



主持人: 朱国鹏

火龙果 *HubHLH* 基因家族的全基因组 分析及其对冬季补光诱导开花 的表达响应

李佳雪^{1,2}, 丁一^{1,2}, 王猛^{1,2}, 李涛^{1,2},
郭攀阳², 刘成立², 韦双双², 黄家权^{1,2},
李洪立³, 胡文斌³, 汤华^{1,2}

(1. 海南大学 三亚南繁研究院/崖州湾种子实验室, 海南 三亚 572025; 2. 海南大学 热带作物学院, 海口 570228; 3. 中国热带农业科学院 热带作物品种资源研究所, 海口 571700)

摘要: 为了获得较完整的候选基因, 探讨 *HubHLH* 基因在火龙果 (*Hylocereus undatus*) 冬季补光诱导开花过程的表达响应, 对火龙果 *HubHLH* 基因家族进行全基因组分析。鉴定出 153 个 *HubHLH* 基因; 这些基因的编码蛋白含有 176 ~ 687 个氨基酸, 分子量大小为 19.28 ~ 74.44 kDa, 等电点 (pI) 为 4.81 ~ 9.88, 均为亲水蛋白, 亚细胞定位预测大多定位到细胞核。为鉴定 *HubHLH* 基因家族的同源性, 本研究将 153 个火龙果 *HubHLH* 和 120 个拟南芥 *AtbHLH* 蛋白进行系统发育比较分析。系统发育比较分析结果: 火龙果 *HubHLH* 基因家族成员被分为 12 个组, 25 个亚族; 对 *HubHLH* 基因家族的保守 motif、基因结构及在染色体的位置分布的分析结果表明, 同一亚族的基因具有相似的基序组成和基因结构。对 *HubHLH* 基因家族的内部复制事件的分析结果表明, 有 78 条片段复制被鉴定为片段重复基因, 说明片段复制是 *HubHLH* 基因家族的主要扩张力量。此外, 基于已测定的关于火龙果冬季补光诱导开花的 4 个时期转录组数据, 筛选到 59 个 *HubHLH* 基因在冬季补光诱导成花过程中有差异表达, 随后对这 59 个 *HubHLH* 基因进行 GO 功能富集, 发现它们在红光或远红光的反应、对光刺激的反应、有性生殖功能、对辐射的反应等功能上均有富集, 说明 *HubHLH* 基因家族可能在冬季补光诱导火龙果成花过程中起到了调控作用。

关键词: 火龙果; bHLH 基因家族; 全基因组; 补光诱导开花; 基因差异表达

中图分类号: S667.9; Q819 文献标志码: A 文章编号: 1674 - 7054(2024)02 - 0198 - 12

李佳雪, 丁一, 王猛, 等. 火龙果 *HubHLH* 基因家族的全基因组分析及其对冬季补光诱导开花的表达响应 [J]. 热带生物学报, 2024, 15(2): 198-209. doi: 10.15886/j.cnki.rdswwb.20220108

火龙果 (*Hylocereus undatus*) 原产中美洲热带地区, 是仙人掌科 (Cactaceae) 量天尺属 (*Hylocereus*) 或蛇鞭柱属 (*Selenicereus*) 植物, 喜光喜热。火龙果在我国的种植面积已达到 6.7 万公顷, 在热区经济中已占有重要地位^[1]。火龙果是长

日照作物, 在我国因冬季光照时长不足, 不能开花, 为此在生产上可采用夜晚补光的方式诱导火龙果开花结果。bHLH 转录因子蛋白, 包含 1 个碱性 DNA 结合区和 1 个螺旋—环—螺旋 (HLH) 区, 能通过与其顺式启动子元件相结合来调节基因表

收稿日期: 2022 - 12 - 06

修回日期: 2023 - 05 - 04

基金项目: 海南省自然科学基金高层次人才项目 (320RC487); 海南省重大科技计划项目 (ZDKJ2021014); 海南省重点研发项目 (ZDYF2022XDNY199)

第一作者: 李佳雪 (1996-), 女, 海南大学热带作物学院 2019 硕士研究生。E-mail: 252249803@qq.com

通信作者: 汤华 (1974-), 男, 教授。研究方向: 作物遗传育种与分子生物学。E-mail: thtiger@163.com

达,参与植物开花、种子发芽、矿质营养与非生物胁迫、应激反应、光信号、激素信号、光和植物激素之间的相互作用等多种生理过程^[2]。

bHLH 基因家族参与调控植物的开花调控,在拟南芥、番茄等植物中已有广泛的研究。在成花起始阶段,拟南芥的 FBH 蛋白(*bHLH* 类型转录因子),可与光周期信号途径的关键基因 *CO* (*CONSTANS*)启动子的 E-box 顺式元件结合,激活光周期开花基因的共转录^[3]。在花发育阶段,与 *CIB1* 同源的 *bHLH* 蛋白,可以作为激活子直接促进成花素基因 *FT* 的转录^[4]。*bHLH* 家族的光敏色素互作因子 PIFs 可以与光信号直接作用,PIF3 在光感受器信号网络中直接与光敏色素反应,PIF4 在高温下直接激活成花素基因 *FT*^[5]。*bHLH48* 和 *bHLH60* 与 PIF7 联合,可以作为赤霉素途径中的正调控因子,正调控 GA 介导的开花^[6]。*bHLH38/100/101* 会干扰 *CO* 的转录,调节 *FT* 的表达,进而影响拟南芥的开花^[7]。*bHLH* 基因家族成员 *SibHHLH22* 的过表达,会促进番茄早开花^[8]。在茉莉酸途径中,JA 激活的 *bHLH* 转录因子 MYC2, MYC3 和 MYC4 冗余调控拟南芥的开花^[9]。拟南芥的 *bHLH* 家族基因 *BEE1* 具有诱导开花起始的功能,*BES1-BEE1-FT* 可以调节光周期并与 *FT* 的激活因子 *CO* 互作来影响植物开花^[10]。*bHLH* 基因家族 *CRY2*、*CIB1* 可以响应蓝光,能将生物钟信号转化为开花信号,促进 *CO* 基因形成蛋白质复合物,诱导花的形成^[11]。除了调节开花时间外,*bHLH* 基因家族成员的 *SPT* 基因可以在花器官形成中起作用^[12]。

目前,*bHLH* 基因家族在拟南芥(*Arabidopsis thaliana*)^[13]、小麦(*Triticum aestivum* L.)^[14]、谷子(*Setaria italica* L.)^[15]、烟草(*Nicotiana tabacum* L.)^[16]、柚子(*Citrus maxima*)^[17] 等植物中进行了全基因组鉴定。火龙果作为一种重要的热带水果,有独特的开花习性,*bHLH* 家族基因可能发挥重要作用,但目前还未见火龙果 *HubHHLH* 基因家族相关系统分析的研究报道。火龙果(*Hylocereus undatus*)全基因组序列的发表,为鉴定火龙果 *HubHHLH* 基因家族提供了条件^[18]。本研究拟利用火龙果(*Hylocereus undatus*)基因组数据,对火龙果 *HubHHLH* 转录因子家族进行全基因组筛选,并

进行理化性质、系统进化、基因表达等方面的比较分析;利用火龙果补光诱导开花的转录组数据,研究 *HubHHLH* 基因家族在火龙果冬季补光诱导开花过程的表达响应,旨在解析火龙果冬季补光诱导开花的分子机制。

1 材料与方法

1.1 火龙果 *HubHHLH* 基因家族的全基因组筛选 分别在 <http://www.pitayagenomic.com/index.php> 网址和 Ensembl plants 核酸数据库(<https://plants.ensembl.org/index.html>)中下载火龙果和拟南芥的基因组及注释文件。从已报道的文章中查找拟南芥 *bHLH* 蛋白质序列作为种子序列,blast 后去掉重复的火龙果基因号,提取蛋白质序列,通过注释信息,筛选得到火龙果候选 *bHLH* 基因。从 PFAM 数据库(<http://pfam.xfam.org/>)中下载 *bHLH* (PF00-010)基因家族的隐马尔科夫模型,在 HMMER 3.0 软件中比对,得到火龙果候选 *bHLH* 基因。取 BLASTP 和 HMM 两种方式获得的 *bHLH* 基因交集,并通过 NCBI conserved Domains(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)数据库、SMART (<http://smart.ebl-heidelberg.de/>)数据库,进一步确定所有候选 *bHLH* 基因的结构域。根据 *bHLH* 家族成员在染色体的位置,将其依次编号为 *HubHHLH1* ~ *HubHHLH153* 用于后续分析。

1.2 火龙果 *HubHHLH* 基因家族的理化特征与亚细胞定位预测 使用 ExPASy 蛋白质组服务器上的 ProtParamtool (<https://web.expasy.org/protparam/>)工具^[19],预测火龙果 *bHLH* 基因家族蛋白的基本理化性质,用 Plant-mPLoc (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>)在线网站,进行亚细胞定位预测^[20]。

1.3 火龙果 *HubHHLH* 基因家族的蛋白序列比对及系统发育分析 用 MEGA 7.0 自带的 Clustal W 工具进行 153 个 *HubHHLHs* 的多序列比对,将比对结果上传到 WEBLOGO 在线程序(<http://weblogo.berkeley.edu>)分析其序列(图 1),展示出共同的序列标识^[21]。同时,用 TBTOOLS 软件的 Trimmer 程序,将获得的比对结果修剪,将火龙果和拟南芥的 *bHLH* 蛋白序列共同构建进化树,选择最大似然法,校验参数 Bootstrap 重复 1000 次,其他使用默

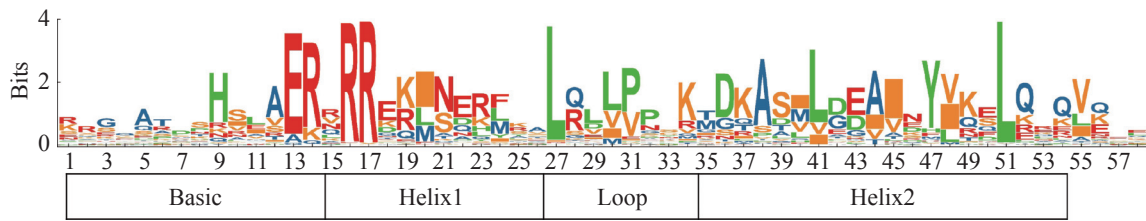


图1 火龙果 bHLH 保守结构域序列 LOGO

图中每个氨基酸的高度代表它在该横坐标位置的保守性,该位点的字母越高,表明此氨基酸的出现频率越高,可以突出序列比对中的保守位置,底部的方框描述了 bHLH 保守域中碱性区域、螺旋区域和环区的位置和边界。

认参数。进化树美化用 ITOL(<http://itol.embl.de/>) 网址绘制^[22]。

1.4 火龙果 *HubHHLH* 基因家族的基因结构与 motif 预测 根据火龙果的基因组注释,提取 *HubHHLH* 的基因结构,采用 MEME(<http://meme-suite.org/tools/meme>) 程序,设置默认参数,预测 *HubHHLH* 蛋白的保守 motif。使用 TBTOOLS 的 Gene Structure view(Advanced)程序展示 153 个 *HubHHLHs* 基因的进化树、保守 motif 以及基因结构的组合图^[23]。

1.5 火龙果 *HubHHLH* 基因家族的染色体定位 基于火龙果基因组,使用 TBTOOLS 提取 153 个 *HubHHLHs* 基因在染色体的物理位置,使用 Gene Density Profile 程序提取染色体的基因密度信息,在 TBTOOLS 的 Gene Location Visualize from GFF 程序中一起作图展示。

1.6 火龙果 *HubHHLH* 基因家族的共线性分析 采用多重共线性扫描工具包(MCScanX)分析基因重复事件,使用默认参数,利用 TBTOOLS 的 Fasta Stats 功能得到火龙果染色体长度信息,利用 TBtools 的 Advanced Circos 程序进行 *HubHHLH* 基因家族共线性的可视化。

1.7 火龙果 *HubHHLH* 基因对冬季补光诱导开花的基因表达响应和 GO、KEGG 分析 利用笔者所在的课题组前期研究的高质量火龙果转录组^[24] 进行分析,筛选出全部注释到 Nr、Swissprot、KOG (evalue<0.00001)、KEGG 蛋白质数据库的 *bHLH* 基因,并提取蛋白质序列,在 SMART 数据库中对,进一步确定候选基因。将候选基因的蛋白质序列在基因组数据库进行本地 BLAST 比对,实现转录组和基因组中的 ID 转换。RNA-Seq 数据中使用 RPKM 法计算基因表达量,可直接用于比较不同样品间的基因表达差异。使用 TBTOOLS 的

Heatmap 工具,展示 *bHLH* 基因在 NL、L0、L1、L2(NL: 火龙果未补光未开花, L0: 补光了未开花, L1: 补光后刚现蕾, L2: 补光后现蕾一周) 4 个时期的表达变化,并设置筛选条件为 FDR<0.05 且 $|\log_2FC|>1$, 展示差异基因的表达模式。利用 eggNOG-mapper 平台,得到火龙果基因功能注释信息,利用 TBTOOLS 整理结果,进行 GO 和 KEGG 富集分析,并以条状图进行可视化。

2 结果与分析

2.1 火龙果 *HubHHLH* 基因的蛋白质理化性质分析 火龙果全基因组筛选得到 153 个 *HubHHLH* 基因家族成员,占总基因的比率约为 0.55%。对其蛋白质序列在 ProtParam tool 上分析,得到蛋白序列长度最小的 *HubHHLH143* 有 176 个氨基酸,长度最大的 *HubHHLH48* 有 687 个氨基酸,蛋白质的分子量范围为 19.28 ~ 74.44 kDa,理论等电点范围为 4.81(*HubHHLH68*)至 9.88(*HubHHLH143*),平均值为 6.86。所有的 *HubHHLHs* 全部为亲水蛋白,总平均亲水性全部为负值。除了 *HubHHLH143* 外,其他基因的不稳定系数均大于 40,为不稳定蛋白。亚细胞定位预测显示,有 140 个基因定位到细胞核,7 个位于细胞质中,6 个位于叶绿体中,符合转录因子主要在细胞核中发挥作用的特点。

2.2 火龙果 *HubHHLH* 蛋白序列比对及系统发育分析 将获得的 153 个火龙果 *HubHHLH* 家族成员进行同源序列比对后,在 WEBLOGO 网址以图形的形式展示 bHLH 保守结构域氨基酸的序列,可以得到所鉴定的 *HubHHLH* 家族成员,具有明显保守的火龙果 *HubHHLH* 结构域,且 *HubHHLH* 的氨基酸数量约为 60。

基于 153 个火龙果 bHLH 和 120 个拟南芥 bHLH 的蛋白序列共同构建有根系统发育树

(图 2), 参考拟南芥中的 bHLH 蛋白亚家族分类结果, 发现 273 个 bHLH 蛋白被分为 25 个亚家族, 除了拟南芥中一个单独的组, 其他拟南芥和火龙果中的 bHLH 共同聚集到 12 个组别 (I - XII), 火龙

果与拟南芥的蛋白分类几乎一致, 说明火龙果 *HubHLH* 基因在长期进化的过程中, 这些蛋白质没有遗失。

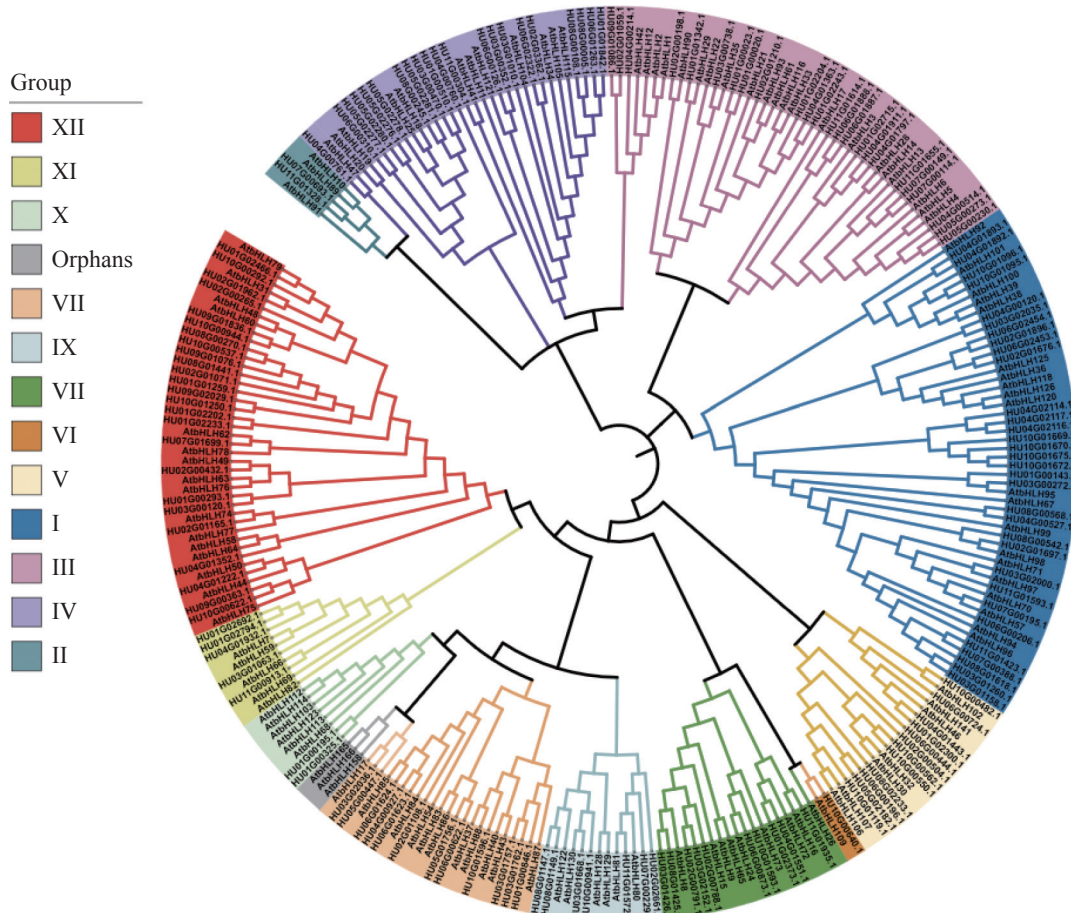


图 2 火龙果和拟南芥 bHLH 蛋白的系统发育树和亚科分类
分别以不同的颜色范围, 展示火龙果和拟南芥 bHLH 蛋白的系统发育树。

2.3 火龙果 *HubHLH* 基因家族的基因结构及 motif 分析 基于 153 个 *HubHLH* 蛋白序列构建的矩形系统发育树(图 3-A), 系统发育树在外结点的自展值几乎都大于 70, 所以分支比较可靠, *HubHLH* 基因被分成 12 组 (I - XII), 包含 25 个亚家族。其中, Group I 中包含的火龙果 bHLH 基因家族成员数最多, 有 32 个, 占基因家族总数的 20.92%, 其次为 Group XII 有 25 个 bHLH (16.34%), 第三为 Group III, 含有 24 个 bHLH (15.69%), Group VI 含有的 bHLH 成员最少, 仅有 1 个 (0.66%)。其他组别 Group II 和 Group X 均只含有 2 个, Group IV 有 21 个, Group V 含有 12 个, Group VII 含有 10 个, Group VIII 有 12 个, Group IX 有 7 个, Group XI 有

5 个。这与拟南芥中 Group VI、Group II 含有 bHLH 基因较少, Group I、Group III、Group XII 含有 bHLH 较多的结果一致^[13]。

使用 TBtools 展示 MEME 在线软件预测的 153 个 *HubHLH* 蛋白的保守 motif(图 3-B), 结果表明, 所有的 *HubHLH* 家族成员都包含保守 Motif 1: AERRRREKJNERFKALRSLVP 和 Motif 2: NVKKTDKASILDEAINYVKELQRQVQELS; 但不同 *HubHLH* 成员所包含的保守 motif 数目及种类是存在差异的, 含有 motif3 ~ motif10 的基因数量依次为 48、10、15、18、10、45、49、15。本研究发现, 同一组中的 *HubHLH* 成员具有相似的 motif 组成, 有些 motif 仅存在于特定的亚家族中,

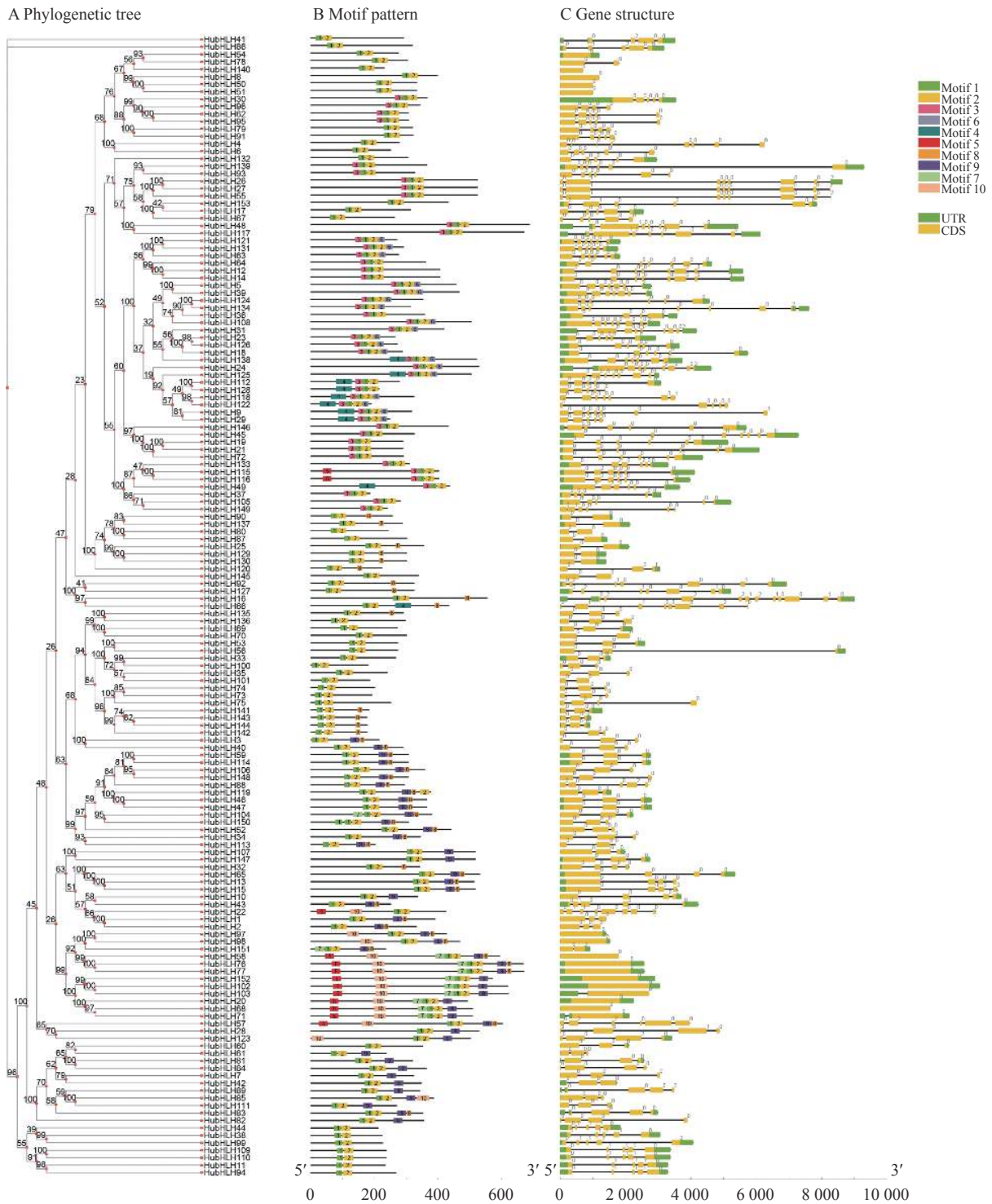


图 3 *HubHLHs* 的 motif 分布和基因结构图

图 3-A 是 *HubHLH* 蛋白序列构建的矩形系统发育树, 每个节点上的数值代表 1000 个重复的自展值; 图 3-B 的彩色矩形框代表 *HubHLH* 蛋白的保守 motif; 图 3-C 中黄色矩形和黑色线条分别表示外显子(CDS, coding sequence)和内含子(introns), 绿色的矩形展示了非翻译区域(UTR, untranslated region), 横坐标的黑线表示相对蛋白质长度。

例如, motif6 只存在Ⅱ成员中, motif5 除了 *HubHLH115* 和 *HubHLH116* 外, 剩余全部在第Ⅲ组。通过比较基因组 DNA 序列, 并根据系统发育分类的顺序, 显示所有 *HubHLH* 的内含子和外显子结

构, 通过比较其数量和位置, 发现 153 个 *HubHLH* 基因具有不同数量的外显子, 从 1 到 11 不等(图 3-C)。同一亚家族的基因外显子数量及位置相近, 同一分类的基因结构也极为相似, 例如

VIIIb 成员都只包含 1 个外显子, IVa 含有 3 或 4 个外显子, IVc 中都有 5 个外显子和 4 个内含子, 这与某些亚家族具有特定的功能有关。

2.4 火龙果 *HubHLH* 基因家族的染色体定位分析 基于火龙果基因组数据库, 创建 *HubHLH* 基因在染色体的分布示意图(图 4), 分布映射结果表明, 大多数 *HubHLH* 基因分布在染色体的末端, 染色体两端的基因密度较大。此外, 153 个 *HubHLH* 基因在染色体 1 至 11 上的分布不均

匀, Chr1 含有最多的 *HubHLH* 基因 (21 个, 约 13.73%), 其次是 Chr4(20 个, 约 13.07%), 然后是 Chr10(19 个, 12.42%), 而 Chr9 含有最少的基因 (5 个, 约 3.27%), Chr2 和 Chr3 都含有 17 个(约 11.11%), Chr6、Chr8、Chr5、Chr11、Chr7 从多到少依次包含 16 个(10.46%)、12 个(7.84%)、10 个(6.54%)、9 个(5.89%)、7 个(4.58%) *HubHLH* 基因。以不同的颜色展示 *HubHLH* 所属的 12 个组别, 显然, 除了少数的基因簇, 每个组别在染色体

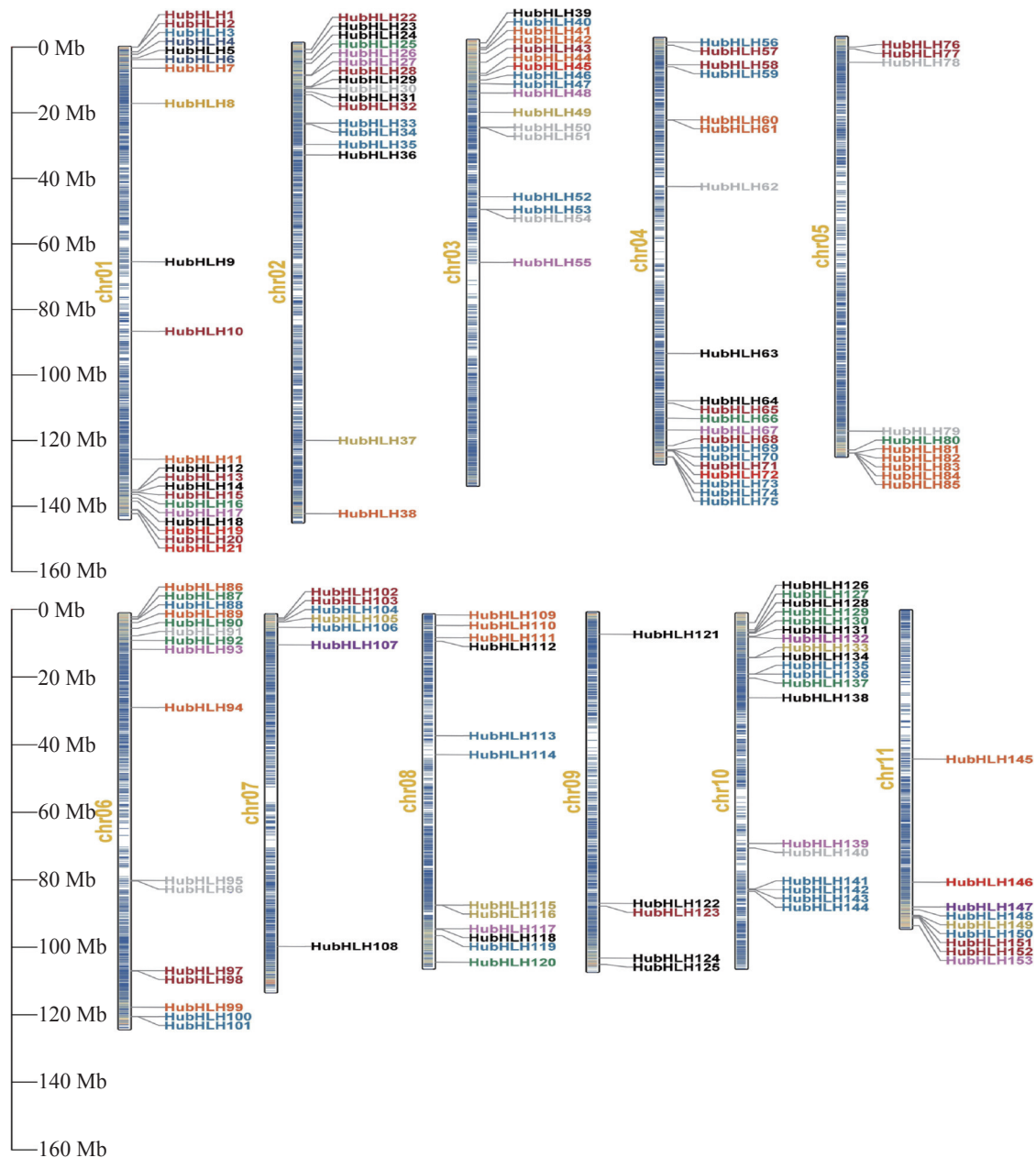


图 4 *HubHLHs* 基因的染色体定位

根据其在火龙果 Chr1 至 Chr11 上, 从上到下的物理位置, 依次命名为 *HubHLH1* ~ *HubHLH153*, 图 4 中左边的刻度尺表示染色体长度, 单位为 Mb。右侧细的垂直条代表火龙果的染色体, 染色体编号显示在每条染色体的左侧, 内部颜色代表染色体上的基因密度, 颜色越白, 表示该位置的基因密度越小, 颜色越红, 表示该位置的基因密度越大。

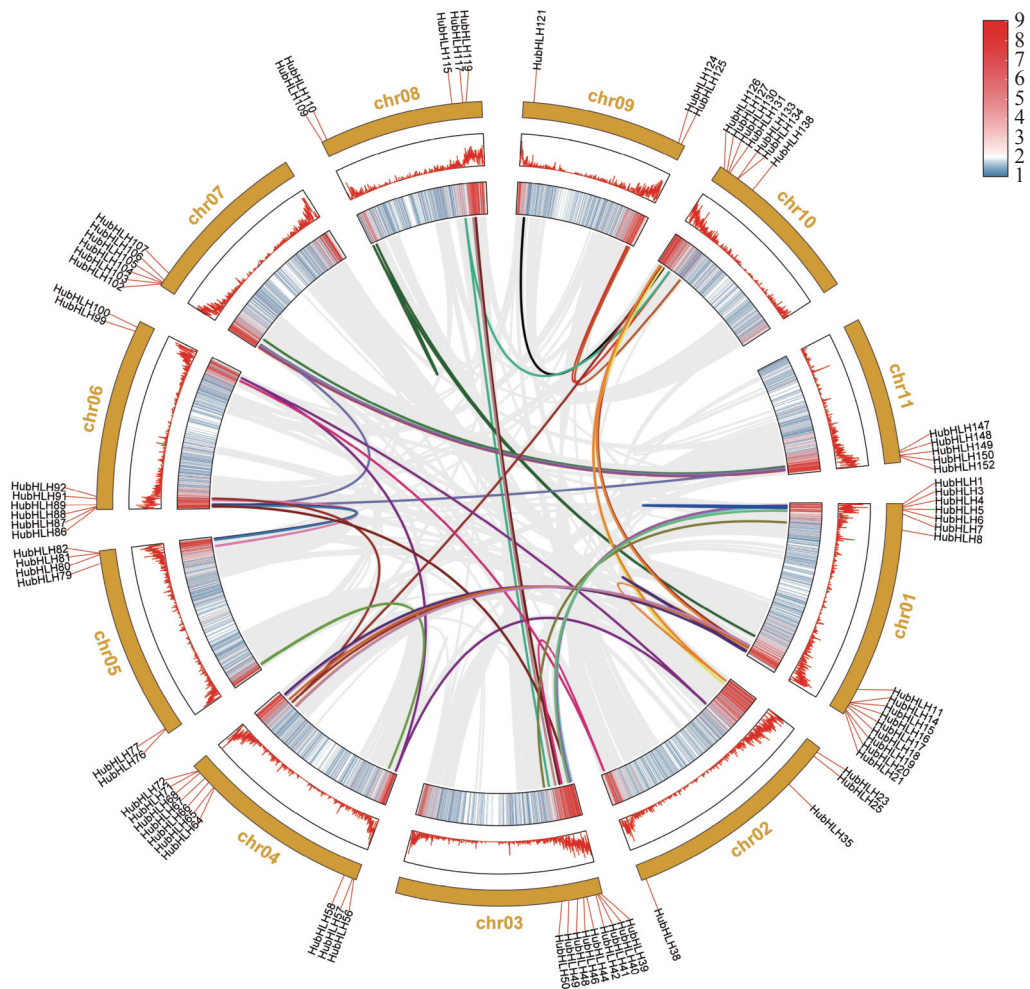


图5 火龙果 *HubHLHs* 基因家族的基因复制和同步分析

HubHLH 基因的染色体位置由圆圈上的短红线连接,背景的灰色线表示火龙果基因组中的所有共线性区块,彩色线表示片段重复的 *HubHLH* 基因对,相同片段重复基因对的两个基因以相同的颜色标记。并将染色体的基因密度分别以热图和线图的形式在外圈展示。

的位置没有特殊的指向,且不同 *HubHLH* 亚家族的一些成员位于同一小染色体区域内,而同一亚家族的某些成员分布在不同的染色体区域,表明 *HubHLH* 基因广泛分布于火龙果基因组中。

2.5 火龙果 *HubHLH* 基因家族的共线性分析

通过 MCScanX 分析火龙果的基因组,共线性结果(图5)显示,在11条染色体上,火龙果 *bHLH* 基因家族,在全基因组水平存在大量的片段重复(segmental duplication, SD)基因对,这将导致火龙果 *bHLH* 基因家族数量上的扩增,图5中共有78个 *HubHLH* 基因聚集在49个片段复制事件。由于基因家族和基因组的扩展是植物进化过程的一部分,说明片段重复事件是驱动 *HubHLH* 基因家族进化的主要力量,这也同样促进了火龙果表

型的多样化和新功能的产生。本研究还发现了7对串联重复基因对(*HubHLH69/HubHLH70*, *HubHLH74/75*, *HubHLH84/85*, *HubHLH95/96*, *HubHLH97/98*, *HubHLH135/HubHLH136*, *HubHLH141/HubHLH142*),可能通过串联复制(tandem duplication, TD)产生,每对都聚集在同一个亚家族中,产生这种现象的原因可能有基因加倍、染色体片段的交换等。

2.6 火龙果冬季补光诱导开花过程的 *HubHLH* 基因表达及 GO、KEGG 分析

对火龙果补光诱导开花的转录组数据进行分析,共注释到142个 *bHLH* 基因,与基因组数据库比对得到87个 *HubHLH* 基因,依据 $FDR < 0.05$ 且 $|\log_2FC| > 1$,挑选差异表达基因,筛选到59个差异表达基因,以热

图展示差异表达情况(图 6)。重点关注补光诱导前后 bHLH 基因的变化,在 L0 至 L1 时期,差异表达 *HubHLH* 基因有 6 个显著下调, 34 个显著上调; 在 NL 和 L1 时期,有 8 个差异表达 *HubHLH* 基因下调, 33 个上调; 在 L1 至 L2 花发育过程, *HubHLH* 差异基因有 15 个下调, 7 个上调。总之, *HubHLH* 基因在补光前后呈现较为明显的上下调趋势, 这些基因在冬季补光诱导开花中可能发挥了作用。

为了解 *HubHLH* 差异表达基因的偏好功能, 对它们进行基于超几何分布的 GO 和 KEGG 分析(图 7)。重点关注生物进程, 发现 *HubHLH* 差异基因, 在红光或远红光的反应、对光刺激的反应、有性生殖功能、对辐射的反应等功能上均有富集。在 KEGG 分析中, 主要富集在转录因子, 信号转导, 环境信息处理, 植物激素信号转导功能通路上。

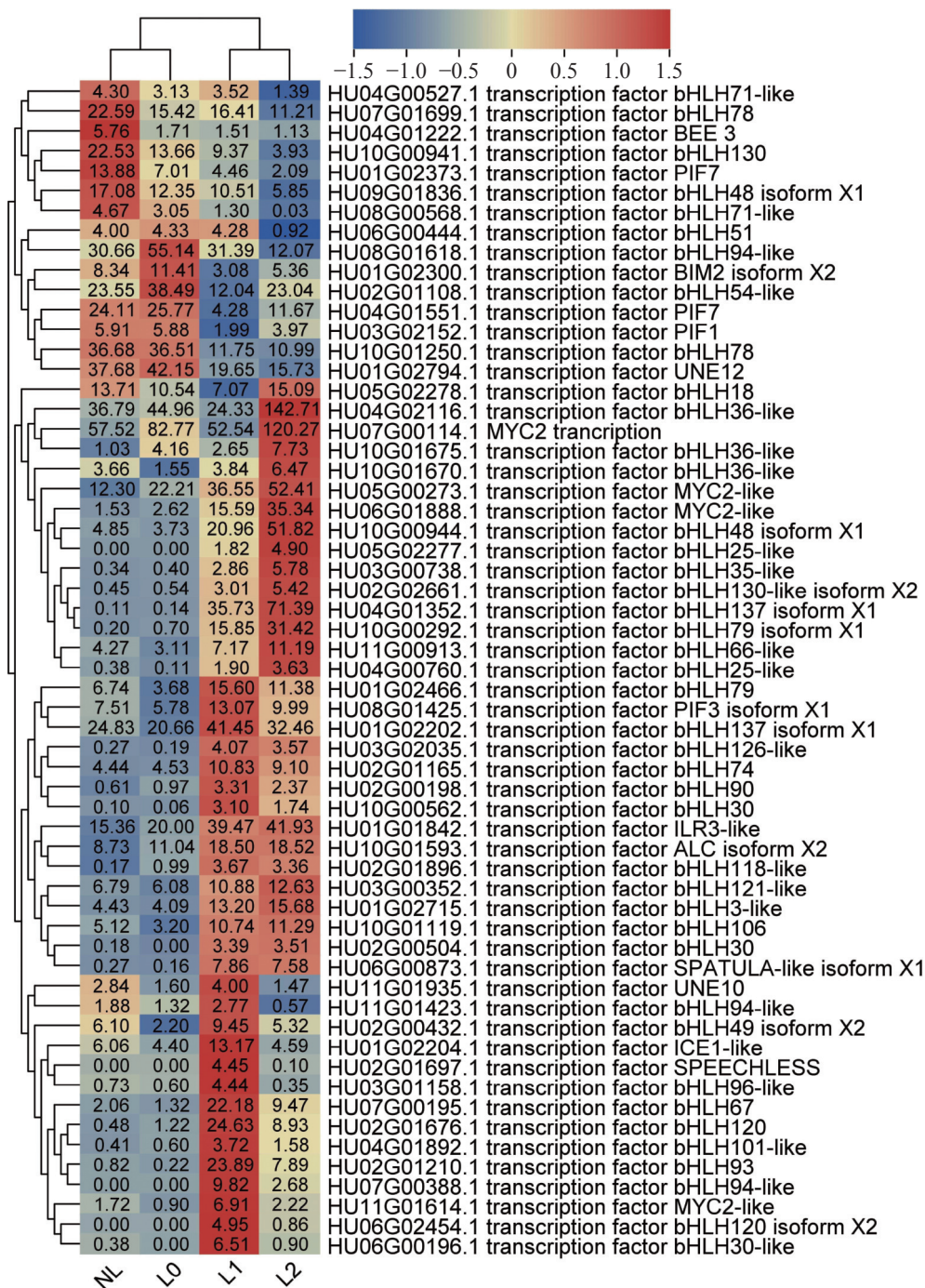


图 6 转录组中 *HubHLH* 基因的表达情况

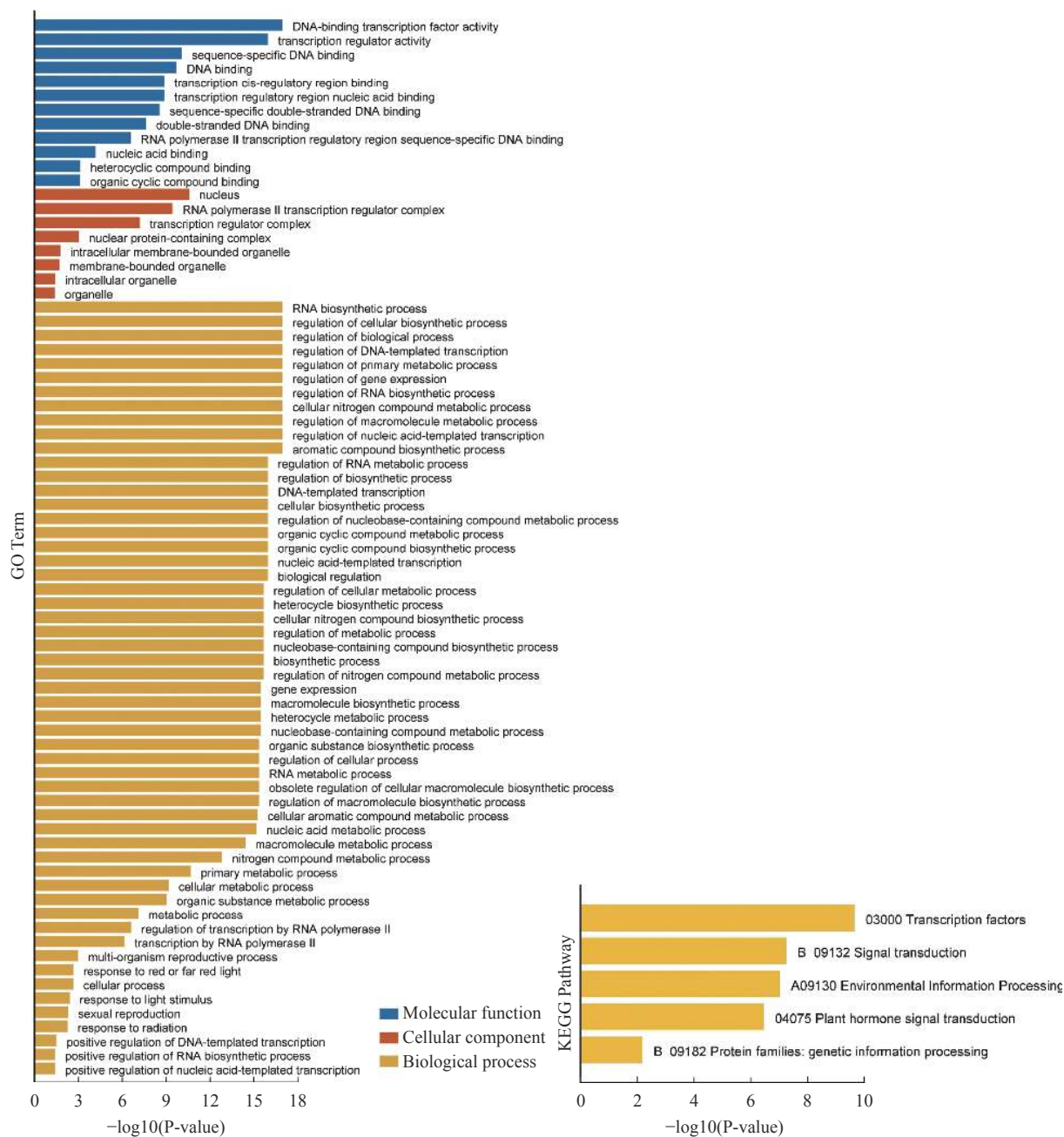


图7 差异 *HubHLH* 基因的 GO 和 KEGG 分析条状图

3 讨论

3.1 火龙果与拟南芥 *HubHLH* 基因家族的系统发育关系 bHLH 作为植物中第二大类转录因子, 近几年在菠萝 (*Ananas comosus* L.)^[25]、香蕉 (*Musa acuminata* L.)^[26]、木薯 (*Manihot esculenta crantz*)^[27] 等热带植物中实现了全基因组鉴定, 分别鉴定出 121、259、148 个 *bHLH* 基因。火龙果全基因组序列的发表, 为进一步检索、克隆和验证火龙果的 *bHLH* 基因提供了条件。同时, 冬季补光诱导开花过程是严密而复杂的过程, bHLH 基因被证实

参与了植物生长发育的很多过程。因此对火龙果进行 bHLH 基因家族的全基因组筛选和分析, 为后续它们对冬季补光诱导开花的响应过程中的功能研究很有必要。本研究首次对火龙果 bHLH 基因家族进行了全基因组鉴定, 得到 153 个火龙果 *HubHLH* 基因, 对它们的理化性质全面预测, 得到火龙果 *HubHLH* 的分子量、氨基酸长度、等电点等理化性质, 与其他植物中的 bHLH 差异不大, 亚细胞定位预测主要分布在细胞核。

目前, 植物中 *bHLH* 基因还未有特别明确的

分类标准, 通常选择与模式植物或亲缘关系相近的植物, 共同构建有根系统发育树进行分类, 本研究选择长日照植物的拟南芥, 将它的 120 个 bHLH 蛋白序列和鉴定到的 153 个火龙果 *HubHLH* 蛋白序列共同构建有根系统发育树, 通过与拟南芥的系统发育关系比较, 火龙果 bHLH 分类结果和拟南芥几乎一致, 都得到了 12 个主要的组别 (I - XII)^[28]。进行火龙果 *HubHLH* 基因家族的保守 motif 和基因结构分析, 153 个 *HubHLH* 基因都含有 motif 1 和 motif 2, 同一分组的基因结构相似, 证明了系统发育分析结果的可靠性。motif 位点的识别有助于鉴定其家族同源性和预测蛋白质的功能, 而系统发育树显示, 一些火龙果中的 *HubHLH* 与拟南芥的 bHLH 基因可能同源, 并具有相似功能, 因此可以通过发现聚集到同一亚家族的拟南芥中功能较为明确的 bHLH 蛋白, 发现火龙果中 bHLH 的可能功能。例如在亚家族 IIIe 中的成员与茉莉酸介导的开花相关, 亚家族 VII 的成员多与光信号有关^[29], 成员较多的 XII 亚家族中有些调节光周期开花^[30], 有些与光和植物激素之间的相互作用有关^[31]。

3.2 火龙果 *HubHLH* 基因的差异表达及 GO、KEGG 富集分析

基于转录组数据展示 *HubHLH* 基因家族成员在火龙果冬季夜晚补光诱导成花过程中 NL、L0、L1、L2 四个时期的差异表达情况, 其中有注释到 JA 途径中调节开花时间的 *MYC2* 基因, 在补光后表达含量急速增加; 还有响应光和温度, 在花形态发生过程中起作用的光稳定 *SPATULA* 基因, 也在 L0-L1 阶段表达出上调趋势, 其他 *HubHLH* 基因对补光诱导也有不同程度的响应。另外, 根据 *HubHLH* 差异基因的 GO 富集结果, 发现在有性生殖, 即火龙果成花过程, 对光反应和光刺激, 即补光过程, 都得到了功能富集。*HubHLH* 基因家族可能在冬季补光诱导火龙果成花过程起到了作用。总之, 开花时间的调控是严格而复杂的, 在冬季通过补光诱导火龙果开花具有重要的生产实践价值, 本研究利用火龙果基因组信息和转录组信息, 对火龙果 *HubHLH* 基因家族进行了全面分析研究, 为火龙果冬季补光诱导开花, 提供了许多潜在的候选基因。

参考文献:

- [1] 刘成立, 王猛, 郭攀阳, 等. 火龙果花和果实的动态发育规律研究[J]. *海南大学学报(自然科学版)*, 2020, 38(2): 147 - 152.
- [2] HAO Y, ZONG X, REN P, et al. Basic helix-loop-helix (bHLH) transcription factors regulate a wide range of functions in *Arabidopsis*[J]. *Int J Mol Sci*, 2021, 22(13): 7152.
- [3] ITO S, SONG Y H, JOSEPHSON-DAY A R, et al. FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator *CONSTANS* in *Arabidopsis*[J]. *Proc Natl Acad Sci U S A*, 2012, 109(9): 3582 - 3587.
- [4] LIU H, YU X, LI K, et al. Photoexcited *CRY2* interacts with *CIB1* to regulate transcription and floral initiation in *Arabidopsis*[J]. *Science*, 2008, 322(5907): 1535 - 1539.
- [5] KUMAR S V, LUCYSHYN D, JAEGER K E, et al. Transcription factor PIF4 controls the thermosensory activation of flowering[J]. *Nature*, 2012, 484(7393): 242 - 245.
- [6] LI Y, WANG H, LI X, et al. Two DELLA-interacting proteins bHLH48 and bHLH60 regulate flowering under long-day conditions in *Arabidopsis thaliana*[J]. *J Exp Bot*, 2017, 68(11): 2757 - 2767.
- [7] CHEN W, ZHAO L, LIU L, et al. Iron deficiency-induced transcription factors bHLH38/100/101 negatively modulate flowering time in *Arabidopsis thaliana*[J]. *Plant Sci*, 2021, 308: 110929.
- [8] WASEEM M, LI N, SU D, et al. Overexpression of a basic helix-loop-helix transcription factor gene, *SlbHLH22*, promotes early flowering and accelerates fruit ripening in tomato (*Solanum lycopersicum* L.)[J]. *Planta*, 2019, 250(1): 173 - 185.
- [9] WANG H, LI Y, PAN J, et al. The bHLH Transcription Factors MYC2, MYC3, and MYC4 are required for jasmonate-mediated inhibition of flowering in *Arabidopsis*[J]. *Mol Plant*, 2017, 10(11): 1461 - 1464.
- [10] WANG F, GAO Y, LIU Y, et al. BES1-regulated *BEE1* controls photoperiodic flowering downstream of blue light signaling pathway in *Arabidopsis*[J]. *New Phytol*, 2019, 223(3): 1407 - 1419.
- [11] LIU Y, LI X, MA D, et al. *CIB1* and *CO* interact to mediate *CRY2*-dependent regulation of flowering[J]. *EMBO Rep*, 2018, 19(10): e45762.
- [12] REYMOND M C, BRUNOUD G, CHAUVET A, et al. A light-regulated genetic module was recruited to carpel development in *Arabidopsis* following a structural change to *SPATULA*[J]. *Plant Cell*, 2012, 24(7): 2812 - 2825.
- [13] HEIM M A, JAKOBY M, WERBER M, et al. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional

- diversity[J]. *Mol Biol Evol*, 2003, 20(5): 735 – 747.
- [14] GUO X J, WANG J R. Global identification, structural analysis and expression characterization of bHLH transcription factors in wheat[J]. *BMC Plant Biol*, 2017, 17(1): 90.
- [15] FAN Y, LAI D, YANG H, et al. Genome-wide identification and expression analysis of the bHLH transcription factor family and its response to abiotic stress in foxtail millet (*Setaria italica* L.)[J]. *BMC Genomics*, 2021, 22(1): 778.
- [16] BANO N, PATEL P, CHAKRABARTY D, et al. Genome-wide identification, phylogeny, and expression analysis of the bHLH gene family in tobacco (*Nicotiana tabacum*) [J]. *Physiol Mol Biol Plants*, 2021, 27(8): 1747 – 1764.
- [17] ZHANG X Y, QIU J Y, HUI Q L, et al. Systematic analysis of the basic/helix-loop-helix (bHLH) transcription factor family in pummelo (*Citrus grandis*) and identification of the key members involved in the response to iron deficiency[J]. *BMC Genomics*, 2020, 21(1): 233.
- [18] CHEN J Y, XIE F F, CUI Y Z, et al. A chromosome-scale genome sequence of pitaya (*Hylocereus undatus*) provides novel insights into the genome evolution and regulation of betalain biosynthesis[J]. *Horticulture Research*, 2021, 8: 164.
- [19] DUVAUD S, GABELLA C, LISACEK F, et al. ExPasy, the Swiss Bioinformatics Resource Portal, as designed by its users[J]. *Nucleic Acids Research*, 2021, 49(W1): W216 – W227.
- [20] CHOU K C, SHEN H B. Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization[J]. *PLoS One*, 2010, 5(6): e11335.
- [21] CROOKS G E, HON G, CHANDONIA J M, et al. Weblogo: a sequence logo generator[J]. *Genome Research*, 2004, 14(6): 1188 – 1190.
- [22] LETUNIC I, BORK P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation[J]. *Bioinformatics*, 2007, 23(1): 127 – 128.
- [23] CHEN C, CHEN H, ZHANG Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194 – 1202.
- [24] XIONG R, LIU C, XU M, et al. Transcriptomic analysis of flower induction for long-day pitaya by supplementary lighting in short-day winter season[J]. *BMC Genomics*, 2020, 21(1): 329.
- [25] MOHAMMAD A, HASSAN J B, BEENISH F, et al. 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[J]. *BMC Genomics*, 2020, 21(1): 735.
- [26] WANG Z, JIA C, WANG J, et al. Genome-wide analysis of basic *Helix*-loop-*Helix* transcription factors to elucidate candidate genes related to fruit ripening and stress in banana (*Musa acuminata* L. AAA group, cv. Cavendish) [J]. *Front Plant Sci*, 2020, 11: 650.
- [27] AN F, XIAO X, CHEN T, et al. Systematic analysis of bHLH transcription factors in cassava uncovers their roles in postharvest physiological deterioration and cyanogenic glycosides biosynthesis[J]. *Frontiers in Plant Science*, 2022, 13: 901128.
- [28] TOLEDO-ORTIZ G, HUQ E, QUAIL P H. The *Arabidopsis* basic/*Helix*-loop-*Helix* transcription factor family[J]. *The Plant Cell*, 2003, 15(8): 1749 – 1770.
- [29] HAO Y, OH E, CHOI G, et al. Interactions between HLH and bHLH factors modulate light-regulated plant development[J]. *Molecular Plant*, 2012, 5(3): 688 – 697.
- [30] SHARMA N, XIN R, KIM D H, et al. NO FLOWERING IN SHORT DAY (NFL) is a bHLH transcription factor that promotes flowering specifically under short-day in *Arabidopsis* [J]. *Development*, 2016, 143(4): 682 – 690.
- [31] WU M, UPRETI S, YAN A, et al. *SPATULA* regulates floral transition and photomorphogenesis in a PHYTOCHROME B-dependent manner in *Arabidopsis* [J]. *Biochemical and Biophysical Research Communications*, 2018, 503(4): 2380 – 2385.

Whole-genome analysis of the *HubHLH* gene family in Pitaya and its differential expression in response to supplementary light-induced flowering in winter

LI Jiaxue^{1,2}, DING Yi^{1,2}, WANG Meng^{1,2}, LI Tao^{1,2}, GUO Panyang², LIU Chengli²,
WEI Shuangshuang², HUANG Jiaquan^{1,2}, LI Hongli³, HU Wenbin³, TANG Hua^{1,2}

(1. Sanya Nanfan Research Institute of Hainan University/ Hainan Yazhou Bay Seed Laboratory, Sanya, Hainan 572025, China; 2. College of Tropical Crops, Hainan University, Haikou, Hainan 570228, China; 3. Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, Hainan 571700, China)

Abstract: In order to obtain relative complete candidate genes, the *HubHLH* gene expression response in Pitaya during the flowering process induced by supplementary light in winter was investigated, and the whole genome of *HubHLH* gene family in Pitaya was analyzed. There were 153 pitaya *HubHLH* genes identified, and their encoding proteins contained 176-687 amino acids with their molecular weight being 19.28 – 74.44 kDa, pI 4.81-9.88. These proteins were all hydrophilic. The prediction of subcellular localization showed that most of the proteins were localized in the nucleus. The pitaya *HubHLH* proteins were compared with 120 *Arabidopsis AtbHLH* proteins for phylogenetic analysis. Phylogenetic analysis showed that the pitaya *HubHLH* proteins were divided into 12 groups and 25 subfamilies. Analysis of conserved motif, gene structure and location distribution in chromosomes showed that genes of the same subfamily had similar motif composition and gene structure. Analysis of the internal replication events of the pitaya *HubHLH* gene family found that 78 genes were identified as fragment repeating genes, indicating that fragment replication was the main expansion force of the pitaya *HubHLH* gene family. In addition, based on the transcriptomic data of the four phases of supplementary light-induced flowering of Pitaya in winter, 59 *HubHLH* genes were differentially expressed during the process of supplementary light induced flowering in winter. GO function enrichment showed that they were enriched in response to red or far-red light, light stimulation, sexual reproduction function, and radiation. The *HubHLH* gene family may play a regulatory role in the flowering process of Pitaya induced by supplementary light in winter. In this study, we conducted a comprehensive genome analysis of the *HubHLH* gene family in Pitaya, obtained relatively complete candidate genes, and preliminarily investigated the expression response of *HubHLH* genes in the winter flowering process induced by supplementary light in Pitaya.

Keywords: pitaya; *HubHLH* gene family; whole genome analysis; supplementary light-induced flowering; differential gene expression

(责任编辑:叶 静)