

• 药材与资源 •

灰毡毛忍冬 *Lm-XL-AP1* 基因克隆、生物信息学和时空表达分析彭美晨¹, 刘湘丹^{1,2}, 徐玉琴³, 王 珊⁴, 刘畅宇¹, 陈 勋¹, 周日宝^{1*}

1. 湖南中医药大学药学院 湖南 长沙 410208

2. 湖南省中药饮片标准化及功能技术研究中心 湖南 长沙 410208

3. 湖南中医药大学第一附属医院 湖南 长沙 410007

4. 湖南中医药大学第二附属医院 湖南 长沙 410000

摘要: 目的 克隆灰毡毛忍冬突变型品种 AP1 基因 (*Lm-XL-AP1*), 并对其进行生物信息学和时空表达分析。方法 通过 RACE 技术克隆 *Lm-XL-AP1* 基因全长, 运用生物信息学的方法进行基因同源性和相似性比较分析, 预测其编码蛋白, 并对其进行各种理化性质分析。运用荧光定量 PCR (qRT-PCR) 检测该基因在灰毡毛忍冬突变型植株不同花期和不同器官的相对表达量。结果 获得 AP1 基因, 开放阅读框 (ORF) 长 729 bp, 编码 242 个氨基酸, 与甘菊中 MADS-box 基因家族 AP1 基因相似性达 80%, 且包含 MADS 和 K-box 的保守序列。AP1 蛋白无跨膜区域, 定位于细胞核中; 在灰毡毛忍冬突变型植株第 6 花期相对表达量最高, 同时茎、叶中也有表达。结论 成功从灰毡毛忍冬突变型植株总 RNA 中克隆到可能参与控制花器官表达的 AP1 基因。

关键词: 灰毡毛忍冬; *Lm-XL-AP1* 基因; 生物信息学; 荧光定量; 克隆

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2018)07-1652-09

DOI: 10.7501/j.issn.0253-2670.2018.07.025

Cloning of AP1 Gene from *Lonicera macranthoides* “Xianglei” and analysis on its Bioinformatics and Space-time expressionPENG Mei-chen¹, LIU Xiang-dan^{1,2}, XU Yu-qin³, WANG Shan⁴, LIU Chang-yu¹, CHEN Xun¹, ZHOU Ri-bao¹

1. College of Pharmacy, Hunan University of Traditional Chinese Medicine, Changsha 410208, China

2. Hunan Traditional Chinese Medicine Piece Standardization and Function Technology Research Center, Changsha 410208, China

3. The First Hospital of Hunan University of Traditional Chinese Medicine, Changsha 410007, China

4. The Second Hospital of Hunan University of Traditional Chinese Medicine, Changsha 410000, China

Abstract: Objective To clone the AP1 (*Lm-XL-AP1*) gene from a Special Variant Varieties of *Lonicera macranthoides* “Xianglei”, and to analyze its bioinformatics and spatio-temporal expression. **Methods** Amplifying the full length of *Lm-XL-AP1* gene by RACE technique, using bioinformatics method to analyze homology and similarity of the gene, predicting the coding protein and analyzing the various physical and chemical properties. The expression of the gene in different parts of *Lonicera macranthoides* Special Variant Varieties was detected by fluorescence quantitative PCR (qRT-PCR). **Results** The AP1 gene, containing a 729 bp ORF that encoding 242 amino acids, was cloned. And the similarity of the gene compared with the AP1 gene from the MADS-box gene family of *Chrysanthemum lavandulifolium* up to 80% (Containing a conserved sequence of MADS and K-box). Without transmembrane domain, AP1 was located in cell nucleus. It is expressed in various organs of *Lonicera macranthoides* Special Variant Varieties. **Conclusion** For the first time, the AP1 gene which may be involved in the control of the expression of floral organ was cloned from the total RNA of *Lonicera macranthoides* Special Variant Varieties.

Key words: *Lonicera macranthoides* Hand. -Mazz.; *Lm-XL-AP1* gene; bioinformatics; fluorescence quantification; cloning

灰毡毛忍冬 *Lonicera macranthoides* Hand. -Mazz. Hand. -Mazz. 多年生藤本植物, 以干燥花蕾或初开的花入药, 《中国药典》2015 年版将其作为山银花的主

收稿日期: 2017-06-19

基金项目: 国家自然科学基金资助项目 (81203007, 81673546); 湖南省自然科学基金资助项目 (2017JJ3237); 湖南省教育厅青年基金项目 (16B193); 湖南省中药学重点学科基金项目 (zy201505); 湖南省“中药学”重点学科建设项目资助 (湘教通 [2011] 76 号); 国家中医药管理局“药用植物学”重点学科资助 (国中医药发 [2009] 30 号); 湖南省教育厅 (14K071); 湖南中医药大学研究生省级创新课题

作者简介: 彭美晨 (1992—), 女, 硕士研究生, 研究方向为中药资源与质量。Tel: 15084736469 E-mail: 810759831@qq.com

***通信作者** 周日宝, 男, 博士, 教授, 主要从事中药资源质量与开发研究。Tel: (0731)88458233

要来源之一。其味甘、性寒，具有清热解毒、疏风散热、抗病毒等功效，其治疗的主要病症有痈肿疔疮、丹毒、喉痹、热毒血痢、风热感冒、温热发病等^[1]，被称为“中药中的青霉素”^[2]。灰毡毛忍冬具有适应性较强、产量高、药材色浅质优、适应性强、抗病害能力强等特点，药用价值较高，临床应用广泛。

灰毡毛忍冬主要成分包括皂苷类^[3-4]、挥发油^[5]、绿原酸及黄酮类等，其中绿原酸、木犀草苷和挥发油为主要指标成分。研究表明^[6-7]灰毡毛忍冬主要指标成分在花蕾期含量较高，花冠展开后下降。20 世纪 90 年代研究工作者在湖南省溆浦县发现灰毡毛忍冬自然突变株，并经过鉴定命名为“湘蕾金银花 *Lonicera macranthoides* ‘Xianglei’ (Lm-XL)”。为灰毡毛忍冬自然突变型（20 世纪 90 年代金银花与山银花没有分列，都统称为金银花），其与野生植株不同，表现为花冠不展开、花蕾期长且花蕾整齐、20 余天后才开始凋谢等优良性状^[8]。而野生品种的灰毡毛忍冬花蕾期短，花冠展开后一两天即凋谢，给其花蕾的采摘带来不便。因此，对灰毡毛忍冬自然突变品种进行蕾期长花冠不展开的机制研究具有重要意义。

植物中的 MADS-box 家族基因是重要的转录调控因子^[9-10]，参与花器官发育和开花时间调节^[11-12]。属于 MADS-box 家族的 AP1 基因为花分生组织和花器官形态特征基因，起到调控花序分生组织转变、调控花器官形态发育，缩短植物营养生长期和促使植物早开花的作用。

本课题组前期对灰毡毛忍冬野生型和自然突变型进行了转录组测序和比较分析，获得多条 MADS-box 家族基因，通过数据分析筛选得到多条

表达差异的 Unigene 序列。本研究以灰毡毛忍冬自然突变型植株总 RNA 为研究对象，对其中注释为 AP1 的 Unigene14703 进行研究，根据其序列设计引物，进行逆转录反应和 RACE 扩增技术，获得完整的开放阅读框，并命名为 Lm-XL-AP1，运用生物信息学的方法对该序列进行同源性分析。应用实时荧光定量 PCR 技术对 AP1 基因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灰毡毛忍冬突变型植株蕾期长花冠不展开的分子机制研究奠定基础。

1 材料与试剂

灰毡毛忍冬自然突变株 *Lonicera macranthoides* “Xianglei” 种植于湖南中医药大学药植园，采收 7 个不同花期的花（分别按照花蕾初期、青绿色花蕾期、浅绿色花蕾期、绿白色花蕾期、白色花蕾期、黄白色花蕾期、花蕾枯萎期 7 个时期采收^[13]）、当年生茎和叶样品^[14]。经湖南中医药大学周日宝教授鉴定为灰毡毛忍冬自然突变株的花、茎、叶样品，分别用封口袋封好，放入盛有液氮的泡沫箱，迅速带回实验室，并保存于 -80 ℃ 冰箱中备用。

多糖多酚试剂盒 (Biospin, 杭州博日科技有限公司); 反转录试剂盒 (Revert Aid™ First Strand cDNA Synthesis Kit, 北京智杰方远科技有限公司); 琼脂糖凝胶回收试剂盒 (康为世纪生物科技有限公司); pEASY®-T1 Cloning Kit (北京全式金生物技术有限公司); RACE 试剂盒 (Clontech 公司, 美国); 感受态细胞; Trans2K DNA Marker (全式金生物技术有限公司); 氨苄青霉素 (北京索莱宝科技有限公司); 主要化学试剂为国产分析纯，均购自布兰哲生物有限公司; 用 Oligo6 软件设计引物，委托上海生物工程有限公司合成，相关引物序列见表 1。

表 1 引物序列
Table 1 Primers sequences

引物	引物序列 (5'→3')	目标 (预期) 扩增长度/bp
Lm-XL-AP1-F	ATGGGGAGAGGGAAAGTACAGC	720
Lm-XL-AP1-R	AAAGCAGCCAAGGTGGCATG	720
5'-RACE 特异性引物	GATTACGCCAAGCTTTGTAAAGAAAGGGGAGCGGCTGAGGC	861
3'-RACE 特异性引物	GATTACGCCAAGCTTCCTCTTGCCTCAGCCGCTCCCCCT	428
q-PCR-F	ACGTGGCTCTGATTGTCTTC	50
q-PCR-R	GAGAAGATCCTTGAGCGGTATG	50
18 S rRNA-F	CTTCGGGATCGGAGTAATGA	—
18 S rRNA-R	GCGGAGTCCTAGAAGCAACA	—

2 方法

2.1 RNA 的提取及 cDNA 链合成

用多糖多酚试剂盒法分别提取灰毡毛忍冬自然突变株的花、茎、叶样品总 RNA, 用 1.5% 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 的纯度。以提取的 RNA 为模板, 使用美国 Thermo 有限公司的反转录 cDNA 第一链合成试剂盒合成第一条链 cDNA, 获得的反转录产物分离纯化后置于 -20 °C 贮存备用。

2.2 核心片段的扩增

通过分析项目组前期实验测得的高通量测序结果, 选择表达量有明显差异的且属于 MADS-box 基因家族的 AP1 基因, 设计特异性引物, 上游引物 Lm-XL-AP1-F, 下游引物 Lm-XL-AP1-R。以“2.1”项下的 cDNA 为模板扩增, PCR 反应体系 (25 μ L): ddH₂O 9.5 μ L, 正向引物和反向引物各 1 μ L, cDNA 1 μ L, 2 $\times$ Taq PCR Master Mix 12.5 μ L。PCR 反应条件: 94 °C 预变性 4 min; 94 °C 变性 45 s, 54 °C 复性 45 s, 72 °C 延伸 60 s, 循环 35 次; 72 °C 延伸 10 min。反应结束后, 产物在 2% 的琼脂糖凝胶上以 150 V, 20 min 条件电泳, 检测目的条带。按照 CW2302 快速琼脂糖凝胶 DNA 回收试剂盒方法回收目的条带。连接和转化按照 *pEASY*[®]-T1 Cloning Kit 试剂盒说明书进行。

用灭菌的白色枪头挑选白色单菌落至 10 μ L 无菌水中, 涡旋混匀并短暂离心。取 1 μ L 混合液进行菌液 PCR, PCR 产物进行琼脂糖凝胶电泳检测。

在超净工作台上将剩下的菌液接种于灭菌的 LB/Amp 液体培养基中, 在恒温震荡培养箱中, 200 r/min, 37 °C 培养过夜。次日取 1 mL 菌液送至铂尚生物技术 (上海) 有限公司测序。

2.3 Lm-XL-AP1 的克隆

根据核心片段序列, 设计 RACE 特异性引物。以“2.1”项方法获得的总 RNA, 严格按照 SMARTer[®] RACE 5'/3' Kit 说明书分别做 5'-RACE 及 3'-RACE。PCR 反应体系为: PCR-Grade H₂O 15.5 μ L; 2 $\times$ SeqAmp Buffer, 25.0 μ L; SeqAmp DNA Polymerase, 1.0 μ L; 混匀并短暂离心, 室温放置, 备用。5'RACE cDNA 或 3'-RACE cDNA, 2.5 μ L; 5'或者 3' GSP (特异性引物 10 μ mol/L) 1 μ L; 10 $\times$ UPM (试剂盒提供) 5 μ L; 轻轻混匀, 并短暂离心。PCR 扩增程序: 94 °C 变性 30 s, 72 °C 延伸 2 min, 5 个循环; 94 °C 变性 30 s, 70 °C 退火 30 s, 72 °C

延伸 2 min, 5 个循环, 94 °C 变性 30 s, 68 °C 退火 30 s, 72 °C 延伸 2 min, 25 个循环。

取 RACE 产物于 2% 琼脂糖凝胶中电泳, 按琼脂糖凝胶回收说明书进行胶回收, 回收目标片段长度约 500 bp 和 800 bp。连接、转化和测序, 与核心序列进行比对, 得 Lm-XL-AP1 基因的 5' 和 3' 端。采用 ContigExpress 软件对测得的 3' 及 5' 端核酸序列进行拼接、分析, 获得基因全长序列。将获得的序列上传到 NCBI ORF Finder 平台 (<http://www.ncbi.nlm.nih.gov/orf/>), 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 片段完整性。

2.4 Lm-XL-AP1 基因生物信息学分析

2.4.1 蛋白理化性质分析 采用 Primer Premier 5.0 软件将 cDNA 全长翻译成氨基酸序列, 并用 ProtParam 软件 (<http://web.expasy.org/protparam/>) 在线预测 Lm-XL-AP1 基因编码的蛋白质的分子式、相对分子质量, 不稳定系数以及等电点等。蛋白亲/疏水性采用 ProtScale 软件 (<http://web.expasy.org/protscale/>) 在线分析; <http://www.sbc.su.se/~miklos/DAS/> 在线预测 Lm-XL-AP1 基因编码蛋白的二级结构; 采用 SWISS-MODEL 软件 (<http://swissmodel.expasy.org/>) 在线分析 Lm-XL-AP1 基因编码蛋白的三级结构。采用 TMHMM 软件 (<http://www.cbs.dtu.dk/services/TMHMM/>) 在线分析蛋白跨膜区及定位信号; SignalP 4.1 Server 软件 (<http://www.cbs.dtu.dk/services/SignalP/>) 在线预测分泌蛋白^[15-16]; WOLF PSORT (<http://www.genscript.com/wolf-psort.html>) 在线预测蛋白质亚细胞定位情况, 编码蛋白的三级结构。将 Lm-XL-AP1 的氨基酸序列于 NCBI 数据库中进行 BLASTP, 分析可能的蛋白功能结构域。将氨基酸序列于 NCBI 进行 BLASTP 比对 (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), 筛选与其同源性较高的物种用 DNAMAN 6.0 软件进行同源性比较并构建系统进化树, 进行聚类分析。

2.4.2 Lm-XL-AP1 基因表达水平分析 按“2.1”项方法提取灰毡毛忍冬突变株茎、叶及 7 个花期的花样品总 RNA 并逆转录合成第一链 cDNA 样品, 18 S rRNA 基因为内参基因^[17], 进行 qRT-PCR 扩增。反应体系 20 μ L, 其中上游引物和下游引物各 1 μ L, 样品 cDNA 模板 1 μ L, iTaq[™] universal SYBR[®]

Green supermix (2×) 10 μL, 补超纯水至 20 μL。
反应条件为 95 ℃、3 min, 95 ℃、3 s, 60 ℃退火 2 min, 39 个循环, 20 min 内从 65 ℃、3 s 升温到 95 ℃、50 s, 每样品重复 3~4 次。

3 结果与分析

3.1 灰毡毛忍冬突变株总 RNA 的提取

用多糖多酚植物总 RNA 提取试剂盒法提取灰毡毛忍冬突变株总 RNA, 琼脂糖凝胶电泳结果如图 1 所示, 28 S、18 S 条带清晰, 说明 RNA 具有较好的完整性。核酸蛋白检测仪测定其 A_{260}/A_{280} 值为 2.13~2.47, A_{260}/A_{280} 值为 1.94~2.06, 说明 RNA 纯度较高, 能满足后续实验要求。

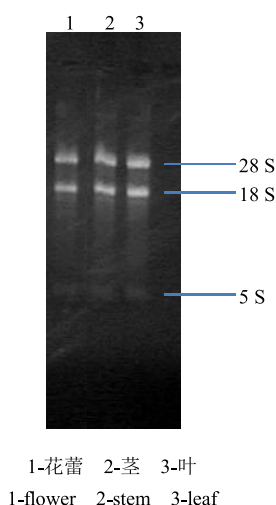


图 1 总 RNA 琼脂糖凝胶电泳图

Fig. 1 Agarose gel electrophoresis of total RNA

3.2 Lm-XL-AP1 核心片段克隆

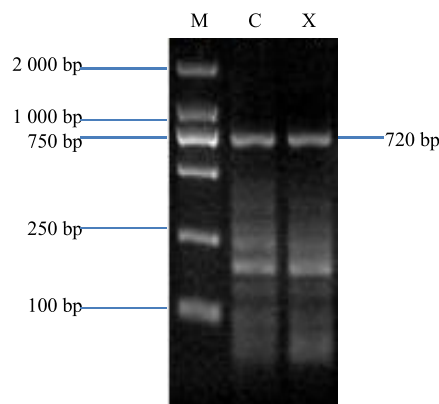
以灰毡毛忍冬突变型品种花总 RNA 反转录产物为模板, 以 Lm-XL-AP1-F 和 Lm-XL-AP1-R 为特异性引物, 进行核心片段的扩增, 结果显示出现一条单一明亮的条带, 长度约 730 bp, 经测序, 核酸序列片段长 720 bp (图 2)。

3.3 Lm-XL-AP1 基因全长的获得

根据核心序列设计引物 (表 1), 分别利用设计合成的 5'-RACE 和 3'-RACE 特异性引物, 进行 PCR 扩增。5'-RACE 扩增结果在 800 bp 左右有一条亮带 (图 3)。3'-RACE PCR 的扩增结果在 500 bp 左右有一条亮带 (图 4), 分别将 PCR 产物回收, 连接后转化, 取阳性克隆子进行测序。

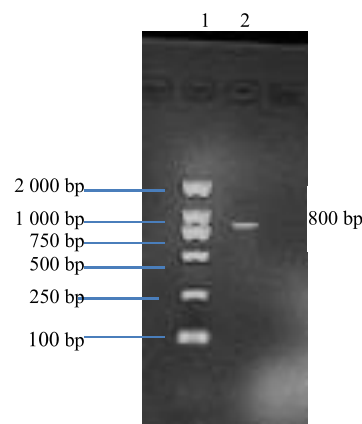
3.4 Lm-XL-AP1 基因及其编码氨基酸序列分析

Lm-XL-AP1 基因全长 1 162 bp 不含内含子序列编码 242 个氨基酸 (GenBank 登录号为 MF139138)



M-Marker C-Lm-AP1 X-Lm-XL-AP1 核心片段
M-Marker C-Lm-AP1 X-Lm-XL-AP1 core fragment

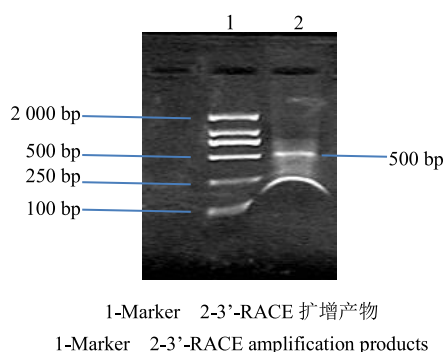
图 2 Lm-XL-AP1 基因 cDNA 核心片段 PCR 产物电泳图
Fig. 2 PCR products of Lm-XL-AP1 cDNA core fragment



1-Marker 2-5'-RACE 扩增产物
1-Marker 2-5'-RACE amplification products

图 3 Lm-XL-AP1 基因 5'-RACE 扩增产物电泳图

Fig. 3 5'-RACE PCR products electrophoresis of Lm-XL-AP1



1-Marker 2-3'-RACE 扩增产物
1-Marker 2-3'-RACE amplification products

图 4 Lm-XL-AP1 基因 3'-RACE 扩增产物电泳图

Fig. 4 3'-RACE PCR products electrophoresis of Lm-XL-AP1

结果见图 5。

利用 ExPASy Proteomics Server 的在线软件对 Lm-XL-AP1 蛋白的理化性质进行预测分析, 推测其分子式为 $C_{1229}H_{1982}N_{356}O_{375}S_8$, 相对分子质量为 28 001.90, 等电点为 8.75, 亲水区域所占比例大于疏水区域。正电残基 (Arg+Lys) 为 37, 负电残基 (Asp+Glu) 为 33。蛋白不稳定系数为 56.77 (>40), 表明 Lm-XL-AP1 蛋白不稳定。脂肪系数为 87.07。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。GRAVY 为 -0.707 (小于 0), 预测其为亲水性蛋白。

通过 WOLF PSORT 软件在线分析 Lm-XL-AP1

蛋白亚细胞定位, 发现其细胞核的定位系数为 14, 故推测 Lm-XL-AP1 蛋白可能定位于细胞核中, 见图 7。通过 SMART 服务器分析, Lm-AP1 第 1~59 个氨基酸为高度保守的 MADS 结构域, 第 93~170 个氨基酸为中度保守的 K-box 结构域 (不含起始密码子), 即 MADS-box 家族成员共有的典型结构域, 见图 8。

3.5 Lm-XL-AP1 基因编码的蛋白质结构与功能预测

SOPMA 软件在线解析 Lm-XL-AP1 蛋白的二级结构, 结果见图 9。由图可知, 此蛋白一共有 121 个氨基酸形成 α -螺旋, 占 50%; 12 个氨基酸形成 β -转角, 占 4.96%; 109 个氨基酸用于形成无规卷曲,

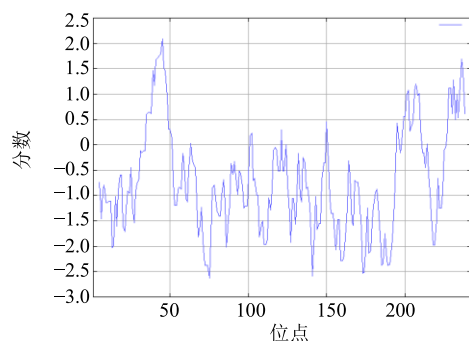
```

1   CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGATATTTA
61  TTAACATCACTTCATTCTTCACCTCTGAGGAAGGGAAGATTTTCATTCCTTTTCTCGG
121 AATTGGGGTTTTGGTTGGTTTCCGAGAAAATTATTAATGGGGAGAGGGAAGTACAG
1   M G R G K V Q
181 CTGAAACGGATAGAGAACAAGATCAACAGACAGGTTACTTTCTCAAAGAGGAGAGGTGGA
8   L K R I E N K I N R Q V T F S K R R G G
241 TTGTTGAAGAAGGCCCATGAGATCTCAGTGCTTTGTGATGCTGACGTGGCTCTGATTGTC
28  L L K K A H E I S V L C D A D V A L I V
301 TTCTCTAACAAGGAAAGCTCTTTGAGTATTCCTGATTCATGCATGGAGAAGATCCTT
48  F S N K G K L F E Y S T D S C M E K I L
361 GAGCGGTATGAAAGATACTCTTACACCGAGAGACAGCTAGTTTCTCATGATCCACAATCT
68  E R Y E R Y S Y T E R Q L V S H D P Q S
421 TCGGGAACGTTACCCTGAATACAACAACCTTAAGGCTAGGGTCGAGCTTTTACAAAGA
88  S G N V T L E Y N K L K A R V E L L Q R
481 AACTATAGGCACTATATAGGAGAAGATCTAGACTCCTTGAGCCTAAAAGACCTCCAAAAT
108 N Y R H Y I G E D L D S L S L K D L Q N
541 TTGAGCAACAGCTTGATACTGCTCTTAAGCGCATACGATCCCATAAAAACCAACTCATG
128 L E Q Q L D T A L K R I R S H K N Q L M
601 TATGAGTCCATCTCCGAGCTTCAGAGAAAGGAAAGAGCAATTCAGGAGCAAAACAGCGCG
148 Y E S I S E L Q R K E R A I Q E Q N S A
661 CTATCAAAGAAGATTAAAGAGAAGGAAAAGACAATGGCGGAGCAAGCTTACTGGGATCAG
168 L S K K I K E K E K T M A E Q A Y W D Q
721 CAAAACCATGCCCAAATTCACCATCATTCTCTTGCCCTCAGCCGCTCCCTTTCTTAAC
188 Q N H A P N S P S F L L P Q P L P F L N
781 ATCGGCGCCGCGCTTACCAGGAGAAGCACTTGAAGAGAGGAGGAATGTTCTCGACCTC
208 I G A G A Y Q G E A L E E R R N V L D L
841 ACTCTGAACCGCTAGTTTCATGCCACCTTGCGTCTTTGCCGAGTGAATGACAAGAAGA
228 T L E P L V S C H L G C F A E *
901 AAAAAGGCTGCTGTTTGATTATAGAATTTGGATTTATACGTGAAGACGTTATAAGATAT
961 ATGGTTGTGCCAGGCAAGATAAAATAAGATGTATAACACTTGATTCTTGTCTGATAA
1 021 TTAATACTAAGTCTGTTTGATAATTAATGTGTACTTGATTTGAATGTGAGTTGACGAA
1 081 GACAAATTATGTATTGTCTATATAAATATATTTAGGAAAAGAACTAGCTTGAGATCAAA
1 141 AAAAAAAAAAAAAAAAAAAAAA

```

图 5 Lm-XL-AP1 基因全长序列

Fig. 5 Full-length sequences of Lm-XL-AP1



分数 > 0 表示疏水性, 分数 < 0 表示亲水性
Score > 0 for hydrophobicity, score < 0 for hydrophilicity

图 6 Lm-XL-AP1 编码蛋白的亲水性分析

Fig. 6 Hydrophilicity analysis on Lm-XL-AP1-coding protein



图 7 Lm-XL-AP1 蛋白亚细胞定位

Fig. 7 Lm-XL-AP1 domain

占 45.04%。在氨基酸的 14~38、63~73、85~112、122~178 位是连续的 α -螺旋。SWISS-MODEL 软件三级结构预测结果显示 Lm-XL-AP1 基因编码蛋白质的三级结构见图 10。

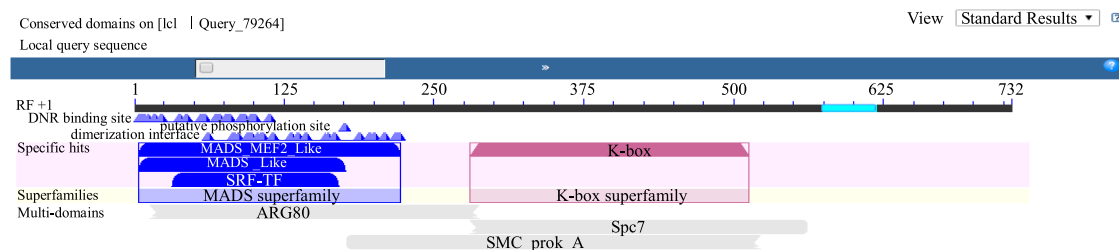


图 8 Lm-XL-AP1 结构域

Fig. 8 Lm-XL-AP1 domain

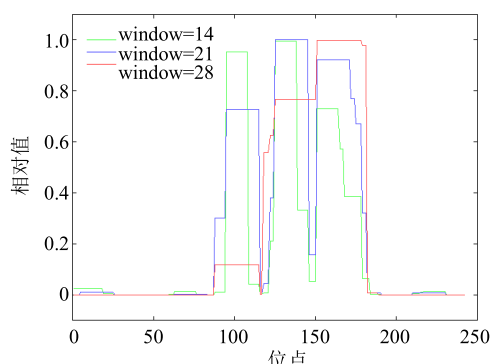


图 9 Lm-XL-AP1 蛋白二级结构预测

Fig. 9 Protein secondary structure prediction chart of Lm-XL-AP1

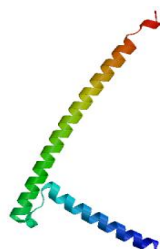


图 10 Lm-XL-AP1 蛋白三级结构预测

Fig. 10 Protein tertiary structure prediction chart of Lm-XL-AP1

3.6 氨基酸序列同源性分析

目前研究者已从葡萄 *Vitis vinifera* L.、枇杷 *Eriobotrya japonica* Thunb.、茶树 *Camellia sinensis* L.、山核桃 *Carya cathayensis* Lecomt.、芒果 *Mangifera indica* L.、龙眼 *Dimocarpus longan* Lour.、芝麻 *Sesamum indicum* Linn. 和牡丹 *Paeonia suffruticosa* Andr. 等植物中克隆了 AP1 同源基因全长, 并对其进行了序列分析和功能分析^[18-20]。

在 NCBI 中, 对 Lm-XL-AP1 基因编码的氨基酸序列进行 BLASTP 比对, 结果显示, 其中与甘菊 *Chrysanthemum lavandulifolium* Fisch. ex Trautv. (80%, JX872512.1) 和甘蓝 *Brassica oleracea* L. (80%, AF126724.1) AP1 同源性最高, 为 80%。其次是灯台树 *Cornus controversa* Hemsl. AP1 (78%, KT965643.1)、山茱萸 *Cornus sanguinea* Sieb. et Zucc. AP1 (78%, KT965642.1)、芝麻 AP1 (78%, KM677186.1)。分析发现 Lm-XL-AP1 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。结果见图 11。

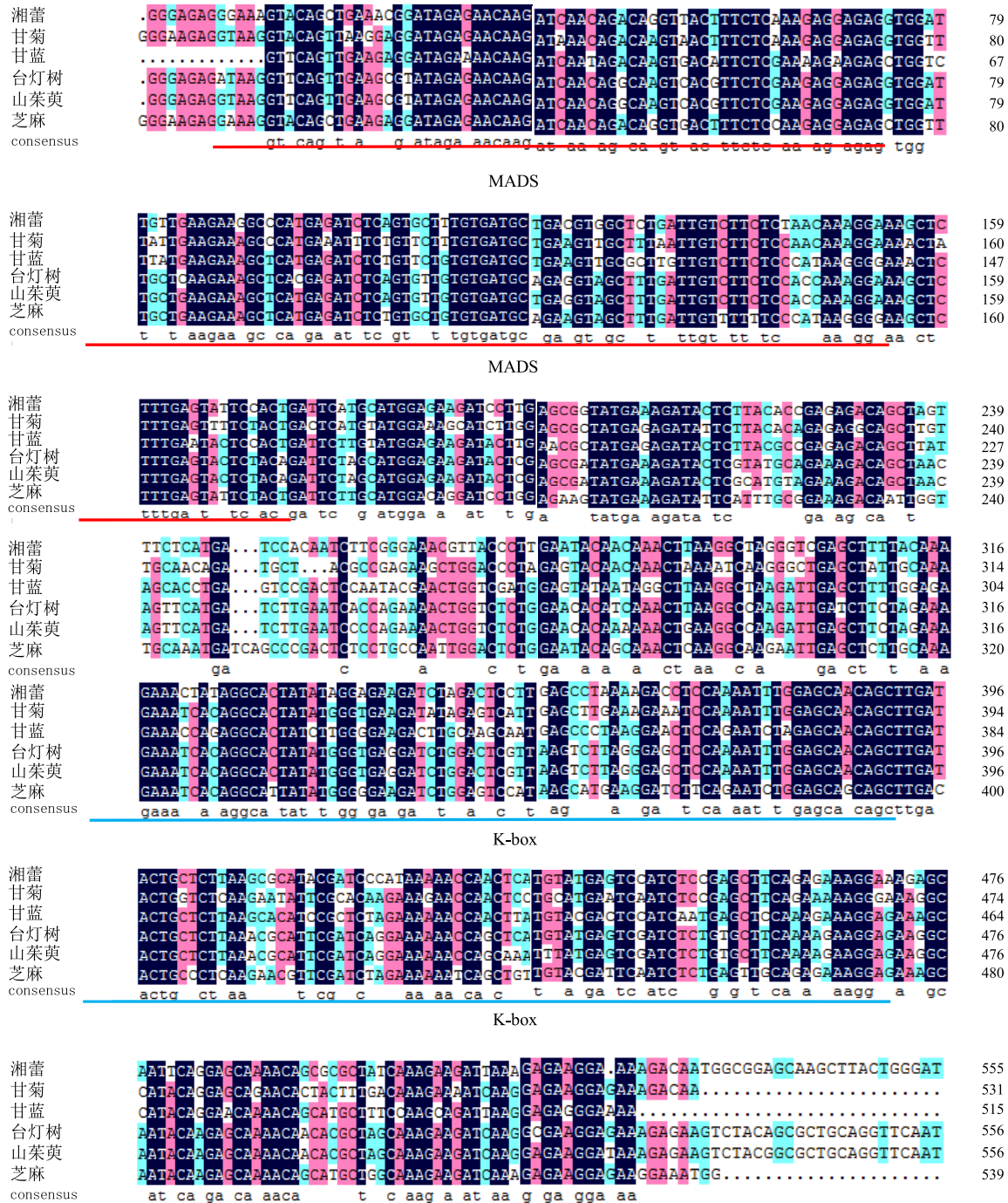


图 11 氨基酸多重序列同源性比对

Fig. 11 Homology alignment of multiple amino acid sequence

3.7 不同来源 AP1 基因的系统进化树分析

为更直观、清晰地分析 Lm-XL-AP1 基因与其他物种 AP1 之间的进化关系,将获得的 Lm-XL-AP1 基因推导的氨基酸序列与其他植物中已报道的属于 AP1 亚家族的成员进行比对,构建系统发育树状图(图 12),其中包括葡萄、灯台树、山茱萸、荔枝 *Litchi chinensis* Sonn.、川桑 *Morus notabilis*

Schneid.、甘蓝、麻风树 *Jatropha carcas* L.、甘菊以及芝麻。从图 12 可以分析得出 AP1 亚家族基因大致归为 3 类, Lm-XL-AP1 蛋白序列与其他被子植物门双子叶植物纲的 AP1 基因有较近的亲缘关系,归为一类,可确定该基因编码 AP1 亚家族基因类蛋白, Lm-XL-AP1 与甘菊亲缘关系最近,与葡萄亲缘关系最远。

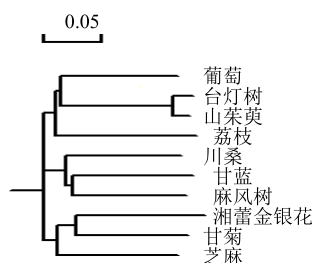


图 12 AP1 基因不同物种系统进化树

Fig. 12 Phylogenetic trees of different species of AP1

3.8 Lm-XL-AP1 基因的时空表达分析

利用实时荧光定量 PCR 检测 AP1 基因在灰毡毛忍冬突变型植株不同花期及不同部位中的表达量差异以及与灰毡毛忍冬野生型花的比较。实验结果表明, 均能检测到符合预期不同花期及不同部位 AP1 基因的相对表达量。从图 13 可以看出 1~6 花期的相对表达量整体呈现一个上升趋势, 6~7 花期呈下降趋势, 其中第 6 花期^[13]黄白色花蕾期相对表达量最多, 说明第 6 花期的花发育器官达到最旺盛, 丰富度也最高。从图 14 可以看出野生型灰毡毛忍冬花蕾中也检测到 AP1 基因, 灰毡毛忍冬突变型植株中花蕾(第 6 花期)丰富度最高, 其茎、叶丰富度较低, 野生型花蕾(第 6 花期)较低。说明 AP1 基因主要在发育器官部位表达, 灰毡毛忍冬突变型植株花蕾中 AP1 基因的表达远高于野生型灰毡毛忍冬花蕾。

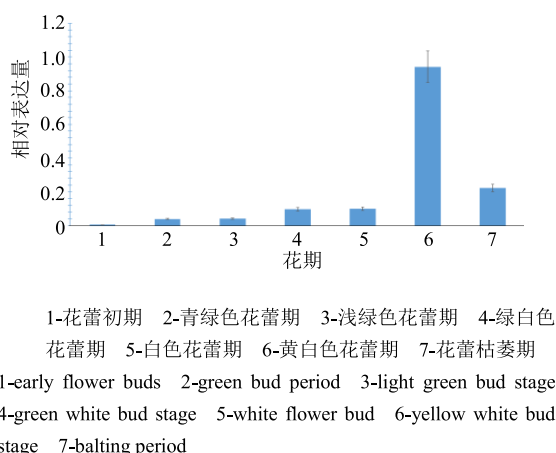


图 13 AP1 基因在突变型植株不同花期中相对表达量

Fig. 13 Relative expression of Lm-XL-AP1 gene in different flowering stages

4 讨论

本研究通过 RACE 技术成功克隆了灰毡毛忍冬自然突变型植株 AP1 基因, 命名为 Lm-XL-AP1, 将其提交到 GenBank 数据库 (GenBank 登录号为

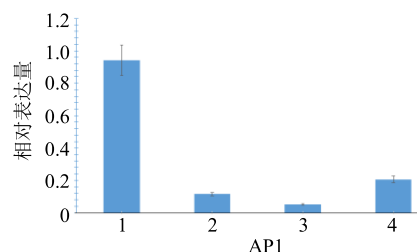


图 14 AP1 基因在突变型植株的花、叶、茎与野生型植株花的相对表达量

Fig. 14 Relative expression of AP1 gene in Lm-XL flower, leaf, stem, and Lm. flower

MF139138) 全长 1 162 bp, 该基因序列的第 1~159 位核苷酸基因为 5'UTR, 第 889~1 162 位核苷酸基因为 3'UTR, 含有 25 bp 的 polyA 结构, 其中包括 1 个 729 bp 的 ORF, 编码 242 个氨基酸。同时, 研究表明, Lm-XL-AP1 基因编码蛋白的二级结构以 α 螺旋结构为主, 此基因第 1~74 个氨基酸为 MADS_MEF2_like, 第 1~59 个氨基酸为 MADS 区, 第 93~170 个氨基酸为 K-box, 第 10~56 位氨基酸为 SRF-TF 区, 第 1~58 位为 ARG80。含有 1 个典型的 MADS 区和 K-box 区结构, 显示其属于 MADS-box 家族基因, 共分为 M、I、K、C 4 个区, 其中 N 端为 59 个氨基酸的 MADS 结构域, 高度保守, 为 M 区, 具有 DNA 结合单元 CArG-box[5'-CC (A/T)6GG-3'], 可与特异 DNA 结合, 也具有蛋白质二聚化功能; 第 93~170 个氨基酸为中度保守的 K-box 结构域, 为 K 区及蛋白因子之间相互作用区域, 与基因的转录有关; M 区和 K 区之间为 I 区, 包含 35 个氨基酸; C 区位于 K 区下游, 包含 22 个氨基酸, 保守性较低。

通过实时荧光定量分析显示, 灰毡毛忍冬自然突变型植株中 Lm-XL-AP1 基因在不同花期的花中, 相对表达量大体呈现先递增后下降趋势, 且在茎、叶中均有表达。因此推测 AP1 基因可能不止调控花发育过程, 还有可能调控茎、叶中某些性状的表达, 丰富了 AP1 基因在同一植物中的分布情况。

研究表明^[21-22], 将苹果中的同源 AP1 基因转入拟南芥中, 可以使拟南芥提早开花。AP1 基因的过量表达, 可以促使植物提早开花, 其与 LFY 基因共同调控植物的花发育过程。单个氨基酸的突变可能会导致该蛋白与 DNA 的结合能力变弱, 从而导致

灰毡毛忍冬自然突变型植株中 AP1 基因在不同花期的花中, 相对表达量大体呈现先递增后下降趋势, 且在茎、叶中均有表达。因此推测 AP1 基因可能不止调控花发育过程, 还有可能调控茎、叶中某些性状的表达, 丰富了 AP1 基因在同一植物中的分布情况。

植物表型发生变化。灰毡毛忍冬作为山银花主要药材来源之一,其用途广、用量大,克隆并找出控制灰毡毛忍冬突变型植株花蕾期长短的相关基因,对于灰毡毛忍冬药用价值方面的研究及开发具有积极的促进作用。今后会进一步应用分子技术,进行Lm-XL-AP1基因功能分析。

高质量RNA的提取和引物的设计为RACE和qRT-PCR的关键。RNA提取的最好结果是在琼脂糖凝胶电泳中28S、18S条带清晰,28S亮度是18S的1~2倍,核酸蛋白检测仪测定 A_{260}/A_{230} 值大于2.00, A_{260}/A_{280} 值在1.80~2.10时RNA质量高,满足后续试验要求;RACE引物长度最好23~28个碱基;GC含量50%~70%,设计3'端引物和5'端引物时,可选择2个片段有一定重叠的引物(100~200 bp的重叠区),将有利于全长序列的拼接。设计qRT-PCR引物长度最好15~25个碱基;GC含量50%左右;引物扩增长度尽量在40~100 bp。

本研究通过RACE技术克隆得到Lm-XL-AP1基因全长序列,并对其相关氨基酸序列、蛋白质和表达量等特性进行预测与分析。发现Lm-XL-AP1具有MADS-box家族基因典型特征,为今后进一步深入研究MADS-box家族基因奠定理论基础,为对于花类药材和其他植物MADS-box家族基因的研究提供了参考,并且在转基因分子育种应用中具有重要意义。通过对灰毡毛忍冬Lm-XL-AP1基因的研究,为从分子水平探索灰毡毛忍冬突变型植株的生理生化研究奠定理论基础,为克隆灰毡毛忍冬突变型植株的不同部位的基因及建立时空表达模式提供了实验基础。

参考文献

- [1] 中国药典[S].一部.2015.
- [2] 周华,张新,刘腾云,等.高通量转录组测序的数据分析与基因发掘[J].江西科学,2012,30(5):607-611.
- [3] 谷筱玉,陈振鹏,陈乾平,等.HPLC-UV-ELSD测定山银花中绿原酸、灰毡毛忍冬皂苷乙和川续断皂苷乙的含量[J].药物分析杂志,2011,31(5):884-887.
- [4] 贾晓东,冯煦,董云发,等.灰毡毛忍冬中皂苷类成分的研究[J].中草药,2007,38(10):1452-1455.
- [5] 唐丽君,周日宝,刘芙蓉,等.超临界CO₂流体萃取法与水蒸气蒸馏法提取灰毡毛忍冬中挥发油的GC-MS比较研究[J].湖南中医药大学学报,2010,30(9):109-113.
- [6] 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.
- [7] 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].植物分类与资源学报,2005,27(3):279-287.
- [8] 蔡嘉洛.毡毛忍冬蕾期延长相关基因AGL15的克隆与表达[D].长沙:湖南中医药大学,2016.
- [9] 郭爽,马宁,杨文才,等.辣椒花器官发育MADS-box基因的克隆与表达分析[J].园艺学报,2010,37(10):1591-1597.
- [10] Colombo M, Masiero S, Vanzulli S, et al. AGL23, a type I MADS-box gene that controls female gametophyte and embryo development in *Arabidopsis* [J]. *Plant J*, 2008, 54(6): 1037-1048.
- [11] 黄方,迟英俊,喻德跃.植物MADS-box基因研究进展[J].南京农业大学学报,2012,35(5):9-18.
- [12] Gunter T, Annette B, Alexandra D R, et al. A short history of MADS-box genes in plants [J]. *Plant Mol Biol*, 2000, 42(1): 115-149.
- [13] 孙梦姍,周日宝,彭晓丹,等.灰毡毛忍冬不同品种开花时期内源乙烯动态变化的研究[J].中成药,2013,35(9):1969-1972.
- [14] Tilly J J, Allen D W, Jack T. The CArG boxes in the promoter of the *Arabidopsis* floral organ identity gene APETALA3 mediate diverse regulatory effects [J]. *Development*, 1998, 125(9): 1647-1657.
- [15] Munster T, Pahnke J, Di Rosa A, et al. Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants [J]. *Proc Natl Acad Sci USA*, 1997, 94(6): 2415-2420.
- [16] Lam H M, Xu X, Liu X, et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection [J]. *Nat Gen*, 2010, 42(12): 1053-1059.
- [17] 蔡嘉洛,朱贻霖,谢舒平,等.灰毡毛忍冬内参基因筛选和Mads-box家族基因AGL15的时空表达分析[J].中草药,2016,47(15):2727-2733.
- [18] 张百成,沙伟,宋璐.毛尖紫萼藓GH394基因克隆、表达载体的构建[J].北方园艺,2012(12):113-116.
- [19] 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.
- [20] 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.
- [21] Becker A, Theissen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants [J]. *Mol Phylogenet Evol*, 2003, 29: 464-489.
- [22] Riechmann J L, Meyerowitz E M. MADS domain proteins in plant development [J]. *Biol Chem*, 1997, 378(10): 1079-1101.