

百合 bHLH 家族系统进化和花香相关 基因克隆及表达分析

武琼¹, 刘佳鑫², 赵瑛杰¹, 梁蕤³, 韩美玲², 李超⁴, 杜方²

(¹山西农业大学园艺学院, 太谷 030801; ²山西农业大学城乡建设学院, 太谷 030801; ³南京农业大学园艺学院, 南京 210095;

⁴山西禾田悦农业技术服务有限公司, 吕梁 033299)

摘要: bHLH 转录因子家族是植物中最大的转录因子家族之一。当前越来越多的植物 bHLH 转录因子家族被鉴定, 然而百合 bHLH 转录因子家族的系统分析尚未见报道。本研究基于百合转录组数据, 共鉴定出 74 个 bHLH 家族蛋白, 均为亲水性蛋白, 93% 为不稳定蛋白, 61% 为酸性蛋白。结构域分析发现有 25 个保守残基的一致性 $\geq 50\%$, 其中 R16、R17、L27、L49 和 L59 位点高度保守。此外, 64 个 bHLH 蛋白能与 DNA 结合, 包含 58 个 E-box 结合物和 46 个 G-box 结合物。系统进化分析将百合 bHLHs 分为 21 个亚家族, 基于进化树发现百合 bHLHs 可能执行信号转导、非生物胁迫、植物生长发育和物质合成等功能。对 3 个可能与罗勒烯和芳樟醇相关的基因 *Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 进行了克隆, 发现它们在品种间存在 97.30%~99.89% 的同源性。表达模式分析结果显示, 三者在花、叶和鳞中均有表达。该研究为开展百合 bHLH 转录因子的功能研究提供了重要的序列信息。

关键词: 百合; bHLH 家族; 鉴定; 基因克隆; 半定量 RT-PCR

Evolutionary of bHLH Transcription Factor Family in *Lilium* and Cloning and Expression Analysis of Floral Scent Related Genes

WU Qiong¹, LIU Jiaxin², ZHAO Yingjie¹, LIANG Rui³, HAN Meiling², LI Chao⁴, DU Fang²

(¹College of Horticulture, Shanxi Agricultural University, Taigu 030801; ²College of Urban and Rural Construction, Shanxi

Agricultural University, Taigu 030801; ³College of Horticulture, Nanjing Agricultural University, Nanjing 210095; ⁴Shanxi

Hetianyue Agricultural Technology Service Company Limited, Lvliang 033299)

Abstract: The bHLH transcription factor families is one of the largest families of transcription factors in plants. More and more plant bHLH transcription factor families have been identified presently. However, systematic analysis of the lily bHLH transcription factor family has not been reported. Based on the transcriptome data of *Lilium*, this study identified a total of 74 bHLH family members. All the bHLH proteins were hydrophilic, 93% of which were unstable proteins, and 61% were acidic proteins. Through the domain analysis, 25 conserved residues were accounted for $\geq 50\%$ consistence, with R16, R17, L27, L49 and L59 being highly conserved. Sixty-four bHLH proteins were predicted to bind to DNA, including 58 E-box and 46 G-box bindings. Based on the phylogenetic analysis, lily bHLHs were assigned into 21 subfamilies, and they were

收稿日期: 2023-12-02 网络出版日期: 2024-01-23

URL: <https://doi.org/10.13430/j.cnki.jpgr.20231202001>

第一作者研究方向为园艺植物种质创新与利用, E-mail: wuqiong04252021@163.com

通信作者: 杜方, 研究方向为园艺植物种质创新与利用, E-mail: df730227@yeah.net

基金项目: 山西省基础研究计划自然科学面上项目(20210302123416); 山西农业大学 2023 年“特优”农业科技支撑工程项目(TYGC23-64); 山西农业大学横向科技项目(2023HX292)

Foundation projects: Natural Science Basic Research Project of Shanxi General Program (20210302123416); Excellent Agricultural Science and Technology Project of Shanxi Agricultural University in 2023 (TYGC23-64); Commercial Project of Shanxi Agricultural University (2023HX292)

annotated with functions on signal transduction, abiotic stress, plant growth and development, and substance synthesis. Three genes *Unigene23213_All*, *CL1682.Contig2_All* and *CL8286.Contig2_All*, which associated with ocimene and linalool biosynthesis, were cloned showing 97.30%-99.89% homology among varieties at each locus. The three genes were expressed in flowers, leaves and scales. The findings of this research provided a valuable information in further study deciphering the functions of lily bHLH transcription factors.

Key words: *Lilium*; bHLH family; identification; gene cloning; semi-quantitative RT-PCR

bHLH(basic helix-loop-helix)是一类广泛存在于动植物和真菌中的转录因子超家族,其在真核生物的正常生长及发育过程中必不可缺^[1]。在植物中,bHLH转录因子主要参与生长发育和各种物质合成及信号转导的调控,同时在植物响应胁迫反应中起着关键作用。首个植物bHLH转录因子于1990年在玉米中发现,是由R基因编码的玉米Lc蛋白^[2]。当越来越多的bHLH家族被鉴定后,1998年Atchley等^[3]分析了动物、植物、酵母等生物中392个bHLH蛋白结构域,发现bHLH结构域包含19个保守氨基酸残基,其中5个位于基本区(Basic),5个位于第1螺旋(Helix1),1个位于环(Loop),8个位于第2螺旋(Helix2)。2003年Toledo-Ortiz等^[4]基于拟南芥bHLH家族的分析,对Atchley等^[3]的结果进行了调整,以“19个保守氨基酸残基中最多有9个不匹配”作为鉴定bHLH的标准。植物bHLH转录因子家族大致分为15~32个亚家族^[5-6]。根据序列相似性和进化关系以及结合DNA的能力,分为A~F六大类群^[7]。根据bHLH结构域的N端序列特征以及结合DNA的能力,分为能与DNA结合和不能与DNA结合的bHLH^[4]。DNA结合蛋白又被分为两个亚类,E-box(包括G-box结合物及非G-box结合物)和非E-box结合物。其中13、16位点分别为谷氨酸(E)、精氨酸(R)的bHLH具有识别并结合E-box序列(CANNTG)的功能,无E13、R16位点特性的bHLH则不能与E-box结合;再者,9位点为赖氨酸(K)或组氨酸(H)且17位点为精氨酸(R)的能与G-box结合,否则不能与G-box结合。髓细胞组织增生蛋白(MYC,myelocytomatosis proteins)是bHLH转录因子家族成员之一,其除了含有bHLH结构域外,在N端还有一个保守的bHLH-MYC_N结构域^[8]。

百合(*Lilium* spp.)是具有重要经济价值的花卉。虽然有学者对百合鳞茎花青素合成相关的bHLH进行了挖掘分析^[9],且有学者克隆了部分***bHLH***基因,但目前尚未见关于百合bHLH家族系统分析的报道。研究人员已在5种百合中克隆到11

个***bHLH***转录因子基因,包括岷江百合(*L. regale* Wilson)中功能未知的***bHLH062***和***BHLH***;亚洲百合Montreux中克隆得到的参与花青素合成的***LhbHLH1***和***LhbHLH2***基因^[10],美丽百合(*L. speciosum* Thunb)中克隆到的***LsbHLH2-rr1***、***LsbHLH2-rr2***、***LsbHLH2-wr1***、***LsbHLH2-wr2***、***LsbHLH2-wy1***、***LsbHLH2-wy2***^[11],以及岷江百合中克隆到的***LrbHLH2-1***、***LrbHLH2-2***基因^[12];与黄瓜花叶病毒有关的岷江百合***bHLH18***和***bHLH100***基因^[13];与灰霉病抗性有关的岷江百合***LrMYC2***^[14]和东方百合Sorbonne***LhMYC2***基因^[15];与花药发育有关的东方百合Siberia***LoUDT1***^[16]和***LoAMS***^[17];与花香相关的Siberia***LoMYC2***^[18]和***LibHLH22***与***LibHLH63***^[19]基因。

本研究基于百合转录组数据库,挖掘百合***bHLH***基因家族,分析其数量和特性并进行分类,同时在东方百合Sorbonne中克隆可能与百合花香相关的***bHLH***基因,并分析其在不同组织器官中的表达规律以预测其潜在的功能,以期深入探讨百合***bHLH***基因的功能提供理论依据。

1 材料与方法

1.1 数据来源

课题组前期利用Illumina平台对东方百合Sorbonne花、叶、鳞、茎皮、根、柱头6个器官的混合样测序获得了一套转录组数据^[20]。下载基于Illumina平台的麝香百合White Europe花粉转录组数据^[21],以及基于Roche454平台的4个基因型百合叶片转录组数据^[22],并将三套转录组数据进行杂交组装,形成一套新的转录组L.-Unigene-All^[20],它包含7个器官(根、鳞、叶、茎皮、花瓣、柱头和花粉)和6个品种/品系(Sorbonne、Star Gazer、White Fox、Connecticut King、Trumpet 061099、White Europe)。

1.2 百合bHLH家族的鉴定

从Pfam 36.0数据库^[23](<http://pfam.xfam.org/>)中下载bHLH结构域的种子文件(PF00010),利用HMMER 3.0^[24]软件构建相应的Profile HMM(数值

表格型隐马可夫模型),并基于此模型检索百合转录组数据库中潜在的 bHLH,获得数据集 1。将百合转录组数据正反向序列提交到 PlantTFDB v5.0^[25] (<http://plantfdb.gao-lab.org/>) 在线网站预测数据库所包含的转录因子家族及数量,并筛选百合 bHLH,获得数据集 2。汇总数据集 1 和 2 的结果确定候选百合 bHLH。随后,将初步筛选的百合 bHLH 蛋白序列提交至 Tertools v2.069^[26] 的 Batch SMART 插件,去除结果中‘HLH’长度不在 44~55 aa 范围内的序列,并将筛选结果提交至 TBtools 的 Blast Zone 进行去冗余,保留最长的序列,并根据“19 个保守氨基酸残基中最多有 9 个不匹配”的原则去掉 19 个保守氨基酸残基中小于 10 个匹配的序列,最终获得百合 bHLHs 转录因子。

1.3 百合 bHLH 家族蛋白特性分析

基于筛选鉴定的百合 bHLHs,用在线软件 ExPASy ProtParam tool^[27] (<https://web.expasy.org/protparam/>) 分析氨基酸数量、相对分子量、理论等电点、不稳定性及疏水性。用在线软件 Cell Ploc 2.0^[28] (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-Ploc/>) 预测亚细胞定位。用在线软件 SOPMA^[29] (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) 分析二级结构。

1.4 百合 bHLH 家族的保守域序列分析及分类

利用序列分析软件 DNAMAN 对百合 bHLH 家族蛋白结构域进行比对,比对结果使用在线软件 WebLogo v2.8.2^[30] (<http://weblogo.berkeley.edu/logo.cgi>) 分析保守碱基。参照 Toledo-Ortiz 等^[4] 的方法,根据 bHLH 与 DNA 的结合模式对百合 bHLH 进行分类。

1.5 系统进化树的构建

利用 MEGA 7.0 软件内置的 Clustal W 程序对候选百合 bHLH 家族、已报道的百合 bHLH 及拟南芥 bHLH 蛋白结构域进行比对分析,将比对结果采用

邻接法(NJ, neighbor joining)构建系统发育树,其中校验参数 Bootstrap 重复 1000 次,模式设置为 Poisson model,缺失值处理方式为 pairwise deletion,其他参数使用默认值。利用在线软件 iTOL^[31] (<https://itol.embl.de/itol.cgi>) 对进化树进行美化。

1.6 *Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 基因的表达模式分析

根据 1.5 的结果选择与已报道的百合花香相关 bHLH 基因(*LoMYC2*、*LibHLH22* 和 *LibHLH63*) 高度同源的基因(*Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All*) 进行进一步研究。

依据课题组东方百合 Sorbonne 花、叶和鳞的表达谱数据^[32],查询 *Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 3 个基因的 RPKM 值并进行 \log_2 转化,将其输入到 Tertools 的 HeatMap 插件中绘制热图。

1.7 *Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 基因的半定量 RT-PCR

根据序列同源性,设计与 *LoMYC2* 同源的 *Unigene23213_All* 的引物,与 *LibHLH63* 同源的 *CL8286.Contig2_All* 的引物,与 *LibHLH22* 同源的 *CL1682.Contig2_All* 的引物来源于文献[19]。所有引物及内参 *LhActin* 交由上海生工生物工程股份有限公司合成,引物序列如表 1 所示。以东方百合 Sorbonne 花、叶、鳞 cDNA 为模板,对 *Unigene23213_All*、*CL1682.Contig2_All*、*CL8286.Contig2_All* 和 *LhActin* 进行半定量 RT-PCR 检测。PCR 扩增体系为 10 μ L,包括 2 \times Rapid Taq Master Mix (诺唯赞生物科技股份有限公司) 3 μ L、cDNA 2 μ L、正/反向引物各 0.5 μ L、ddH₂O 4 μ L。反应程序为 95 $^{\circ}$ C 3 min; 95 $^{\circ}$ C 15 s, 退火(温度见表 1) 15 s, 72 $^{\circ}$ C 1 min, 35 个循环; 72 $^{\circ}$ C 延伸 5 min。将第一次 PCR 反应产物作为模板进行二次 PCR 扩增后,取 5 μ L PCR 产物,在 1% 琼脂糖凝胶上检测。

表 1 引物信息

Table 1 Information of the primers

引物名称 Primer name	上游引物(5'-3') Forward primer (5'-3')	下游引物(5'-3') Reverse primer (5'-3')	目标长度 (bp) Target length	退火温 度($^{\circ}$ C) T _m
<i>Unigene23213_All</i>	ATGAAGAGGAACCTCGTCCC	TCAGGCTCCTCAGAGATC	912	58
<i>CL1682.Contig2_All</i>	ATGGAGACAACGCCAGCTC	TCAGAGCTCCGCTTTAACTGGTTC	1089	60
<i>CL8286.Contig2_All</i>	ACTCTATCGCAAGATTTCAG	CTGAGTAAATTCTGTGTGGC	1211	55
<i>LhActin</i>	TGCTGGATTCTGGTGATGGT	ATAGGTGGTCTCGTGGATGC	383	58

1.8 Unigene23213_All、CL1682.Contig2_All 和 CL8286.Contig2_All基因的克隆

采用表1中的引物,以东方百合 Sorbonne 的花为模板进行基因克隆,PCR反应体系与程序同1.7。利用 Simgen 公司的凝胶DNA回收试剂盒回收目的片段,回收后连接至Takara公司pMD™19-T载体上,然后转化到全式金公司 Trans5 α 大肠杆菌感受态细胞,涂入蓝白斑筛选的平板,挑取白斑进行菌液PCR验证后送上海生工生物工程股份有限公司进行测序。利用NCBI网站中的Open Reading Frame Finder 工具确定测序结果的CDS序列。利用DNAMAN对获得序列和已知的同源基因进行序列比对。

2 结果与分析

2.1 百合 bHLH 家族成员鉴定

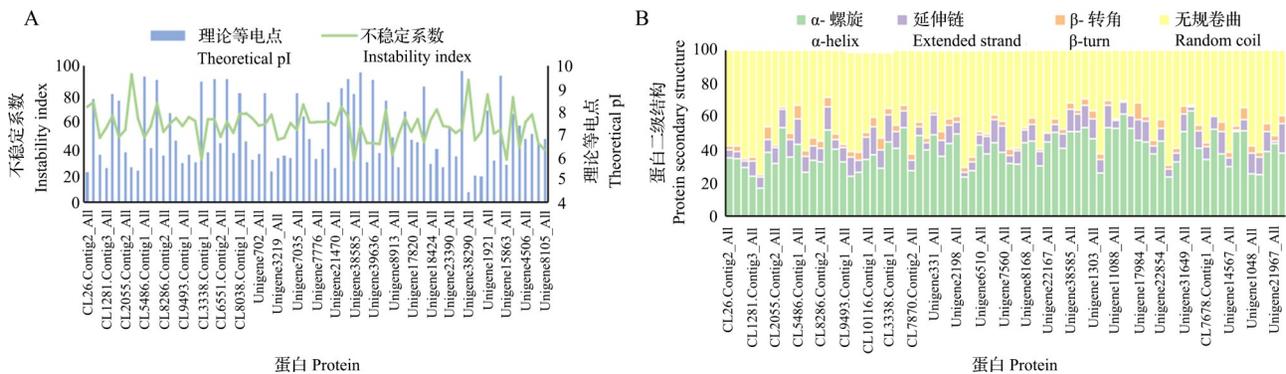
使用 Pfam 与 PlantTFDB 对东方百合 Sorbonne 的 bHLH 转录因子进行初步筛选,其中从 Pfam 中得到的数据集1包含120个候选序列,从 PlantTFDB 中得到的数据集2包含132个候选序列,数据集2中的132个序列包含了数据集1中所得到的120个序列。可见 PlantTFDB 数据库中的信息更为全面,从 PlantTFDB 中得到的数据集2即为候选百合 bHLH。将这132个候选序列输入 TBtools 的 Batch SMART 插件,去掉结构域不完整的序列,用 Blast Zone 插件

去除冗余序列,最后根据 Toledo-Ortiz 等^[4]制定的标准保留结构域中最少有10个氨基酸残基保守的序列,最终获得74个百合 bHLHs。

2.2 百合 bHLH 家族蛋白特性分析

百合 bHLHs 可编码110~597个氨基酸,预测相对分子量在12462.43~67174.13 Da之间。理论等电点范围在4.42~10.10,其中酸性蛋白(理论等电点<7)有45个(61%),碱性蛋白(理论等电点>7)有29个(39%)(图1A)。不稳定指数(II)分析表明除 CL3338.Contig1_All、Unigene38585_All、Unigene8913_All、Unigene15863_All 和 Unigene8105_All 等5个为稳定蛋白外(不稳定指数<40),其余均为不稳定蛋白(93%)(图1A)。所有 bHLH 蛋白的亲水性值均为负值,范围在-0.865~-0.054,表明其都具有亲水特性。亚细胞定位预测显示,除 Unigene8105_All 定位在叶绿体,其余 bHLH 蛋白均定位于细胞核。

利用 SOPMA 在线网站预测百合 bHLH 蛋白二级结构, α -螺旋占比为17.11%~63.64%,延伸链为0.81%~17.66%, β -转角占比为0~10.00%,无规则卷曲程度为28.42%~74.95%。二级结构分析发现,百合 bHLH 家族蛋白中 α -螺旋和无规卷曲所占比例较大,其中62.16%的蛋白二级结构中无规卷曲所占比例最大,37.84%的蛋白二级结构中 α -螺旋所占比例最大(图1B)。



图中只显示25个百合 bHLH 蛋白名称
Only 25 lily bHLH protein names were shown
图1 百合 bHLH 转录因子家族蛋白特性分析

Fig.1 Characterization of lily bHLH transcription factor family proteins

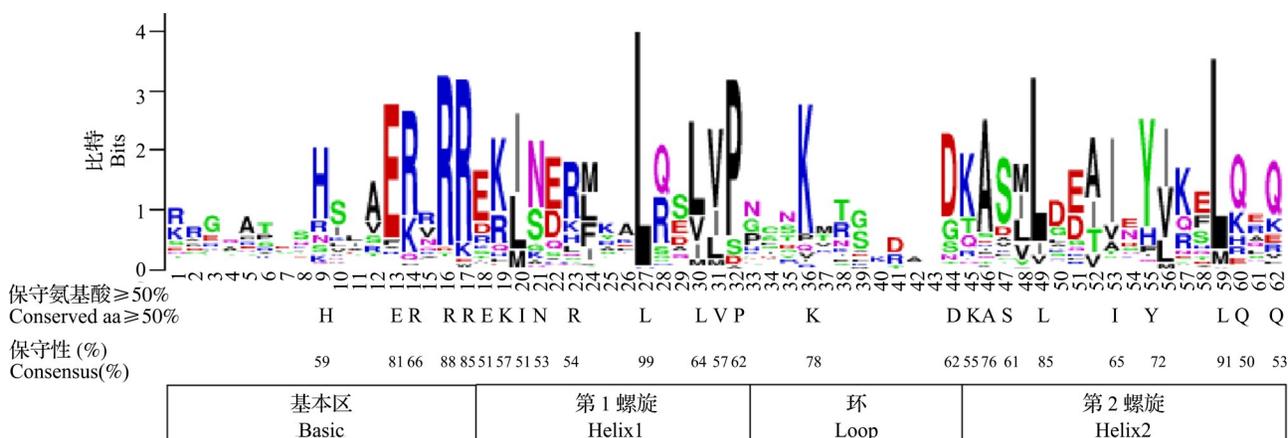
2.3 百合 bHLH 结构域保守性氨基酸残基的分布及特征

74个百合 bHLH 结构域比对显示,百合中氨基酸残基一致性 $\geq 50\%$ 的残基有25个(图2)。其中5

个分布在基本区(H9、E13、R14、R16、R17),9个分布在第1螺旋区(E18、K19、I20、N21、R23、L27、L30、V31、P32),环区2个(K36、D44),其余9个在第2螺旋区(K45、A46、S47、L49、I53、Y55、L59、Q60、

Q62)。可见,保守氨基酸残基主要分布在两个螺旋区。其中 R16、R17、L27、L49 与 L59 位点表现出极高的保守性($\geq 85\%$)。此外,100%的 20、48、56 和 59

位点,99%的 24、27、30、31 和 49 位点为疏水性氨基酸(A、F、I、L、M、P、V、W 或 Y)。



天冬氨酸(D)和谷氨酸(E)为红色,天冬酰胺(N)和谷氨酰胺(Q)为紫色,赖氨酸(K)、精氨酸(R)和组氨酸(H)为蓝色,半胱氨酸(C)、甘氨酸(G)、丝氨酸(S)、苏氨酸(T)和酪氨酸(Y)为绿色,丙氨酸(A)、缬氨酸(V)、亮氨酸(L)、异亮氨酸(I)、脯氨酸(P)、色氨酸(W)、苯丙氨酸(F)和甲硫氨酸(M)为黑色

Aspartic acid (D) and glutamic acid (E) were shown in red, asparagine (N) and glutamine (Q) were shown in purple, lysine (K), arginine (R) and histidine (H) were shown in blue, cysteine (C), glycine (G), serine (S), threonine (T) and tyrosine (Y) were shown in green, alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophan (W), phenylalanine (F), and methionine (M) were shown in black

图2 百合 bHLH 转录因子家族保守氨基酸分析

Fig.2 Conserved amino acid analysis of lily bHLH transcription factor family

2.4 基于DNA结合能力的百合 bHLH 分类

基于DNAMAN对百合 bHLH 家族蛋白结构域的比对结果发现,有 58 个 bHLH 具有保守的 E13/R16 位点,为 E-box 结合蛋白。在能与 E-box 结合的 bHLH 中,46 个 bHLH 具有保守的 K/H9 和 R17 位点,为 G-box 结合蛋白。由于氨基酸 E13 或 R16 的缺失,以及在基本区中碱性氨基酸少于 6 个,有 10 个 bHLH 为不能与 DNA 结合的蛋白(表 2)。

表 2 预测百合 bHLH 蛋白结构域的 DNA 结合特性

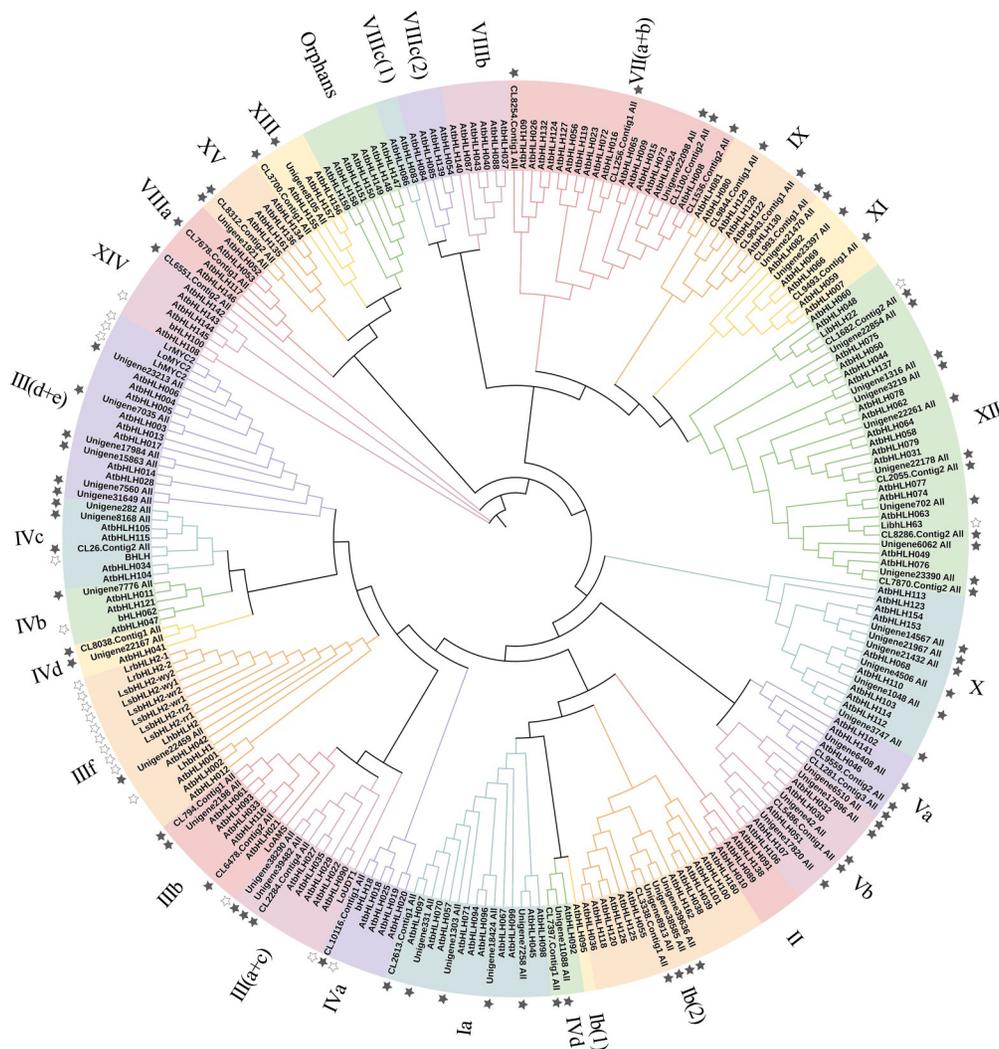
Table 2 Prediction of DNA binding properties of lily bHLH protein domains

结合模式 Binding pattern	数量 Number
能与 DNA 结合 DNA binding	64
能与 E-box 结合 E-box binding	58
能与 G-box 结合 G-box binding	46
不能与 G-box 结合 Non-G-box binding	12
不能与 E-box 结合 Non-E-box binding	6
不能与 DNA 结合 Non-DNA binding	10

2.5 百合 bHLH 转录因子家族进化分析

构建拟南芥与百合 bHLH 家族的系统进化树,研究其进化关系(图 3)。根据拟南芥的进化分类可将所鉴定的百合 bHLHs 划分为 21 个亚族,其主要分布在 XII(12 个)、III(d+e)(6 个)和 X(6 个) 3 个亚家族中,其中已报道的百合 bHLH 分别集中在 III(a+c)、IIIb、III(d+e)、IIIc、IVa、IVb、IVc、XII 和 XIV 9 个亚族。与拟南芥相比,百合 bHLH 没有聚集到 Ib(1)、II、VIIIb、VIIIc(1)和 VIIIc(2) 5 个亚家族中。在多数情况下,拟南芥和百合在同一个亚家族中具有不同数量的 bHLH。

同一亚家族成员可能具有相类似的功能,推测与拟南芥 bHLH 聚在相同亚族的百合 bHLH 发挥着相同或相近的作用。在百合 21 个 bHLHs 亚族中, Vb 和 VIIIa 亚族同源的 AtbHLH 功能未知,且这 2 个亚族蛋白在 NCBI 中未能对比到任何功能明确的 bHLH,因此 Vb 和 VIIIa 亚族功能尚待研究。其余亚族可能参与信号转导、非生物胁迫、植物生长发育和物质合成等过程。其中 III(d+e)家族基因可能与花香物质芳樟醇的合成相关^[18], XII 家族基因可能参与花香物质罗勒烯和芳樟醇的合成^[19]。



灰色实心五角星标注的是本研究鉴定的百合 bHLHs, 灰色空心五角星标注的是已报道的百合 bHLH

The solid gray pentagram indicates the lily bHLHs identified in this paper, and the hollow gray pentacle indicates reported lily bHLH

图3 百合与拟南芥 bHLH 结构域系统发育树

Fig.3 Phylogenetic analysis of bHLH domain from *Lilium* and *Arabidopsis thaliana*

2.6 *Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 基因在花、叶和鳞中的表达

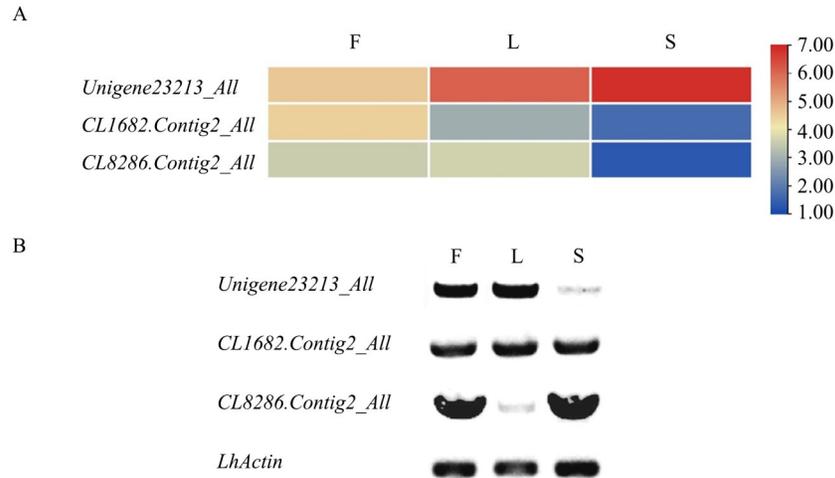
基于 Sorbonne 花、叶和鳞 3 个器官表达谱所做热图可知(图 4A), *Unigene23213_All* 在花、叶和鳞中的表达相较于 *CL1682.Contig2_All* 和 *CL8286.Contig2_All* 在 3 个器官中的表达量高。*Unigene23213_All* 在鳞和叶中高表达, 在花中低表达。*CL1682.Contig2_All* 在花中表达量最高, 在叶和鳞中表达量低。*CL8286.Contig2_All* 在花和叶中的表达量近似, 在鳞中低表达。

半定量 RT-PCR 琼脂糖凝胶电泳图显示(图 4B), *Unigene23213_All*、*CL1682.Contig2_All* 与 *CL8286.Contig2_All* 在 Sorbonne 的花、叶和鳞中均有表达。*Unigene23213_All* 在 Sorbonne 花和叶中的条带明亮, 鳞中暗淡, 推测其在 Sorbonne 花和叶中的表达量远高于鳞。*CL1682.Contig2_All* 在 Sorbonne 花、叶和鳞中的条带清晰明亮, 在 3 个器官中的表达

水平几乎一致。*CL8286.Contig2_All* 在 Sorbonne 花和鳞中明显高表达, 在叶中微弱表达。

2.7 *Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 基因的序列分析

前人研究^[15]表明, Sorbonne 中的 *LhMYC2* 与岷江百合中 *LrMYC2* 的氨基酸序列一致, 两者都与灰霉病抗性有关。序列对比发现(图 5), *LhMYC2* 和 *LrMYC2* 与调控花香的 Siberia *LoMYC2* 氨基酸序列同源性高达 99.89%, 存在 3 个单氨基酸多态性(SAP, single amino acid polymorphism)位点, *LhMYC2*、*LrMYC2* 和 *LoMYC2* 为同源基因, 该基因可能同时参与灰霉病和花香的调控。Local Blast 发现, *LhMYC2* 与转录组数据库(GeneBank No. SRP084220)中的 *Unigene23213_All* 同源, 推测 *Unigene23213_All* 即为 *LhMYC2*, 且因 *LhMYC2* 已在 Sorbonne 中克隆, 则本研究不再重复克隆。

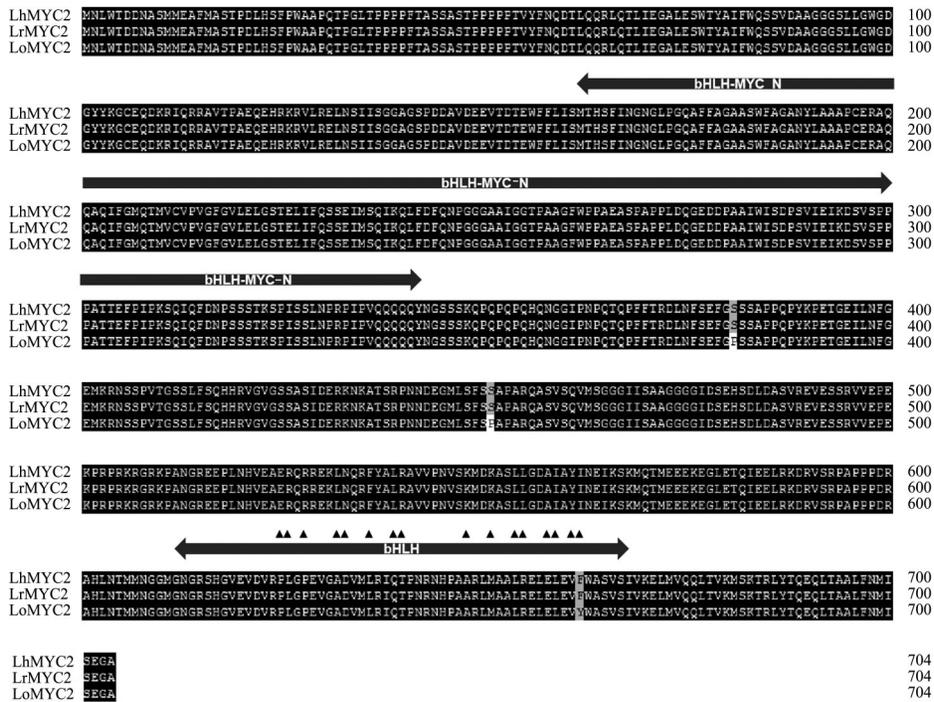


A: *Unigene23213_All*, *CL1682.Contig2_All*, *CL8286.Contig2_All*在花、叶和鳞中的表达情况; B: *Unigene23213_All*, *CL1682.Contig2_All*, *CL8286.Contig2_All*在花、叶、鳞中的半定量分析; F: 花, L: 叶, S: 鳞

A: Expression pattern of *Unigene23213_All*, *CL1682.Contig2_All*, *CL8286.Contig2_All* in flower, leaf and scale; B: Semi-quantitative analysis of *Unigene23213_All*, *CL1682.Contig2_All*, *CL8286.Contig2_All* in flower, leaf and scale; F: Flower, L: Leaf, S: Scale

图4 *Unigene23213_All*, *CL1682.Contig2_All* 和 *CL8286.Contig2_All* 的空间表达

Fig.4 The spatial expression of *Unigene23213_All*, *CL1682.Contig2_All* and *CL8286.Contig2_All*



黑色三角表示 bHLH 结构域的保守位点; 下同

The black triangle represents conserved sites in the bHLH domain; The same as below

图5 *LhMYC2*, *LrMYC2* 与 *LoMYC2* 氨基酸序列比对图

Fig.5 Amino acid sequence comparison of *LhMYC2*, *LrMYC2* and *LoMYC2*

CL1682.Contig2_All 和 *CL8286.Contig2_All* 分别与 Siberia 中调控花香物质罗勒烯和芳樟醇合成的 *LibHLH22* 和 *LibHLH63* 同源。以 Sorbonne 花 cDNA 为模板对 *CL1682.Contig2_All* 和 *CL8286.Contig2_All* 进行 PCR 扩增, 凝胶电泳检测结果与预期大小一致(图 6)。经克隆测序后发现, *CL1682.*

Contig2_All CDS 区为 1089 bp, 编码 362 个氨基酸, 与 *LibHLH22* 同源性为 99.45%, 存在 2 个 SAP 位点(图 7A)。*CL8286.Contig2_All* CDS 区为 1116 bp, 编码 371 个氨基酸, 与 *LibHLH63* 同源性为 97.30%, 在序列 3' 端比 *LibHLH63* 多 8 个氨基酸(图 7B)。序列比对发现, *CL1682.Contig2_All* 与 *LibHLH22* 和 *CL8286.*

*Contig2_All*与*LibHLH63*两两之间bHLH结构域序列完全一致。推测Sorbonne中*CL1682.Contig2_All*和*CL8286.Contig2_All*可能也与罗勒烯和芳樟醇的合成有关。

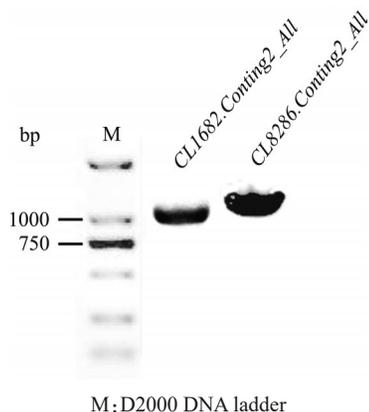
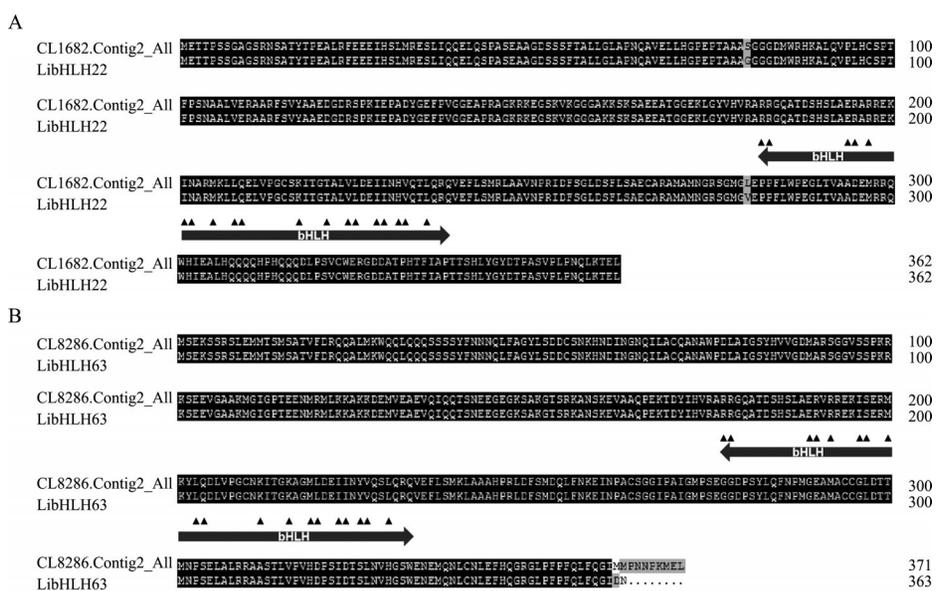


图6 *CL1682.Contig2_All*和*CL8286.Contig2_All*基因的凝胶电泳图

Fig.6 Agarose gel electrophoresis diagram of *CL1682.Contig2_All* and *CL8286.Contig2_All* genes

3 讨论

作为植物中最大的转录因子家族之一, bHLH在植物整个生命周期的各种生物过程中起着关键作用。本研究共鉴定百合bHLH转录因子74个,比单子叶植物水稻(167个)^[33]和玉米(208个)^[34]所挖掘的bHLH转录因子少,可能是因为玉米和水稻bHLH家族的挖掘都是基于基因组数据进行的,而本研究是基于转录组数据进行的。所鉴定的百合bHLH多为酸性蛋白、不稳定蛋白,且均为亲水性蛋白。多数百合bHLH蛋白在二级结构分析中无规卷曲所占比例最大。以上结果与生菜^[35]、甜橙^[36]、云南红梨^[37]等多种植物的bHLH家族蛋白性质相似。亚细胞定位预测除Unigene8105_All定位在叶绿体外,其余bHLHs均定位于细胞核中,推测Unigene8105_All可能参与叶绿体某些功能转录调控,其余百合bHLH蛋白可能通过调控细胞核基因的转录来发挥功能。



A: *CL1682.Contig2_All*与*LibHLH22*蛋白序列比对结果; B: *CL8286.Contig2_All*与*LibHLH63*蛋白序列比对结果
A: Protein sequence alignment between *CL1682.Contig2_All* and *LibHLH22*; B: Protein sequence alignment between *CL8286.Contig2_All* and *LibHLH63*

图7 *CL1682.Contig2_All*与*LibHLH22*和*CL8286.Contig2_All*与*LibHLH63*氨基酸序列比对图

Fig.7 Amino acid sequence comparison between *CL1682.Contig2_All* and *LibHLH22*, *CL8286.Contig2_All* and *LibHLH63*

百合bHLH蛋白结构域中有25个氨基酸位点的保守性在50%及以上,其中E13和R16/R17与DNA结合能力有关^[7],螺旋区域的L27和L59在二聚化过程中发挥关键作用^[6]。此结果与玉米^[34]、谷子^[38]、石蒜^[39]等多数植物的bHLH蛋白结构域存在相同的保守残基,表明bHLH转录因子家族在不同植物之间具有保守性。再者有9个位点的氨基酸多

为疏水性氨基酸,而疏水性氨基酸是保证bHLH蛋白二级结构稳定性的关键^[40]。

根据bHLH与DNA的结合模式和序列同源性,可对bHLH进行不同的分类。研究表明,74个百合bHLH蛋白中有58个(78.38%)能与E-box相结合,46个(62.16%)具有G-box结合功能,该结果与芒果(83.91%/65.52%)^[41]和拟南芥(74.15%/60.54%)^[4]的

分析结果相似。根据与拟南芥 bHLHs 的系统进化关系,将百合 bHLH 划分为 21 个亚族,而甜橙^[36]、菠萝^[42]、石蒜^[39]分别被分为 16 个、18 个、20 个亚族,这可能是不同物种间存在的差异。根据进化树分析结果,鉴定的 68 个百合 bHLH 功能得到了预测,它们可能参与信号转导、非生物胁迫、植物生长发育和物质合成等过程。

Sorbonne 和 Siberia 都属于东方百合品种,基于与 Siberia *LibHLH22* 和 *LibHLH63* 序列的同源性设计引物克隆获得了 Sorbonne *CL1682.Contig2_All* 和 *CL8286.Contig2_All* 基因,但其氨基酸序列与 *LibHLH22* 和 *LibHLH63* 的同源性分别为 99.45% 和 97.30%,推测同一基因在 Sorbonne 和 Siberia 间存在序列差异。表达热图与半定量结果显示 *Unigene23213_All*、*CL1682.Contig2_All* 与 *CL8286.Contig2_All* 在 Sorbonne 花、叶和鳞中都表达。Cui 等^[15]发现 *LhMYC2* 在 Sorbonne 花、叶、茎和根中表达。本研究发现与其同源的 *Unigene23213_All* 在鳞片中也表达。前人发现 *LibHLH22* 和 *LibHLH63* 在花的各个组织如花瓣、花柱、花药、花托、柱头、花丝和子房中均有表达^[19]。在本研究中,*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 除在 Sorbonne 花中表达外,也在叶和鳞中表达。可见,*Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 这 3 个基因具有非器官特异性,说明它们除在花中参与花香调控外,可能在其他器官中也发挥着作用。不过,*Unigene23213_All*、*CL1682.Contig2_All* 和 *CL8286.Contig2_All* 在 Sorbonne 各器官中的表达量需进一步采用实时荧光定量 PCR 进行验证。

综上所述,本研究系统分析了东方百合 Sorbonne 转录组数据,共鉴定 74 个百合 bHLH 转录因子,分为 21 个亚族。通过与拟南芥 bHLHs 的比较分析,预测百合 bHLHs 可能与信号转导、非生物胁迫、植物生长发育和物质合成等功能相关。明确了 3 个可能与罗勒烯和芳樟醇相关的基因 *Unigene23213_All*、*CL1682.Contig2_All* 与 *CL8286.Contig2_All*, 并发现三者 Sorbonne 花、叶和鳞中均有表达。总体上,百合中还有大量功能未知的 bHLH 基因需要进一步研究,本研究为进一步系统而全面了解百合 bHLH 转录因子提供了基础资料。

参考文献

[1] Massari M E, Murre C. Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Molecular Cellular Biology*, 2000, 20(2): 429-440

- [2] Spelt C, Quattrocchio F, Mol J N, Koes R. *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *The Plant Cell*, 2000, 12(9): 1619-1631
- [3] Atchley W R, Terhalle W, Dress A. Positional dependence, cliques, and predictive motifs in the bHLH protein domain. *Journal of Molecular Evolution*, 1999, 48: 501-516
- [4] Toledo-Ortiz G, Huq E, Quail P H. The *Arabidopsis* basic/helix-loop-helix transcription factor family. *The Plant Cell*, 2003, 15(8): 1749-1770
- [5] Pires N, Dolan L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular Biology Evolution*, 2010, 27(4): 862-874
- [6] Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García J F, Bilbao-Castro J R, Robertson D L. Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae. *Plant Physiology*, 2010, 153(3): 1398-1412
- [7] Atchley W R, Fitch W M. A natural classification of the basic helix-loop-helix class of transcription factors. *Proceedings of the National Academy of Sciences*, 1997, 94(10): 5172-5176
- [8] Fernández-Calvo P, Chini A, Fernández-Barbero G, Chico J M, Gimenez-Ibanez S, Geerinck J, Eeckhout D, Schweizer F, Godoy M, Franco-Zorrilla J M. The *Arabidopsis* bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses. *The Plant Cell*, 2011, 23(2): 701-715
- [9] 范文广, 李保豫, 田辉, 李欣, 任海伟, 曹莹莹, 姜欣彤, 柴佳靖, 陈少青. 兰州百合鳞茎花青素合成相关 MYB 与 bHLH 转录因子的筛选分析. *食品与发酵工业*, 2024, 50(2): 57-66
- Fan W G, Li B Y, Tian H, Li X, Ren H W, Cao Y Y, Jiang X T, Chai J J, Chen S Q. Screening and analysis of MYB and bHLH transcription factors associated with anthocyanin synthesis in *Lilium davidii* var. unicolor bulbs. *Food and Fermentation Industries*, 2024, 50(2): 57-66
- [10] Nakatsuka A, Yamagishi M, Nakano M, Tasaki K, Kobayashi N. Light-induced expression of basic helix-loop-helix genes involved in anthocyanin biosynthesis in flowers and leaves of Asiatic hybrid lily. *Scientia Horticulturae*, 2009, 121(1): 84-91
- [11] Suzuki K, Tasaki K, Yamagishi M. Two distinct spontaneous mutations involved in white flower development in *Lilium speciosum*. *Molecular Breeding*, 2015, 35(10): 1-14
- [12] Yamagishi M. A novel R2R3-MYB transcription factor regulates light-mediated floral and vegetative anthocyanin pigmentation patterns in *Lilium regale*. *Molecular Breeding*, 2016, 36(1): 1-14
- [13] Sun D Y, Zhang X G, Zhang Q Y, Ji X T, Jia Y, Wang H, Niu L X, Zhang Y L. Comparative transcriptome profiling uncovers a *Lilium regale* NAC transcription factor, *LrNAC35*, contributing to defence response against cucumber mosaic

- virus and tobacco mosaic virus. *Molecular Plant Pathology*, 2019, 20(12): 1662-1681
- [14] Cui Q, Liu Q, Gao X, Yan X, Jia G X. Transcriptome-based identification of genes related to resistance against *Botrytis elliptica* in *Lilium regale*. *Canadian Journal of Plant Science*, 2018, 98(5): 1058-1071
- [15] Cui Q, Gao X, Wang L J, Jia G X. Ectopic expression of *LhMYC2* increases susceptibility to *Botrytis cinerea* in *Arabidopsis thaliana*. *Canadian Journal of Plant Science*, 2020, 101(3): 328-340
- [16] Yuan G Z, Wu Z, Liu X Y, Li T, Teng N J. Characterization and functional analysis of LoUDT1, a bHLH transcription factor related to anther development in the lily oriental hybrid Siberia (*Lilium* spp.). *Plant Physiology Biochemistry*, 2021, 166: 1087-1095
- [17] Sui J J, Cao X, Yi M F, Wu J, He J N. Isolation and characterization of *LoAMS* gene in anther development of lily (*Lilium* oriental hybrids). *New Zealand Journal of Crop Horticultural Science*, 2020, 48(4): 257-269
- [18] 王红. 百合 MYC2 转录因子在花香形成中的作用研究. 广州: 华南农业大学, 2017
Wang H. The roles of MYC2 transcription factor in the fragrance biosynthesis of *Lilium* 'Siberia'. Guangzhou: South China Agricultural University, 2017
- [19] Feng Y, Guo Z Y, Zhong J, Liang Y L, Zhang P, Sun M. The *LibHLH22* and *LibHLH63* from *Lilium* 'Siberia' can positively regulate volatile terpenoid biosynthesis. *Horticulturae*, 2023, 9(4): 459
- [20] Du F, Wu Y, Zhang L, Li X W, Zhao X Y, Wang W H, Gao Z S, Xia Y P. De novo assembled transcriptome analysis and SSR marker development of a mixture of six tissues from *Lilium* Oriental hybrid 'Sorbonne'. *Plant Molecular Biology Reporter*, 2015, 33: 281-293
- [21] Obermeyer G, Fagner L, Lang V, Weckwerth W. Dynamic adaptation of metabolic pathways during germination and growth of lily pollen tubes after inhibition of the electron transport chain. *Plant Physiology*, 2013, 162(4): 1822-1833
- [22] Shahin A, van Kaauwen M, Esselink D, Bargsten J W, van Tuyl J M, Visser R G, Arens P. Generation and analysis of expressed sequence tags in the extreme large genomes *Lilium* and *Tulipa*. *BMC Genomics*, 2012, 13: 1-13
- [23] Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar G A, Sonnhammer E L, Tosatto S C, Paladin L, Raj S, Richardson L J. Pfam: The protein families database in 2021. *Nucleic Acids Research*, 2021, 49(D1): D412-D419
- [24] Finn R D, Clements J, Eddy S R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research*, 2011, 39: W29-W37
- [25] Tian F, Yang D C, Meng Y Q, Jin J P, Gao G. PlantRegMap: Charting functional regulatory maps in plants. *Nucleic Acids Research*, 2020, 48(D1): D1104-D1113
- [26] Chen C, Chen H, Zhang Y, Thomas H R, Frank M H, He Y, Xia R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, 2020, 13(8): 1194-1202
- [27] Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C. ExPasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Research*, 2021, 49 (W1): W216-W227
- [28] Chou K C, Shen H B. Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE*, 2010, 5(6): e11335
- [29] Geourjon C, Deleage G. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, 1995, 11 (6): 681-684
- [30] Crooks G E, Hon G, Chandonia J M, Brenner S E. WebLogo: A sequence logo generator. *Genome Research*, 2004, 14(6): 1188-1190
- [31] Letunic I, Bork P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics*, 2007, 23(1): 127-128
- [32] Du F, Fan J M, Wang T, Wu Y, Grierson D, Gao Z S, Xia Y P. Identification of differentially expressed genes in flower, leaf and bulb scale of *Lilium* oriental hybrid 'Sorbonne' and putative control network for scent genes. *BMC Genomics*, 2017, 18(1): 1-14
- [33] Li X X, Duan X P, Jiang H X, Sun Y J, Tang Y P, Yuan Z, Guo J K, Liang W Q, Chen L, Yin J Y. Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and *Arabidopsis*. *Plant Physiology*, 2006, 141(4): 1167-1184
- [34] Zhang T T, Lv W, Zhang H S, Ma L, Li P H, Ge L, Li G. Genome-wide analysis of the basic Helix-Loop-Helix (bHLH) transcription factor family in maize. *BMC Plant Biology*, 2018, 18: 1-14
- [35] 隋心意, 赵小刚, 毛欣, 李亚灵, 温祥珍. 生菜 bHLH 转录因子家族的鉴定与生物信息学分析. *中国农学通报*, 2023, 39 (3): 104-110
Sui X Y, Zhao X G, Mao X, Li Y L, Wen X Z. Identification and bioinformatics analysis of bHLH transcription factor family in *Lactuca sativa* L.. *Chinese Agricultural Science Bulletin*, 2023, 39(3): 104-110
- [36] 孟富宣, 杨玉皎, 段元杰, 吕陟远, 陆晓英, 郭淑萍, 刘海刚. 甜橙 bHLH 转录因子家族的鉴定及生物信息学分析. *西南林业大学学报: 自然科学版*, 2020, 40(5): 73-86
Meng F X, Yang Y J, Duan Y J, Lv Z Y, Lu X Y, Guo S P, Liu H G. Identification and bioinformatics analysis of the bHLH transcription factor family in *Citrus sinensis*. *Journal of Southwest Forestry University: Natural Sciences Edition*, 2020, 40(5): 73-86
- [37] 孟富宣, 周军, 辛培尧, 唐军荣, 董娇, 付海辉, 周林涛, 徐世宏. 云南红皮梨 bHLH 转录因子的生物信息学分析. *基因组学与应用生物学*, 2013, 32(5): 652-659
Meng F X, Zhou J, Xin P Y, Tang J R, Dong J, Fu H H,

- Zhou L T, Xu S H. Bioinformatics analysis of bHLH transcription factor in Yunnan red skin pear (*Pyrus pyrifolia* Naki). *Genomics and Applied Biology*, 2013, 32(5): 652-659
- [38] 孙颖琦, 孟亚轩, 赵心月, 王凤霞, 瓮巧云, 赵治海, 刘颖慧, 袁进成. 谷子 bHLH 转录因子家族基因鉴定及生物信息学分析. *种子*, 2021, 40(12): 45-55
- Sun Y Q, Meng Y X, Zhao X Y, Wang F X, Weng Q Y, Zhao Z H, Liu Y H, Yuan J C. Identification and bioinformatics analysis of transcription factor family genes bHLH in Foxtail Millet. *Seed*, 2021, 40(12): 45-55
- [39] Wang N, Shu X C, Zhang F J, Wang Z. Transcriptome-wide characterization of bHLH transcription factor genes in *Lycoris radiata* and functional analysis of their response to MeJA. *Frontiers in Plant Science*, 2023, 13: 975530
- [40] Atchley W R, Zhao J P. Molecular architecture of the DNA-binding region and its relationship to classification of basic helix-loop-helix proteins. *Molecular Biology Evolution*, 2007, 24(1): 192-202
- [41] 郑斌, 文定青, 武红霞, 邹明宏, 刘恒, 王松标, 赵巧丽. 芒果果实 bHLH 家族转录因子的生物信息学分析. *热带作物学报*, 2019, 40(2): 289-299
- Zheng B, Wen D Q, Wu H X, Zou M H, Liu H, Wang S B, Zhao Q L. Bioinformatics analysis of bHLH transcription factor family in Mango (*Mangifera indica* Linn.). *Chinese Journal of Tropical Crops*, 2019, 40(2): 289-299
- [42] Aslam M, Jakada B H, Fakher B, Greaves J G, Niu X P, Su Z X, Cheng Y, Cao S J, Wang X M, Qin Y. Genome-wide study of pineapple (*Ananas comosus* L.) bHLH transcription factors indicates that cryptochrome-interacting bHLH2 (AcCIB2) participates in flowering time regulation and abiotic stress response. *BMC Genomics*, 2020, 21(1): 1-13