

过路黄 MYB 转录因子家族成员的挖掘及鉴定

杨敬天¹, 黎 华³, 赵 欢², 龙 芸², 张咏祀², 刘小红^{2*}

1. 绵阳师范学院生态安全与保护四川省重点实验室, 四川 绵阳 621000
2. 西华师范大学生命科学院, 四川 南充 637009
3. 北川羌族自治县国有林场, 四川 绵阳 622761

摘要:目的 分析过路黄 *Lysimachia christinae* MYB 转录因子家族相关成员的结构和功能, 挖掘与黄酮类物质生物合成相关的基因信息。方法 基于过路黄全长转录组数据库, 利用 ProtParam、GenStrip、MEME、WebLogo、SOPMA、Swiss-Model 等在线网站分析过路黄 MYB 转录因子家族蛋白的理化性质、亚细胞定位、保守基序和结构域、蛋白高级结构等。利用 MEGA 软件构建系统发育树, 分析过路黄 MYB 转录因子家族 (LcMYB) 与拟南芥 *Arabidopsis thaliana* MYB 蛋白的系统发育关系。结果 挖掘到 51 个 LcMYB 基因, 预测 51 条具有 MYB 转录因子保守结构域的蛋白序列; 根据结构域可分为 R1-LcMYB、R2R3-LcMYB 和 R3-LcMYB 3 个亚类, 且所有 LcMYB 基序中都含有 3 个保守的色氨酸; 过路黄 MYB 家族蛋白均为亲水性蛋白, 热稳定性较高且富含碱性氨基酸, 大部分蛋白以无规则卷曲为主, 其中 R2R3-LcMYB39、R2R3-LcMYB47、R2R3-LcMYB48 与 R2R3-LcMYB50 蛋白是以 α -螺旋为主; 与拟南芥 MYB 转录因子家族共同构建的进化树发现, 过路黄 MYB 家族在进化上包括 3 个大类。结论 在过路黄物种中获得了 51 个 MYB 转录因子家族成员的基因的结构和功能信息, 其中有 2 个成员与黄酮类物质的生物合成有关, 该研究结果为后续克隆利用黄酮类物质生物合成相关基因提供了分子基础。

关键词: 过路黄; MYB; 生物信息学; 黄酮类; 转录因子

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2022)16-5149-10

DOI: 10.7501/j.issn.0253-2670.2022.16.024

Excavation and identification of members of MYB transcription factor family in *Lysimachia christinae*

YANG Jing-tian¹, LI Hua³, ZHAO Huan², LONG Yun², ZHANG Yong-si², LIU Xiao-hong²

1. Sichuan Key Laboratory of Ecological Security and Protection, Mianyang Normal University, Mianyang 621000, China
2. College of Life Science, China West Normal University, Nanchong 637009, China
3. State-owned Forest Farm of Beichuan Qiang Minority Autonomous County, Mianyang 622761, China

Abstract: Objective To analyze the structure and function of the members of MYB transcription factor family in *Lysimachia christinae*, and excavate some information of the genes related to flavonoid biosynthesis. **Methods** Based on transcriptome database of *L. christinae*, the physicochemical properties, subcellular localization, conserved motifs and domains and protein structures of the LcMYB were analyzed by online websites, including ProtParam, GenStrip, MEME, WebLogo, SOPMA, Swiss-Model, etc. The MYB phylogenetic tree was constructed by MEGA to analyze the phylogenetic relationship between *L. christinae* and *Arabidopsis thaliana*. **Results** A total of 51 LcMYB genes were identified, 51 MYB protein sequences with conserved domain were predicted. LcMYB can be divided into three classes (R1-LcMYB, R2R3-LcMYB and R3-LcMYB) according to the conserved domain, and all the LcMYB motifs contained three conserved tryptophan. All the MYB proteins of *L. christinae* were hydrophilic and rich in basic amino acids, with high thermal stability. Most of the proteins were mainly in random curl, among them, R2R3-LcMYB39, R2R3-LcMYB47, R2R3-LcMYB48 and R2R3-LcMYB50 were mainly in α -helix. MYB

收稿日期: 2022-02-03

基金项目: 西华师范大学校基本科研业务费(19B033); 青海民族大学校级理工科项目(2019XJY02); 绵阳师范学院科研启动项目(QD2019A13)

作者简介: 杨敬天(1983—), 男, 四川绵阳人, 博士, 讲师, 主要从事植物系统进化与药用植物学研究。

Tel: 13699629310 E-mail: yjtdc83@163.com

*通信作者: 刘小红(1975—), 男, 重庆彭水人, 博士, 教授, 主要从事植物分子遗传学研究。Tel: 13330761109 E-mail: 350783409@qq.com

family of *L. christinae* could be divided into three groups on the basis of evolutionary tree with the MYB family of *A. thaliana*.

Conclusion The structure and function information of 51 genes associated with MYB transcription factor family members was obtained in *L. christinae*, thereinto, two genes were involved in the biosynthesis of flavonoids, which provided molecular basis for the subsequent cloning of genes related to the biosynthesis of flavonoids.

Key words: *Lysimachia christinae* Hance; MYB; bioinformatics; flavonoids; transcription factor

过路黄 *Lysimachia christinae* Hance 为报春花科 (Primulaceae) 珍珠菜属 *Lysimachia* L. 植物, 其新鲜或干燥全草又名金钱草, 在我国分布较为广泛, 主要产于四川、云南、贵州、陕西、河南、湖北、湖南、浙江等省区^[1-2]。中医认为过路黄味酸、苦、微寒, 其归肝、胆、肾、膀胱经, 具有利湿退黄、利尿通淋及解毒消肿等功效, 主要用于治疗湿热黄疸、石淋、热淋、痈肿疔疮、小便涩痛、肝胆结石和尿路结石等症状^[3-4]。过路黄的化学成分有黄酮类、酚类、多糖类、三萜皂苷类、挥发油、醌类等, 其中以黄酮类化合物槲皮素和山柰素为主要有效成分^[5-6]。药理学研究表明, 这些化学成分具有抗肿瘤、抗氧化、保肝、保护内皮、治疗痛风等多种药理活性^[7-9]。因此, 研究过路黄黄酮生物合成途径具有重要价值。

植物中黄酮类化合物主要通过苯丙烷代谢途径合成。苯丙氨酸依次由苯丙氨酸解氨酶 (phenylalanine ammonia-lyase, PAL)、肉桂酸 4-羟化酶 (cinnamate 4-hydroxylase, C4H) 和 4-香豆酰 CoA 连接酶 (4-coumarate:CoA ligase, 4CL) 催化形成 4-香豆酰 CoA (4-coumarinyl-CoA)^[10]。再以 1 分子的 4-香豆酰 CoA 和 3 分子丙二酰 CoA 为底物, 经黄酮类化合物合成途径的查耳酮合成酶 (chalcone synthetase, CHS)、查耳酮异构酶 (chalcone isomerase, CHI) 和黄烷酮 3-羟化酶 (flavanone 3-hydroxylase, F3H) 等多种酶共同催化完成^[11-12]。植物中发现很多转录因子家族 (如 MYB、bHLH、bZIP、MADS-box、WRKY 和 WDR 等) 参与调控黄酮类化合物的合成^[12-13]。已有研究表明拟南芥 *Arabidopsis thaliana* L. MYB 可以增加黄酮等次级代谢产物的含量, 从而提高对虫害、干旱和高盐胁迫的抵御能力^[14-15]。其中 MYB 转录因子是植物最大的转录因子家族之一, 其 DNA 结合域中含有 4 个不完全重复的 R1、R2、R3 和 R4 序列^[16]。每个 MYB 结构域 (R) 由 51~53 个氨基酸组成, 编码 3 个 α -螺旋, 其第 2 和第 3 个螺旋通过 3 个色氨酸残基组成螺旋-转角-螺旋 (HTH) 的结构, 参与结合 DNA

大沟, 从而提高或抑制下游基因的转录^[17]。根据 R 的重复数目, 可将 MYB 家族分为 4 类: R1-MYB 型 (含 MYB-related)、R2R3-MYB 型、3R-MYB 型和 4R-MYB 型, 其中, R2R3-MYB 型 MYB 因子在植物 MYB 家族中所占比例最大^[18-19]。

目前, 对过路黄的研究主要集中于化学成分、药理活性等方面, 有关过路黄分子生物学方面的研究鲜有报道。课题组前期对过路黄全长转录组进行了测序, 在此基础上, 本实验进一步利用生物信息学对过路黄 MYB 转录因子蛋白 (LcMYB) 的分类、理化性质、亚细胞定位、保守基序、结构域、系统发育树、氨基酸高级结构及生物学功能注释等进行分析, 以期过路黄 MYB 家族参与调控黄酮类化合物的作用机制研究、高产优质过路黄育种和种植提供参考依据。

1 材料与方法

1.1 序列来源

采集田间自然生长于西华师范大学校园内的过路黄幼嫩植株 (30°49'25"N, 106°03'43"E), 液氮速冻后, 分别提取根、茎和叶的总 RNA, 等质量混匀后, 由派森诺生物科技有限公司进行 mRNA 纯化、反转录构建全长 cDNA 文库, 再作三代高通量转录组测序, 对测序结果去除接头和低质量的 reads 后获得 clean reads。基于转录组测序数据, 在 PFAM、GO、NR、Swiss-Prot、eggNOG 和 KEGG 6 个等数据库初步注释 7 个 MYB 转录因子相关基因, 将这些基因编码蛋白逐个进行 NCBI Blast 和 SMART 预测, 去除冗余序列后, 最后得到 51 条具有完整 MYB 保守结构域的序列。从 TAIR 数据库中下载已知并鉴定功能的拟南芥 AtMYB 家族的编码序列 (CDS) 和氨基酸序列, 使用在线软件 ExPasy 中的 Translate 对 LcMYB 家族各成员的全长 cDNA 序列进行分析, 从而得到其对应的编码蛋白的氨基酸序列。利用 SMART 检测候选蛋白质序列, 选择具有 MYB 结构域的氨基酸序列, 剔除不含 MYB 基序的蛋白序列和冗余序列, 最后得到过路黄 MYB 蛋白的氨基酸序列。

1.2 LcMYB 理化性质分析

采用网络在线分析软件 ExPASy 中的 ProtParam 对过路黄 MYB 蛋白的氨基酸数目、相对分子质量、等电点 (pI)、脂肪族氨基酸数和蛋白质疏水性等理化性质进行分析。采用在线 GenStrip 软件预测 LcMYB 蛋白的亚细胞定位。

1.3 LcMYB 保守结构域分析

采用 MEME 软件预测 LcMYB 蛋白的保守基序 (Motif), 基序的查找数量设为 10, 其他参数为默认值。采用 ClustalW 软件分别对过路黄 MYB 序列中 3 个 SANT 序列进行比对, 再用 WebLogo 在线软件绘制 LcMYB 蛋白保守结构域可视化图。

1.4 LcMYB 空间结构和生物学功能分析

采用 SOPMA 在线软件预测 LcMYB 蛋白二级结构, 并采用 Swiss-Model 在线软件预测 LcMYB 蛋白的三级结构。采用 Blast2GO 软件注释 LcMYB 蛋白的生物学功能。

1.5 聚类分析

利用在线软件 Clustalw 对 51 条过路黄 MYB 转录因子蛋白和与其结构相似的 65 条拟南芥 MYB 转录因子的氨基酸序列进行比对。然后利用 MEGA7.0 软件以邻接法 (neighbor-joining method) 绘制进化树, 参数选择 P-distance、pairwise deletion 和 Bootstrap method 为 1000, 其他参数选择默认值。

2 结果与分析

2.1 过路黄 MYB 基因挖掘以及蛋白序列分析

从过路黄全长转录组数据库中筛选出 54 条 LcMYB 核苷酸序列, 经与网络数据库进行同源比对, 最后获得 51 条具有完整 MYB 保守结构域的序列, 分别命名为 LcMYB1~LcMYB51, 其中属于 R1-MYB 类转录因子的序列有 19 条, 长度分布在 420~2125 bp; 属于 R2R3-MYB 类序列有 31 条, 长度分布在 650~5340 bp; 属于 R3-MYB 类序列仅 1 条, 长度为 2114 bp。

2.2 过路黄 MYB 转录因子家族蛋白理化性质及亚细胞定位

采用 ProtParam 在线软件分析过路黄 LcMYB 转录因子家族蛋白的理化性质 (表 1), 结果表明, 预测的 51 个 LcMYB 蛋白在氨基酸数目、蛋白相对分子质量、pI、脂肪系数、不稳定性系数和总平均亲水性等方面均存在显著差异。LcMYB 蛋白的氨基酸残基 (aa) 长度变化范围

为 57~989 aa, 平均长度为 299.1 aa。其中, R1 型 LcMYB 蛋白的 aa 长度为 57~632 aa, 平均长度为 259.79 aa; R2R3 型 LcMYB 蛋白的长度为 120~989 aa, 平均长度为 321.03 aa; 而 R3 型 LcMYB 蛋白的长度为 367 aa。LcMYB 蛋白理论相对分子质量 (M_w) 范围为 6 588.31~111 684.07, 平均值为 33 185.75。其中, R1 型 LcMYB 蛋白的 M_w 为 6 588.31~70 486.65, 平均为 29 100.87; R2R3 型 LcMYB 蛋白的 M_w 为 14 320.34~111 684.07, 平均值为 35 430.66; 而 R3 型 LcMYB 蛋白的 M_w 为 41 206.18。在上述 2 种理化性质中, 氨基酸残基序列长度最短和 M_w 最小的是 LcMYB4, 氨基酸序列最长和 M_w 最大的是 LcMYB47 (表 1)。

另外, LcMYB 蛋白家族的理论 pI 范围在 5.05 (LcMYB1)~10.44 (LcMYB33), 平均值为 7.98, 其中 35 个 LcMYB 蛋白的 pI 值大于 7, 其余 16 个 LcMYB 蛋白的 pI 值小于 7, 表明大多数过路黄 LcMYBs 蛋白为偏碱性。所有 LcMYB 家族蛋白的总平均亲水性为负值, 说明其均为亲水性蛋白。R1 型 LcMYB 蛋白中仅 LcMYB 4 (不稳定性指数为 33.93)、LcMYB 14 (37.11) 和 LcMYB 19 (28.77) 3 种蛋白的不稳定性指数小于 40, 为稳定蛋白, 其余 R1 型蛋白均为不稳定蛋白; 所有 R2R3 型和 R3 型 LcMYB 蛋白的不稳定性指数均大于 40, 为不稳定蛋白 (表 1)。

利用 ProtComp 在线网站分析 LcMYB 蛋白的亚细胞定位, 结果显示, 39 种 LcMYB 蛋白定位于细胞核, 8 种 LcMYB 蛋白定位于叶绿体, 剩余的 4 种 LcMYB 蛋白定位于过氧化物酶体, 说明 LcMYB 家族成员功能可能有所差异或具有多种功能 (表 1)。

2.3 过路黄 MYB 转录因子家族蛋白保守序列分析

采用 HEME 在线软件分析过路黄 MYB 转录因子蛋白家族的氨基酸序列, 结果表明, 共鉴定出 10 个保守序列, 长度为 8~50 aa, 且不同 LcMYB 蛋白含有的序列数不尽相同。其中, LcMYB21、LcMYB26、LcMYB29、LcMYB32、LcMYB35、LcMYB36、LcMYB38 和 LcMYB42 8 个 MYB 蛋白的氨基酸序列中所含的保守序列数目最多, 为 6 个; 而 LcMYB1、LcMYB2、LcMYB5、LcMYB6、LcMYB18 和 LcMYB49 6 个 MYB 蛋白的保守序列数最少, 仅为 1 个; 其余 37 个 MYB 转录因子所含的序列数为 2~5 个。

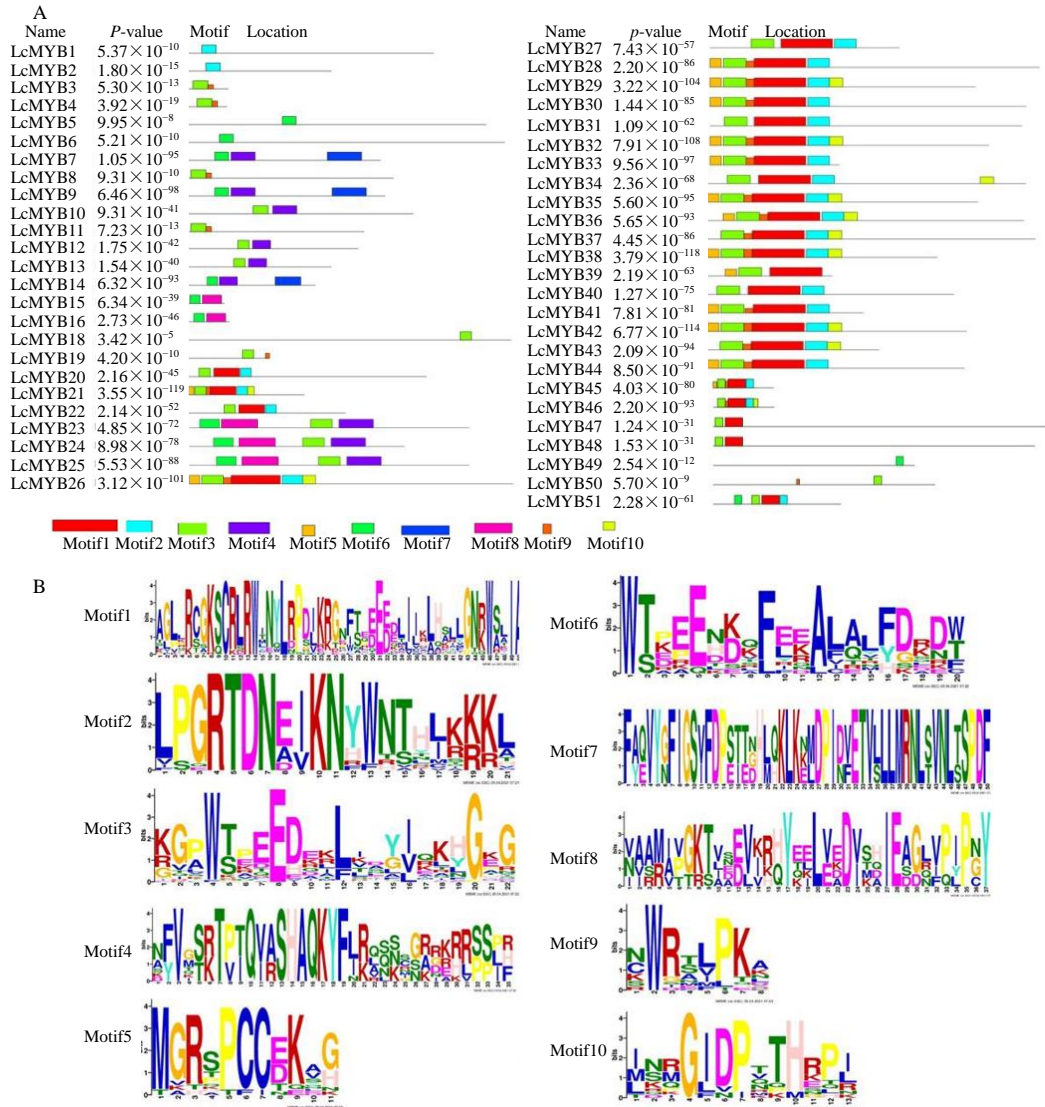
表1 过路黄 MYB 转录因子的理化性质及亚细胞定位

Table 1 Physicochemical property and subcellular localization of MYB in *L.christinae*

编号	类型	长度/aa	相对分子质量	pI	脂肪系数	总平均亲水性	不稳定系数	亚细胞定位
LcMYB1	R1	360	40 682.60	5.05	72.33	-0.676	57.56	叶绿体
LcMYB2	R1	210	23 882.87	5.19	72.43	-0.643	54.59	叶绿体
LcMYB3	R1	59	6 982.95	6.57	79.32	-0.827	70.85	细胞核
LcMYB4	R1	57	6 588.31	9.22	53.16	-1.400	33.93	细胞核
LcMYB5	R1	437	48 857.19	9.30	75.01	-0.687	50.09	过氧化物酶体
LcMYB6	R1	464	53 380.90	9.76	74.03	-0.749	50.97	过氧化物酶体
LcMYB7	R1	282	30 947.95	9.08	71.56	-0.541	52.32	细胞核
LcMYB8	R1	301	33 072.95	9.58	76.38	-0.526	41.76	细胞核
LcMYB9	R1	289	31 860.98	9.11	68.48	-0.584	53.10	细胞核
LcMYB10	R1	330	36 103.28	6.87	60.33	-0.673	56.45	过氧化物酶体
LcMYB11	R1	258	28 665.77	10.08	71.20	-0.720	60.62	细胞核
LcMYB12	R1	333	36 032.08	6.46	58.02	-0.664	58.02	过氧化物酶体
LcMYB13	R1	280	30 982.54	7.99	65.39	-0.667	59.10	叶绿体
LcMYB14	R1	249	27 949.35	6.08	67.71	-0.630	37.11	细胞核
LcMYB15	R1	71	8 408.39	8.08	61.83	-1.049	69.91	细胞核
LcMYB16	R1	81	9 135.14	8.07	60.25	-0.869	64.61	细胞核
LcMYB17	R1	84	10 287.76	9.46	52.26	-1.079	95.27	细胞核
LcMYB18	R1	632	70 486.65	9.35	73.42	-0.652	53.59	细胞核
LcMYB19	R1	159	18 608.88	6.34	66.73	-1.134	28.77	叶绿体
LcMYB20	R2R3	466	52 434.97	5.67	75.13	-0.674	65.80	叶绿体
LcMYB21	R2R3	227	26 091.71	8.32	75.20	-0.630	57.65	细胞核
LcMYB22	R2R3	308	3 474.81	7.76	70.91	-0.695	48.93	叶绿体
LcMYB23	R2R3	286	31 397.48	8.80	54.62	-0.651	52.83	细胞核
LcMYB24	R2R3	220	25 714.90	8.63	60.73	-0.819	46.15	细胞核
LcMYB25	R2R3	286	32 539.37	7.78	62.03	-0.770	54.51	细胞核
LcMYB26	R2R3	336	37 598.27	5.78	79.55	-0.490	47.19	细胞核
LcMYB27	R2R3	184	20 540.25	9.93	65.27	-0.789	43.20	细胞核
LcMYB28	R2R3	320	35 755.94	6.66	67.34	-0.722	50.73	细胞核
LcMYB29	R2R3	258	29 388.75	9.71	85.78	-0.520	41.10	细胞核
LcMYB30	R2R3	307	34 691.71	6.60	65.80	-0.755	55.91	细胞核
LcMYB31	R2R3	303	33 390.08	7.62	74.59	-0.438	57.26	细胞核
LcMYB32	R2R3	271	30 417.42	8.89	72.36	-0.604	49.84	细胞核
LcMYB33	R2R3	126	14 843.23	10.44	69.68	-0.964	55.97	细胞核
LcMYB34	R2R3	306	35 342.34	5.73	70.39	-0.744	60.65	细胞核
LcMYB35	R2R3	260	29 553.18	5.50	73.54	-0.737	50.45	细胞核
LcMYB36	R2R3	304	34 338.21	7.17	63.52	-0.922	49.43	细胞核
LcMYB37	R2R3	315	35 737.58	5.44	64.06	-0.721	45.26	细胞核
LcMYB38	R2R3	221	25 245.80	8.97	76.38	-0.718	48.58	细胞核
LcMYB39	R2R3	120	14 320.34	9.00	67.42	-0.962	66.73	细胞核
LcMYB40	R2R3	237	27 632.28	9.15	68.35	-0.834	56.06	细胞核
LcMYB41	R2R3	150	16 983.46	9.51	70.93	-0.715	62.72	细胞核
LcMYB42	R2R3	249	28 573.26	8.60	70.88	-0.792	51.04	细胞核
LcMYB43	R2R3	165	19 181.92	9.80	73.33	-1.009	43.26	细胞核
LcMYB44	R2R3	247	28 229.65	8.42	65.95	-0.857	51.51	细胞核
LcMYB45	R2R3	174	20 331.91	9.34	57.76	-1.045	61.31	细胞核
LcMYB46	R2R3	176	20 148.16	9.72	90.34	-0.726	44.06	细胞核
LcMYB47	R2R3	989	111 684.07	5.21	78.00	-0.825	48.32	细胞核
LcMYB48	R2R3	924	104 152.62	5.44	73.48	-0.880	47.76	细胞核
LcMYB49	R2R3	579	65 384.06	9.14	68.55	-0.694	44.57	叶绿体
LcMYB50	R2R3	638	73 232.63	8.65	65.82	-1.003	55.67	叶绿体
LcMYB51	R3	367	41 206.18	8.09	63.81	-0.789	57.86	细胞核

进一步分析发现，在预测的 10 个保守序列中，51 种 LcMYB 蛋白至少含有 motif1、motif2、motif3 中的一个，推测这 3 个氨基酸序列是 MYB 保守结构域的主要部分。其中，R1 型 LcMYB 蛋白含有

motif2、motif3、motif4、motif6、motif7、motif8 和 motif9；R2R3 型 LcMYB 蛋白含 10 种保守序列且均含有 motif3；R3 型 LcMYB 蛋白含有 motif1、motif2、motif3 和 motif6（图 1）。



A-LcMYB 转录因子的基序分布 B-LcMYB 转录因子的各基序的氨基酸排列
 A-Distribution of motifs of LcMYB B-Amino acid sequence logos of motifs of LcMYB

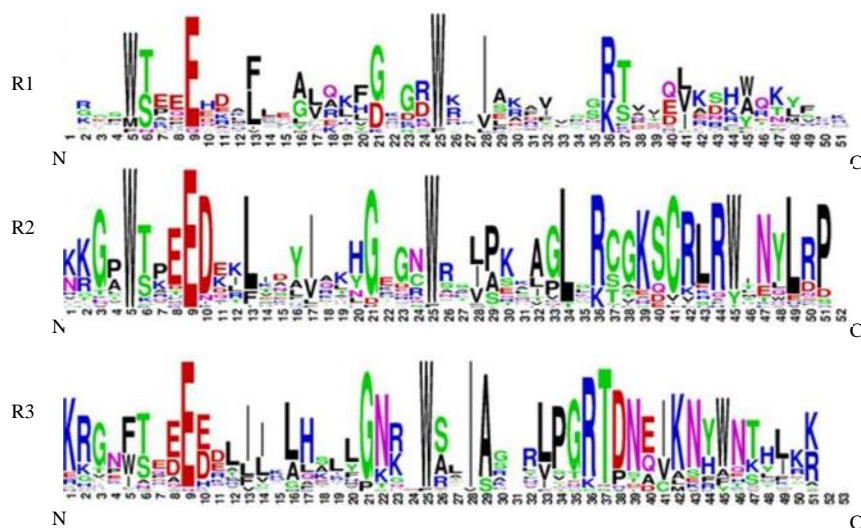
图 1 过路黄 MYB 转录因子基序预测

Fig. 1 Prediction of motifs of MYB in *L. christinae*

2.4 过路黄 MYB 转录因子家族蛋白保守结构域分析

利用 WebLogo 网站对 51 个 LcMYB 蛋白的保守结构域进行分析。结果表明，LcMYB 蛋白每个重复的结构域 (R) 均为 51 aa，序列含有 3 个色氨酸 (W)，分别位于序列的第 5、25、45 位，且两相邻色氨酸的间距相同，为 19 aa。进一步分析发现，在 R1、R2 和 R3 结构域的色氨酸均以第 25 位最为保守，R1 结构域中的第 5 个氨基酸 W 存在甲硫氨

酸 (M) 取代情况，而第 45 个氨基酸的 W 位由酪氨酸 (Y)、丙氨酸 (A) 和 W 共同组成；R2 结构域中的第 5 个氨基酸 W 部位高度保守，但第 3 个 W 也存在 Y 取代的情况；而 R3 结构域中的第 1 个 W 部位由苯丙氨酸 (F)、异亮氨酸 (I) 和 W 共同组成，第 45 位的 W 也存在 Y 取代的现象。此外，MYB 结构域中还包括其它较为保守的氨基酸残基，如第 9 位的谷氨酸 (E)、第 21 位的甘氨酸 (G) 和第 36 位精氨酸 (R) (图 2)。



R1、R2 和 R3 分别表示 R1-MYB、R2R3-MYB 和 3R-MYB
R1, R2 and R3 represent R1-MYB, R2R3-MYB and R3-MYB, respectively

图 2 过路黄 MYB 转录因子保守结构域分析

Fig. 2 Analysis of conserved domains of MYB in *L. christinae*

2.5 过路黄 MYB 转录因子家族进化树分析

采用 MEGA7.0 软件绘制过路黄和拟南芥 MYB 蛋白家族系统进化树 (图 3), 结果显示, 来源过路黄的 MYB 转录因子分为 3 类, 分别聚在系统进化树的 3 个不同分支下 (Group I、Group II、Group III), 其中 R1-MYB 类转录因子主要是分布在 I、II 类, 占过路黄总 R1-MYB 类转录因子的 37.25%, R2R3-MYB 类转录因子主要被聚在第 III 类中, 占过路黄总 R2R3-MYB 类转录因子的 60.78%。R2R3-MYB 类转录因子, 结构多样导致其在进化树上分布范围最广。第 III 大类被分为 2 个亚类, R1-LcMYB5 最先被分化, 其中 R1-LcMYB5 与 AT1G77450、AT1G32450 聚在一起。

从过路黄和拟南芥 MYB 转录因子家族系统进化树 (图 3) 中可以看出大部分过路黄 MYB 转录因子可与拟南芥聚在一起, 说明过路黄 MYB 转录因子和拟南芥 MYB 转录因子可以相互匹配, 如 LcMYB19 与 AT1G15720、LcMYB13 与 AT1G70000、LcMYB23 与 AT5G04760、LcMYB16 与 AT2G21650、LcMYB24 与 AT3G11280、LcMYB50 与 AT5G45420、LcMYB17 与 AT2G46410、LcMYB51 与 AT3G09370、LcMYB22 与 AT1G17950、LcMYB2 与 AT3G23250、LcMYB3 与 AT3G46130、LcMYB41 与 AT5G57620、LcMYB39 与 AT1G63910、LcMYB43 与 AT1G22640、LcMYB26 与 AT1G34670、LcMYB29 与 AT4G28110、LcMYB45 与 AT1G16490 和

LcMYB33 与 AT5G49330 分别聚在一起, 根据系统进化的序列相似性, 表明其进化关系较近, 说明其与拟南芥的亲缘关系比较接近, 在蛋白结构和序列上更为相似, 可以根据此来预测 LcMYB 的生物学功能。在该进化树中存在着不同结构域但同源性较高的情况, 如 R2R3-LcMYB20 与 R1-LcMYB1 相邻; 同时存在着相同结构域同源性高的情况, 如 R2R3-LcMYB21 与 R2R3-LcMYB38、R2R3-LcMYB28 与 R2R3-LcMYB30 分别相邻。总体表明, MYB 家族蛋白与拟南芥 MYB 家族蛋白同源性高, 在相邻或较近进化关系上的蛋白有可能具有相似的功能。

2.6 过路黄 MYB 转录因子家族蛋白高级结构分析

利用 SOPMA 预测过路黄 MYB 蛋白质的二级结构, 结果显示, 除 LcMYB15 蛋白不含延伸链外, 其余 50 个 LcMYB 蛋白均具有 α -螺旋、 β -转角、无规卷曲和延伸链 4 种构型。其中部分 R1-LcMYB 蛋白 (LcMYB6、LcMYB17) 及部分 R2R3-LcMYB 蛋白 (LcMYB39、LcMYB47、LcMYB48、LcMYB50) 是以 α -螺旋为主, β -转角、延伸链以及无规则卷曲分布在蛋白序列中; 而其余 44 种 LcMYB 蛋白则都以无规则卷曲为主, α -螺旋、 β -转角、延伸链散布于整个蛋白序列 (表 2)。

蛋白质的三级结构与其功能密切相关, 使用 Swiss-Model 对过路黄的氨基酸序列进行三级结构同源建模 (图 4), 结果表明, 过路黄 MYB 家族

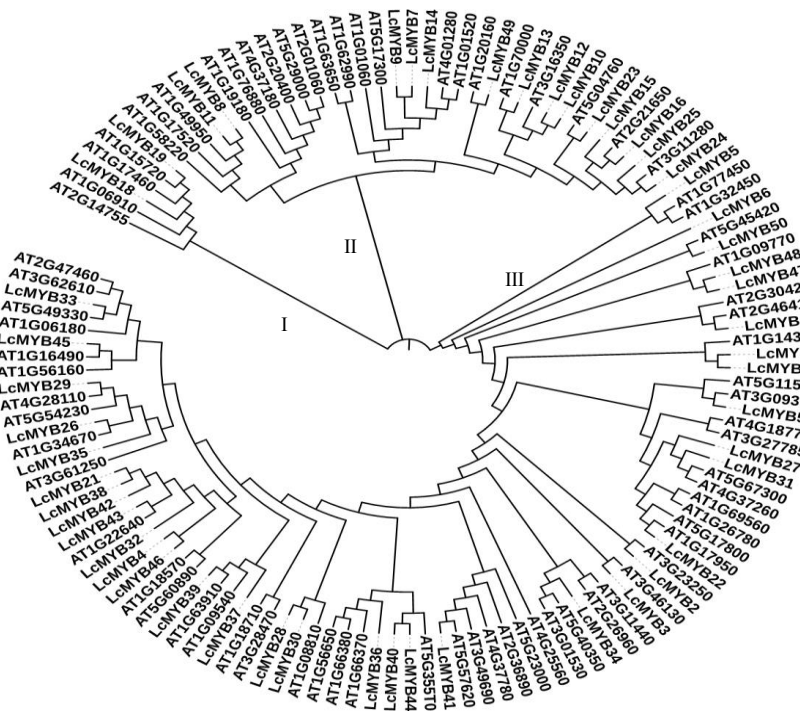


图3 过路黄 MYB 转录因子家族进化树

Fig. 3 Phylogenetic tree of MYB in *L. christinae*

表2 过路黄 MYB 转录因子家族蛋白的二级结构

Table 2 Secondary structure of MYB proteins in *L.christinae*

编号	α-螺旋/%	β-转角/%	无规则卷曲/%	延伸链/%	编号	α-螺旋/%	β-转角/%	无规则卷曲/%	延伸链/%
LcMYB1	38.89	2.22	51.39	7.50	LcMYB27	41.30	4.35	45.65	8.70
LcMYB2	29.05	6.67	54.29	10.00	LcMYB28	31.56	1.88	60.00	6.56
LcMYB3	35.59	8.47	40.68	15.25	LcMYB29	34.50	8.53	43.41	13.57
LcMYB4	21.05	7.02	61.40	10.53	LcMYB30	28.99	2.61	61.89	6.51
LcMYB5	40.05	1.37	53.09	5.49	LcMYB31	27.72	3.63	60.73	7.92
LcMYB6	53.45	4.09	33.62	8.84	LcMYB32	25.09	5.90	57.20	11.81
LcMYB7	27.30	4.26	61.70	6.74	LcMYB33	34.13	11.90	43.65	10.32
LcMYB8	44.52	3.32	42.52	9.63	LcMYB34	34.64	2.61	58.17	4.58
LcMYB9	27.68	6.23	57.89	9.00	LcMYB35	36.54	3.85	53.85	5.77
LcMYB10	16.06	2.73	69.39	11.82	LcMYB36	24.67	5.92	61.51	7.89
LcMYB11	34.11	3.49	52.33	10.08	LcMYB37	38.10	4.76	51.43	5.71
LcMYB12	16.22	3.60	66.67	13.51	LcMYB38	31.67	5.88	52.49	9.95
LcMYB13	17.14	4.29	61.07	17.50	LcMYB39	48.33	5.83	39.17	6.67
LcMYB14	29.72	6.02	54.22	10.04	LcMYB40	26.58	5.06	64.56	3.80
LcMYB15	56.34	5.63	38.03	0	LcMYB41	36.67	8.67	47.33	7.33
LcMYB16	49.38	2.47	46.91	1.23	LcMYB42	27.31	5.62	56.22	10.84
LcMYB17	54.76	3.57	33.33	8.33	LcMYB43	43.03	9.09	44.24	3.64
LcMYB18	19.30	3.01	64.87	12.82	LcMYB44	28.74	6.88	58.70	5.67
LcMYB19	28.93	2.52	59.75	8.81	LcMYB45	44.25	9.20	43.10	3.45
LcMYB20	34.98	4.08	53.22	7.73	LcMYB46	32.39	6.82	47.16	13.64
LcMYB21	28.63	5.29	56.83	5.29	LcMYB47	53.69	2.22	40.24	3.84
LcMYB22	32.14	6.28	47.73	13.31	LcMYB48	54.55	2.92	37.12	5.41
LcMYB23	21.33	2.80	65.38	10.49	LcMYB49	45.42	3.97	40.93	9.67
LcMYB24	26.82	5.45	50.91	16.82	LcMYB50	63.17	4.23	28.06	4.55
LcMYB25	21.33	3.15	59.44	16.08	LcMYB51	28.07	5.45	61.04	5.45
LcMYB26	41.96	3.57	52.68	1.79					

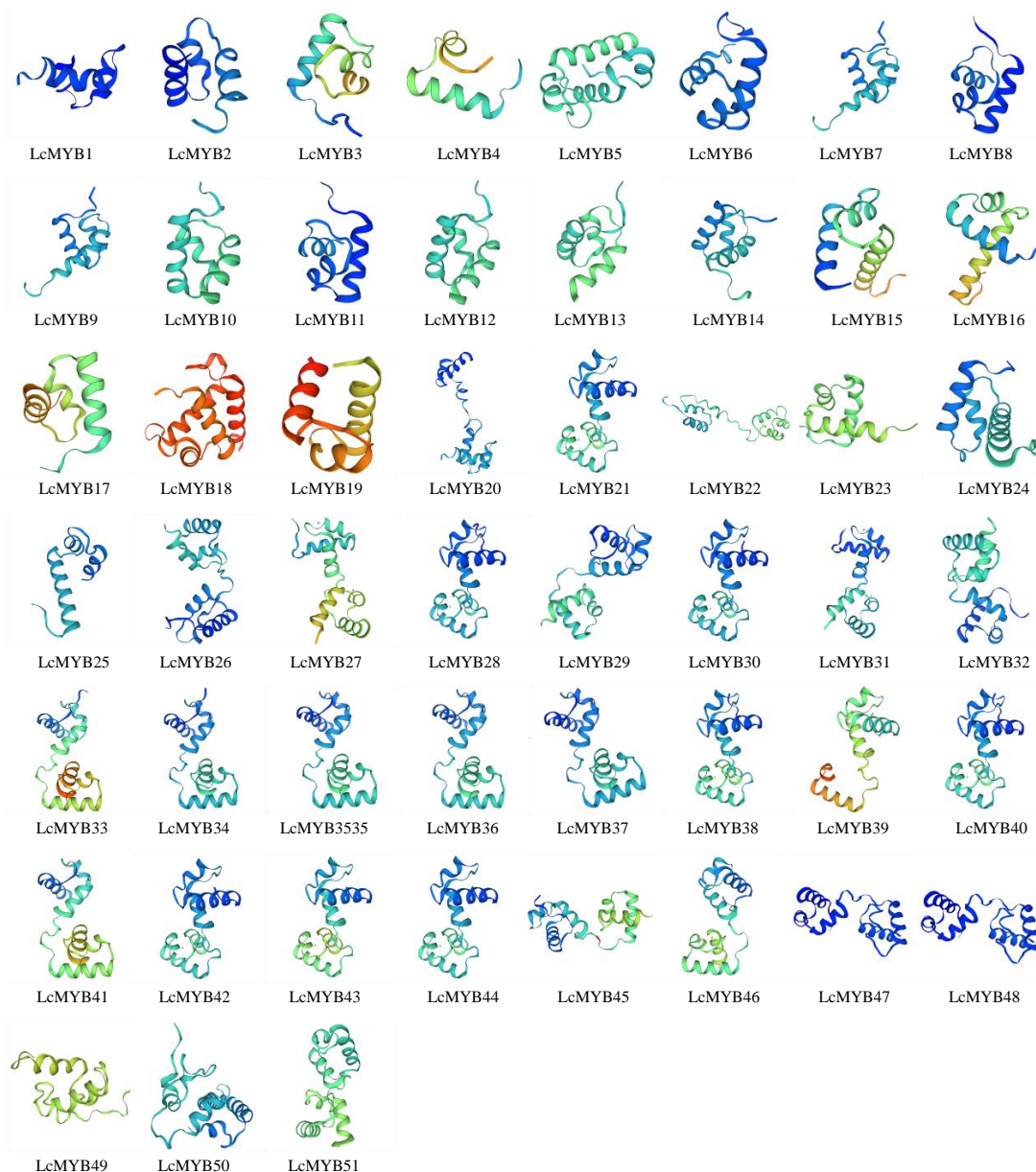


图 4 过路黄 MYB 蛋白的三级结构

Fig. 4 Tertiary structures of MYB proteins in *L. christinae*

蛋白均具有 α -螺旋、无规则卷曲、 β -折叠和 β -转角，这与二级结构预测相符合。R2R3-LcMYB 亚族蛋白三级结构相似性较高，R2、R3 结构分别有 3 个 α -螺旋，能明显看到 HTH 结构，主要是无规则卷曲长度有所不同。对于 R1-LcMYB 亚族蛋白，主要在 α -螺旋数量上存在差异。

2.7 过路黄 MYB 转录因子蛋白的功能注释

Blast2GO 软件处理结果表明，51 条 LcMYB 序列被注释到分子功能、生物过程和细胞组成 3 个 GO 类别的 24 小组。分子功能注释均为 binding，LcMYB 序列数量为 36 条，表明所筛选的 MYB 转录因子通过

与目的基因的启动子区结合，调控生物体的生长发育。生物过程中注释的生物过程和细胞过程丰度最高，均为 16 条；其次是代谢过程、大分子代谢过程、细胞内代谢过程、初级代谢过程、细胞内大分子代谢过程和有机物代谢过程，丰度均为 13 条；生物调节、细胞过程调节和生物过程调节丰度也较高，为 8 条，其他基因均较小。值得注意的是，参与芳香族化合物生物合成的有 8 条，分别为 LcMYB10、LcMYB12、LcMYB21、LcMYB23、LcMYB24、LcMYB25、LcMYB7 和 LcMYB9。细胞组成中注释最多的为细胞组分、细胞器及内腔（图 5）。

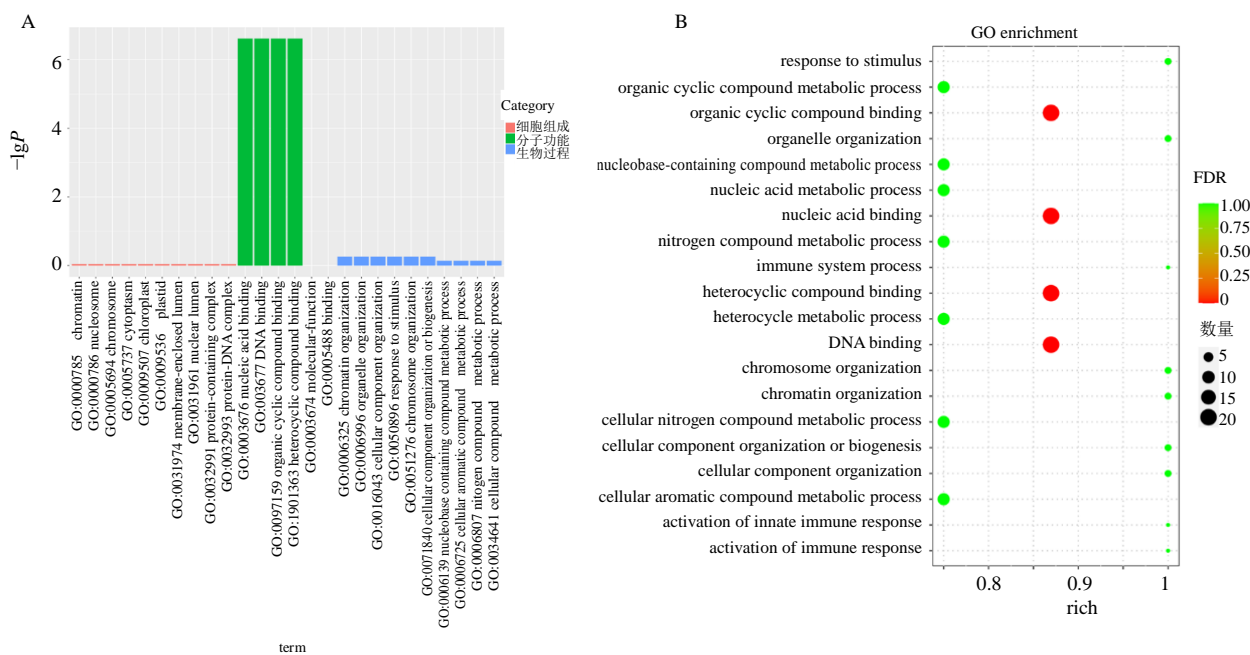


图5 过路黄 MYB 转录因子家族蛋白的 GO 注释 (A) 和富集分析 (B)

Fig. 5 GO annotation (A) and enrichment analysis (B) of MYB family proteins in *L. christinae*

3 讨论

3.1 LcMYB 的理化性质及细胞定位

本实验首次从过路黄全长转录组中分别鉴定出 19 个 R1 型 LcMYB、31 个 R2R3 型 LcMYB 和 1 个 R3 型 LcMYB 基因, 相较于拟南芥、大豆 *Glycine max* (Linn.) Merr.、水稻 *Oryza setiva* L.、柑橘 *Citrus sinensis* Blanco、葡萄树 *Vitis vinifera* L.、毛果杨 *Populus trichocarpa* Torr. & Gray 等基因组中相应的 3 种类型 MYB 基因均偏少, 但比细叶百合的 23 个 MYB 家族成员多, 这可能与各植物基因组倍数以及基因的复制相关^[20]。此外, 过路黄中 R2R3 型 LcMYB 基因所占比例最大, 与上述物种中 R2R3-MYB 基因在 MYB 家族中的分布比例相同^[17]。在少数物种中还发现了 R4 型 MYB 蛋白, 但在过路黄全长转录组中尚未发现。

3.2 LcMYB 的保守结构域

通过对其翻译的氨基酸序列中保守结构域基序分析, 表明 R2R3-LcMYB 结构域比 R1-LcMYB 结构域保守。其中, R2-LcMYB 与 R3-LcMYB 基序中都含有 3 个保守的色氨酸, 色氨酸之间的间隔氨基酸数目分别为 20、20 与 20、19。R3-LcMYB 基序第 1 个和第 3 个色氨酸处还包括疏水氨基酸苯丙氨酸, 这与黑果枸杞 *Lycium ruthenicum* L. R3-MYB 基序一致^[21], 表明 MYB 基因家族 N 端在不同植物间高度保守, 对于维持 HTH 构型有着重要的作用^[22]; 其中 R2-LcMYB 基序中 2 个色氨酸之间的间隔与杜

海描述大豆植物 R2R3-MYB 一致^[23], 说明过路黄 R2-LcMYB 基序的色氨酸非常保守; 而 R3-LcMYB 基序中色氨酸比大豆中少 1 个氨基酸残基, 表明过路黄 R2R3-LcMYB 类转录因子中的 R3-LcMYB 基序发生氨基酸缺失, 该区域可能是其植物进化和功能分化的热点区; 而 R1-LcMYB 基序只有 2 个保守的色氨酸且间隔 19 个氨基酸, 第 3 个色氨酸被组氨酸 (H)、赖氨酸 (K) 以及精氨酸 (R) 取代。

3.3 LcMYB 高级结构分析

分别对 R1-LcMYB 和 R2R3-LcMYB 蛋白进行二级结构预测, 分析表明所有蛋白都具有 α -螺旋、 β -转角、 β -折叠及无规则卷曲结构, 且大部分以无规则卷曲为主, 只有 R2R3-LcMYB3、R2R3-LcMYB4 与 R1-LcMYB15 蛋白是以 α -螺旋为主; 从 2 个亚族中选取具有代表性的序列 R2R3-LcMYB9、R2R3-LcMYB4、R1-LcMYB2、R1-LcMYB15 进行三级结构同源建模, 发现 R2R3-LcMYB 亚族中的蛋白质三级结构相似性较高, 在无规则卷曲长度上略有不同; 而 R1-LcMYB 亚族蛋白, 主要是在 α -螺旋数量上差异较大。通过对过路黄 MYB 转录因子的保守结构域基序分析以及氨基酸高级结构的预测, 为研究过路黄属及报春花科其他植物的 MYB 转录因子的结构和功能进化奠定基础。

3.4 聚类分析

过路黄与拟南芥 MYB 蛋白的聚类分析结果表明, 大部分 LcMYB 都与拟南芥处于同一分支, 表明过路

黄和拟南芥 MYB 成员具有相近的进化起源,但一些分支中缺少拟南芥或过路黄 MYB 成员,证明这 2 个物种的 R2R3-MYB 基因可能存在多样性。R3-MYB 转录因子被认为具有相同或相似功能^[21];本研究用所得 47 条过路黄 MYB 类转录因子蛋白与 22 条拟南芥 MYB 蛋白作系统进化树,结果显示过路黄和拟南芥 MYB 基因被聚为 3 个大类,其中 R2R3-MYB 分布广泛,在 II 类中占过路黄总 R2R3-MYB 类转录因子的 60.9%, R1-MYB 主要被聚在第 I 类中,占过路黄总 R1-MYB 类转录因子的 37.3%;在该进化树中存在着不同结构域与相同结构域相邻的情况。总体表明,过路黄 MYB 家族蛋白与拟南芥 MYB 家族蛋白同源性高。其中, R2R3-LcMYB28 与 R2R3-LcMYB30 相邻; R2R3-LcMYB40 与 R2R3-LcMYB44 相邻且在 R2R3-LcMYB34 基础上进化而来。研究表明拟南芥中黄酮类物质主要受 AtMYB11、AtMYB12 和 AtMYB111 以及 AtMYB75/PAP1、AtMYB90、AtMYB114 调控^[24-25]。因此可以推测 R2R3-LcMYB33 及 R2R3-LcMYB36 可能具有相似的功能,参与过路黄中黄酮类次生代谢物的生物合成,这将为过路黄 MYB 转录因子的功能研究提供参考,进而为过路黄 MYB 基因调控黄酮类化合物合成的功能鉴定与作用机制研究、高产优质过路黄育种和种植提供科学依据。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国药典 [S]. 一部. 2020: 229.
- [2] 万德光, 彭成, 赵军宁. 四川道地中药材志 [M]. 成都: 四川科学技术出版社, 2005: 367-375.
- [3] 莫刘基, 邓家泰, 张金梅, 等. 几种中药对输尿管结石排石机理的研究(摘要) [J]. 新中医, 1985, 17(6): 53-54.
- [4] 金丽鑫. 金钱草对照药材标定技术标准的研究 [D]. 成都: 成都中医药大学, 2012.
- [5] 王佰灵, 罗伦, 戈振凯, 等. 基于网络药理学和分子对接探讨金钱草治疗痛风的作用机制 [J]. 天然产物研究与开发, 2021, 33(5): 859-867.
- [6] 王宇杰, 孙启时. 金钱草的化学成分研究 [J]. 中国药物化学杂志, 2005, 15(6): 357-359.
- [7] Kim H A, Lee D S, Lee H, et al. *Lysimachia christinae* Hance as an anticancer agent against breast cancer cells [J]. *Food Sci Nutr*, 2020, 8(10): 5717-5728.
- [8] Wu N H, Ke Z Q, Wu S, et al. Evaluation of the antioxidant and endothelial protective effects of *Lysimachia christinae* Hance (Jin Qian Cao) extract fractions [J]. *BMC Complement Altern Med*, 2018, 18(1): 128.
- [9] Wang J M, Zhang Y Y, Zhang Y S, et al. Protective effect of *Lysimachia christinae* against acute alcohol-induced liver injury in mice [J]. *Biosci Trends*, 2012, 6(2): 89-97.

- [10] 伍小方, 高国应, 左倩, 等. FtMYB1 转录因子调控苦荞毛状根黄酮醇合成的机理研究 [J]. 植物遗传资源学报, 2020, 21(5): 1270-1278.
- [11] 赵莹, 杨欣宇, 赵晓丹, 等. 植物类黄酮化合物生物合成调控研究进展 [J]. 食品工业科技, 2021, 42(21): 454-463.
- [12] Falcone Ferreyra M L, Rius S P, Casati P. Flavonoids: biosynthesis, biological functions, and biotechnological applications [J]. *Front Plant Sci*, 2012, 3: 222.
- [13] Kocábek T, Mishra A K, Matoušek J, et al. The R2R3 transcription factor HlMYB8 and its role in flavonoid biosynthesis in hop (*Humulus lupulus* L.) [J]. *Plant Sci*, 2018, 269: 32-46.
- [14] Onkokesung N, Reichelt M, van Doorn A, et al. Modulation of flavonoid metabolites in *Arabidopsis thaliana* through overexpression of the MYB75 transcription factor: Role of kaempferol-3, 7-dihydroxy in resistance to the specialist insect herbivore *Pieris brassicae* [J]. *J Exp Bot*, 2014, 65(8): 2203-2217.
- [15] Nakabayashi R, Yonekura-Sakakibara K, Urano K, et al. Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids [J]. *Plant J*, 2014, 77(3): 367-379.
- [16] Ogata K, Kanei-Ishii C, Sasaki M, et al. The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation [J]. *Nat Struct Biol*, 1996, 3(2): 178-187.
- [17] Liu J Y, Osbourn A, Ma P D. MYB transcription factors as regulators of phenylpropanoid metabolism in plants [J]. *Mol Plant*, 2015, 8(5): 689-708.
- [18] Du H, Zhang L, Liu L, et al. Biochemical and molecular characterization of plant MYB transcription factor family [J]. *Biochemistry (Mosc)*, 2009, 74(1): 1-11.
- [19] Dubos C, Stracke R, Grotewold E, et al. MYB transcription factors in *Arabidopsis* [J]. *Trends Plant Sci*, 2010, 15(10): 573-581.
- [20] 王静文, 谭萌萌, 孙绍营, 等. 基于转录组信息的细叶百合 MYB 转录因子家族分析 [J]. 分子植物育种, 2021, <https://kns.cnki.net/kcms/detail/46.1068.S.2021.1213.1536.009.html>
- [21] 严莉, 王翠平, 陈建伟, 等. 基于转录组信息的黑果枸杞 MYB 转录因子家族分析 [J]. 中国农业科学, 2017, 50(20): 3991-4002.
- [22] Jin H, Martin C. Multifunctionality and diversity within the plant MYB-gene family [J]. *Plant Mol Biol*, 1999, 41(5): 577-585.
- [23] 杜海. 植物 MYB 转录因子家族的分子进化机制及调控类黄酮生物合成 MYB 基因的鉴定 [D]. 雅安: 四川农业大学, 2013.
- [24] Qi X W, Fang H L, Chen Z Q, et al. Ectopic expression of a R2R3-MYB transcription factor gene LjaMYB12 from *Lonicera japonica* increases flavonoid accumulation in *Arabidopsis thaliana* [J]. *Int J Mol Sci*, 2019, 20(18): 4494.
- [25] Wang X C, Wu J, Guan M L, et al. *Arabidopsis* MYB4 plays dual roles in flavonoid biosynthesis [J]. *Plant J*, 2020, 101(3): 637-652.

[责任编辑 时圣明]