

本文引用: 付学森, 刘紫璇, 王玲, 龙雨青, 曾娟, 周日宝, 刘湘丹. 灰毡毛忍冬 MADS-box 基因家族鉴定与 CMB1 基因克隆[J]. 湖南中医药大学学报, 2024, 44(3): 383-394.

灰毡毛忍冬 MADS-box 基因家族鉴定与 CMB1 基因克隆

付学森¹, 刘紫璇¹, 王玲¹, 龙雨青¹, 曾娟¹, 周日宝^{1,2,3,4*}, 刘湘丹^{1,2,3,4*}

1. 湖南中医药大学药学院, 湖南长沙 410208; 2. 湘产大宗道地药材种质资源及规范化种植重点研究室,

湖南长沙 410208; 3. 湖南省普通高等学校中药现代化研究重点实验室, 湖南长沙 410208;

4. 湖南省中药饮片标准化及功能工程技术研究中心, 湖南长沙 410208

[摘要] 目的 鉴定灰毡毛忍冬 *Lonicera macranthoides* MADS-box 家族基因并进行生物信息学分析与表达模式验证, 克隆湘蕾型和野生型灰毡毛忍冬 MADS-box 家族成员 CMB1 全长。方法 基于转录组数据, 利用在线工具对灰毡毛忍冬 MADS-box 进行生物信息学分析, 并利用 qRT-PCR 验证 MADS-box 基因在不同品种中的表达模式, 通过 RT-PCR、RACE 技术克隆湘蕾型和野生型灰毡毛忍冬 CMB1 基因全长。结果 28 个灰毡毛忍冬 MADS-box 蛋白长度为 89-359 aa, 碱性蛋白占比 85.7%, 不稳定蛋白占比 96.4%, 亲水性蛋白占比 96.4%, 均定位于细胞核, 均为无跨膜结构的非分泌蛋白。转录组显示, 与野生型比较, 湘蕾型中 28 个 MADS-box 基因有 17.86% 表达下调, MIKC^C 型基因中有 18.75% 表达下调。克隆得到在两个品种的灰毡毛忍冬花中高度特异性表达的 CMB1 基因全长, 均包含一个 738 bp 的 ORF, 编码 245 个氨基酸。CMB1 基因在两个品种的花中表达量存在显著差异 ($P < 0.01$), 且与茎、叶相比, 在花中高度特异性表达 ($P < 0.01$)。结论 基于灰毡毛忍冬转录组数据, 鉴定了 28 个 MADS-box 家族基因, 克隆得到湘蕾型与野生型灰毡毛忍冬 CMB1 基因全长, 为进一步研究灰毡毛忍冬花发育、优良表型形成的分子机制提供研究基础与理论依据。

[关键词] 灰毡毛忍冬; MADS-box 基因家族; 表达分析; CMB1 基因; 基因克隆; 表型变异

[中图分类号] R28

[文献标志码] A

[文章编号] doi:10.3969/j.issn.1674-070X.2024.03.006

Identification of the MADS-box gene family and cloning of the CMB1 gene in *Lonicera macranthoides*

FU Xuesen¹, LIU Zixuan¹, WANG Ling¹, LONG Yuqing¹, ZENG Juan¹,

ZHOU Ribao^{1,2,3,4*}, LIU Xiangdan^{1,2,3,4*}

1. School of Pharmacy, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China; 2. Key Research Laboratory of Germplasm Resources and Standardized Planting of Genuine Regional Medicinal Materials Produced in Hunan Province, Changsha, Hunan 410208, China; 3. Key Laboratory of Modern Research of Chinese Medicine, Education Department of Hunan Province, Changsha, Hunan 410208, China; 4. Hunan Engineering Technology Research Center for Standardization and Function of Chinese Herbal Decoction Pieces, Changsha, Hunan 410208, China

[Abstract] **Objective** To identify the MADS-box gene family of *Lonicera macranthoides*, analyze its bioinformatics, validate

[收稿日期] 2023-11-27

[基金项目] 国家自然科学基金项目 (82373992); 湖南省自然科学基金项目 (2021JJ30497, 2021JJ30515); 国家现代农业产业技术体系 (CARS-21); 湖南中医药大学研究生创新训练项目 (2022CX79, 2023CX152); 2020 年湖南省一流专业建设点: 中药资源与开发; 湖南中医药大学重点学科中药学科 (校行发规字 [2023]2 号)。

[通信作者] * 刘湘丹, 女, 博士, 教授, 博士研究生导师, E-mail: paeonia_dd@126.com; 周日宝, 男, 博士, 教授, 博士研究生导师, E-mail: 1057323510@qq.com。

the expression patterns, and clone the full-length sequences of *CMB1*, a member of the MADS-box family of *Lonicera macranthoides* of Xianglei and wild types. **Methods** Based on the transcriptome data, the bioinformatics of *Lonicera macranthoides* MADS-box was analyzed with online tools, the expression patterns of MADS-box genes in different cultivars were verified using qRT-PCR, and the full-length sequences of *Lonicera macranthoides* *CMB1* gene of the Xianglei and wild types were cloned by RT-PCR and RACE techniques. **Results** The sequence lengths of 28 MADS-box proteins of *Lonicera macranthoides* ranged from 89 aa to 359 aa, with alkali proteins accounting for 85.7%, unstable proteins accounting for 96.4%, and hydrophilic proteins accounting for 96.4%, which were localized in the nucleus, and were non-secretory proteins without transmembrane structure. The transcriptome showed that compared with the wild type, 17.86% of the 28 MADS-box genes in the Xianglei type and 18.75% of the genes of MIKC^C type were downregulated. The full-length sequences of *CMB1* genes, which were highly specifically expressed in two varieties of flowers of *Lonicera macranthoides*, were cloned. Both of them contained a 738 bp ORF, encoding 245 amino acids. There was a significant difference in expression level of *CMB1* gene between the two varieties of flowers ($P<0.01$), and it was highly and specifically expressed in flowers compared to that in stems and leaves ($P<0.01$). **Conclusion** Based on the transcriptome data of *Lonicera macranthoides*, 28 MADS-box family genes were identified, and the full-length sequences of *CMB1* gene of Xianglei and wild types were cloned, which provides the research foundation and theoretical basis for the further study of molecular mechanism of development and superior phenotype formation of *Lonicera macranthoides*.

[**Keywords**] *Lonicera macranthoides*; MADS-box gene family; expression analysis; *CMB1* gene; gene clone; phenotypic variation

MADS-box 是一类古老的基因家族,广泛存在于植物、动物和真菌中。MADS-box 基因家族分为Type I型和Type II型两个谱系^[1]。植物中Type I型分为M α 、M β 、M γ 3个亚家族,Type II型MADS-box基因分为MEF2型和MIKC型,MEF2型基因主要存在于真菌和动物中,而MIKC型基因为植物特有。MIKC型基因又分为MIKC^C型和MIKC^{*}型,植物中大多数MIKC型基因为MIKC^C型,包括12个亚家族SVP、FLC、TM3、AP1/FUL、SEP、AGL6、AG、AGL12、ANR1、AP3/PI、AGL15、BS^[2-3]。MADS-box基因影响植物生长发育过程和各个器官的形态建成,如根系、果实发育,对花器官生长发育的影响尤为显著,具体表现为可调控开花时间,改变花器官表型和花序形态等。此外,MADS-box基因家族也会参与植物抗逆境生理过程^[4-5]。从营养生长到生殖生长的转变是开花植物发育的关键过程,SEP亚族成员在调控花芽分化过程中发挥重要作用,同时在花的发育过程中参与到四轮花器官(花萼、花冠、雌蕊、雄蕊)的形态建成、开花时间等过程,对植物的花发育过程有重要作用^[6-8]。

灰毡毛忍冬 *Lonicera macranthoides* Hand.-Mazz 为忍冬科忍冬属(*Lonicera*)植物,以干燥花蕾或者初开的花入药,是中药山银花的基源之一。本药性寒,味甘,归肺、心、胃经,有清热解毒、疏散风热的功效,

用于治疗痈肿疔疮、喉痹、丹毒、热毒血痢、温病发热等疾病^[9-10]。现代研究证明,山银花具有抗菌抗炎、解热抗毒、保肝利胆的功效,广泛用于防治传染类疾病,如流感、手足口病、病毒性肺炎等^[11]。目前,收载山银花的中成药有24个,且山银花在银翘伤风胶囊、清热银花糖浆、风热清口服液等中成药中均为君药^[12]。山银花主要在湖南、重庆、贵州等地区生产,灰毡毛忍冬是山银花的主要来源,其中湖南省隆回县是全国最大的山银花生产区,年产量占全国总产量53%^[13]。野生型灰毡毛忍冬的花蕾不整齐,花蕾期短,花冠展开后一两天即凋谢,需分批采收,若采收不及时易造成药材品质下降和资源浪费。20世纪90年代末,研究者在湖南溆浦发现灰毡毛忍冬自然突变株,该突变株经多代无性繁殖和筛选,培育出了优良的无性品种,命名为“湘蕾金银花”*Lonicera macranthoides* “Xianglei”(LmXL),即湘蕾型灰毡毛忍冬,表现为花冠不展开、花蕾整齐和花蕾期长达20余天,可一次性采收^[14]。研究表明,灰毡毛忍冬中主要成分绿原酸含量随花冠的展开而下降^[15-16],湘蕾型所具有的良好表型,为山银花药材的采收、加工节约了成本,保障了药材品质的稳定性,有利于资源高效利用。但湘蕾型灰毡毛忍冬优良表型形成的分子机制仍未被揭晓。

目前,本课题组已经从灰毡毛忍冬中克隆得到 MADS-box 家族中 *API*^[17]、*AGL19*^[18]、*AGL15*^[14] 基因,龙丽君等^[19]克隆得到 *SVP* 基因。然而,尚没有研究对灰毡毛忍冬 MADS-box 家族基因进行系统地挖掘与鉴定,也未见 SEP 亚族基因在灰毡毛忍冬中报道。因此,本文以湘蕾型、野生型灰毡毛忍冬的花为材料进行转录组测序。基于 RNA-seq 结果,对灰毡毛忍冬 MADS-box 家族基因进行鉴定,分析两个灰毡毛忍冬品种中 MADS-box 家族基因的表达情况,并克隆 SEP 亚族成员 *CMB1*,以期为进一步研究灰毡毛忍冬花发育奠定研究基础,为探讨湘蕾型灰毡毛忍冬花冠形态变异机制提供参考依据。自 SEP 亚族基因被发现以来,该亚族成员的生物学功能多于花中体现^[20]。因此,本研究只对 *CMB1* 基因在品种间花的相对表达量和在品种内茎、叶与花的相对表达量做统计学分析。

1 材料与方法

1.1 样品来源与处理

两个灰毡毛忍冬品种均采自湖南省隆回县,经湖南中医药大学刘湘丹教授鉴定为忍冬科灰毡毛忍冬 *Lonicera macranthoides* Hand.-Mazz. 的野生型和湘蕾型。按课题组前期花期分类标准,将灰毡毛忍冬花样品分类,即花期 1~7^[21],详见图 1。各个花期花样品分装密封,经液氮速冻后于 -80 °C 冰箱保存,送至上海百趣生物医学科技有限公司进行转录组测序。

1.2 主要试剂

Biospin 多糖多酚植物总 RNA 提取试剂盒(杭州博日科技有限公司,批号:BSC65S1);RevertAid First Strand cDNA Synthesis Kit(美国 ThermoFisher Scientific 公司,批号:K1622);TranStart® Green qPCR Su-

perMix UDG(北京全式金生物科技有限公司,批号:AQ111);2×Taq MasterMix(Dye)(康为世纪科技有限公司,批号:CW0682);DNA 凝胶回收试剂盒、Trelief™ 5α 感受态细胞(北京擎科新业生物技术有限公司,批号:GE0101、TSC01);SMARTer® RACE 5'/3' Kit(美国 Clontech 公司,批号:634858)。

1.3 方法

1.3.1 灰毡毛忍冬 MADS-box 家族基因筛选与鉴定 在 Pfam 数据库(<http://pfam-legacy.xfam.org/>)下载 MADS-box 家族基因保守结构域 SRF-TF(PF00319)隐马尔可夫模型(Hidden Markov Model, HMM)文件,以此文件为种子文件利用 TBtools 软件在灰毡毛忍冬转录组数据中筛选 MADS-box 家族蛋白序列(E-value<1×10⁻⁵)。从 TAIR 数据库(<https://www.arabidopsis.org/>)下载拟南芥 MADS-box 家族基因蛋白序列为参考,使用 TBtools 软件中的生物大分子序列比对搜索工具(Basic Local Alignment Search Tool, BLAST)进行比对。将 BLAST 与 HMM 筛选出的蛋白质序列结果取交集,通过 SMART(<https://smart.embl.de/>)、NCBI-CDD(<https://www.ncbi.nlm.nih.gov/Structure/index.shtml>)、Pfam 数据库对筛得的序列逐一验证,删除结构域不完整的蛋白序列及冗余序列。

1.3.2 灰毡毛忍冬 MADS-box 家族基因生物信息学分析 通过 ProtParam 网站(<https://web.expasy.org/protparam/>)预测其蛋白理化性质,包括编码蛋白的分子式、相对分子质量、不稳定系数及等电点等;采用 ProtScale 网站(<https://web.expasy.org/protscale/>)预测蛋白亲/疏水性;WOLF PSORT 网站(<https://wolfpsort.hgc.jp/>)预测蛋白质亚细胞定位;TMHMM 2.0 网站(<http://www.cbs.dtu.dk/services/TMHMM/>)分析蛋白质跨膜结构;SignalP 4.1 Server 网站(<http://www.cbs.dtu.dk/services/SignalP/>)

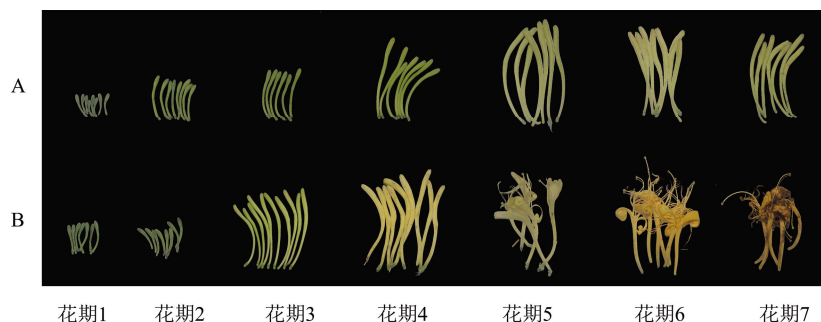


图 1 湘蕾型和野生型灰毡毛忍冬不同时期花样品

注:A.湘蕾型灰毡毛忍冬;B.野生型灰毡毛忍冬。

dk/services/SignalP-4.1/)进行信号肽预测;运用SOPMA网站(https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_server.html)预测蛋白二级结构;SWISS-MODEL网站(<https://swissmodel.expasy.org/>)预测蛋白三级结构;MEME网站(<https://meme-suite.org/meme/tools/meme>)分析灰毡毛忍冬 MADS-box 家族蛋白保守基序(motif),查找基序数设定为10,其他参数保持默认值。通过TBtools软件对motif结果进行可视化分析;拟南芥、灰毡毛忍冬 MADS-box 家族蛋白序列取合集,通过TBtools软件的音乐 Wrapper 程序进行多序列比对,再经过 trimAL Wrapper 程序修剪后,以最大似然法构建系统发育树。

1.3.3 灰毡毛忍冬 MADS-box 家族基因表达分析与验证 以两个灰毡毛忍冬品种 MADS-box 家族基因在花中的 TPM 值为变量,通过TBtools软件绘制热图。选取9个 MADS-box 家族表达差异基因进行实时荧光定量 PCR 验证。通过 Biospin 多糖多酚植物总 RNA 提取试剂盒提取两个灰毡毛忍冬花品种的总 RNA,并定量至 100 ng/ μ L;通过逆转录获得 cDNA 为模板;以 18S 基因为内参进行 qRT-PCR 验证。反应体系 20 μ L:cDNA 1 μ L,上、下游引物各 0.4 μ L, 2xTansStart® Green qPCR SuperMix UDG 10 μ L, Nuclease-free water 8.2 μ L。扩增程序:50 $^{\circ}$ C 2 min; 94 $^{\circ}$ C 预变性 10 min;94 $^{\circ}$ C 变性 5 s,55 $^{\circ}$ C 退火 15 s, 72 $^{\circ}$ C 延伸 10 s;40 个循环。每个样品 3 次生物学重复, 2^{- $\Delta\Delta$} 法计算基因的相对表达量。

1.3.4 灰毡毛忍冬 CMB1 基因的克隆 以 1.3.3 项中 cDNA 为模板进行核心片段 PCR 扩增。反应体系 25 μ L;2xTaq MasterMix (Dye) 12.5 μ L,ddH₂O 9.5 μ L, CBM1-F 1 μ L,CMB1-R 1 μ L,cDNA 1 μ L。反应条件:94 $^{\circ}$ C 预变性 5 min;94 $^{\circ}$ C 变性 30 s,55 $^{\circ}$ C 退火 30 s,72 $^{\circ}$ C 延伸 1 min 30 s,循环 40 次;72 $^{\circ}$ C 延伸 10 min。1.2%琼脂糖凝胶电泳纯化 PCR 产物,目的条带经 DNA 凝胶回收试剂盒回收并连接至 pEASY®-T1 载体,热激转化至 Trelief™ 5 α 感受态细胞并涂布于含有 Amp,IPTG,X-gal 的 LB 固体培养基进行蓝白斑筛选,挑取白色单菌落进行菌落 PCR 鉴定,将阳性菌液过夜培养后送至上海生工生物工程股份

有限公司测序。根据核心片段测序结果,设计 RACE 特异性引物。按照 SMARTer® RACE 5'/3' Kit 说明书获得 5'和 3'RACE-Ready cDNA 并进行 RACE 扩增。目的条带回收和测序等过程同前。通过 ContingExpress 软件对 5'端和 3'端进行序列拼接,得到两个灰毡毛忍冬品种的 CMB1 基因全长,根据拼接 cDNA 全长序列设计验证引物进行全长验证。反应体系及条件同核心片段克隆,其中退火温度调整至 60 $^{\circ}$ C。扩增目的条带回收和测序等过程同前。引物序列详见表 1。

2 结果

2.1 灰毡毛忍冬 MADS-box 基因鉴定与蛋白理化性质

通过 BLAST、HMM 分别鉴定出 51、35 条 MADS-box 蛋白序列,两种方法所得结果取交集,经过结构域检验和冗余序列删除后,最终从灰毡毛忍冬转录组中得到 28 条 MADS-box 家族蛋白序列。结合拟南芥 MADS-box 蛋白名与 NCBI-Blastp 结果对其命名,详见表 2。通过 ProtParam、ProtScale 网站在线分析蛋白质理化性质,序列长度 89 aa(LmGLO-2)~359 aa(LmAGL65);相对分子质量为 10 172.82(LmGLO-2)~41 162.39(LmAGL65);等电点为 6.21(LmAGL65)~10.01(LmSVP),其中酸性蛋白 4 个,碱性蛋白 24 个,碱性蛋白占比 85.7%;蛋白不稳定系数 25.01(LmTM6)~73.39(LmSVP),其中稳定蛋白 1 个(不稳定系数 < 40),不稳定蛋白 27 个(不稳定系数 > 40),不稳定蛋白占比 96.4%;脂肪系数为 66.71(LmAGL80-2)~98.5(LmAGL12)。MADS-box 蛋白总平均疏水指数 (GRAVY) 为 -0.883(LmTM6)~-0.284(LmAGL12),其中亲水性蛋白 27 个(GRAVY < -0.5),占比 96.4%,两性蛋白 1 个(-0.5 < GRAVY < 0.5)。

2.2 灰毡毛忍冬 MADS-box 蛋白的亚细胞定位、信号肽和跨膜结构

WOLF PSORT 亚细胞定位预测显示,灰毡毛忍冬 MADS-box 蛋白均定位于细胞核中。SignalP 4.1 Server 网站预测可知 MADS-box 蛋白不存在含信号肽序列,为非分泌型蛋白。TMHMM 2.0 网站分析蛋白跨膜结构,结果表明 MADS-box 蛋白不具跨膜区域,且所有氨基酸均在膜外。

表1 引物序列表

引物名称	序列(5'-3')	目标长度/bp	用途
qLmAGL80-1	正向:AAGGATCGCGAAAGCTAGTG 反向:TACCGGTCAAACATTGGTACAT	95	qRT-PCR 验证
qLmAGL62-1	正向:GCAATAGGCTCGTTGAGATAGG 反向:GACCCATGCTTCCAGGATTT	78	同上
qLmGLO-1	正向:AAGCTGACTGGGAAGAGATTTG 反向:TCAGCCTTCATGTGCCTTAG	116	同上
qLmAGL65	正向:GGCTTCCCGTAATGAGAAT 反向:ACCAGACTAGCTAGGAAGAGAG	110	同上
qLmAGL12	正向:GAAGGCTAAGGAAGCTCTCTGTG 反向:GGTTCCTTTGGTGGCTAGTT	97	同上
qLmSCO1-1	正向:TGATGCTGAGGTTGCTCTTAT 反向:GTGCTTCCCATAGCGTTCTA	100	同上
qLmCMB1	正向:CAAGAGACTGGTCTGCTGATG 反向:CATTCCAGTGGCTCGAAGAA	104	同上
qLmAP1-1	正向:CTGGATCCGAATCCCAGATAAG 反向:CAAGATCTTCTCCACGTAGTT	105	同上
qLmAP1-2	正向:ACGTGGCTCTGATTGTCTTC 反向:CATACCGCTCAAGGATCTTCTC	87	同上
18S	正向:CTTCGGGATCGGAGTAATGA 反向:GCGGAGTCTAGAAGCAACA	118	内参基因
CMB1	正向:GATAGAGAACAAAATAAACAGGCAG 反向:ATTAACAACACAACACTACCCAAAAA	805	核心片段克隆
LmCMB1-3'	正向:GAATGGTCTTCTGAAAAAGGCCTATG	957	RACE 扩增
LmXLCMB1-3'	正向:GATAGAGAACAAAATAAACAGGCAG	998	同上
CMB1-5'	反向:ATTAACAACACAACACTACCCAAAAA	867	同上
LmCMB1	正向:AAGAGAGTAGTAGGTAGAGAGAGAGAGAGAT 反向:AAAAAGGATTCATTTTCATTAATTGTTCTTCAAATTACT	1 064	全长验证
LmXLCMB1	正向:AAGAGAGACTAGTAGGTAGAGAGAGAGAGAGAG 反向:GCAGGGCTAAAAGGATTCATTTTCATATGTCT	1 062	同上

2.3 灰毡毛忍冬 MADS-box 蛋白基序、二级、三级结构

利用 MEME 网站在灰毡毛忍冬 MADS-box 蛋白中共鉴定出 10 个 Motif, 长度为 11~50 aa。28 个 MADS-box 蛋白分别含有 2~7 个 Motif。除 LmGLO-1 外, 均含有 Motif 1; 除 LmAGL80-1、LmAGL80-2、LmAGL63-2、LmAGL65 外, 均含有 Motif 3; 除 LmAGL62-1、LmAGL62-3、LmAGL62-3、LmSVP、LmSCO1-3、LmGLO2-2 外, 均含有 Motif 2。通过 TBtools 软件对 10 个 Motif 进行可视化, 其中 Motif 1、Motif 3 属于 MADS 结构域, Motif 2、Motif 4 属于 K-box 结构域, 表明 MEME 网站预测结果可信。详见图 2。

灰毡毛忍冬 MADS-box 蛋白二级结构中 α 螺旋占比 26.97%(LmGLO-2)~69.32%(LmSCO1-2), 均值为 54.01%; β 转角占比 2.24%(LmAGL42-1)~

17.98%(LmGLO-2), 均值为 5.45%; 无规则卷曲占比 7.69%(LmAGL80-1、LmAGL80-2)~30.34%(LmGLO-2), 均值为 12.77%; 延伸链占比 11.33%(LmAP1-1)~52.37%(LmAGL65), 均值为 27.77%; α 螺旋占比最高, 为灰毡毛忍冬 MADS-box 蛋白二级结构的主要结构组成元件, 延伸链为次要原件, 详见表 3。SWISS-MODEL 软件预测的灰毡毛忍冬 MADS-box 蛋白三级结构可见 α 螺旋、 β 转角、无规则卷曲、延伸链, 分布比例与二级结构预测结构相符, 详见图 3。

2.4 灰毡毛忍冬 MADS-box 蛋白系统进化

将 106 个拟南芥 MADS-box 蛋白和 28 个灰毡毛忍冬 MADS-box 蛋白通过最大似然法构建系统发育树, 详见图 4。28 个灰毡毛忍冬 MADS-box 蛋白可分为 5 个 Type I 型与 23 个 Type II 型。其中, Type I 型中 M α 亚族 3 个、M γ 亚族 2 个、M β 亚族

表 2 灰毡毛忍冬 MADS-box 家族蛋白的理化性质

蛋白名称	基因 ID	氨基酸长度/aa	相对分子质量	理论等电点	不稳定系数	脂肪系数	平均疏水指数
LmAGL80-1	TRINITY_DN38649_c0_g2	234	26 776.05	9.48	53.40	67.14	-0.658
LmAGL80-2	TRINITY_DN38649_c0_g1	234	26 840.16	9.48	53.45	66.71	-0.654
LmAGL62-1	TRINITY_DN49310_c1_g5	193	21 986.14	8.83	54.50	73.73	-0.741
LmAGL62-2	TRINITY_DN49310_c1_g4	193	21 887.16	9.32	56.34	73.26	-0.720
LmAGL62-3	TRINITY_DN4185_c0_g1	242	28 038.02	6.68	52.93	80.08	-0.621
LmAGL65	TRINITY_DN53229_c2_g2	359	41 162.39	6.21	60.09	79.30	-0.673
LmSCO1-1	TRINITY_DN45858_c3_g2	214	24 378.85	7.71	45.85	84.72	-0.711
LmSCO1-2	TRINITY_DN44423_c0_g1	176	20 516.97	9.68	53.77	88.64	-0.580
LmSCO1-3	TRINITY_DN53368_c1_g4	179	20 670.59	9.37	33.07	73.07	-0.323
LmSVP	TRINITY_DN51230_c2_g1	158	18 495.38	10.01	73.39	83.29	-0.454
LmAGL12	TRINITY_DN59132_c0_g1	200	22 956.90	7.68	49.93	98.50	-0.284
LmAGL15	TRINITY_DN43585_c2_g1	244	27 815.79	8.67	60.67	82.34	-0.647
LmAGL42-1	TRINITY_DN44261_c5_g2	223	25 856.61	8.47	69.51	87.94	-0.673
LmAGL42-2	TRINITY_DN50865_c0_g1	146	17 062.69	9.59	54.61	93.49	-0.644
LmAGL11	TRINITY_DN52309_c1_g1	243	28 120.17	9.31	53.30	85.51	-0.652
LmAGL8	TRINITY_DN48224_c1_g1	244	27 781.03	9.56	54.08	86.35	-0.499
LmAP1-1	TRINITY_DN48323_c1_g1	150	17 522.18	9.64	59.68	98.20	-0.625
LmAP1-2	TRINITY_DN49917_c3_g1	242	28 001.90	8.75	56.77	87.07	-0.707
LmTM6	TRINITY_DN54663_c1_g1	229	26 689.18	8.90	25.01	71.92	-0.883
LmGLO-1	TRINITY_DN55933_c3_g6	141	16 553.00	6.31	48.69	88.58	-0.718
LmGLO-2	TRINITY_DN55933_c1_g1	89	10 172.82	9.84	64.04	76.63	-0.581
LmDEFH21	TRINITY_DN39025_c0_g1	277	31 796.22	6.67	55.61	74.26	-0.694
LmAG1	TRINITY_DN57150_c1_g1	159	18 168.76	9.76	50.56	78.55	-0.774
LmSEP1	TRINITY_DN53368_c0_g4	246	28 039.09	9.01	58.04	77.68	-0.591
LmSEP3-1	TRINITY_DN50642_c0_g1	242	27 827.58	8.80	44.97	77.81	-0.772
LmSEP3-2	TRINITY_DN50642_c0_g2	154	17 939.66	9.61	45.59	89.94	-0.613
LmSEP4	TRINITY_DN52649_c6_g1	245	28 318.48	8.66	52.01	82.45	-0.600
LmCMB1	TRINITY_DN52649_c6_g2	230	26 589.32	8.01	54.17	86.52	-0.597

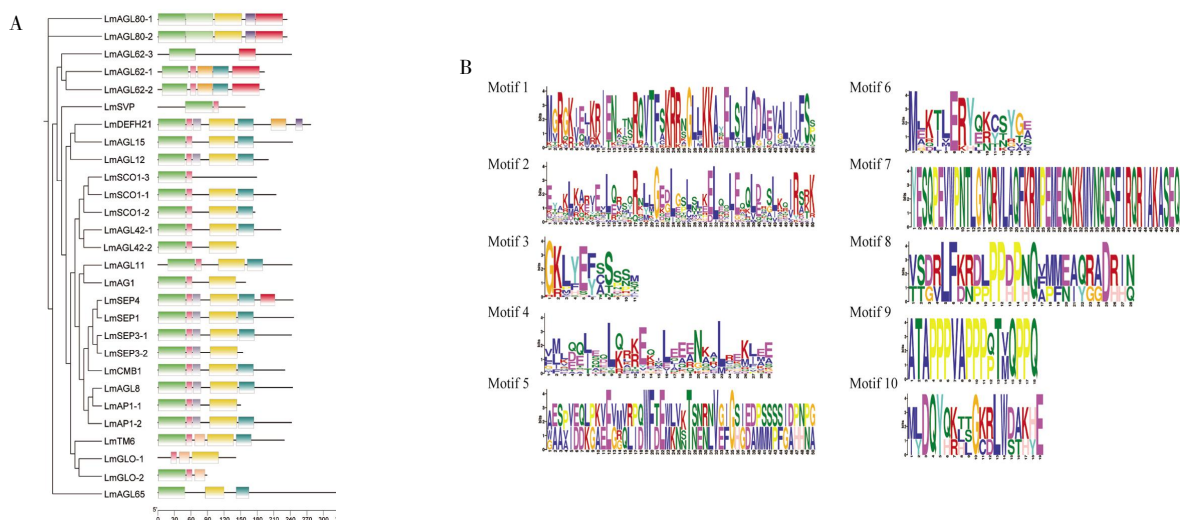


图 2 灰毡毛忍冬 MADS-box 家族基因蛋白基序分析

注:A.MADS-box 家族基因 Motif 分布;B.MADS-box 家族基因 Motif 的氨基酸分布。

成员 0 个;Type II 型中,MIKC^{*} 型只有 1 个成员,MIKC^c 型中除 AGL6、FLC 亚族外,在其余 11 个亚族均有分布;SVP 亚族 1 个、TM3 亚族 4 个、AP1/FUL

亚族 3 个,SEP 亚族 5 个、AG 亚族 1 个、AGL12 亚族 1 个、ANR1 亚族 1 个、AP3/PI 亚族 3 个、BS 亚族 1 个,其中 SEP 亚族成员最多。

表3 灰毡毛忍冬 MADS-box 家族蛋白的二级结构

蛋白名称	α 螺旋/%	β 转角/%	无规则卷曲/%	延伸链/%
LmAGL80-1	54.27	2.56	7.69	35.47
LmAGL80-2	53.85	2.56	7.69	35.90
LmAGL62-1	47.15	5.70	13.99	33.16
LmAGL62-2	45.60	5.70	13.99	34.72
LmAGL62-3	40.50	5.37	19.42	34.71
LmAGL65	32.31	3.90	11.42	52.37
LmSCO1-1	61.21	2.80	9.35	26.64
LmSCO1-2	69.32	3.98	8.52	18.18
LmSCO1-3	57.54	6.15	13.41	22.91
LmSVP	34.18	11.39	25.95	28.48
LmAGL12	58.00	4.00	11.50	26.50
LmAGL15	51.64	2.87	10.25	35.25
LmAGL42-1	60.54	2.24	9.42	27.80
LmAGL42-2	62.33	4.79	13.70	19.18
LmAGL11	60.08	3.70	11.11	25.10
LmAGL8	55.33	4.51	9.02	31.15
LmAP1-1	66.67	6.67	15.33	11.33
LmAP1-2	61.16	4.55	9.50	24.79
LmTM6	56.33	10.04	14.85	18.78
LmGLO-1	59.57	4.96	12.06	23.40
LmGLO-2	26.97	17.98	30.34	24.72
LmDEFH21	52.35	4.69	10.11	32.85
LmAG1	67.30	6.29	11.32	15.09
LmSEP1	53.66	4.47	9.76	32.11
LmSEP3-1	57.44	4.13	9.92	28.51
LmSEP3-2	61.04	6.49	15.58	16.88
LmSEP4	49.80	4.90	8.98	36.33
LmCMB1	56.09	5.22	13.48	25.22

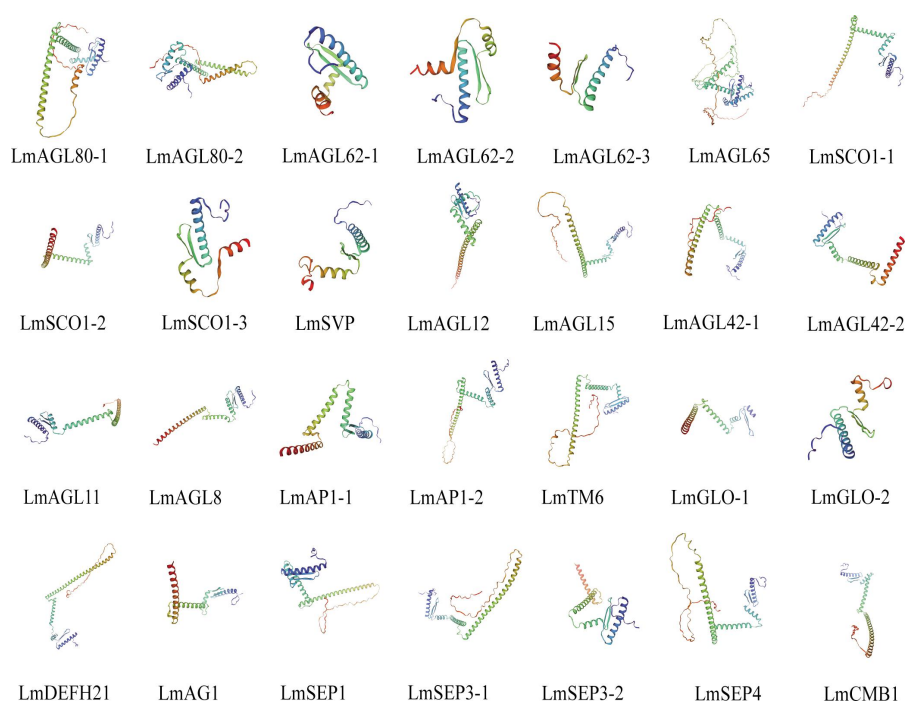


图3 灰毡毛忍冬 MADS-box 家族蛋白三级结构

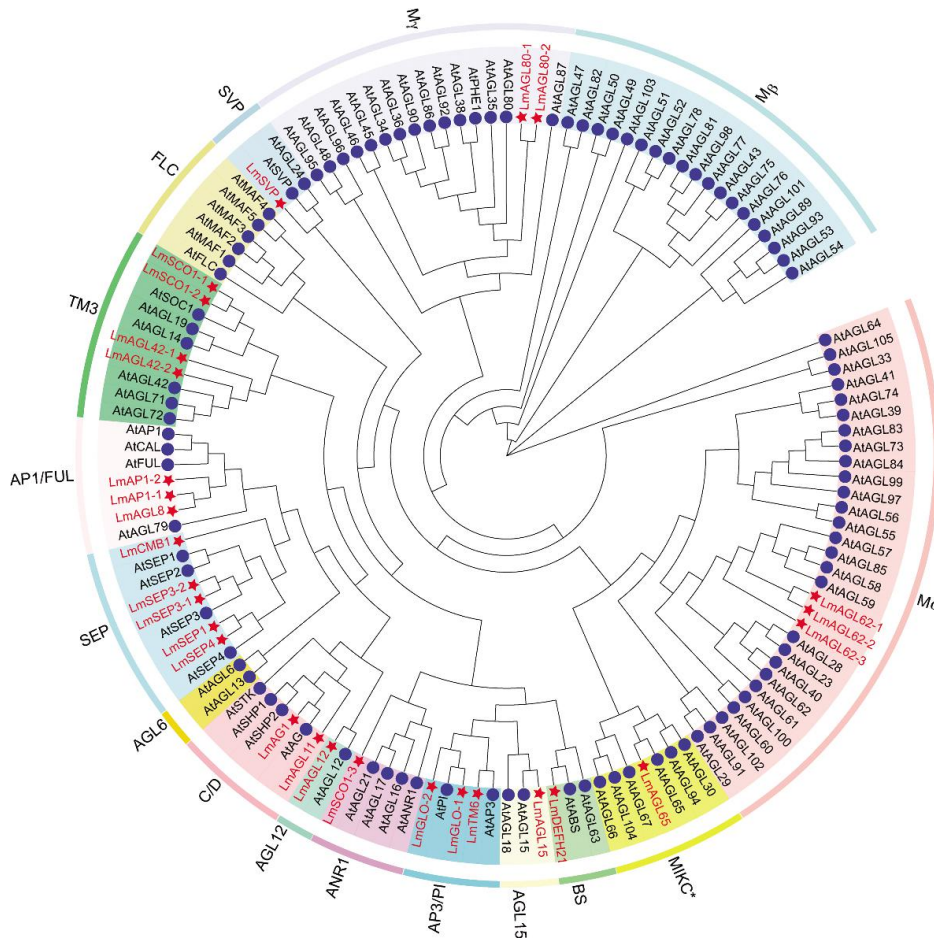


图4 灰毡毛忍冬和拟南芥 MADS-box 家族基因系统进化树

2.5 灰毡毛忍冬 MADS-box 家族基因表达情况

对两个品种的灰毡毛忍冬花期4转录组结果中28个MADS-box家族基因表达量做聚类热图,详见图5。转录组结果显示,在第4花期,与野生型比较,湘蕾型中有2个基因(*LmAGL80-1*、*LmAGL65*)表达显著上调,3个基因(*LmAPI-1*、*LmGLO-1*、*LmSCO1-1*)表达显著下调。除此之外,还有8个基因相对表达量在湘蕾型中高于野生型,15个基因相对表达量低于野生型。在湘蕾型 Type I 型中3个基因的表达量高于野生型,2个基因相对表达量低于野生型;Type II 中MIKC^a 仅有1个基因(*LmAGL65*)表达量高于野生型;MIKC^c 中7个基因相对表达量高于野生型,16个基因相对表达量低于野生型。湘蕾型中,AP1/FUL亚族3个基因相对表达量低于野生型,其中*LmAPI-1*表达显著下调;3个AP3/PI亚族基因相对表达量均高于野生型;SEP亚族基因中*LmSEP4*相对表达量高于野生型,其余4个基因相对表达量均低于野生型。总体而言,与野生型比较,湘蕾型中28个MADS-box基因有17.86%表达下调,MIKC^c型基因中有18.75%表达下调。

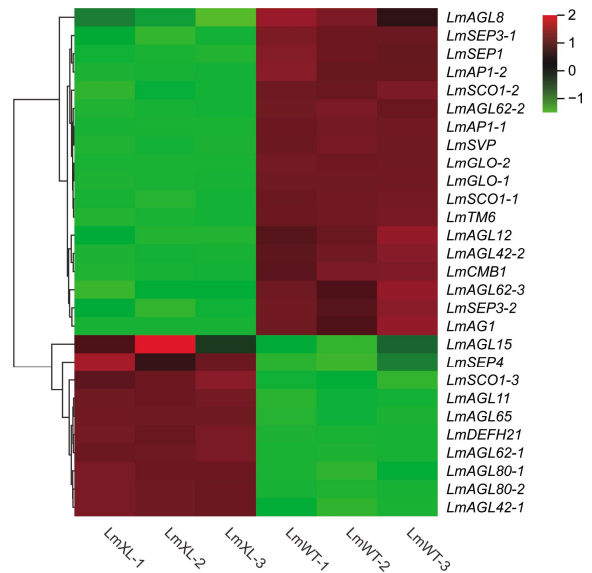


图5 灰毡毛忍冬 MADS-box 基因表达量热图

注:LmXL:湘蕾型灰毡毛忍冬;LmWT:野生型灰毡毛忍冬;1、2、3分别表示3次生物学重复。

选取9个代表性基因(*LmAGL12*、*LmAGL65*、*LmAGL80-1*、*LmAPI-1*、*LmAPI-2*、*LmTM6*、*LmGLO-1*、*LmCMB1*、*LmSCO1-1*)通过qRT-PCR对转录组结果进行验证。结果显示,qRT-PCR结果与RNA-seq趋

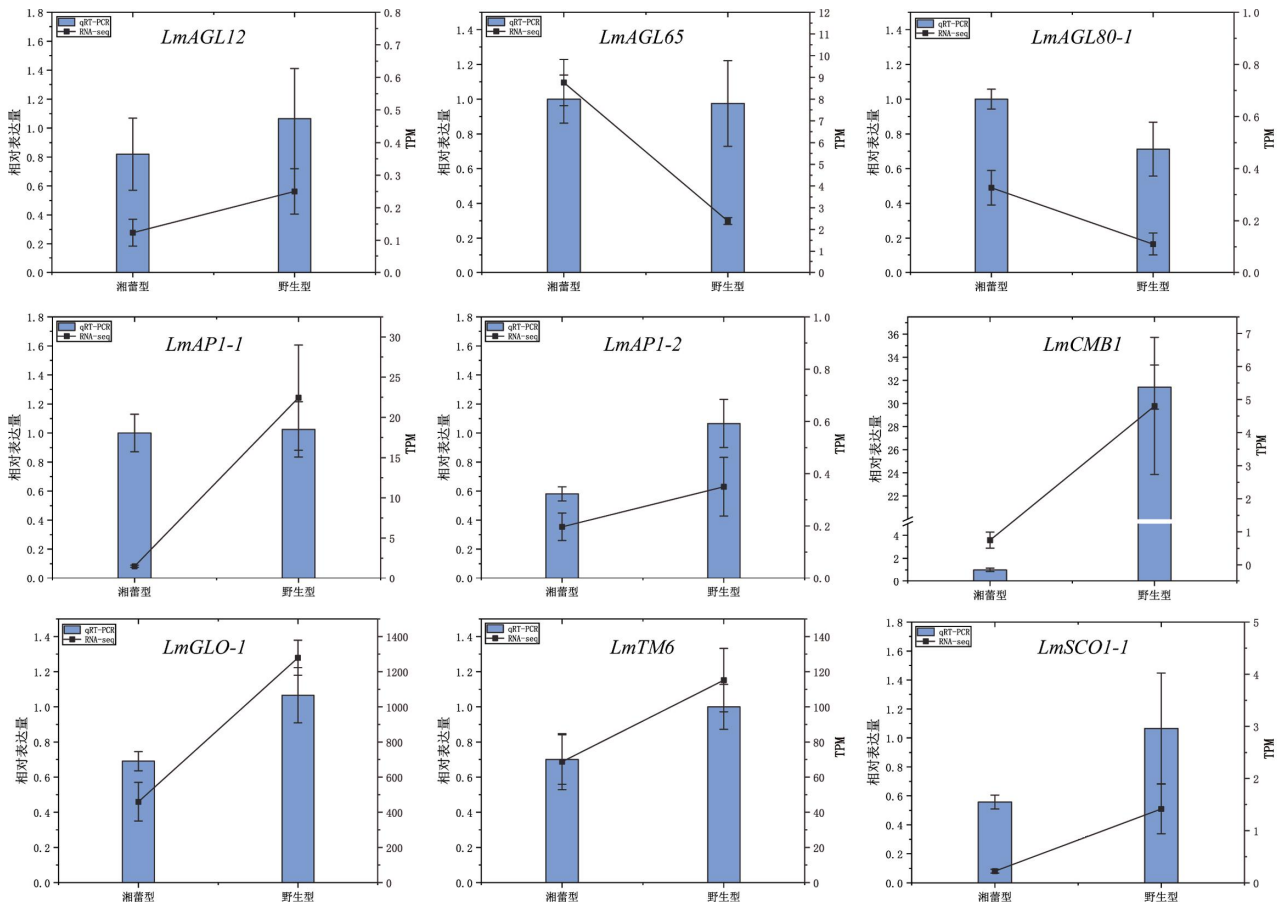


图6 两个品种的灰毡毛忍冬 MADS-box 基因表达模式 qRT-PCR 验证

势一致,湘蕾型中除 *LmAGL65*、*LmAGL80-1* 表达量高于野生型,其余 7 个基因表达量均低于野生型 ($P<0.05$)。详见图 6。

2.6 两个品种的灰毡毛忍冬 *CMB1* 基因的克隆

与野生型花比较,湘蕾型花中 *CMB1* 基因表达量显著下调 ($P<0.01$)。进一步检测 *CMB1* 基因在野生型和湘蕾型灰毡毛忍冬花、茎、叶相对表达量,结果表明。两个品种的灰毡毛忍冬 *CMB1* 基因在花中的表达量都显著高于叶和茎 ($P<0.01$),详见图 7。

基于 *CMB1* 在花中高度特异性表达的特征,及两个品种的灰毡毛忍冬花的极显著表达差异,进一步对两个品种的灰毡毛忍冬 *CMB1* 基因进行克隆研究。以灰毡毛忍冬花总 RNA 反转录产物为模板扩增两个品种的灰毡毛忍冬 *CMB1* 核心片段,经测序与转录组 Unigene 序列匹配后,通过 RACE 技术得到两个灰毡毛忍冬品种的 *CMB1* 基因的 5'、3'端序列。RACE 产物经测序后通过 Conting Express 拼接,得到两个品种的灰毡毛忍冬 *CMB1* 基因全长序列。最终利用全长验证引物对两个品种的灰毡毛忍冬 *CMB1* 基因进行全长验证,凝胶电泳图,详见图 8A。

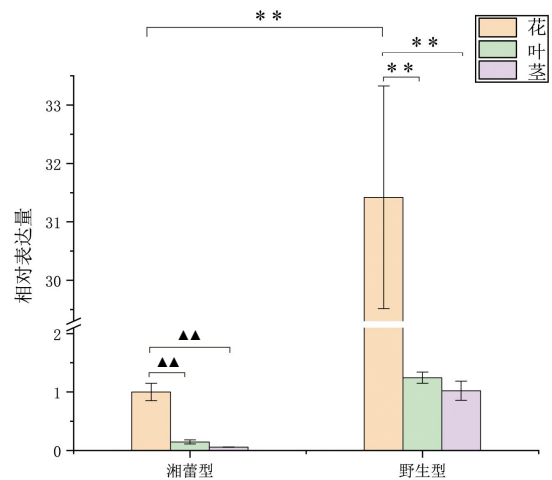


图7 两品种灰毡毛忍冬 *CMB1* 基因不同部位表达量
注:与野生型花比较,** $P<0.01$;与湘蕾型花比较,** $P<0.01$ 。

测序显示,野生型和湘蕾型 *CMB1* 基因全长分别为 1 064 bp、1 062 bp, 通过 DNAMAN 比对分析,PCR产物测序结果与拼接结果一致,上传至 NCBI 数据库分别命名为 *LmCMB1*(GenBank: OR 687631.1)、*LmXLCMB1*(GenBank: OR 687632.1)。

*LmCMB1*与*LmXLCMB1*均包含一个738bp的ORF,碱基无差异,编码 245 个氨基酸,翻译蛋白无差异,

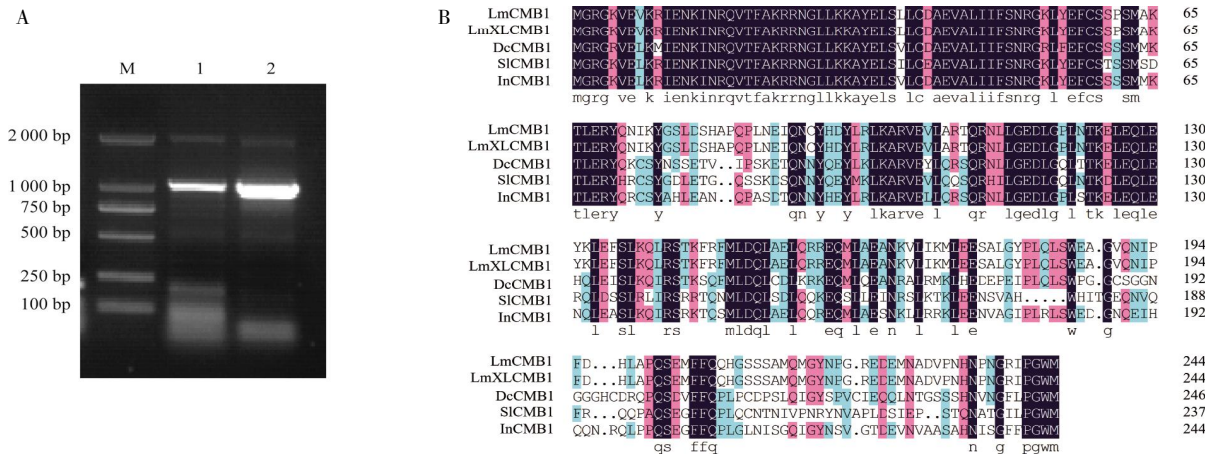


图 8 两个品种的灰毡毛忍冬 *CMB1* 基因全长克隆

注:A.灰毡毛忍冬 *CMB1* 基因凝胶电泳图;M.2 000 DNA Marker;1.Lm *CMB1*;2.LmXL *CMB1*;B.两个品种的灰毡毛忍冬 *CMB1* 蛋白与其他物种多序列比对图。

与黄石斛 *Dendrobium catenatum*(XP 020679397.1)、番茄 *Solanum lycopersicum*(NP 001362848.1)、牵牛 *Ipomoea nil*(XP 019199885.1)*CMB1* 基因编码蛋白序列比对结果显示总相似性达 77.51%。详见图 8B。

3 讨论

MADS-box 家族基因与植物生长发育关系密切,尤其在花器官发育过程中起重要作用。一般情况下,花萼由 A 类基因调控,花冠由 B 类和 A 类基因调控,雄蕊由 B 类和 C 类基因调控,雌蕊由 C 类基因调控,胚珠由 D 类和部分 C 类基因调控,以上四轮花器官发育均有 E 类基因参与,即调控花器官发育有 ABCDE 类模型^[22]。目前,研究者已从部分药用植物中挖掘得到 MADS-box 家族基因信息,如甘松^[23](转录组数据,20 个)、当归^[24](转录组数据,29 个)、忍冬^[25](基因组数据,48 个)。本研究从灰毡毛忍冬中共鉴定出 28 个 MADS-box 基因家族成员,其中 A 类基因 3 个、B 类基因 3 个、C 类基因 2 个、E 类基因 5 个,并克隆了湘蕾型与野生型灰毡毛忍冬 E 类基因 *CMB1* 的全长序列。E 类基因是植物花器官发育过程的协同因子,贯穿于整个花器官的发育过程。紫薇 *LiCMB1* 基因在花芽分化、花萼分化中起关键作用^[26],桃 *PpCMB1* 基因抑制 *PpDAM5* 基因对花芽萌发的抑制作用,正调控花芽萌发,并在桃花雄蕊和萼片的发育中期起积极的调控作用^[27]。关键多序列比对结果总相似达 77.51%,推测灰毡毛忍冬 *CMB1* 基因可能与其同源基因功能相似,参与了花芽分化、

花器官形态建成等过程。两个品种的灰毡毛忍冬 *CMB1* 蛋白虽然在理化性质与结构上并无差异,但在花的表达量上存在显著差异($P<0.01$)。因此,可能在基因调控网络与 PPI 网络中功能相同但作用存在差异,进而影响花器官的表型。

灰毡毛忍冬以花蕾或初开的花入药,因此,本研究对灰毡毛忍冬花冠展开前 1 个花期,即花期 4 的 MADS-box 基因表达量进行分析。湘蕾型中,28 个 MADS-box 家族基因的相对表达量总体上有 64.3% 低于野生型,有 17.86% 较野生型表达下调,MIKC^c 型中 72.7% 的基因相对表达量低于野生型,有 18.75% 较野生型表达下调。基于 ABCDE 类模型,经过 qRT-PCR 验证,直接调控花冠发育的 B 类基因中, *LmGLO-1* 表达下调,参与花冠发育的 A 类基因 *LmAPI-1* 表达下调, E 类基因中 *LmCMB1* 基因表达下调。已有研究表明,山茶 B 类基因 *CjDEF* 在烟草中过表达后,部分转基因植株的花瓣出现湘蕾型灰毡毛忍冬类似表型,即花瓣在发育过程中始终无法展开^[28]。在烟草中过表达水稻 *API* 基因可使烟草花期延长^[29]。紫苏 *PjAPI* 基因过表达拟南芥植株在表型上表现出早花、株高增加、生长加快的特点^[30]。因此,推测 *LmGLO-1*、*LmTM6*、*LmAPI-1*、*LmCMB1* 和灰毡毛忍冬的花冠发育与蕾期长短有关。此外,ABCDE 类模型上游基因开花整合因子 *LmSCO1* 在湘蕾型灰毡毛忍冬表达下调。*SCO1* 可直接整合植物开花启动中的年龄途径、赤霉素途径信号控制开花时间^[31-32]。据此推测,灰毡毛忍冬 *LmSCO1* 在调控花蕾长短时可能也

起到重要作用。目前,有关 *AGL12*、*AGL80*、*AGL65* 基因研究较少,*AGL12* 为少数优先在植物根部组织中表达的 *MADS-box* 基因,调控根系发育,*AGL80* 主要调控中心细胞和胚乳发育,*AGL65* 与花粉管发育有关^[33-35]。灰毡毛忍冬 *LmAGL12*、*LmAGL80*、*LmAGL65* 基因的功能有待进一步研究。

MADS-box 基因家族与植物的药用价值紧密相关。例如,菊 *FUL* 同源基因 *CmFL1* 参与菊花开花时间和小花形成的调控过程^[36],红花 A 类基因 *CtMADS24* 可直接影响红花花期长短、花萼的发育^[37]。除直接影响中药材外观性状外,*MADS-box* 家族基因对药用植物内含活性成分含量和抵抗逆境生长也具有调控作用,如人参 *PgMADS17-04* 基因既促进人参皂苷的生物合成,又能响应冷胁迫^[38]。黄花蒿 *AaSEPI* 调控分泌型腺毛的发育进而影响青蒿药材中青蒿素的含量^[39]。灰毡毛忍冬花的发育对山银花的产量、品质有至关重要的影响。湘蕾型灰毡毛忍冬花冠不展开、花蕾整齐、蕾期长的优良表型的产生,明显提高了灰毡毛忍冬农业生产效率。灰毡毛忍冬 *MADS-box* 基因家族中 *LmAGL15* 已证实与蕾期延长有关^[14],同家族 *LmAPI*^[17]、*LmAGL19*^[18] 基因均可能与湘蕾灰毡毛忍冬优良表型的变异有关。植物花器官的发育是一个复杂的过程,ABCDE 类模型各个基因之间存在复杂的调控关系,单个基因表达变化往往能够引起连锁反应,湘蕾型灰毡毛忍冬优良表型产生的分子机制仍需进一步探讨。探究湘蕾型灰毡毛忍冬花冠优良表型的分子机制对山银花药材的产业发展有重要意义,通过基因编辑技术应用到其他花类药材或园艺植物,对后续创新表型优良的种质具有指导意义与产业价值。

参考文献

- [1] WANG Y S, ZHANG J L, HU Z L, et al. Genome-wide analysis of the *MADS-box* transcription factor family in *Solanum lycopersicum*[J]. International Journal of Molecular Sciences, 2019, 20(12): 2961.
- [2] ALVAREZ-BUYLLA E R, LILJEGREN S J, PELAZ S, et al. *MADS-box* gene evolution beyond flowers: Expression in pollen, endosperm, guard cells, roots and trichomes[J]. The Plant Journal, 2000, 24(4): 457-466.
- [3] SCHILLING S, KENNEDY A, PAN S R, et al. Genome-wide analysis of *MIKC-type MADS-box* genes in wheat: Pervasive duplications, functional conservation and putative neofunctionalization[J]. The New Phytologist, 2020, 225(1): 511-529.
- [4] ZHAO P X, ZHANG J, CHEN S Y, et al. Arabidopsis *MADS-box* factor *AGL16* is a negative regulator of plant response to salt stress by downregulating salt-responsive genes[J]. The New Phytologist, 2021, 232(6): 2418-2439.
- [5] THEIßEN G, RÜMPLER F, GRAMZOW L. Array of *MADS-box* genes: Facilitator for rapid adaptation?[J]. Trends in Plant Science, 2018, 23(7): 563-576.
- [6] LIU C, XI W Y, SHEN L S, et al. Regulation of floral patterning by flowering time genes[J]. Developmental Cell, 2009, 16(5): 711-722.
- [7] LIN Z Y, ZHU G F, LU C Q, et al. Functional conservation and divergence of *SEPALLATA-like* genes in floral development in *Cymbidium sinense*[J]. Frontiers in Plant Science, 2023, 14: 1209834.
- [8] ZHANG C L, WEI L D, YU X M, et al. Functional conservation and divergence of *SEPALLATA-like* genes in the development of two-type florets in marigold[J]. Plant Science, 2021, 309: 110938.
- [9] 朱香梅, 李 晴, 石雨荷, 等. 山银花的研究进展及其质量标志物预测分析[J]. 世界中医药, 2022, 17(13): 1860-1868.
- [10] 国家药典委员会. 中华人民共和国药典: 2020年版. 一部[M]. 北京: 中国医药科技出版社, 2020: 320.
- [11] 路俊仙, 张才波. 金银花与灰毡毛忍冬治疗流感病毒 FM1 肺炎小鼠作用比较[J]. 中华中医药杂志, 2022, 37(10): 6036-6040.
- [12] 刘 萍, 贾 凡, 胡 翔, 等. 山银花资源的功能特性与综合利用[J]. 食品与机械, 2023, 39(9): 219-227.
- [13] 刘 森, 曾德超, 苗 蕾, 等. 湖南省金银花产业发展现状、问题与对策研究[J]. 企业技术开发, 2018, 37(9): 37-39.
- [14] 易刚强, 蔡嘉洛, 朱贻霖, 等. 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[J]. 中草药, 2016, 47(4): 640-647.
- [15] 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.
- [16] 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.
- [17] SHAN W, PENG M C, XUN C, et al. Molecular cloning and spatiotemporal expression of *APETALA1-like* gene in *Lonicera macranthoides*[J]. Journal of Genetics, 2018, 97(5): 1281-1288.
- [18] LONG Y Q, LIU X, ZENG J, et al. Cloning and expression of *AGL19* gene in two *Lonicera macranthoides* varieties[J]. Journal of Genetics, 2021, 100: 13.

- [19] 龙丽君, 刘思思, 曾慧杰, 等. 灰毡毛忍冬 MADS-box 家族基因 SVP 克隆及表达分析[J]. 中草药, 2023, 54(16): 5350-5357.
- [20] DRENI L, FERRÁNDIZ C. Tracing the evolution of the SEPALATA subfamily across angiosperms associated with neo- and sub-functionalization for reproductive and agronomically relevant traits[J]. *Plants*, 2022, 11(21): 2934.
- [21] 刘畅宇, 陈 勋, 陈 娅, 等. 不同品种灰毡毛忍冬 ACS3 基因克隆、表达及生物信息学分析[J]. 中草药, 2019, 50(9): 2154-2164.
- [22] THEISSEN G, BECKER A, ROSA A D, et al. A short history of MADS-box genes in plants[J]. *Plant Molecular Biology*, 2000, 42(1): 115-149.
- [23] 崔 琪, 俸明康, 丰日落, 等. 基于转录组信息的甘松 MADS-box 转录因子家族分析[J]. 中草药, 2023, 54(3): 898-906.
- [24] 崔秀文, 刘 迪, 黄天苗, 等. 当归 MADS-box 生物信息学及 SOC1 克隆与表达分析[J]. 中草药, 2023, 54(5): 1551-1560.
- [25] LIN Y, QI X W, WAN Y, et al. Genome-wide analysis of the MADS-box gene family in *Lonicera japonica* and a proposed floral organ identity model[J]. *BMC Genomics*, 2023, 24(1): 447.
- [26] 尚林雪, 王 群, 张国哲, 等. 紫薇 LiCMB1 基因的克隆及表达特性分析[J]. 浙江农林大学学报, 2023, 40(2): 330-337.
- [27] 张新昊, 沈红艳, 文滨滨, 等. 桃 MADS-box 家族 PpCMB1 基因调控花发育分子机制[J]. 植物生理学报, 2021, 57(6): 1211-1217.
- [28] 聂紫艳. 山茶花 B 功能 MADS-box 基因的表达模式和功能分析[D]. 合肥: 安徽农业大学, 2021.
- [29] CHUNG Y Y, KIM S R, FINKEL D, et al. Early flowering and reduced apical dominance result from ectopic expression of a rice MADS box gene[J]. *Plant Molecular Biology*, 1994, 26(2): 657-665.
- [30] 郭 曦, 胡秋芸, 徐华祥, 等. 紫苏 PfAP1 基因的克隆与功能研究[J]. 植物生理学报, 2022, 58(8): 1598-1606.
- [31] 宋 杨, 窦连登, 张红军. 高等植物成花诱导调控的分子和遗传机制[J]. 植物生理学报, 2014, 50(10): 1459-1468.
- [32] 王 莹, 穆艳霞, 王 锦. MADS-box 基因家族调控植物花器官发育研究进展[J]. 浙江农业学报, 2021, 33(6): 1149-1158.
- [33] MONTIEL G, GAUDET M, LAURANS F, et al. Overexpression of MADS-box gene AGAMOUS-LIKE 12 activates root development in *Juglans* sp. and *Arabidopsis thaliana*[J]. *Plants*, 2020, 9(4): 444.
- [34] BEMER M, WOLTERS-ARTS M, GROSSNIKLUS U, et al. The MADS domain protein DIANA acts together with AGAMOUS-LIKE80 to specify the central cell in *Arabidopsis* ovules[J]. *The Plant Cell*, 2008, 20(8): 2088-2101.
- [35] ADAMCZYK B J, FERNANDEZ D E. MIKC* MADS domain heterodimers are required for pollen maturation and tube growth in *Arabidopsis*[J]. *Plant Physiology*, 2009, 149(4): 1713-1723.
- [36] ZHAO K K, JIA D W, ZHANG X, et al. FUL homologous gene CmFL1 is involved in regulating flowering time and floret numbers in *Chrysanthemum morifolium*[J]. *Plant Science*, 2023, 336: 111863.
- [37] WANG Y F, GE H S, AHMAD N, et al. Genome-wide identification of MADS-box family genes in safflower (*Carthamus tinctorius* L.) and functional analysis of CtMADS24 during flowering[J]. *International Journal of Molecular Sciences*, 2023, 24(2): 1026.
- [38] 霍慧敏. 吉林人参 MADS-box 基因家族的系统分析及 Pg-MADS17-04 基因功能的初步研究[D]. 长春: 吉林农业大学, 2021.
- [39] CHEN T T, LIU H, LI Y P, et al. AaSEPALATA1 integrates jasmonate and light-regulated glandular secretory trichome initiation in *Artemisia annua*[J]. *Plant Physiology*, 2023, 192(2): 1483-1497.

(本文编辑 田梦妍)