

郑玲,代琪芳,陈雪冬. 森林草莓 bHLH 家族基因的鉴定与分析[J]. 江苏农业科学,2022,50(2):47-52.  
doi:10.15889/j.issn.1002-1302.2022.02.008

# 森林草莓 bHLH 家族基因的鉴定与分析

郑玲,代琪芳,陈雪冬

(洛阳师范学院生命科学学院,河南洛阳 471934)

**摘要:** bHLH 转录因子在植物的生长发育、生物合成及胁迫反应等方面起着重要作用。利用生物信息学方法对森林草莓全基因组序列进行 bHLH 家族的搜索和鉴定,共鉴定出 112 个森林草莓的 bHLH 基因,这些基因分别被定位在 7 条染色体上,但分布不均匀,在 5 号染色体上分布得最多,3 号染色体上分布得最少。进化树结果显示,森林草莓的 bHLH 家族可以进一步被分为 12 个亚家族,其中第 I、III 亚族的成员较多。森林草莓 bHLH 家族基因的表达模式分析结果显示,有 73 个 *FvbHLH* 基因与花的发育相关,有 28 个 *FvbHLH* 基因与果实的发育相关,其中 *FvbHLH97/36/55/39/42/52/26/44/68/54* 对森林草莓花的整个发育过程都尤为重要,而 *FvbHLH57/59/66/78/100/15/17/46* 则在果实发育中起到主要的调控作用。

**关键词:** 森林草莓; bHLH; 进化树; 花发育; 果实发育

**中图分类号:** S668.401 **文献标志码:** A **文章编号:** 1002-1302(2022)02-0047-06

转录因子又称为反式作用因子,是一类能够结合在顺式作用元件上调控基因表达的 DNA 结合蛋白<sup>[1]</sup>。目前,人们已经从植物中鉴定出很多转录因子,其中 bHLH(碱性螺旋-环-螺旋)是第二大类转录因子<sup>[2]</sup>。bHLH 转录因子因含有高度保守的 bHLH 结构域而被命名, bHLH 结构域约含有 60 个保守氨基酸序列,由 basic 区域、HLH(helix-loop-helix)区域组成<sup>[3]</sup>。位于 N 端的 basic 区域是 DNA 结合的碱性区域,约含有 15 个氨基酸,其中有 6 个是碱性氨基酸残基; HLH 区域位于 C 末端,主要由疏水氨基酸组成,并参与形成同源或异源二聚体<sup>[4]</sup>。

bHLH 转录因子在真核生物中最为广泛存在,其功能也具有多样性。在酵母和其他单细胞真核生物中, bHLH 参与染色体分离和代谢调控等过程<sup>[5]</sup>。 bHLH 转录因子可参与植物的物质合成和信号转导过程<sup>[6]</sup>,并在植物生长发育<sup>[7]</sup>及抗旱<sup>[8]</sup>、耐盐<sup>[9]</sup>等胁迫应答反应中起着关键作用。植物中最早被发现的 bHLH 转录因子(即玉米中由 R 基因编码的 Lc 蛋白)的主要功能为参与花青素的生物合成<sup>[10]</sup>。随后,拟南芥的 *AtbHLH1/2/42*、水稻的 *Os bHLH013/016* 等基因都被证明与花青素的合成调

控有关<sup>[11-13]</sup>,拟南芥 *AtbHLH40* 则与赤霉素的合成转导有关<sup>[14]</sup>。在响应植物胁迫方面,拟南芥 *AtbHLH112* 是盐、干旱及渗透胁迫的正调控因子<sup>[15]</sup>; *AtbHLH116* 基因是耐寒性的调节子,能够提高植物的抗冻能力<sup>[16]</sup>。在植物生长发育方面,拟南芥的 *AtbHLH9*、*AtbHLH65* 基因均是调控植物生长发育的关键信号枢纽<sup>[17]</sup>; *AtbHLH95* 基因主要在胚乳中表达,控制胚胎发育,从而导致植物生长迟缓<sup>[18]</sup>; 拟南芥的 *AtbHLH21* 和水稻的 *Os bHLH164* 对花粉粒的形成及绒毡层细胞的发育有着关键的调控作用<sup>[19-20]</sup>。

森林草莓(*Fragaria vesca*)也称野草莓,是蔷薇科草莓属的多年生草本植物,属于二倍体植物<sup>[21]</sup>,具有相当高的食用价值和药用价值。目前,国内外在分子生物学方面对拟南芥、水稻等植物 bHLH 家族的研究已经较为成熟,然而还没有关于森林草莓 bHLH 转录因子鉴定和分析的报道。本研究利用生物信息学方法对森林草莓 bHLH(*FvbHLH*)家族的成员进行鉴定,并对其理化性质、系统进化、基因在染色体上的分布、基因结构及表达模式等方面进行全面分析,以期将来深入研究该家族基因的结构和功能提供重要参考。

## 1 材料与方法

### 1.1 森林草莓 bHLH 家族的鉴定

本研究对象是森林草莓(*Fragaria vesca*)。2020

收稿日期:2021-03-28

基金项目:河南省高等学校重点科研项目(编号:20A180018);洛阳师范学院博士科研启动项目(编号:180141051235)。

作者简介:郑玲(1981—),女,河南南阳人,硕士,副教授,主要从事植物学研究。E-mail:linda2266@qq.com。

年3—6月在洛阳师范学院生命科学学院生物信息学实验室(214)电脑上进行软件操作和信息处理。森林草莓 bHLH 家族的蛋白质序列数据来自 Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) 与 PlantTFDB (<http://planttfdb.cbi.pku.edu.cn/>) 数据库,由于2个数据库中搜索到的结果会存在差异,因此将2类搜索结果进行序列比对后,手动去除冗余序列,随后利用 InterPro (<http://www.ebi.ac.uk/interpro/>) 在线软件鉴定搜索的候选基因是否具有 bHLH 结构域,进而获得 bHLH 家族的所有成员。

### 1.2 森林草莓 bHLH 家族的理化性质分析

通过在线软件 ExPASy (<https://web.expasy.org/protparam/>) 得到森林草莓 bHLH 蛋白序列的基本理化性质信息,如理论等电点、氨基酸数量、分子量及总水平疏水性等。

### 1.3 森林草莓 bHLH 家族系统进化树的构建

从 PlantTFDB 数据库下载森林草莓和拟南芥的 bHLH 序列,随后通过 MEGA 6.0 的 Clustal W 工具进行多序列比对,将获得的比对结果用 MEGA 6.0 软件构建系统进化树<sup>[22]</sup>。同时采用相同方法单独构建森林草莓的 bHLH 家族进化树。

### 1.4 森林草莓 bHLH 基因在染色体上的定位

根据 Phytozome 数据库中搜索到的 *FvbHLH* 基因在染色体上的位置信息,利用 MapInspect 软件绘制基因在染色体上的分布图<sup>[23]</sup>。

### 1.5 森林草莓 bHLH 基因在花、果实发育过程中的作用分析

从文献[24-25]的附件中分别下载森林草莓花发育、果实发育过程中所有转录因子的表达量,进一步筛选出 bHLH 基因在各组织中的转录表达量。用 MeV4 软件分别对检测到的 *FvbHLH* 基因表达量进行聚类分析<sup>[26]</sup>。

## 2 结果与分析

### 2.1 森林草莓 bHLH 家族基因的鉴定及理化特性分析

通过 PlantTFDB、Phytozome 两大数据库共鉴定出 112 个候选 bHLH 家族基因,利用 InterPro 进行验证,结果显示,这 112 个基因编码的蛋白质都具有完整的 bHLH 结构域。同时利用 ExPASy 在线分析工具得到这 112 个 *FvbHLH* 家族基因编码蛋白的理论等电点、氨基酸数、原子组成及脂溶指数等理化信

息。根据基因在染色体上的位置,将这 112 个转录因子依次命名为 *FvbHLH1* ~ *FvbHLH112*。

结果显示,各蛋白质分子中的氨基酸数量在 90 (*FvbHLH64*) ~ 1 496 (*FvbHLH27*) 个之间,平均为 434 个;蛋白质分子量在 10 296.6 u (*FvbHLH64*) ~ 166 739.0 u (*FvbHLH27*) 之间,平均为 47 959.6 u;理论等电点在 4.35 (*FvbHLH9*) ~ 10.94 (*FvbHLH106*) 之间,并且有 70 个蛋白质的等电点在 4~7 之间,说明森林草莓 bHLH 家族蛋白的氨基酸序列的等电点大多在酸性范围内。有 73 个蛋白质分子的不稳定指数均大于 50,说明 bHLH 蛋白质的状态不稳定。对蛋白质的亲疏水性进行分析发现,除 *FvbHLH76* 外,其余蛋白质的总水平疏水性均为负值,表明森林草莓 bHLH 家族蛋白大多属于亲水蛋白。

### 2.2 森林草莓 bHLH 家族基因进化树的构建

选取 111 个拟南芥 bHLH 家族成员作为参考来构建森林草莓 bHLH 家族的系统进化树(图 1)。根据拟南芥 bHLH 转录因子的分类系统,可将森林草莓的 112 个 bHLH 成员划分为 12 个亚族。其中第 VI 亚族成员最少,只有 1 个 *FvbHLH*、2 个 *AtbHLH*;其次是第 II 亚族,有 1 个 *FvbHLH*、3 个 *AtbHLH*;第 III 亚族的成员数最多,包含 19 个 *FvbHLH*、17 个 *AtbHLH*;第 I 亚族的基因数也较多,有 15 个 *FvbHLH*、17 个 *AtbHLH*。其余亚家族的成员数量不一。

### 2.3 森林草莓 bHLH 家族基因在染色体上的定位

利用 MapInspect 软件绘制出基因在染色体上的分布情况,由图 2 可以看出,在森林草莓的全部 bHLH 基因中,*FvbHLH104* ~ *FvbHLH112* 尚未定位到染色体上,其余 103 个基因在 7 条染色体上均有分布,但分布不均匀。其中 5 号染色体上分布的基因最多,有 22 个(*FvbHLH55* ~ *FvbHLH76*);其次是 2、4 号染色体,都各有 17 个基因(*FvbHLH12* ~ *FvbHLH28* 和 *FvbHLH38* ~ *FvbHLH54*);3 号染色体上分布的基因较少,为 9 个(*FvbHLH29* ~ *FvbHLH37*),其余染色体上的基因分布数量为 11~15 个不等。多条染色体上出现了由 2 个或多个基因聚集在一起形成的基因簇。

### 2.4 森林草莓 bHLH 基因在花发育过程中的作用分析

用 MeV4 软件对检测到的 *FvbHLH* 家族成员在花发育过程中相关基因的表达量进行聚类分析。由

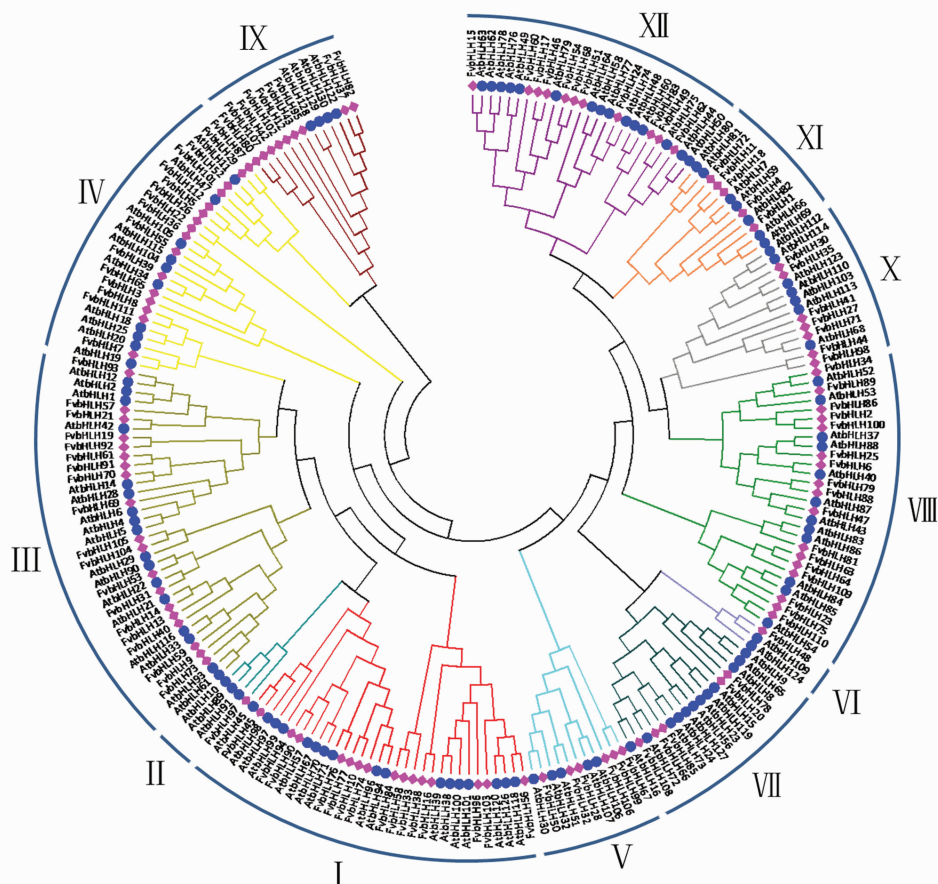


图1 森林草莓 bHLH 蛋白家族的系统进化树

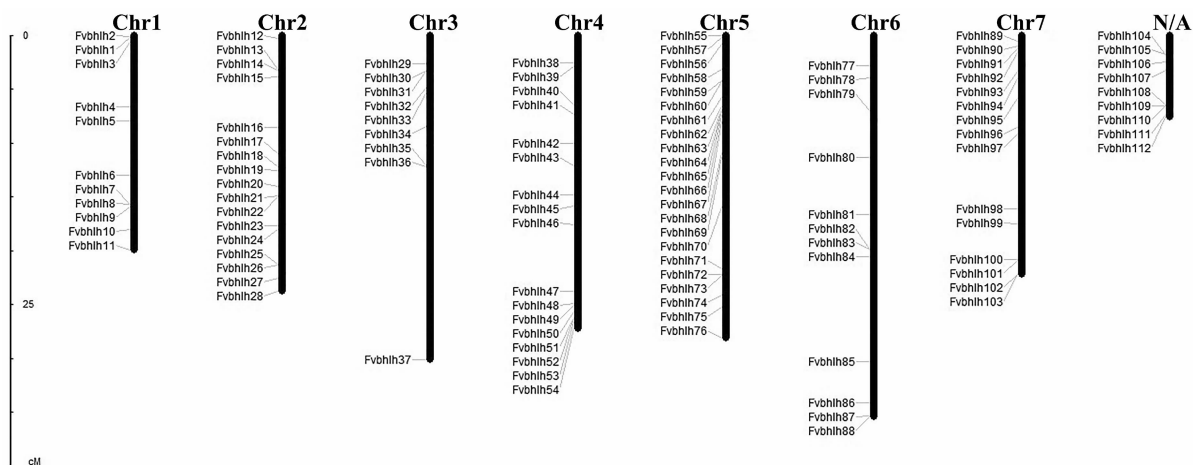


图2 FvbHLH 家族基因的染色体分布

图3可以看出,在森林草莓的花瓣(flower)、花被(perianth)、花托(receptacle)、花药(anther)及小孢子(microspore)、花粉(pollen)和雌蕊(carpel)发育的过程中,共有73个FvbHLH家族基因表达。

结果显示,有10个基因(*FvbHLH97/36/55/39/42/52/26/44/68/54*)在所有组织中的表达量均较高。*FvbHLH90/19/9/107/80/67/22/1/11/18/15/51/24/46*这14个基因除了在水粉中不表达之外,

在其余组织中均大量表达。同样的,*FvbHLH57/93/37/43/78/71*也仅在小孢子、花粉中的表达量较低。此外,有些基因的表达模式显示出组织特异性,例如*FvbHLH53*在花药、花粉和雌蕊中高表达;*FvbHLH13*在小孢子、花药中高表达;*FvbHLH86*在雌蕊、花药中高表达;*FvbHLH89/34*在花药、花粉中高表达;*FvbHLH16*只在花粉中大量表达;*FvbHLH12/4*只在花药发育过程中的表达水平较高;*FvbHLH10/*

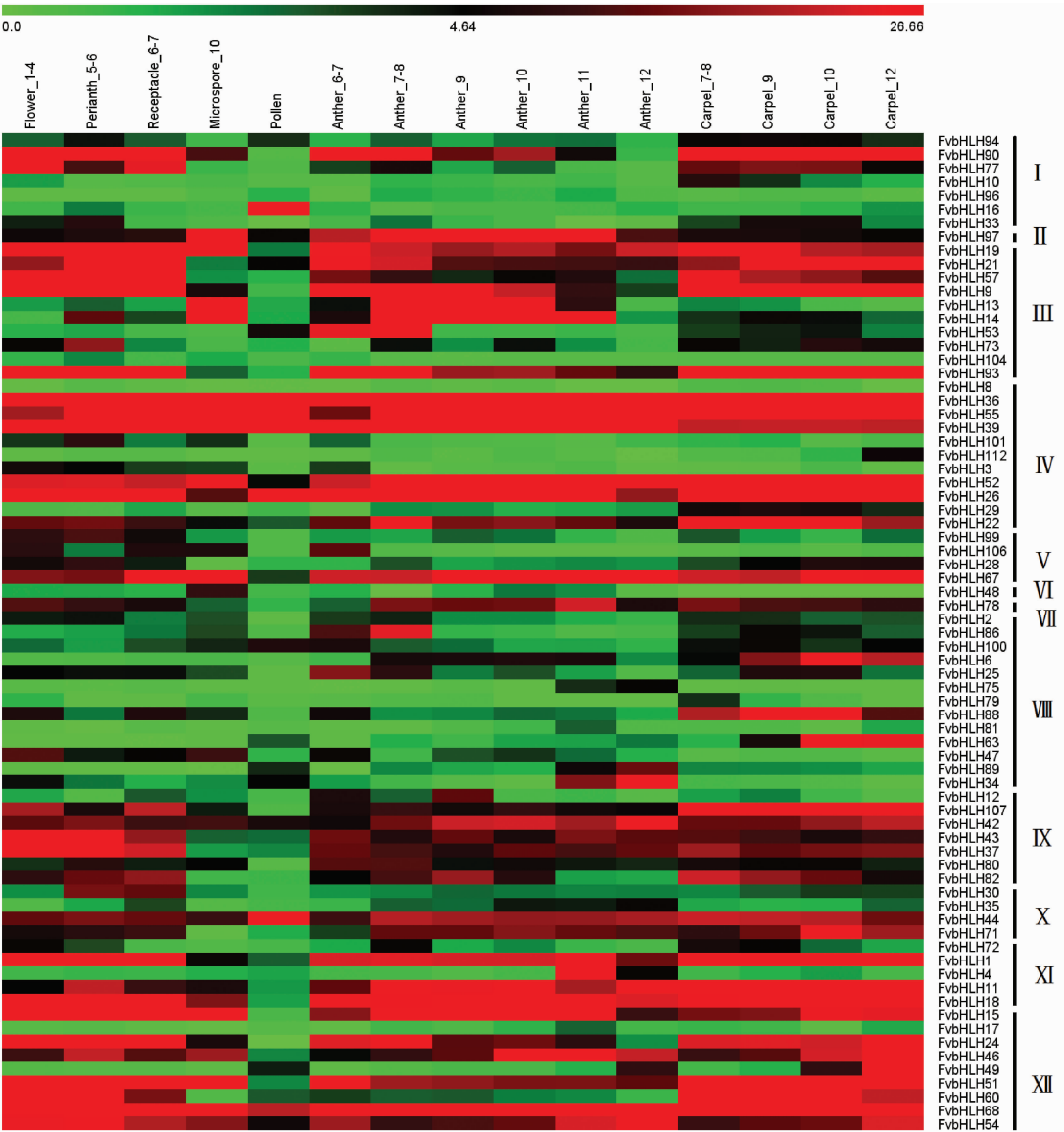


图3 *FvbHLH* 基因在花发育过程中的表达分析

112/88/63/29 只在雌蕊中高表达。

2.5 森林草莓 bHLH 家族基因在果实发育过程中的作用分析

通过 MeV4 软件对检测到的 *FvbHLH* 家族成员在果实发育过程中相关基因的表达量进行聚类分析。由图 4 可以看出,共有 28 个 *FvbHLH* 家族基因在森林草莓胚胎 (embryo)、胚珠 (ovule)、种皮 (ghost)、种子 (seed) 及子房壁 (wall)、果皮 (cortex)、木髓 (pith) 的发育过程中表达。

研究发现,在胚珠、种子和子房壁发育的过程中,有 14 个基因即 *FvbHLH90/96/58/40/57/91/16/66/78/100/81/15/17/46* 均显示出高表达,其中 *FvbHLH96/57/91/16/66/78/100/81/15/17/46* 在种皮整个发育阶段中的表达量也很高,还有部分基因

的表达水平则在种皮的不同发育阶段存在较大差异,例如 *FvbHLH10/91/33* 的表达量持续增加,而 *FvbHLH58/66/15* 的表达量却逐渐减少。在果皮、木髓中, *FvbHLH77/57/59/112/99/66/78/100/79/64/30/15/17/46* 共 14 个基因的表达水平都很高,有些基因则只在这 2 种组织的特定发育阶段表达,例如 *FvbHLH90*、*FvbHLH84* 分别只在果皮、木髓发育的 1~2 阶段显示出较高的表达量, *FvbHLH72* 只在果皮发育的 4~5 阶段高表达。在胚胎发育过程中, *bHLH* 基因的表达量较少,只有 *FvbHLH90/10/59/101/33/88/81/72* 共 8 个基因的表达量较高。

3 讨论

目前,还未见关于森林草莓 bHLH 转录因子的



图4 *FvbHLH* 基因在果实发育过程中的表达分析

报道,本研究利用生物信息学方法鉴定得到 112 个森林草莓 *bHLH* 转录因子,对其理化性质进行全面分析发现,*bHLH* 转录因子大多是不稳定的亲水蛋白,且等电点大多处于酸性范围。在拟南芥、水稻中的 *bHLH* 蛋白等电点也存在这种现象,表明 *bHLH* 转录因子的保守结构域中虽然存在 basic 碱性区域,但酸性氨基酸也同样富集于蛋白质分子中<sup>[27]</sup>。

在模式植物拟南芥中,研究者将 133 个 *bHLH* 基因划分为 12 个亚家族<sup>[28]</sup>;水稻中的 167 个 *bHLH* 转录因子被分为 22 个亚家族<sup>[29]</sup>;159 个马铃薯<sup>[30]</sup>、230 个大白菜<sup>[31]</sup> *bHLH* 家族分别被划分为 21、24 个亚族。本研究参照拟南芥的分类方式,将森林草莓的 112 个 *bHLH* 转录因子分成 12 个亚族,其中第 II、VI 亚族的成员最少,均只有 1 个 *FvbHLH* 基因;第 III 亚族的成员最多,为 19 个。在拟南芥 *bHLH* 亚家族中较少的也是第 II、VI 亚族,第 III 亚族的 *bHLH* 成员也较多。通过染色体定位分析发现,*bHLH* 基因在各号染色体上的分布都不均匀,并且形成了基因簇,说明它们可能属于串联复制。

已有大量研究发现,*bHLH* 转录因子与植物的胁迫反应有关,包括抗盐、干旱、缺铁等抗胁迫过程。进化树分析发现,森林草莓 *FvbHLH30* 与拟南芥 *AtbHLH112*<sup>[15]</sup> 聚在同一分支,推测其在抗盐、耐旱和抗渗透胁迫方面也有重要作用;*FvbHLH59* 与 *AtbHLH116*<sup>[16]</sup> 位于同一分支,因此 *FvbHLH59* 可能

与植物的抗低温调控有关。*bHLH* 家族在植物信号转导、物质合成方面也有关键作用。森林草莓 *FvbHLH93/57/21* 与拟南芥 *AtbHLH1/2/42*<sup>[11-12]</sup> 的进化关系较近,推测它们与花青素的合成有关;*FvbHLH6* 与 *AtbHLH40*<sup>[14]</sup> 聚在同一分支,推测 *FvbHLH6* 也参与了赤霉素的合成转导。

表达模式分析结果显示,在花的发育过程中显示出高表达的部分基因主要集中于第 III、IV、IX 及 XI、XII 亚族,其中 *FvbHLH97/36/55/39/42/52/26/44/68/54* 的表达量在森林草莓花的整个发育过程中都很高,表明这些基因对于花的发育尤为重要。*FvbHLH57/59/66/78/100/15/17/46* 几乎在果实发育的全过程表达,说明这 8 个基因与森林草莓果实的发育调控有关。草莓是一种重要的小型浆果,其果实主要由花托发育而来<sup>[24]</sup>。由此推测,在花托发育中起较大调控作用的 *FvbHLH19/21/57/9* 等基因可能对改善森林草莓果实的品质起着重要的作用。

本研究对森林草莓 *bHLH* 家族成员的理化性质、进化关系、染色体分布及组织表达模式等方面进行了全面分析,初步分析了 *bHLH* 基因在森林草莓花和果实发育过程中的作用,可为提高其经济价值提供理论基础和研究方向,并为进一步了解森林草莓 *bHLH* 转录因子结构与功能之间的关系和在植物生长发育过程中的调控机制提供重要的依据。

#### 参考文献:

[1] Wu S, Gallagher K L. Transcription factors on the move[J]. Current



- Opinion in Plant Biology,2012,15(6):645–651.
- [2] Feller A, Machemer K, Braun E L, et al. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors [J]. Plant Journal,2011,66(1):94–116.
  - [3] Feng X M,Zhao Q,Zhao L L,et al. The cold – induced basic helix – loop – helix transcription factor gene *MdCibHLH1* encodes an ICE – like protein in apple[J]. BMC Plant Biology,2012,12(1):1–14.
  - [4] Massari M E,Murre C. Helix – loop – helix proteins: regulators of transcription in eucaryotic organisms [J]. Molecular and Cellular Biology,2000,20(2):429–440.
  - [5] Robinson K A, Lopes J M. SURVEY AND SUMMARY: *Saccharomyces cerevisiae* basic helix – loop – helix proteins regulate diverse biological processes [J]. Nucleic Acids Research,2000,28(7):1499–1505.
  - [6] Fairchild C D,Schumaker M A,Quail P H. HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction [J]. Gene Dev,2000,14(18):2377–2391.
  - [7] Penfield S,Josse E M,Kannangara R,et al. Cold and light control seed germination through the bHLH transcription factor SPATULA [J]. Curr Biol,2005,15(22):1998–2006.
  - [8] Abe H,Yamaguchi – Shinozaki K,Urao T,et al. Role of *Arabidopsis* MYC and MYB homologs in drought – and abscisic acid – regulated gene expression [J]. Plant Cell,1997,9(10):1859–1868.
  - [9] Jiang Y Q,Deyholos M K. Comprehensive transcriptional profiling of NaCl – stressed *Arabidopsis* roots reveals novel classes of responsive genes [J]. BMC Plant Biology,2006,6:25.
  - [10] Ludwig S R,Habera L F,Dellaporta S L,et al. Lc,a member of the maize R gene family responsible for tissue – specific anthocyanin production,encodes a protein similar to transcriptional activators and contains the myc – homology region [J]. Proceedings of the National Academy of Sciences of the United States of America,1989,86(18):7092–7096.
  - [11] Nesi N,Debeaujon I,Jond C,et al. The TT8 gene encodes a basic helix – loop – helix domain protein required for expression of DFR and BAN genes in *Arabidopsis* siliques [J]. Plant Cell,2000,12(10):1863–1878.
  - [12] Ramsay N A,Walker A R,Mooney M,et al. Two basic – helix – loop – helix genes (MYC – 146 and GL3) from *Arabidopsis* can activate anthocyanin biosynthesis in a white – flowered *Mathiola incana* mutant [J]. Plant Molecular Biology,2003,52(3):679–688.
  - [13] Sakamoto W,Ohmori T,Kageyama K,et al. The *purple leaf* (*Pl*) locus of rice;the *Pl<sup>u</sup>* allele has a complex organization and includes two genes encoding basic helix – loop – helix proteins involved in anthocyanin biosynthesis [J]. Plant and Cell Physiology,2001,42(9):982–991.
  - [14] Arnaud N,Girin T,Sorefan K,et al. Gibberellins control fruit patterning in *Arabidopsis thaliana* [J]. Genes & Development,2010,24(19):2127–2132.
  - [15] Liu W,Tai H,Li S,et al. *bHLH22* is important for drought and osmotic stress resistance in *Arabidopsis* and in the repression of ABA catabolism [J]. New Phytologist,2014,201(4):1192–1204.
  - [16] Chinnusamy V,Ohta M,Kanrar S,et al. ICE1:a regulator of cold – induced transcriptome and freezing tolerance in *Arabidopsis* [J]. Genes Development,2003,17(8):1043–1054.
  - [17] Paik I,Kathare P K,Kim J I,et al. Expanding roles of PIFs in signal integration from multiple processes [J]. Molecular Plant,2017,10(8):1035–1046.
  - [18] Kondou Y,Nakazawa M,Kawashima M,et al. RETARDED GROWTH OF EMBRYO1,a new basic helix – loop – helix protein,expresses in endosperm to control embryo growth [J]. Plant Physiology,2008,147(4):1924–1935.
  - [19] Sorensen A M,Krüber S,Unte U S,et al. The *Arabidopsis* ABORTED MICROSPORES (AMS) gene encodes a MYC class transcription factor [J]. Plant Journal for Cell & Molecular Biology,2003,33(2):413–423.
  - [20] Jung K H,Han M J,Lee Y S,et al. Rice undeveloped Tapetum1 is a major regulator of early tapetum development [J]. The Plant Cell,2005,17(10):2705–2722.
  - [21] 周鹤莹,张 玮,张 卿,等. 森林草莓‘Hawaii4’高效遗传转化系统的建立 [J]. 北京农学院学报,2015,30(1):10–14.
  - [22] Tamura K,Steicher G,Peterson D,et al. MEGA6: molecular evolutionary genetics analysis version 6.0 [J]. Molecular Biology and Evolution,2013,30(12):2725–2729.
  - [23] Zhang S P,Liu M M,Miao H,et al. Chromosomal mapping and QTL analysis of resistance to downy mildew in *Cucumis sativus* [J]. Plant Disease,2013,97(2):245–251.
  - [24] Hollender C A,Kang C,Darwish O,et al. Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks [J]. Plant Physiol,2014,165(3):1062–1075.
  - [25] Kang C,Darwish O,Geretz A,et al. Genome – scale transcriptomic insights into early – stage fruit development in woodland strawberry *Fragaria vesca* [J]. Plant Cell,2013,25(6):1960–1978.
  - [26] Howe E,Holton K,Nair S,et al. MeV: multi experiment viewer [M]. Boston:Springer,2010:267–277.
  - [27] 应炎标,朱友银,郭卫东,等. 樱桃 bHLH 转录因子家族基因鉴定及表达分析 [J]. 分子植物育种,2018,16(14):4559–4568.
  - [28] 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]. Molecular Biology and Evolution,2003,20(5):735–747.
  - [29] Li X,Duan X,Jiang H,et al. Genome – wide analysis of basic/helix – loop – helix transcription factor family in rice and *Arabidopsis* [J]. Plant Physiology,2006,141(4):1167–1184.
  - [30] Sun H,Fan H J,Ling H Q. Genome – wide identification and characterization of the bHLH gene family in tomato [J]. BMC Genomics,2015,16(1):9.
  - [31] Song X,Huang Z,Duan W,et al. Genome – wide analysis of the bHLH transcription factor family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) [J]. Molecular Genetics and Genomics,2014,289(1):77–91.