAVLCLDAEVGVIVF.
TTTTCAAATACAGGGAGGCTCTTTGAATTTTCAAGTTGTAGTATGCGAGAAATCTTTCAAGATACAACAGGTGTTTAGATCCCCAGA          450
SNTGRLRFESSCSMQRILSRYNRC LDSPET.
GACCCAGCCGCTTTAACTAAGGCAAGGCAAGGAACTTAAAGAGGTTGATTTCTAAAGAGGAATAGCAAGCTCAAGTCAAAACA          540
Q T A L T N Y K A E E P K E V D I L K E E I A K L K S K Q L.
ATTGCAGTTTTTGGGTAAGGATCTAAATGGCCCTGAGCTTAAAGAAATTCACCAGCTGGAGCAGCTACTAATGAAGGATTGTTTTCTGT          630
Q F L G K D L N G L S L K E I H Q L E Q L L N E G L F S V K.
GAAAGAGAGGAAGGACAACTACTAATGGAAGAACTAGAGCACTCAAGATTACAGGAA CAGCGTGTATTGTGGAAAATGAGGCTTTGCG          720
E R K E Q L L M E E L E H S E L Q E Q R V I V E N E A L R R
CAGAGAGGTTGAAGAGCTTCGGCAATTTGGTACCATCGACTGAAA CCTTACCTTGAATACTACCTGTG CAGGACAATAA          810
E V E L R H L V P S T E N S V Q P Y L E Y Y P V Q D N N Y.
TTATTCTATGCCAAAACTGTTCCCTGAGTCCAGCAGAGCTGTGTAACCTCATAGACGCAATAAA GATTGAGACA CCACTTGCACATT          900
S M P K H C S L S P D R V C N S I D D N K D S D T T L H L G.
GGGGCTTCCATATGGTTTGTATCGAAAGAGGAAGACTCCCGAGGGAGAAACTCACTCCAGCACCTCTGAGAGTCAAATGGGCATAGAAAAC          1090
L P Y G L Y R K R K T L P E G E T H S P T S E S Q M G I E T S.
AAGTGAATGAAATGAAATGTAATGTAATCTTGGAGCTTTTTTGTGAAAGACCAATCGGCTCAAAAATTTTAAACAATTTTAGATG          1180
E M K L K C
ACTTTCAAGATGATGAAAAAAAAGACACGGTGGAGAATTTTCTTTTGGGGATGTGTTAAAGGGTGTGGATATGGGAAAAAAGGGT          1270
AAATATAGTTTAAAGTAGTTCTCTTAAAGGTGATGTTCTGTTGAATGTTGAGACTATGAAATATGGGCATATGCAAACACTACAGGCT          1360
TTTTCTGTGTTTTGATGTTCTGTTTGGCTGCACAGTAGAAATACAGAATGATTTTCAAGTGTATTCTTCAAACACAGTCTCAITTCCT          1450
TTTTGGACATCAITTAGCATTATAACGTAGAAAAA          1497

```

图 5 *Lm*-AGL15 基因核苷酸序列
Fig. 5 Nucleotide sequences of *Lm*-AGL15 gene

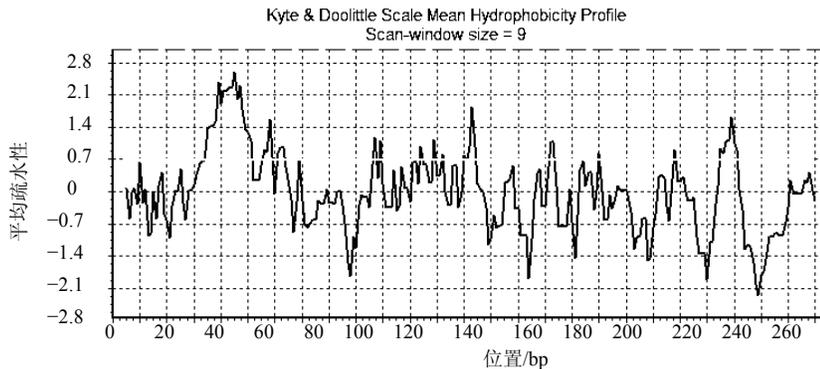


图 6 *Lm*-AGL15 编码蛋白的亲水性分析
Fig. 6 Hydrophily analysis on *Lm*-AGL15-coding protein



图 7 灰毡毛忍冬 AGL15 蛋白三级结构同源建模

Fig. 7 Homologous modeling of tertiary structure of AGL15 protein in *L. macranthoides*

3.5 灰毡毛忍冬 AGL15 基因编码蛋白质的相似性分析

Lm-AGL15 基因编码的氨基酸序列经 NCBI 的 BLASTp 分析, 获得与其相似性较高的蛋白, 其中与麻风树 *Jatropha curcas* L.、葡萄 AGL15 基因

相似性最高, 达 67%。使用 GENEDOC 软件将该氨基酸序列与比对上的前 5 条不同物种的蛋白质序列进行比对, 结果如图 8 所示, 相似性由高到低分别是葡萄 AGL15 基因 [AGL15 isoform X1, 731414618], 麻风树 *Jatropha curcas* L. AGL15 基因 [AGL15, 802556094], 甜橙 *Citrus sinensis* (L.) Osbeck. AGL15 基因 [AGL15-like, 568858840], 梅 *Prunus mume* Apricot. AGL15 基因 [AGL15, 645251155], 野草莓 *Fragaria vesca* Duch. AGL15 基因 [AGL15, 470143144]。分析发现 *Lm*-AGL15 基因编码的氨基酸序列具有 Mads-box 家族特有的 MADS 与 K-box 保守序列, 可进一步确定其为 MADS-box 家族基因^[11-15]。

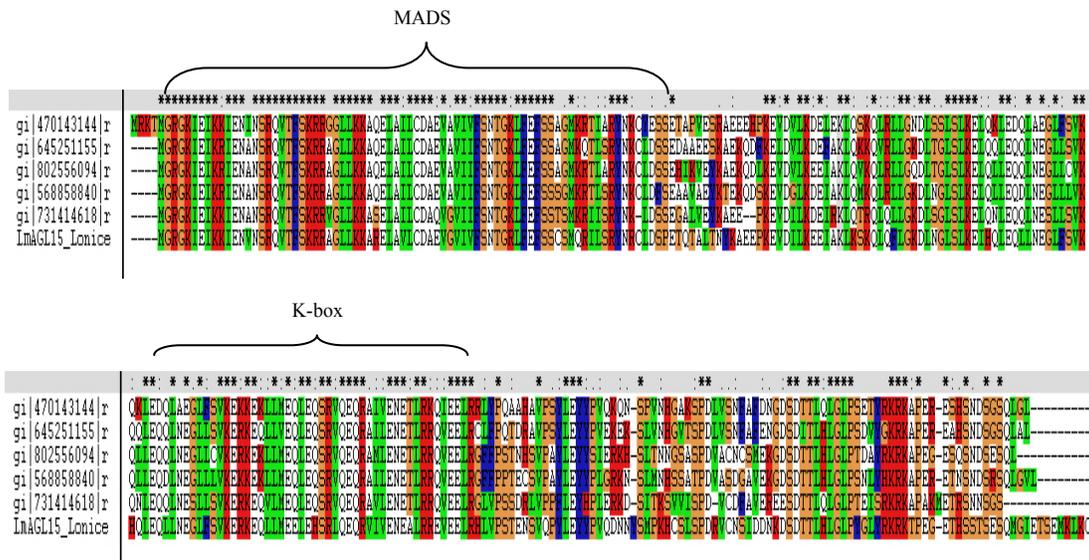


图 8 *Lm*-AGL15 编码蛋白与其他物种 AGL15 基因编码蛋白的相似性分析

Fig. 8 Similarity analysis on *Lm*-AGL15-coding protein and other AGL15 proteins

3.6 不同物种来源 AGL15 基因的系统进化分析

为了分析灰毡毛忍冬 *Lm*-AGL15 基因与其他 AGL15 基因的进化关系, 将获得的 *Lm*-AGL15 基因推导的氨基酸序列与其他植物中已报道的属于 AGL15 亚家族的成员进行比对, 构建系统发育树状图 (图 9), 其中包括葡萄、龙眼 *Dimocarpus longan* Lour.、白梨 *Pyrus bretschneideri* Rehd.、雷蒙德式棉 *Gossypium raimondii* Ulbr.、苹果 *Malus pumila* Mill.、梅、麻风树、橙子以及野草莓。

从图 9 中可以分析得出 AGL15 亚家族基因大致归为 3 类, *Lm*-AGL15 蛋白序列与其他被子植物门双子叶植物纲的 AGL15 基因有较近的亲缘关系归为一类, 可确定该基因编码 AGL15 亚家族基因

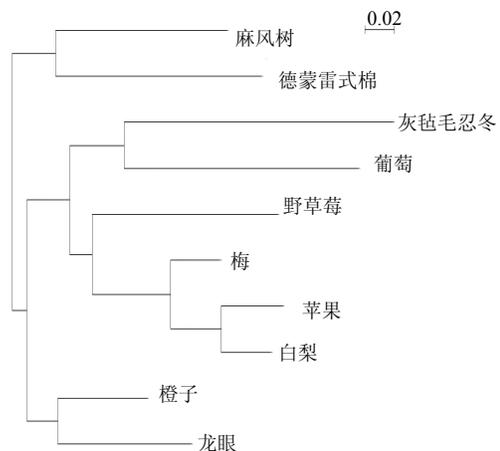


图 9 *Lm*-AGL15 基因与已知的 AGL15 基因的进化树分析
Fig. 9 Phylogenetic tree for *Lm*-AGL15 and these from other species

类蛋白。*Lm*-AGL15 和葡萄 AGL15 的距离最近，而 *Lm*-AGL15 和龙眼 AGL15 的距离最远^[16-19]。

3.7 灰毡毛忍冬 *Lm*-AGL15 基因的表达

PCR 结果显示湘蕾型灰毡毛忍冬植株中的不同部位，均能检测到符合预期大小的 *Lm*-AGL15 基因片段，同时常规型灰毡毛忍冬花蕾中也检测到 *Lm*-AGL15 基因片段，湘蕾型灰毡毛忍冬植株中花蕾亮度最高，茎、叶亮度基本一致。说明 *Lm*-AGL15 基因主要在花器官部位表达，同时湘蕾型灰毡毛忍冬植株花蕾中 *Lm*-AGL15 基因的表达远高于常规型灰毡毛忍冬花蕾中的表达（图 10）。

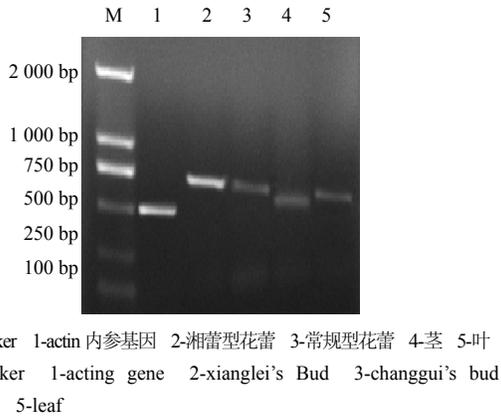


图 10 灰毡毛忍冬不同品种及部位中 *Lm*-AGL15 基因的表达
Fig. 10 Expression of *Lm*-AGL15 genes in different parts and varieties of *L. macranthoides*

4 讨论

灰毡毛忍冬作为山银花主要药材来源，其生物活性物质高、产量高的特点被人们广泛熟知与接受，但如何解决灰毡毛忍冬蕾期短、采摘不及时是灰毡毛忍冬生产和应用中一个棘手的问题。因此克隆并研究湘蕾灰毡毛忍冬控制蕾期长短的相关基因，对于灰毡毛忍冬药用价值的开发及利用具有积极的促进作用。本研究通过对灰毡毛忍冬 MADS-box 基因家族中的 *Lm*-AGL15 基因的研究，为从基因水平探索灰毡毛忍冬的生理生化研究奠定基础。

MADS-box 基因参与植物花发育不同环节的调控，该基因家族可能形成复杂的复合物来调控花器官的特征以及控制花发育的过程。研究通过同源基因克隆，利用多聚酶链式反应(PCR)结合 3'-RACE、5'-RACE 技术，克隆了灰毡毛忍冬 MADS-box 基因 AGL15 的 cDNA 全长序列，命名为 *Lm*-AGL15 基因，将其提交到 GenBank 数据库，获得基因登录号为 KR028478。该基因的核苷酸序列全长为 1 399 bp，该基因序列的第 1~222 位核苷酸为基因的

5'UTR，第 1 018~1 399 位核苷酸为基因的 3'UTR，含有 16 bp 的 PolyA 结构，其中包括一个 795 bp 的完整开放阅读框 (ORF) 编码 264 个氨基酸。同时研究表明，*Lm*-AGL15 基因编码蛋白的二级结构以 α -螺旋为主，*Lm*-AGL15 在结构上有很高的同源性，与 MADS-box 家族基因相比都含有 MADS 与 K-box 两个保守的功能区域，且分成 4 个区，N 端是 60 个氨基酸的 MADS 结构域，蛋白的中部为 69 个氨基酸 K-box，MADS 区和 K 区之间为 27 个氨基酸的 I 区，C 端区域包含 57 个氨基酸，该基因还含有 MADS-box 基因 D 类基因特有的 AG motif I 和 AG motif II，表明 *Lm*-AGL15 是灰毡毛忍冬的 AGAM (AG) 类基因。经 BLASTp 分析发现其与多条 AGL15 基因具有较高相似性，且含有一个相同的 CArG (C-[A/T]rich-G) 基序的 DNA 结合元件，此结构推测能够激活或抑制特异基因的表达^[20-23]。

同时研究发现胚珠发育时，AG 基因得以表达，在拟南芥 *Arabidopsis thaliana* (L.) Heynh. 的研究中，AG 同源基因均有类似的表达特征，拟南芥 AGL15 基因有控制拟南芥蕾期长短的作用，当该基因过表达时，拟南芥花蕾期延长且不易开花。这些基因在将要产生雄蕊和心皮原基的分生组织、雄蕊和心皮原基、雄蕊和心皮以及胚珠中表达。在对葡萄 AGL15 基因的研究中发现其在胚珠中特异表达，表明葡萄 AGL15 有类似 AGAMOUS 基因的功能。定量 PCR 分析表明，*Lm*-AGL15 基因在湘蕾型花蕾中高度表达且远高于常规型花蕾中的表达量，与 AGL15 基因在拟南芥花中的表达类似，与相同属葡萄 AGL15 基因的表达一致，推测 *Lm*-AGL15 基因在灰毡毛忍冬蕾期延长发育的过程中具有重要的调控作用。这也丰富了 D 类基因在其他植物中的分布^[24]。

因此获得 *Lm*-AGL15 基因的 cDNA 全长序列，并利用生物信息学对 *Lm*-AGL15 基因的功能域和结构进行分析和预测，发现 *Lm*-AGL15 具有 MADS-box 家族基因典型特征，为对于探明一些蕾期延长型突变的分子机制和利用基因调控技术进行蕾期延长育种具有重要意义。同时为开展该基因的表达研究，进一步确定其功能奠定了基础^[25]。

参考文献

[1] 中国药典 [S]. 一部. 2010.
[2] 汪 治, 黄 云. 湘蕾金银花的选育及推广初报 [J].

- 中草药, 2004, 27(12): 896-987.
- [3] 杨苗, 马逾英, 周娟, 等. 细毡毛忍冬采收期动态研究 [J]. 时珍国医国药, 2009, 20(1): 34-35.
- [4] 耿世磊, 宁熙平, 吴鸿, 等. 山银花不同发育阶段花结构与绿原酸含量变化关系研究 [J]. 云南植物研究, 2005, 27(3): 279-287.
- [5] 李敬, 谷慧英, 王志敏, 等. 拟南芥成花关键基因调控网络研究进展 [J]. 生物技术通报, 2014, 31(12): 1-8.
- [6] 吴宏清, 王磊, 何欣, 等. 白木香倍半萜合成酶基因 *As-SesTPS* 的克隆及生物信息学与表达分析 [J]. 中草药, 2014, 45(1): 94-101.
- [7] 杨晓娜. 植物组织总 RNA 提取方法的比较 [J]. 科技视界, 2014, 34: 157.
- [8] 龙松华, 乔瑞清, 陈信波, 等. RACE 法克隆亚麻 COMT 基因及生物信息学分析 [J]. 生物技术通报, 2013(12): 78-83.
- [9] 韩慧, 王希东, 姚正培, 等. 鹰嘴豆 LMY 基因 cDNA 克隆及生物信息学分析 [J]. 新疆农业大学学报, 2014, 37(4): 321-326.
- [10] 何欣, 叶伟, 高晓霞, 等. 白木香倍半萜合成酶 *As-SesTPS1* 基因的克隆、生物信息学和表达分析 [J]. 中草药, 2015, 46(5): 733-739.
- [11] 唐杰, 魏军亚, 刘国银, 等. 杧果 MADS-box 基因保守区片段的克隆与序列分析 [J]. 中国南方果树, 2015(3): 22-27.
- [12] Zhang Y F, Cao G Y, Qu L J. Characterization of Arabidopsis MYB transcription factor gene *AtMYB17* and its possible regulation by *LEAFY* and *AGL15* [J]. *J Genet Genom*, 2009, 2: 99-107.
- [13] 蔡小钿, 王金发. 植物 MADS 盒基因的功能和调节机理 [J]. 植物生理与分子生物学, 2000(36): 274-281.
- [14] Theißen G, Strater T, Fischer A. Structural characterization, chromosomal localization and phylogenetic evaluation of two pairs of AGAMOUS-like MADS-box genes from maize. [J]. *Gene* 2005, 156: 155-166.
- [15] Shiu MDL, Adamczyk BJ, and Fernandez DE Expression of MADS-box genes during the embryonic phase in Arabidopsis [J]. *Plant Mol Biol*, 2005, 58(1): 89-107.
- [16] 李鱼, 张金林, 王锁民, 等. 小花碱茅 HKT2; 1 基因全长 cDNA 的克隆与生物信息学分析 [J]. 草业学报, 2013, 22(2): 140-149.
- [17] 刘虎岐, 刘应保, 宋云鹏, 等. 淡色库蚊氯菊酯抗性相关基因 *PR-OP* 全长 cDNA 的克隆及生物信息学分析 [J]. 西北农林科技大学学报: 自然科学版, 2010, 38(9): 109-117.
- [18] 康桂娟, 曾日中, 聂智毅, 等. 巴西橡胶树 NAC 转录因子 *HbNAC1* 基因的克隆及生物信息学分析 [J]. 中国农学通报, 2012, 28(34): 1-11.
- [19] 申鹏, 毛娟, 刘爱平, 等. 拟南芥转录因子 *CBF2* 的克隆与生物信息学分析 [J]. 中国农学通报, 2014, 30(4): 90-95.
- [20] 武安泉, 王俊生, 张玉玺, 等. 小麦 *Cyp450* 基因的电子克隆与生物信息学分析 [J]. 中国农学通报, 2013, 29(18): 38-44.
- [21] 胡尚连, 曹颖, 黄胜雄, 等. 慈竹 4CL 基因的克隆及其生物信息学分析 [J]. 西北农林科技大学学报: 自然科学版, 2009, 37(8): 204-210.
- [22] 尚洁, 陈阿丽, 潘伟志. 宁夏枸杞查耳酮合成酶基因 cDNA 片段的克隆及生物信息学分析 [J]. 中草药, 2013, 44(20): 122-127.
- [23] 张晓琳, 柴晓杰, 张婷, 等. 盐藻 CDPK 基因的克隆与生物信息学分析 [J]. 核农学报, 2013, 27(4): 418-424.
- [24] 刘松, 李毅, 赵德刚. 无籽刺梨 (*Rosa kweichowensis* var. *sterilis*) *RksAGL* 基因克隆及表达分析 [J]. 基因组学与应用生物学, 2015, 34(3): 579-586.
- [25] 李节法, 田义柯, 王彩虹, 等. 梨贝壳杉烯氧化酶基因 *PpKO* 的克隆及生物信息学分析 [J]. 园艺学报, 2010, 37(10): 1575-1582.