



### 甜橙bHLH转录因子家族的鉴定及生物信息学分析

孟富宣 杨玉皎 段元杰 吕陟远 陆晓英 郭淑萍 刘海刚

### Identification and Bioinformatics Analysis of the bHLH Transcription Factor Family in *Citrus sinensis*

Meng Fuxuan, Yang Yujiao, Duan Yuanjie, Lv Zhiyuan, Lu Xiaoying, Guo Shuping, Liu Haigang

引用本文:

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[J]. 西南林业大学学报, 2020, 40(5):73–86. doi: 10.11929/j.swfu.201906018

Meng Fuxuan, Yang Yujiao, Duan Yuanjie, Lv Zhiyuan, Lu Xiaoying, Guo Shuping, Liu Haigang. Identification and Bioinformatics Analysis of the bHLH Transcription Factor Family in *Citrus sinensis* [J]. *Journal of Southwest Forestry University(Natural Science)*, 2020, 40(5):73–86. doi: 10.11929/j.swfu.201906018

在线阅读 View online: <https://doi.org/10.11929/j.swfu.201906018>

---

### 您可能感兴趣的其他文章

#### Articles you may be interested in

#### 荷花转录组测序及花青素苷合成相关基因表达分析

Transcriptome Sequencing and Expression Analysis of Anthocyanin Synthesis Related Genes in Lotus  
西南林业大学学报. 2018, 38(2): 61 <https://doi.org/10.11929/j.issn.2095-1914.2018.02.010>

#### 基于流式细胞术的15个油茶品种基因组测定研究

Estimation of Genome Sizes of 15 *Camellia oleifera* Varieties Using Flow Cytometry Analysis  
西南林业大学学报. 2018, 38(5): 46 <https://doi.org/10.11929/j.issn.2095-1914.2018.05.008>

#### 产香真菌的鉴定及挥发性成分分析

Identification and Volatile Components Analysis of Aroma-producing Fungi  
西南林业大学学报. 2017, 37(3): 118 <https://doi.org/10.11929/j.issn.2095-1914.2017.03.019>

#### 青藏高原贡觉县森林规划设计因子的地统计学分析

Geo-statistical Analysis of the Forest Planning and Design Factors in Gongjue County of Qinghai-Tibetan Plateau  
西南林业大学学报. 2017, 37(3): 146 <https://doi.org/10.11929/j.issn.2095-1914.2017.03.023>

#### 基于代谢组学分析技术的木材形成机理研究

Metabolic Profiling for Revealing the Mechanisms of Wood Formation  
西南林业大学学报. 2019, 39(1): 1 <https://doi.org/10.11929/j.swfu.201811009>

#### 拟南芥AtIDD4与3 × Flag融合蛋白表达植株的构建及其鉴定

Construction and Identification of AtIDD4 and 3 × Flag Fusion Protein Expressing in *Arabidopsis thaliana*  
西南林业大学学报. 2019, 39(4): 89 <https://doi.org/10.11929/j.swfu.201812024>

DOI: 10.11929/j.swfu.201906018

引文格式: 孟富宣, 杨玉皎, 段元杰, 等. 甜橙 bHLH 转录因子家族的鉴定及生物信息学分析 [J]. 西南林业大学学报 (自然科学), 2020, 40(5): 73–86.

## 甜橙 bHLH 转录因子家族的鉴定及生物信息学分析

孟富宣 杨玉皎 段元杰 吕陟远 陆晓英 郭淑萍 刘海刚

(云南省农业科学院热区生态农业研究所/元谋干热河谷植物园, 云南 楚雄 651399)

**摘要:** bHLH 转录因子家族是第 2 大转录因子家族, 广泛存在于微生物、植物以及动物中。以甜橙为试验材料, 通过全基因组数据库搜索鉴定, 对 *bHLH* 基因进行了理化性质、系统发育、染色体分布、保守基序、蛋白互作生物信息学分析, 探讨它们在不同组织的表达情况。结果表明: 共鉴定出 126 个 *bHLH* 基因, 其中 112 个不均匀的分布于 9 条染色体上, 14 个分布位置尚未确定; 系统发育分析表明甜橙 *bHLH* 基因家族可分为 16 个亚家族, 并且给定亚家族中的大多数基因具有相似的基因结构和保守基序。在甜橙中, 多个 *bHLH* 基因相互作用共同调控其开花、光的形态发生和花色素苷合成等过程; 大多数的 *CsbHLH* 基因在愈伤组织、叶、花和果实中均有表达。本研究为 *CsbHLH* 基因的功能研究奠定了基础, 并有助于 *bHLH* 基因在其他植物物种中的功能研究。

**关键词:** 甜橙; bHLH; 全基因组; 转录因子; 鉴定; 生物信息学

中图分类号: S666.4

文献标志码: A

文章编号: 2095-1914(2020)05-0073-14

## Identification and Bioinformatics Analysis of the bHLH Transcription Factor Family in *Citrus sinensis*

Meng Fuxuan, Yang Yujiao, Duan Yuanjie, Lv Zhiyuan,  
Lu Xiaoying, Guo Shuping, Liu Haigang

(Institute of Tropical Eco-agricultural Sciences, Yunnan Academy of Agricultural Sciences/  
Yuanmou Dry-Hot Valley Botanical Garden, Chuxiong Yunnan 651399, China)

**Abstract:** The bHLH TFs (basic Helix-Loop-Helix transcription factors) family is widely found in microorganisms, plants and animals, and is the second largest family of transcription factors. In this study, *Citrus sinensis* was used as the experimental material, the physicochemical properties, phylogeny, chromosome distribution, conserved motif and protein interaction bioinformatics of *bHLH* gene were analyzed by genome-wide database search, and their expression in different tissues was discussed. Studies have shown that 126 *bHLH* genes were identified. Among them, 112 *bHLH* genes are distributed on 9 chromosomes, 14 *bHLH* genes are located in uncertain locations. Phylogenetic analysis indicated that the *C. sinensis* *bHLH* gene family can be divided into 16 subfamilies, and most genes in a given subfamily have similar gene structures and conserved motifs. Multiple *bHLH* gene interactions in *C. sinensis* regulate the development of flower organogenesis, light morphogenesis and anthocyanin synthesis. The tissue expression pattern of *CsbHLH* gene indicates that most members are in callus, leaves, flowers and fruits are expressed. This study laid the foundation for the functional study of the *CsbHLH* gene and

收稿日期: 2019-06-10; 修回日期: 2019-07-23

基金项目: 云南省科技厅科技计划青年项目 (2018FD131) 资助。

第 1 作者: 孟富宣 (1989—), 女, 硕士, 助理研究员。研究方向: 果树遗传育种。Email: 616967560@qq.com。

通信作者: 刘海刚 (1982—), 男, 硕士, 副研究员。研究方向: 作物育种及栽培。Email: 1939440231@qq.com。

contributed to the functional study of the *bHLH* gene in other plant species.

**Key words:** *Citrus sinensis*; bHLH; genome-wide; transcription factor; identification; bioinformatics

转录因子 (TF), 又称为反式作用因子, 通常包含 4 个功能区, 即 DNA 结合域、转录调节域、核定位信号和寡聚位点<sup>[1-6]</sup>。目前的研究中, 发现植物中的转录因子家族已经超过 60 个, 根据 DNA 结合域中赖氨酸和精氨酸残基的数量, 把转录因子分为四类, 即锌指型 (ZF)、螺旋转螺旋型 (HLH)、碱性螺旋环螺旋型 (bHLH) 和碱性亮氨酸拉链 (bZIP)。迄今为止, 在高等植物中发现的最常见的转录因子家族是 WD40、MYB、WRKY、bHLH 和 bZIP 家族<sup>[7]</sup>。

bHLH TFs 广泛存在于真核生物中, 是继 MYB 家族之后的第 2 大家族<sup>[8-9]</sup>。bHLH TFs 的 DNA 结合结构域包含约 60 个氨基酸, 在转录因子序列的 N 端包含 10~15 个氨基酸的基本区域和约 40 个氨基酸的螺旋-环-螺旋区域, 并且主要负责 TFs 与特定 DNA 序列的结合<sup>[9-13]</sup>。研究表明, 同一个 bHLH 转录因子的 2 个  $\alpha$ -螺旋或不同的 bHLH TF 的  $\alpha$ -螺旋之间相互作用形成同源或异源二聚体, 这些二聚体可以针对基因启动子的不同部分来调节目标基因<sup>[14-15]</sup>。与高度保守的 BHLH 结构域不同, TF 序列的其余部分通常差异非常大。在靶基因启动子中, bHLH TFs 主要与 E-box 序列 (5'-CANNTG-3') 结合, 最常见的形式是重复的 G-box (5'-CACGTG-3')。DNA 结合结构域中几个保守的氨基酸决定了转录因子对不同 E-boxes 核心序列位点的特异性<sup>[3,16]</sup>。Atchley 等人根据动物 bHLH TFs 的进化关系和序列相似性将动物 bHLH TFs 分为 6 组 (A-F)<sup>[10,17]</sup>。Heim 等人发现多数植物的 bHLH TFs 属于 B 组<sup>[11]</sup>。随着分子生物学的快速发展, bHLH TF 家族的多样性已被揭示, 植物 bHLH TFs 基因在拟南芥 (*Arabidopsis thaliana*)<sup>[14]</sup>、水稻 (*Oryza sativa*)<sup>[13]</sup>、番茄 (*Solanum lycopersicum*)<sup>[18]</sup>、短柄草 (*Brachypodium sylvaticum*)<sup>[19]</sup>、花生 (*Arachis hypogaea*)<sup>[20]</sup> 和丹参 (*Salvia miltiorrhiza*)<sup>[21]</sup> 等多种植物中已经得到鉴定。高等植物的生长和形态发生是一个十分复杂的过程, 是通过 DNA 和蛋白质之间的相互作用来实现的。已发现雌性藤本植物开花过程是由 SPT 和 HEC 编码基因共同调控的<sup>[22]</sup>。在拟南芥中, 光形态发生的调节主要通过光敏色素调控基因相互作用来实现的<sup>[23]</sup>。在转录水平上, 一般通过 TF 和

相关基因的联合调控来实现高等植物逆境应答的调控。Abe 等人发现拟南芥中 RD22 的产生主要是由干旱胁迫和脱落酸 (ABA) 的诱导。RD22 启动子区含有 MYC (bHLH) 和 MYB 转录因子识别位点, 相应的转录因子与这些位点特异结合以调节基因表达<sup>[24-25]</sup>。

bHLH TFs 还参与植物次生代谢, 由 R 基因编码的玉米 (*Zea mays*) Lc 蛋白是植物中首次报道的 bHLH TF, 并且发现其在玉米花青素代谢途径中调控至少 2 个结构基因<sup>[26]</sup>。Goodrich 等人从金鱼草 (*Antirrhinum majus*) 中克隆得到 *Delila* 基因, 该基因编码具有与玉米 R 基因家族编码的蛋白质相似的螺旋-环-螺旋结构域的蛋白质, 随后证实 *Ddila* 基因的表达与花色苷的积累密切相关<sup>[27]</sup>。参与拟南芥花色苷合成调控的 bHLH TFs 属于第 III 亚组, 其中包括 TT8、EGL3、GL3 和 MYC1 TF 家族<sup>[11,28]</sup>。

柑橘 (*Citrus reticulata*) 是具有重要经济价值的多年生木本果树作物之一, 是世界第一大水果, 全世界的总产量达 1.2 亿 t, 居所有水果之首<sup>[29]</sup>。柑橘主要在我国南方地区种植, 栽培面积达 229 万  $\text{hm}^2$ , 产量达到 2944 万 t。本研究以甜橙 (*Citrus sinensis*) 为试验材料, 通过全基因组数据库搜索鉴定, 共鉴定出 126 个 *CsbHLH* 基因, 并对其进行了理化性质、系统发育、染色体分布、保守基序、蛋白互作生物信息学分析, 并分析了它们在不同组织的表达情况。为 *CsbHLH* 基因的功能研究奠定了基础, 并有助于 *bHLH* 基因在其他植物物种中的功能研究。

## 1 材料与方法

### 1.1 甜橙 bHLH 转录因子家族的鉴定

从柑橘基因组数据库中下载了 *CsbHLH* 序列 (<http://citrus.hzau.edu.cn/orange/>)。使用 Pfam 数据库 (<http://pfam.janelia.org/search/sequence>) 中的隐马尔可夫模型 (HMM) 文件进行本地 BLAST (Evalue-5) 搜索, 以筛选所有候选 *CsbHLH* 基因序列。保留含有已知保守结构域的候选基因, 通过对 Pfam, NCBI Conserved Domains (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 和 SMART (<http://smart.embl-heidelberg.de/>) 数据库的核对

以确定 bHLH 结构域 (PF00010) 的存在<sup>[30]</sup>。利用 BLSATP 方法在拟南芥基因组数据库 (TAIR; <https://www.arabidopsis.org/>) 获得甜橙 *bHLH* 同源基因的相关注释信息, 参数为默认值<sup>[31]</sup>。拟南芥 *bHLH* 基因序列分别从拟南芥信息资源 (TAIR) 网站下载。对 *CsbHLH*、*AtbHLH* 序列进行生物信息学分析, 并使用 ExPASy ([http://www.expasy.ch/tools/pi\\_tool.html](http://www.expasy.ch/tools/pi_tool.html)) 计算理化参数, MW、pI 和 GRAVY 等。

### 1.2 甜橙 *bHLH* 基因的系统发育和多重序列比对分析

用 MEGA 6.0 中的邻距法 (NJ, Neighbor-joining) 构建系统发育树, 其中校验参数 Bootstrap=1 000。利用 MEGA 6.0 构建 *CsbHLH* 基因的个体系统发育树和来自 2 个物种 (*AtbHLH* 和 *CsbHLH*) 的 *bHLH* 基因的综合系统发育树, 并对甜橙 *bHLH* 基因家族进行分类。使用 ClustalX 2.11 对 *CsbHLH* 基因的蛋白质序列进行多重序列比对, 参数设定为默认参数。

### 1.3 甜橙 *bHLH* 基因保守基序分析

利用在线 MEME 系统 (<http://meme.sdsc.edu/meme/itro.html>) 来识别保守基序, 参数设置为: 重复次数“任意”, 最高基序数“20”, 基序长度“6-200”, 其他均为默认值。

### 1.4 甜橙 *bHLH* 基因的染色体位置和基因重复

通过使用来自支架的染色体定位和注释信息来探索 *FRO* 基因在染色体上的分布, 并且使用公开可用的数据库和 Map Draw 软件来根据物理位置显示精确的基因位置。还通过使用植物基因组

复制数据库 (PGDD, <http://chibba.agtec.uga.edu/duplication/index/locus>) 在 *FRO* 基因家族中进行基因重复分析。

### 1.5 甜橙 *bHLH* 蛋白功能联系网络

利用 String 蛋白互作数据库 (<http://string-db.org/>) 对 126 个 bHLH 蛋白的互作网络进行分析, 物种参数选择模式植物拟南芥。

### 1.6 甜橙 *bHLH* 转录因子的表达分析

甜橙 RNA-seq 表达数据来自 Citrus sinensis Annotation project 数据库, 利用 Matrix2png interface 软件对甜橙 *bHLH* 家族基因表达模式进行分析, 绿、黑、红 3 个颜色表示基因的表达强度, 绿色越强代表表达量越小, 红色越强代表表达量越大, 黑色为中间值。

## 2 结果与分析

### 2.1 甜橙中 *bHLH* 基因家族成员的鉴定及理化性质分析

研究者利用是否含有保守结构域 *bHLH* 基序 (PF00010), 在拟南芥<sup>[14]</sup> 中鉴定出 *bHLH* TFs 144 条, 鉴于此, 利用该保守结构域鉴定柑橘中的 *bHLH* 基因, 然后通过 NCBI BlastP 和 Citrus Sinensis Annotation Project 数据库 (<http://citrus.hzau.edu.cn/orange/index.php>) 在柑橘基因组中搜索验证, 最终获得 126 个 *bHLH* 基因 (表 1)。其内含子数目介于 0~12 之间, 其中 *CsbHLH002*、*CsbHLH012*、*CsbHLH041*、*CsbHLH068*、*CsbHLH124*、*CsbHLH126* 不含有内含子, *CsbHLH087* 最多含 12 个内含子。

表 1 甜橙中鉴定出的 *CsbHLH* 基因家族成员及其基本特性

Table 1 The *CsbHLH* gene family members in *C. sinensis* and their basic characteristics

基因	基因ID	NCBI登录号	染色体位点	外显子	氨基酸/aa	分子量/Da	等电点 PI	不稳定指数	脂肪系数	总平均疏水指数 GRAVY
<i>CsbHLH001</i>	<i>CsIg02580</i>	XP_024949282.1	Chr1:1893837~1895996	4	460	51 670.48	5.48	51.55	74.61	-0.633
<i>CsbHLH002</i>	<i>CsIg05540</i>	XP_006464873.1	Chr1:5277679~5279047	1	252	28 567.92	6.16	65.68	73.21	-0.667
<i>CsbHLH003</i>	<i>CsIg10425</i>	XP_024952817.1	Chr1:12442899~12445293	5	79	9 398.80	9.16	57.81	86.33	-0.653
<i>CsbHLH004</i>	<i>CsIg15360</i>	XP_006465638.1	Chr1:18574008~18577203	9	339	36 896.97	6.42	54.15	65.34	-0.602
<i>CsbHLH005</i>	<i>CsIg19150</i>	XP_006466558.2	Chr1:22522400~22525018	4	430	47 562.56	6.02	48.40	65.81	-0.724
<i>CsbHLH006</i>	<i>CsIg20060</i>	XP_006466564.1	Chr1:23201140~23203265	3	310	34 635.16	6.60	64.90	72.00	-0.416
<i>CsbHLH007</i>	<i>CsIg20320</i>	XP_015388414.1	Chr1:23405583~23406505	3	225	25 656.06	9.93	53.02	96.53	-0.367
<i>CsbHLH008</i>	<i>CsIg20840</i>	XP_006466366.1	Chr1:23818371~23820098	4	353	39 510.40	5.10	63.42	78.50	-0.492
<i>CsbHLH009</i>	<i>CsIg20990</i>	XP_006466576.1	Chr1:23955611~23957664	4	369	41 149.63	8.41	49.51	77.97	-0.554



续表 1

基因	基因ID	NCBI登录号	染色体位点	外显子	氨基酸/aa	分子量/Da	等电点PI	不稳定指数	脂肪系数	总平均疏水指数GRAVY
<i>CsbHHLH010</i>	<i>Cs1g21000</i>	XP_006466388.1	Chr1:23965192 ~ 23967805	5	362	40 502.31	5.96	55.55	75.61	-0.562
<i>CsbHHLH011</i>	<i>Cs1g22470</i>	XP_006466588.1	Chr1: 25181102 ~ 25184389	9	409	44 401.44	6.41	56.02	75.26	-0.489
<i>CsbHHLH012</i>	<i>Cs1g22560</i>	XP_006466734.1	Chr1: 25261293 ~ 25262096	1	267	29 841.36	9.38	59.19	69.44	-0.488
<i>CsbHHLH013</i>	<i>Cs2g03700</i>	XP_006467505.1	Chr2:1728317 ~ 1731722	4	241	26 867.89	5.51	54.49	61.45	-0.667
<i>CsbHHLH014</i>	<i>Cs2g04900</i>	XP_006467661.1	Chr2:2559126 ~ 2562431	2	264	28 804.72	7.65	50.42	83.94	-0.428
<i>CsbHHLH015</i>	<i>Cs2g05960</i>	XP_006467790.1	Chr2:3293286 ~ 3296194	5	188	21 073.61	6.09	43.41	75.27	-0.610
<i>CsbHHLH016</i>	<i>Cs2g07180</i>	XP_006467938.1	Chr2:4275314 ~ 4278527	7	487	52 830.38	6.01	60.47	54.76	-0.589
<i>CsbHHLH017</i>	<i>Cs2g07210</i>	XP_024950450.1	Chr2:4295638 ~ 4296379	3	140	15 916.23	5.08	49.39	94.00	-0.366
<i>CsbHHLH018</i>	<i>Cs2g10270</i>	XP_006468467.1	Chr2:7589154 ~ 7590485	2	307	34 764.32	5.37	66.88	74.66	-0.476
<i>CsbHHLH019</i>	<i>Cs2g12690</i>	XP_024951187.1	Chr2:9520564 ~ 9522328	6	242	26 894.30	6.39	47.98	66.90	-0.503
<i>CsbHHLH020</i>	<i>Cs2g13030</i>	XP_024950357.1	Chr2:9779034 ~ 9781958	7	293	31 956.78	7.06	57.22	66.89	-0.486
<i>CsbHHLH021</i>	<i>Cs2g14440</i>	XP_006468882.1	Chr2:11222379 ~ 11226720	5	487	53 696.11	5.37	45.14	75.34	-0.489
<i>CsbHHLH022</i>	<i>Cs2g15630</i>	XP_006469029.1	Chr2:12416439 ~ 12418852	3	269	30 016.07	5.84	53.87	92.42	-0.329
<i>CsbHHLH023</i>	<i>Cs2g16420</i>	XP_006469145.1	Chr2:13242703 ~ 13245731	7	440	49 109.79	6.43	55.24	65.41	-0.771
<i>CsbHHLH024</i>	<i>Cs2g22030</i>	XP_015382932.1	Chr2:19306489 ~ 19309128	5	250	27 465.15	5.54	55.86	74.20	-0.502
<i>CsbHHLH025</i>	<i>Cs2g23280</i>	XP_006469875.2	Chr2:21500008 ~ 21501176	3	241	27 358.47	6.55	52.54	82.57	-0.447
<i>CsbHHLH026</i>	<i>Cs2g29330</i>	XP_024951438.1	Chr2:28861753 ~ 28863635	3	164	18 253.94	8.79	52.51	98.17	-0.289
<i>CsbHHLH027</i>	<i>Cs2g29350</i>	XP_015383298.1	Chr2:28885790 ~ 28888079	4	164	18 394.11	7.77	49.41	92.07	-0.366
<i>CsbHHLH028</i>	<i>Cs3g02100</i>	XP_006470882.1	Chr3:1572941 ~ 1577080	7	491	54 634.35	6.58	40.78	63.38	-0.726
<i>CsbHHLH029</i>	<i>Cs3g04590</i>	XP_024951843.1	Chr3:5772184 ~ 5775239	2	436	49 364.36	5.09	51.77	84.33	-0.489
<i>CsbHHLH030</i>	<i>Cs3g10900</i>	XP_006471682.1	Chr3:15123575 ~ 15130509	8	336	35 889.09	5.43	62.21	70.68	-0.450
<i>CsbHHLH031</i>	<i>Cs3g13290</i>	XP_006471948.1	Chr3:17727120 ~ 17729051	3	317	35 816.89	5.19	59.28	80.00	-0.582
<i>CsbHHLH032</i>	<i>Cs3g15920</i>	XP_024952420.1	Chr3:19863437 ~ 19864938	5	366	40 518.74	5.20	54.00	66.67	-0.669
<i>CsbHHLH033</i>	<i>Cs3g16330</i>	XP_006472332.1	Chr3:20187103 ~ 20191770	10	426	46 433.95	5.44	50.74	64.08	-0.610
<i>CsbHHLH034</i>	<i>Cs3g16770</i>	XP_006472392.1	Chr3:20548171 ~ 20549868	4	349	39 464.37	4.67	49.12	72.69	-0.513
<i>CsbHHLH035</i>	<i>Cs3g17300</i>	XP_006472481.1	Chr3:21051819 ~ 21053135	2	260	28 922.22	7.09	52.46	70.12	-0.713
<i>CsbHHLH036</i>	<i>Cs3g17500</i>	XP_006472510.1	Chr3:21211346 ~ 21215276	4	364	39 840.25	4.60	49.76	69.15	-0.545
<i>CsbHHLH037</i>	<i>Cs3g17780</i>	XP_006472540.1	Chr3:21415344 ~ 21417381	3	329	36 441.78	5.58	67.77	75.96	-0.472
<i>CsbHHLH038</i>	<i>Cs3g22920</i>	XP_024952054.1	Chr3:25368735 ~ 25373055	6	285	31 290.27	8.29	57.12	76.28	-0.605
<i>CsbHHLH039</i>	<i>Cs3g23320</i>	XP_006473302.1	Chr3:25638556 ~ 25641498	6	291	31 349.70	4.92	62.35	70.41	-0.466
<i>CsbHHLH040</i>	<i>Cs3g23330</i>	XP_015384271.1	Chr3:25644032 ~ 25646613	3	266	29 740.76	6.43	55.01	75.90	-0.472
<i>CsbHHLH041</i>	<i>Cs3g23430</i>	XP_015384300.1	Chr3:25711673 ~ 25712987	1	272	30 022.85	8.94	49.32	67.17	-0.535
<i>CsbHHLH042</i>	<i>Cs3g25190</i>	XP_006473548.1	Chr3:26908223 ~ 26912406	7	431	46 494.84	7.66	66.23	68.63	-0.472
<i>CsbHHLH043</i>	<i>Cs4g01460</i>	XP_006473971.1	Chr4:376893 ~ 380473	7	469	49 182.31	5.83	46.75	62.26	-0.583
<i>CsbHHLH044</i>	<i>Cs4g04210</i>	XP_006474281.1	Chr4:2355251 ~ 2359605	5	246	28 152.11	5.19	53.52	83.29	-0.520

续表 1

基因	基因ID	NCBI登录号	染色体位点	外显子	氨基酸/aa	分子量/Da	等电点PI	不稳定指数	脂肪系数	总平均疏水指数GRAVY
<i>CsbHLH045</i>	<i>Cs4g06760</i>	XP_015384588.1	Chr4:4223600 ~ 4227746	10	519	56 318.99	4.96	51.52	80.42	-0.462
<i>CsbHLH046</i>	<i>Cs4g11870</i>	XP_006475252.1	Chr4:8699000 ~ 8702167	7	459	49 915.73	6.81	61.01	53.77	-0.697
<i>CsbHLH047</i>	<i>Cs4g12300</i>	XP_006475303.1	Chr4:9345421 ~ 9348274	4	236	25 950.56	8.36	57.32	74.45	-0.553
<i>CsbHLH048</i>	<i>Cs4g14510</i>	XP_006475472.1	Chr4:12668673 ~ 12671590	8	553	61 863.45	6.36	65.72	80.16	-0.413
<i>CsbHLH049</i>	<i>Cs4g16200</i>	XP_006475662.1	Chr4:15751784 ~ 15755268	9	363	39 483.94	7.13	68.76	61.32	-0.614
<i>CsbHLH050</i>	<i>Cs5g01450</i>	XP_006476285.1	Chr5:408083 ~ 410083	2	519	58 218.63	6.33	39.45	78.81	-0.447
<i>CsbHLH051</i>	<i>Cs5g03530</i>	XP_006476561.1	Chr5:1849188 ~ 1853394	8	618	69 137.94	5.84	46.13	82.49	-0.461
<i>CsbHLH052</i>	<i>Cs5g04000</i>	XP_006476619.1	Chr5:2205830 ~ 2208463	6	241	26 542.57	5.33	60.52	82.95	-0.624
<i>CsbHLH053</i>	<i>Cs5g04070</i>	XP_006476632.1	Chr5:2266207 ~ 2267696	2	91	10 315.60	9.09	79.13	95.38	-0.652
<i>CsbHLH054</i>	<i>Cs5g04620</i>	XP_006476705.1	Chr5:2704897 ~ 2707957	6	278	30 034.23	5.52	54.21	61.83	-0.704
<i>CsbHLH055</i>	<i>Cs5g05300</i>	XP_006476761.1	Chr5:3156670 ~ 3158286	3	181	20 682.94	9.17	41.46	90.39	-0.406
<i>CsbHLH056</i>	<i>Cs5g09950</i>	XP_006477344.1	Chr5:6849061 ~ 6852592	4	246	28 292.90	6.32	71.94	81.59	-0.705
<i>CsbHLH057</i>	<i>Cs5g11380</i>	XP_006477612.1	Chr5:8510879 ~ 8514543	6	419	46 069.66	8.12	63.10	55.27	-0.735
<i>CsbHLH058</i>	<i>Cs5g15670</i>	XP_015385626.2	Chr5:14135657 ~ 14138349	4	272	30 761.32	9.23	52.51	89.89	-0.398
<i>CsbHLH059</i>	<i>Cs5g22140</i>	XP_024954143.1	Chr5:24659321 ~ 24663962	8	143	15 946.46	9.19	52.88	86.01	-0.319
<i>CsbHLH060</i>	<i>Cs5g24590</i>	XP_006478879.1	Chr5:27436872 ~ 27443482	7	371	39 283.07	8.44	61.40	58.89	-0.608
<i>CsbHLH061</i>	<i>Cs5g30020</i>	XP_015386328.1	Chr5:31961958 ~ 31967683	11	748	82 353.73	5.68	42.15	80.99	-0.341
<i>CsbHLH062</i>	<i>Cs5g30170</i>	XP_006479569.1	Chr5:32063029 ~ 32065340	2	502	55 660.62	6.16	47.56	72.17	-0.485
<i>CsbHLH063</i>	<i>Cs5g30420</i>	XP_006479602.1	Chr5:32293153 ~ 32298441	6	294	31 303.42	5.85	55.82	79.35	-0.352
<i>CsbHLH064</i>	<i>Cs5g31400</i>	NP_001306996.1	Chr5:33059625 ~ 33065039	9	694	76 823.72	4.99	63.41	77.81	-0.590
<i>CsbHLH065</i>	<i>Cs5g34550</i>	XP_006480179.1	Chr5:35349068 ~ 35355417	9	552	60 168.33	6.36	59.43	49.75	-0.739
<i>CsbHLH066</i>	<i>Cs6g01760</i>	XP_006480339.1	Chr6:1164034 ~ 1170166	9	735	78 877.70	5.92	54.48	59.56	-0.654
<i>CsbHLH067</i>	<i>Cs6g08850</i>	XP_006481077.1	Chr6:10772274 ~ 10776602	6	762	83 838.39	8.53	36.54	83.95	-0.421
<i>CsbHLH068</i>	<i>Cs6g08890</i>	XP_015386693.1	Chr6:10827652 ~ 10828161	1	169	19 014.44	7.07	56.25	60.65	-0.615
<i>CsbHLH069</i>	<i>Cs6g09670</i>	XP_006481155.1	Chr6:11433090 ~ 11434775	4	155	17 004.12	4.42	40.35	103.68	-0.159
<i>CsbHLH070</i>	<i>Cs6g17010</i>	XP_006482067.1	Chr6:17611685 ~ 17613334	2	238	26 927.67	7.01	61.37	91.39	-0.405
<i>CsbHLH071</i>	<i>Cs6g17670</i>	XP_006482386.1	Chr6:18109071 ~ 18110995	5	291	32 206.78	6.33	43.70	63.44	-0.675
<i>CsbHLH072</i>	<i>Cs6g18610</i>	XP_006482245.1	Chr6:18652793 ~ 18654749	6	272	30 898.58	6.72	47.97	61.69	-0.733
<i>CsbHLH073</i>	<i>Cs6g19800</i>	XP_024955479.1	Chr6:19501582 ~ 19505259	8	321	35 727.62	5.93	62.05	58.97	-0.811
<i>CsbHLH074</i>	<i>Cs6g21120</i>	XP_006482579.1	Chr6:20344571 ~ 20346011	2	91	10 360.86	9.09	84.39	108.24	-0.413
<i>CsbHLH075</i>	<i>Cs7g02010</i>	XP_006482912.1	Chr7:637174 ~ 640227	8	427	46 736.13	8.25	47.49	55.76	-0.731
<i>CsbHLH076</i>	<i>Cs7g02770</i>	XP_015387149.2	Chr7:1067554 ~ 1070613	7	277	31 200.57	8.80	40.00	78.56	-0.488
<i>CsbHLH077</i>	<i>Cs7g04080</i>	XP_006483196.1	Chr7:1974323 ~ 1975567	2	93	10 447.76	7.94	93.56	101.72	-0.490
<i>CsbHLH078</i>	<i>Cs7g04700</i>	XP_006483289.1	Chr7:2407291 ~ 2411028	6	269	29 716.89	6.81	50.89	55.09	-0.911
<i>CsbHLH079</i>	<i>Cs7g08710</i>	XP_006484318.1	Chr7:5476982 ~ 5480292	8	471	52 989.34	6.20	45.08	87.37	-0.415
<i>CsbHLH080</i>	<i>Cs7g12020</i>	XP_006484263.1	Chr7:7982999 ~ 7986150	8	333	37 216.46	5.80	59.31	67.69	-0.848
<i>CsbHLH081</i>	<i>Cs7g12170</i>	XP_006484281.1	Chr7:8084872 ~ 8089360	8	530	56 656.02	7.67	55.02	57.28	-0.678
<i>CsbHLH082</i>	<i>Cs7g12800</i>	XP_006484405.1	Chr7:8762153 ~ 8764535	2	339	37 898.27	6.36	71.74	74.01	-0.776
<i>CsbHLH083</i>	<i>Cs7g16660</i>	XP_015387638.1	Chr7:12161801 ~ 12164688	8	553	60 023.36	6.06	50.93	61.28	-0.733
<i>CsbHLH084</i>	<i>Cs7g18810</i>	XP_015387762.2	Chr7:14753834 ~ 14759064	11	548	60 195.57	7.60	60.54	56.02	-0.673

续表 1

基因	基因ID	NCBI登录号	染色体位点	外显子	氨基酸/aa	分子量/Da	等电点PI	不稳定指数	脂肪系数	总平均疏水指数GRAVY
<i>CsbHLH085</i>	<i>Cs7g19780</i>	XP_006485105.1	Chr7:15844910 ~ 15848727	4	559	61213.50	5.27	46.09	71.36	-0.600
<i>CsbHLH086</i>	<i>Cs7g19870</i>	NP_001275808.1	Chr7:16034375 ~ 16037129	8	370	41 596.16	6.60	43.17	60.68	-0.802
<i>CsbHLH087</i>	<i>Cs7g25850</i>	XP_006485716.1	Chr7:26446270 ~ 26451930	13	565	61 688.58	9.01	52.29	56.81	-0.863
<i>CsbHLH088</i>	<i>Cs7g26900</i>	XP_006485801.1	Chr7:27606591 ~ 27608367	6	264	30 100.91	6.85	59.09	63.90	-0.663
<i>CsbHLH089</i>	<i>Cs7g28800</i>	XP_024957129.1	Chr7:29250518 ~ 29253102	10	294	33 375.59	6.18	49.69	82.48	-0.519
<i>CsbHLH090</i>	<i>Cs7g30210</i>	XP_006486517.1	Chr7:30315562 ~ 30317075	4	158	17 586.99	5.42	45.21	102.41	-0.263
<i>CsbHLH091</i>	<i>Cs7g30860</i>	XP_006486294.1	Chr7:30783226 ~ 30784310	2	279	30 988.10	8.85	50.87	81.18	-0.487
<i>CsbHLH092</i>	<i>Cs8g02500</i>	XP_006486853.2	Chr8:956307 ~ 958807	7	582	66 061.74	5.31	40.88	74.33	-0.701
<i>CsbHLH093</i>	<i>Cs8g03200</i>	XP_015388301.1	Chr8:1473900 ~ 1476267	5	316	35 293.30	8.37	38.39	83.92	-0.380
<i>CsbHLH094</i>	<i>Cs8g04230</i>	XP_024957794.1	Chr8:2207811 ~ 2209323	5	355	39 010.83	5.08	57.61	66.28	-0.784
<i>CsbHLH095</i>	<i>Cs8g04720</i>	XP_006486976.1	Chr8:2570230 ~ 2572077	4	271	28 457.50	7.62	43.63	60.52	-0.522
<i>CsbHLH096</i>	<i>Cs8g05450</i>	XP_024957714.1	Chr8:3021961 ~ 3025148	3	332	36 258.00	7.12	55.09	73.73	-0.497
<i>CsbHLH097</i>	<i>Cs8g06155</i>	XP_024957810.1	Chr8:3698024 ~ 3709725	6	506	56 718.42	8.22	44.99	78.85	-0.467
<i>CsbHLH098</i>	<i>Cs8g09040</i>	XP_015388486.1	Chr8:7283947 ~ 7287154	6	341	37 655.01	6.93	60.58	55.81	-1.005
<i>CsbHLH099</i>	<i>Cs8g15540</i>	XP_015388836.2	Chr8:18732957 ~ 18734368	4	367	41 201.67	5.85	42.62	75.94	-0.414
<i>CsbHLH100</i>	<i>Cs8g15580</i>	XP_006488632.1	Chr8:18749757 ~ 18751172	4	324	35 957.83	4.70	37.15	78.89	-0.537
<i>CsbHLH101</i>	<i>Cs8g15600</i>	XP_006488003.1	Chr8:18760668 ~ 18762319	4	313	35 039.51	4.92	35.71	80.06	-0.426
<i>CsbHLH102</i>	<i>Cs9g02055</i>	XP_006488826.1	Chr9:700263 ~ 701402	2	208	22 495.05	11.68	51.20	77.50	-0.540
<i>CsbHLH103</i>	<i>Cs9g06130</i>	XP_006489302.1	Chr9:3717457 ~ 3719074	3	196	21 995.65	8.67	53.00	112.35	-0.034
<i>CsbHLH104</i>	<i>Cs9g08460</i>	XP_006489519.1	Chr9:5668973 ~ 5670220	3	234	26 616.01	6.40	37.01	76.67	-0.718
<i>CsbHLH105</i>	<i>Cs9g10720</i>	XP_006489748.1	Chr9:8051079 ~ 8055585	5	214	24 461.45	5.52	63.87	67.10	-0.943
<i>CsbHLH106</i>	<i>Cs9g13930</i>	XP_006490036.1	Chr9:13382817 ~ 13385552	6	362	39 832.09	8.97	49.58	59.50	-0.741
<i>CsbHLH107</i>	<i>Cs9g13980</i>	XP_006490044.1	Chr9:13457192 ~ 13462072	7	406	44 880.51	6.41	49.24	68.77	-0.850
<i>CsbHLH108</i>	<i>Cs9g14850</i>	XP_024948017.1	Chr9:14306245 ~ 14307759	3	355	39 877.78	5.56	48.54	78.85	-0.517
<i>CsbHLH109</i>	<i>Cs9g16390</i>	XP_006490279.1	Chr9:15786930 ~ 15788499	3	245	27 731.30	6.18	39.80	88.65	-0.572
<i>CsbHLH110</i>	<i>Cs9g16400</i>	XP_024948035.1	Chr9:15795755 ~ 15800802	4	177	19 652.50	6.96	66.03	96.21	-0.324
<i>CsbHLH111</i>	<i>Cs9g16410</i>	XP_006490280.1	Chr9:15803130 ~ 15804325	3	240	27 205.81	7.18	57.37	90.54	-0.599
<i>CsbHLH112</i>	<i>Cs9g18490</i>	XP_006490681.1	Chr9:17652125 ~ 17655055	3	337	36 927.47	5.07	61.99	84.18	-0.443
<i>CsbHLH113 orange1.lt05043</i>	XP_015382245.2	ChrUn:58 ~ 2440	5	164	18 315.86	6.22	58.36	98.11	-0.340	
<i>CsbHLH114 orange1.lt05049</i>	XP_024950208.1	ChrUn:20416 ~ 22480	3	164	18 396.24	9.05	48.90	97.44	-0.353	
<i>CsbHLH115 orange1.lt04234</i>	XP_006494600.1	ChrUn:76606 ~ 80879	4	361	39 829.33	4.73	55.14	79.47	-0.570	
<i>CsbHLH116 orange1.lt00185</i>	XP_006490847.1	ChrUn:79952 ~ 85540	6	464	50 763.57	8.51	60.71	55.13	-0.643	
<i>CsbHLH117 orange1.lt03288</i>	XP_006493805.1	ChrUn:105807 ~ 108426	5	424	48 095.93	5.40	57.49	76.34	-0.659	
<i>CsbHLH118 orange1.lt03044</i>	XP_024949351.1	ChrUn:161136 ~ 166537	11	730	81 007.77	5.77	47.77	86.52	-0.309	
<i>CsbHLH119 orange1.lt02482</i>	XP_006493102.1	ChrUn:172555 ~ 174155	3	220	25 382.15	9.11	46.57	88.14	-0.442	
<i>CsbHLH120 orange1.lt02367</i>	XP_006492985.1	ChrUn:330505 ~ 333114	8	481	52 708.85	6.22	48.75	63.70	-0.658	
<i>CsbHLH121 orange1.lt03173</i>	XP_024949417.1	ChrUn:405118 ~ 408316	7	408	45 947.03	6.90	62.55	57.65	-0.953	
<i>CsbHLH122 orange1.lt01783</i>	XP_006492435.1	ChrUn:917553 ~ 921576	8	508	56 338.16	9.59	62.54	62.40	-0.751	
<i>CsbHLH123 orange1.lt00363</i>	XP_006491098.1	ChrUn:1209554 ~ 1213572	8	656	73 463.98	5.11	54.09	82.65	-0.420	
<i>CsbHLH124 orange1.lt00550</i>	XP_015389693.1	ChrUn:2597120 ~ 2599250	1	515	57 142.18	5.38	61.44	78.19	-0.544	
<i>CsbHLH125 orange1.lt00575</i>	XP_024948228.1	ChrUn:2799007 ~ 2802182	4	437	49 594.70	5.92	67.63	84.76	-0.378	
<i>CsbHLH126 orange1.lt01021</i>	XP_006491734.1	ChrUn:15916761 ~ 15919426	1	685	74 655.74	5.45	53.09	65.93	-0.621	

根据基因在染色体上的位置，将甜橙 126 个 *bHLH* 基因分别命名为 *CsbHLH001* ~ *CsbHLH126*，编码氨基酸数从 79 (*CsbHLH003*) ~ 762 (*CsbHLH067*)，平均长度为 359 个氨基酸。通过 Ex-pasy 预测其分子量大小为 9.40 ~ 83.84 kD，等电点为 4.42 (*CsbHLH069*) ~ 11.68 (*CsbHLH102*)。从以上理化性质来看，虽然 bHLH TFs 的保守结构域中存在碱性区域，但酸性氨基酸富集于蛋白质分子中，因此 65.87% 的甜橙 bHLH 家族蛋白等电点小于 7，大部分呈弱酸性，此现象与拟南芥<sup>[14]</sup> 及水稻<sup>[13]</sup> 相同。甜橙中 bHLH 家族蛋白总平均亲水性在 -1.005 ~ -0.034 之间，均为负值，说明该家族蛋白均属于亲水性蛋白；不稳定指数介于 35.17 ~ 93.56 之间，其中 94.44% 是不稳定蛋白 ( $II > 40$ )，只有 7 个属于稳定蛋白 ( $II < 40$ )；脂肪指数介于 49.75 ~ 112.35 之间；通过在线软件 TargetP 和 CELLO v2.5 对 126 个甜橙 bHLH 蛋白进行亚细胞定位，结果显示大部分分布在细胞质、细胞核、叶绿体以及线粒体基质上。甜橙 126 条 bHLH 蛋白经 SignalP 预测分析，均不存在信号肽，属于非

分泌型蛋白。

## 2.2 甜橙 *bHLH* 基因的染色体定位与分布

bHLH 转录因子在甜橙连锁群上的分布如图 1 所示。*CsbHLH001* ~ *CsbHLH112* 在 9 条染色体中呈不均匀分布，“ChrUn”表示的是 *CsbHLH113* ~ *CsbHLH126* 基因存在于连锁群上，但不能定位到连锁群的骨架，还需进一步的基因组测序信息才能确定；其中 7 号染色体上分布最多为 17 条，4 号染色体最少为 7 条。一般认为，基因之间的距离小于 100 kb 的即为串联复制，分析发现，甜橙染色体上存在基因串联复制现象的有 Chr1 上的 *CsbHLH010* 和 *CsbHLH011*；Chr2 上的 *CsbHLH016* 和 *CsbHLH017*，*CsbHLH026* 和 *CsbHLH027*；Chr3 上的 *CsbHLH039*、*CsbHLH040* 和 *CsbHLH041*；Chr5 上的 *CsbHLH052* 和 *CsbHLH053*；Chr6 上的 *CsbHLH067* 和 *CsbHLH068*；Chr7 上的 *CsbHLH080* 和 *CsbHLH081*；Chr8 上的 *CsbHLH099*、*CsbHLH100* 和 *CsbHLH101*；Chr9 上的 *CsbHLH106* 和 *CsbHLH107*，*CsbHLH109* 和 *CsbHLH110*。

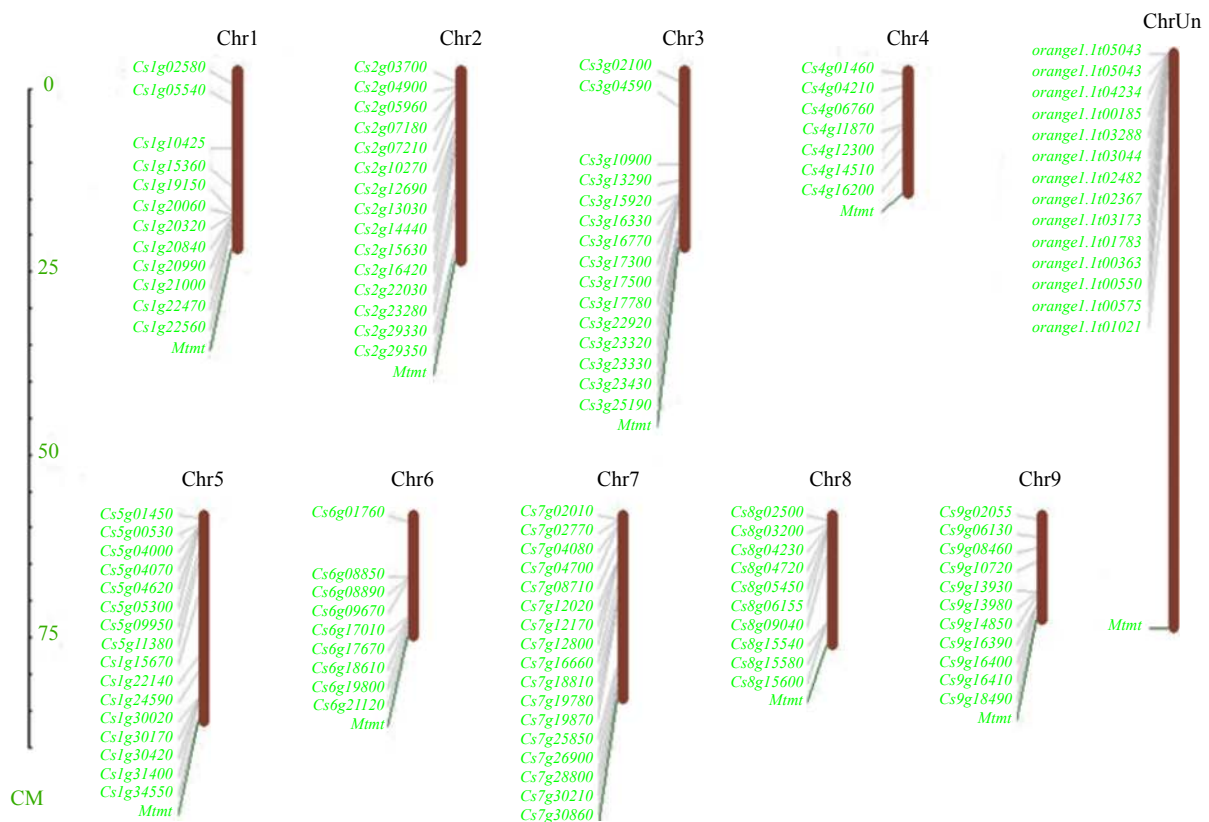


图 1 甜橙 *bHLH* 基因在染色体上的定位

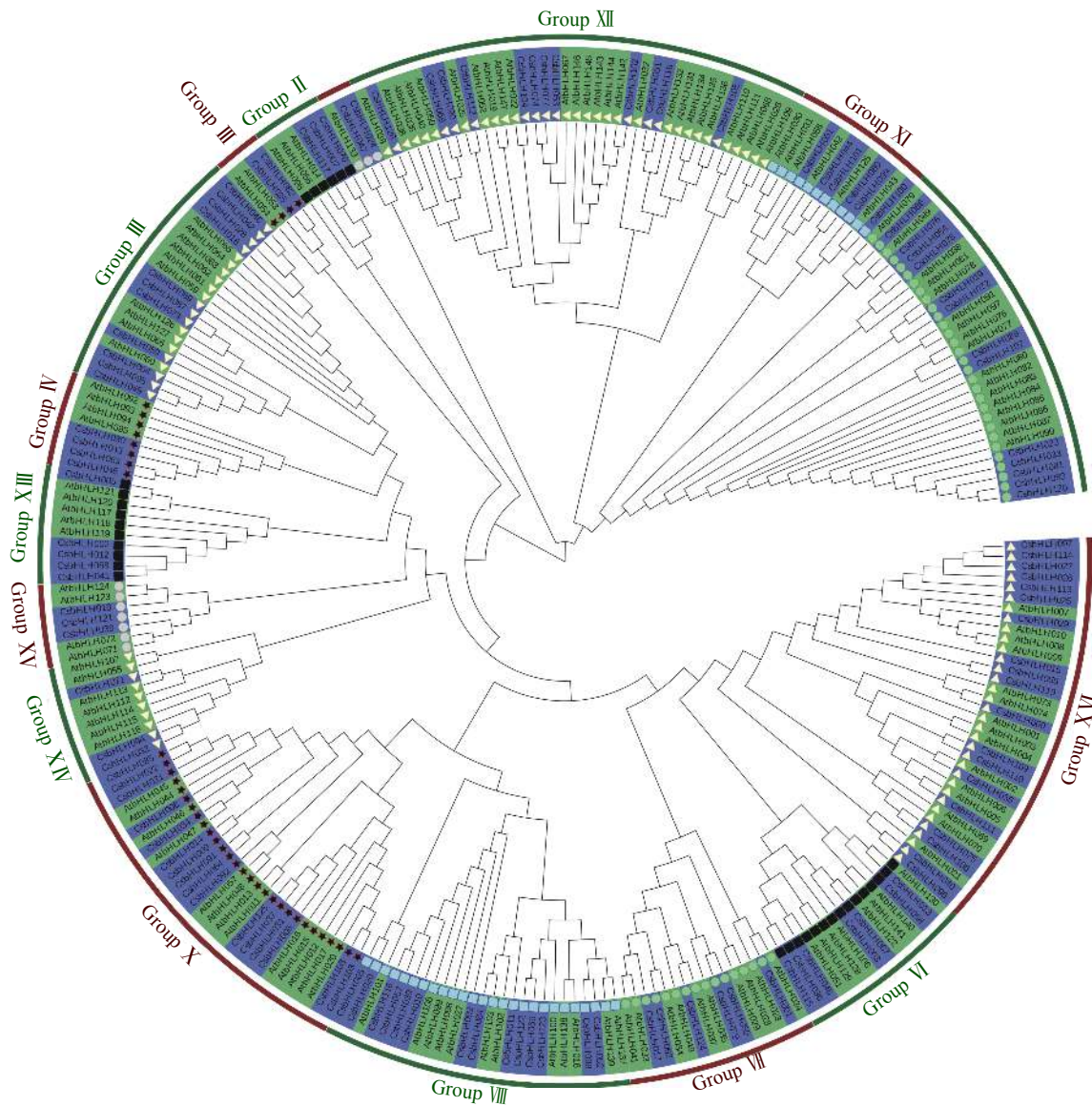
Fig. 1 Chromosomal locations of *C. sinensis* *bHLH* gene family



### 2.3 甜橙与拟南芥 *bHLH* 基因家族的进化关系

为了研究甜橙 *bHLH* 基因家族的进化关系, 利用鉴定出的 126 条甜橙 *bHLH* 蛋白序列与 144 条拟南芥 *bHLH* 蛋白序列构建进化树 (图 2)。根据拟南芥中 *bHLH* TFs 的分类系统和甜橙 *bHLH* TFs 进化树的分类情况, 全部成员分为 16 个亚族, 其中第 XVI 亚族中的 *bHLH* 成员数量最多, 共 19 个, 第 I、V 亚族中的成员数量最少, 仅为

2 个。甜橙与拟南芥在第 I、II、III、IV、V、VI、XI、XIII 和 XV 共 9 个亚族中的成员数量相同或相似, 差异最大的是第 XII 亚族, 甜橙 *bHLH* 数量不及拟南芥的 50%, 而第 I、III 和 XVI 亚族中的 *bHLH* 成员数量均多余拟南芥, 以上结果表明甜橙 *bHLH* 家族基因历史演变的过程中发生了基因扩增或收缩。



不同分支的 *bHLH* 转录因子分别用颜色不同的正方形□、圆形○、三角形△和五角星☆表示。

图 2 甜橙和拟南芥 *bHLH* 基因家族的进化树

Fig. 2 Phylogenetic tree based on *C. sinensis* and *arabidopsis bHLH* transcription factors

### 2.4 甜橙 *bHLH* 家族基因保守基序分析

通过多序列比对和 SMART 分析发现, 甜橙 126 个蛋白序列均含有 *bHLH* 结构域。根据鉴定

出的甜橙 *CsbHLH* 基因的保守结构域序列, 构建甜橙系统进化拓扑图, 得到的结果与图 2 中所示的进化关系是基本一致的。甜橙 *bHLH* 保守基序

进行分析结果如图 3 所示，每一个亚家族之间的转录因子尽管它们的长度差别很大，但是它们的保守基序长度及位置都很相似。motif1-10 在图中用不同的颜色表示，10 个保守基序的详细信息如表 2 所示。除 bHLH 保守结构域外，bHLH 转录因子家族还含有其他保守型的结构域，但 Motif1 和 Motif2 总是相邻，两者共同构成了 bHLH 结构域。这些保守元件是转录因子发挥其调控作用所必需的，对于维持蛋白质的三维构象和与靶基因的结合至关重要。

### 2.5 甜橙 bHLH 蛋白互作关系

甜橙 bHLH 蛋白互作预测分析绘制出调控图如 (图 4) 所示。在植物生长和次生代谢中 bHLH TFs 起重要作用，而且多种植物胁迫反应中也有 bHLH TFs 的参与。在甜橙中 bHLH104 与 ILR3 可以协同作用调控 *PYE* 基因的表达，而 *PYE* 是植物铁吸收相关的重要转录因子，拟南芥中转录调控因子 MYC2 和 MYC4 协同调控了拟南芥的开花诱导，甜橙中 MYC2 和 MYC4 也具有协同调控作用。ICE1 与 SPCH、MUTE、FAM 和 bHLH093 相互作用，正调控气孔的发育。AMS、DYT1 和 AT2G31210 之间具有协同作用，且 AMS 和 DYT1 对植物花器官发生、形成、生长发育有调控作用，因此三者可能协同调控甜橙花器官的发生、形成及生长发育。甜橙中光敏色素互作蛋白 PIL5、PIL6、PIF3 和 PIF7 共同作用来调节光形态发生。

### 2.6 甜橙 bHLH 基因在不同器官表达情况

对甜橙 126 个 bHLH 基因在愈伤组织、花、叶、果实、混合样 (Mixture1: 幼果果肉果皮, Mixture2: 生长中期的果肉果皮, Mixture3: 成熟期果肉果皮) 的表达芯片数据进行了分析。结果如图 5 所示, *CsbHLH013*、*CsbHLH047*、*CsbHLH062*、*CsbHLH105*、*CsbHLH02054*、*CsbHLH043*、*CsbHLH126*、*CsbHLH063*、*CsbHLH059*、*CsbHLH075*、*CsbHLH061*、*CsbHLH115*、*CsbHLH073* 13 个基因在愈伤组织、花、叶、果实及混合样均大量表达。其中 *CsbHLH043* (LRL1) 在愈伤组织中的表达量最大, *CsbHLH124* (MYC4) 在花和叶片中的表达量均最大, *CsbHLH115* 在果实中的表达量最大, 在组织中高度表达的 *CsbHLH* 基因可能具有相似的作用。

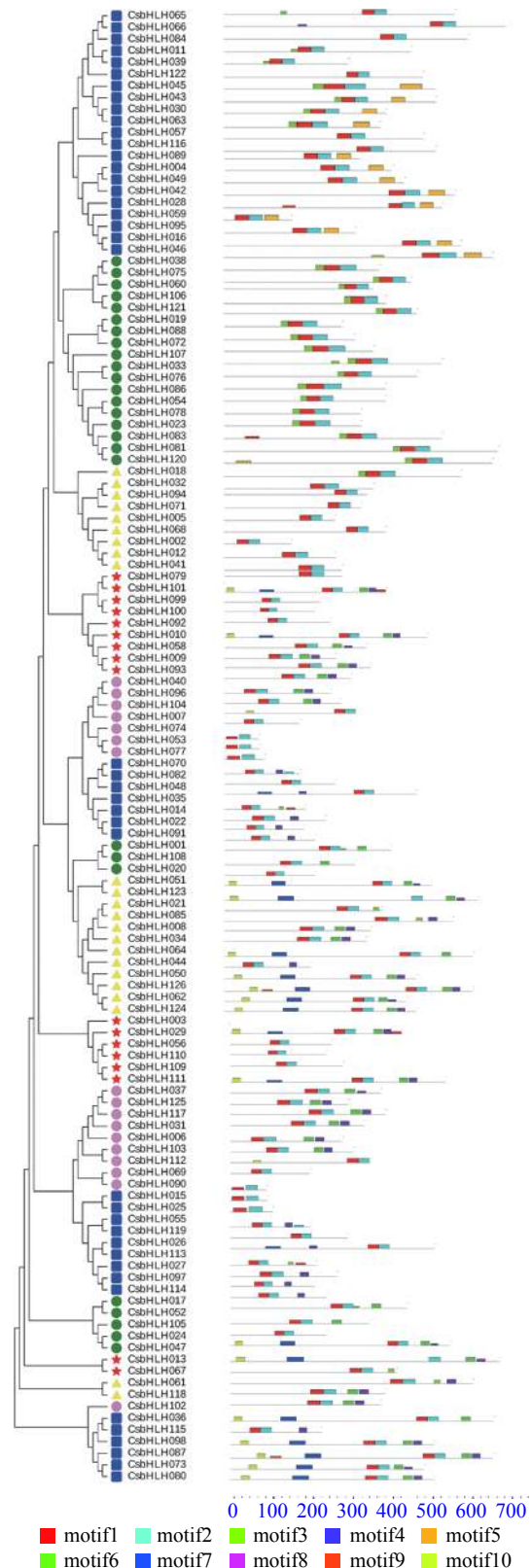


图 3 甜橙 bHLH 转录因子的保守基序分析  
Fig. 3 The conserved motifs analysis of *C. sinensis* bHLH transcription factors

表2 甜橙 *bHLH* 基因的保守基序

Table 2 The conserved motifs of *C. sinensis* *bHLH* genes

序号	最佳匹配	期望值	含该基序的 CsbHLH蛋白数	含该基序的CsbHLH蛋白
1	TDSHSIAE[RK]RRRE[KR][ILM][NS][ED]R [LFM]KAL[QR]SLVP[NG]X	$1.5 \times 10^{-169}$	116	19,86,78,54,120,2,23,83,30,88,43,107,81,33,45,66,63,121,106,84,7 6,38,94,65,32,71,62,41,57,64,126,75,122,12,50,124,11,10,5,20,116, 85,21,93,99,68,2,60,39,117,44,14,34,8,92,87,37,80,82,73,112,58,22 9,31,100,125,101,91,35,48,7,79,3,1,6,108,70,97,27,114,113,46,40, 26,51,59,16,96,18,104,25,56,24,15,105,109,55,110,47,29,4,119,49, 118,61,111,98,42,36,95,89,52,53,115,13
2	[KS]K[TM]DKAS[IMV]LD[ED]A[IV]NY[IV] K[EF]LQXQV[EQK]XL[SE]XK	$2.3 \times 10^{-184}$	121	45,43,11,39,9,41,12,116,84,66,120,93,86,81,42,83,64,63,33,30,28,2 3,2,5,59,10,46,78,54,101,16,49,122,68,5,4,57,92,76,88,38,85,21,12 1,62,51,95,94,19,72,32,100,79,107,8,126,87,20,1,65,106,80,75,71,1 08,89,73,124,6,99,44,96,123,112,40,103,114,97,31,27,125,56,26,34 58,82,113,117,110,98,74,47,14,35,60,91,77,69,24,18,53,48,37,7,10 9,90,70,29,22,15,119,36,105,25,104,13,111,55,52,102,61
3	PP[KS][DV]YI[HR]VRA[RK]RGQA	$2.8 \times 10^{-115}$	27	120,86,83,54,78,81,23,19,76,30,33,72,63,107,38,88,121,106,60,75, 43,45,85,2,47,65,35
4	[KE][IL][IV][SE][AV]LE[ES]LGL[E][VI][LIV] [NH][A][NS][VI][ST]T[VF]	$8.3 \times 10^{-84}$	40	112,55,103,119,69,126,117,34,27,79,25,85,6,93,92,44,50,114,26,8, 91,124,37,125,90,22,70,15,96,31,9,40,113,110,14,17,56,58,80,48
5	[KS]LME[EK]DX[EG][EA][AP][KM][QK]DL [RQ]S[RK]GLCL[VM]P[IV]S[CL][TA]TA[VI] [LV][PA][DE][VI][EDV]V[KR][IV]X[GED]S	$4.6 \times 10^{-77}$	13	43,46,63,16,30,49,45,42,28,4,95,59,89
6	D[AV]LI[KR][IV]H[CS][PE][KR][KR]PG [VI]WL[TCS][GN]A[HQ]E[LA][DQ]S[KY]V[ CF][SE]R[AS][KL][LE]A[KQ]S[AH][GS]IQT[ VL]VC[IFV]P[HT][LA][CD]GVVELG [EN][PL][PT]V[MIV][NE][IL][SHRT][TDE][S	$2.7 \times 10^{-80}$	29	10,9,126,58,117,125,93,40,50,37,31,6,96,51,123,108,103,8,1,124,9 2,112,34,79,64,52,62,69,29
7	CF][SE]R[AS][KL][LE]A[KQ]S[AH][GS]IQT[ VL]VC[IFV]P[HT][LA][CD]GVVELG [EN][PL][PT]V[MIV][NE][IL][SHRT][TDE][S	$1.6 \times 10^{-65}$	7	64,51,123,124,50,126
8	LM]GS[TS][LI]EV[NIV]L[IT][CS]G[LV][ND Q]R[NQ]FM[FL][HY]	$8.8 \times 10^{-64}$	9	27,26,114,113,119,55,15,25,97
9	GD[RK]V[IF][FH]SVQ[SA][KQ]A[IK][NI]SR [IL]GVET[ES]RV[QC][EQ]RL[KH]	$1.8 \times 10^{-44}$	6	27,114,26,113,55,119
10	[LQ]RLQLXVAVRSEQWXYA[IV]FW[KQ]	$1.3 \times 10^{-38}$	11	123,124,51,50,61,118,126,79,64,62,92

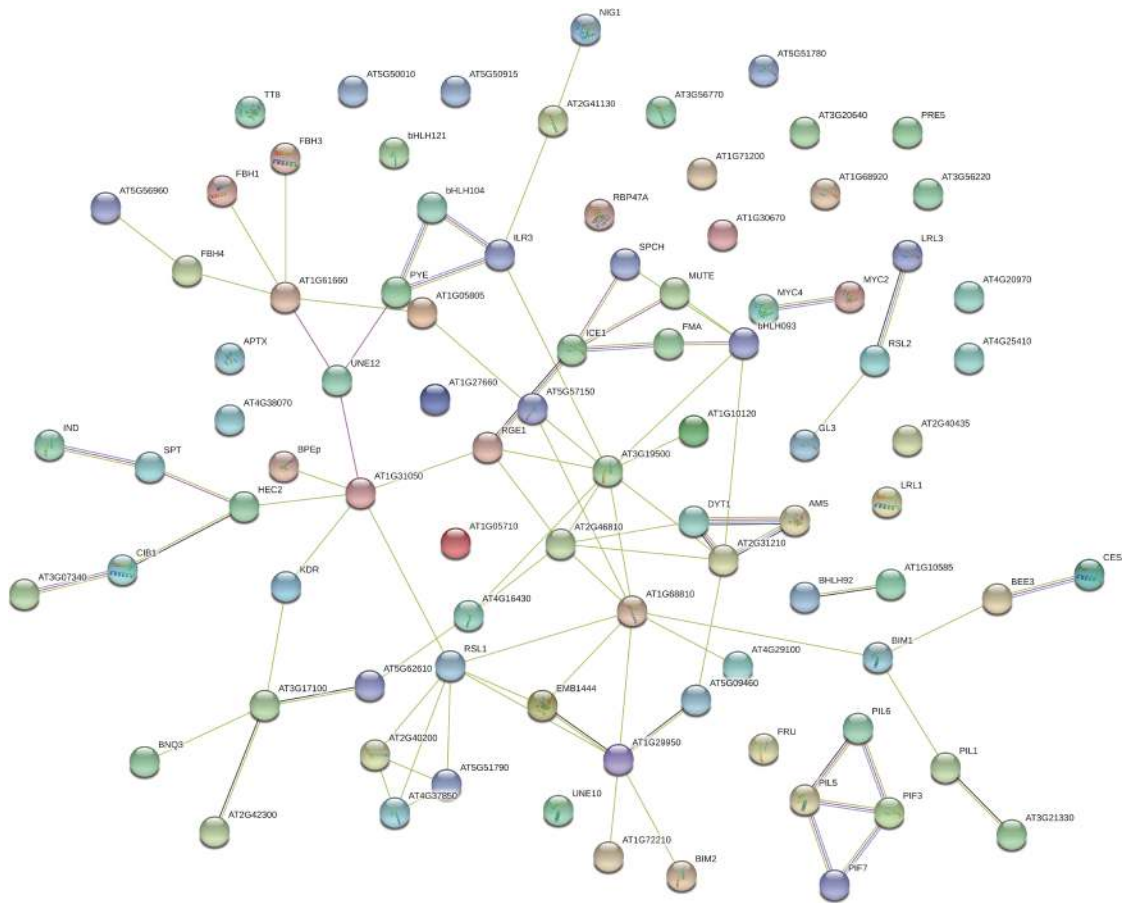


图4 甜橙 CsbHLHs 蛋白功能关系网络

Fig. 4 CsbHLHs protein function control network of *C. sinensis*



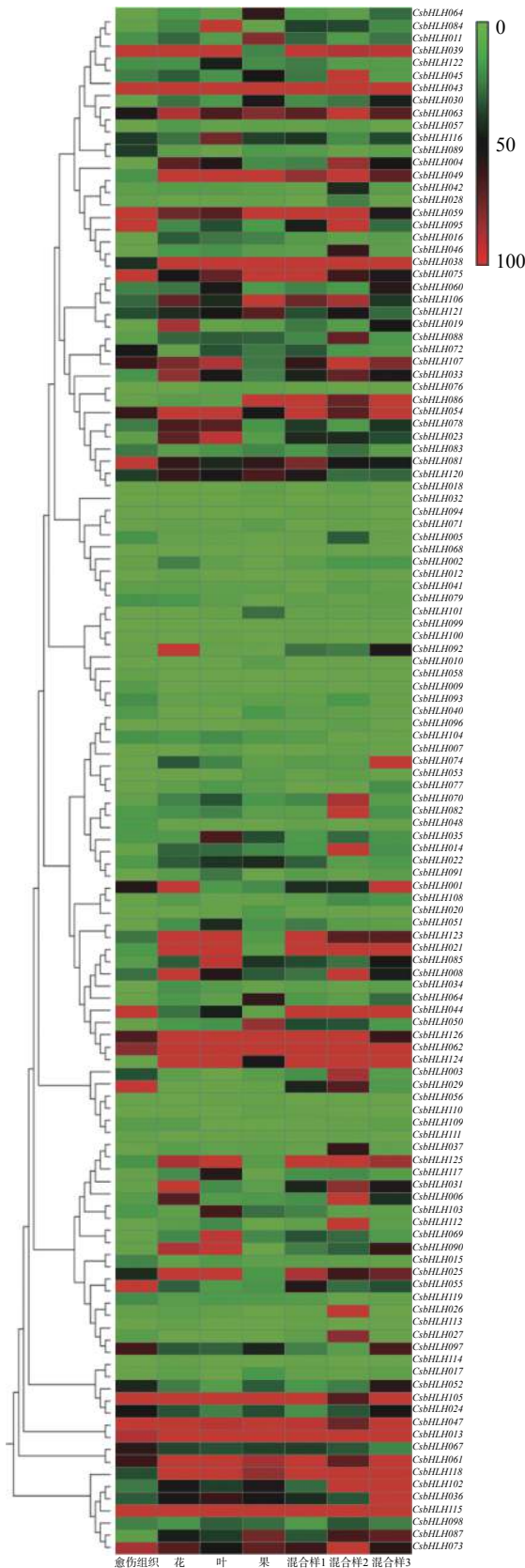


图 5 甜橙 *bHLH* 基因在不同组织的表达分析  
 Fig. 5 Expression analysis of *bHLH* gene in different tissues of *C. sinensis*

### 3 结论与讨论

*bHLH* 家族在真核生物转录因子家族中是一个大家族，并在生物的生长发育、生理及新陈代谢过程中发挥着一系列的重要调节作用<sup>[32]</sup>。已经在诸如苹果 (*Malus domestica*)<sup>[33]</sup>、樱桃 (*Cerasus pseudocerasus*)<sup>[34]</sup>、短柄草<sup>[19]</sup> 和拟南芥<sup>[14]</sup> 等植物中鉴定了 *bHLH* 家族的转录因子，并对其生长与生物和非生物胁迫之间的关系也做了研究。耿晶晶的研究报道鉴定出 56 个甜橙 *bHLH* 家族成员，并分析了其在低温胁迫下的表达模式<sup>[35]</sup>。在本研究中，利用生物信息学的手段对甜橙中的 *bHLH* TFs 进行鉴定，并对其理化性质、保守结构域、系统进化关系、染色体上的位置分布及在不同组织中的表达情况等方面进行了分析，为 *bHLH* 基因家族在基因表达调控、结构和功能等研究提供参考，为揭示 *bHLH* 家族基因参与植物生长发育机制的调控提供参考。

通过分析验证从甜橙中鉴定出了 126 个 *bHLH* 基因 (表 1)，根据基因在染色体上的位置，将柑橘 126 个 *bHLH* 基因分别命名为 *CsbHLH001* ~ *CsbHLH126*，*bHLH* 保守结构域是由 HLH 区和碱性氨基酸区和两部分组成。研究者发现，在番茄 *bHLH* 转录因子家族中同样存在高度保守的 Glu9、His5、Arg13 及 Arg10 等氨基酸残基，在 HLH 区也存在高度保守的氨基酸残基 Leu23 和 Leu64<sup>[18]</sup>。HLH 区的保守氨基酸残基与二聚体形成相关，而碱性氨基酸区保守氨基酸残基与靶基因的结合紧密相关。

基因结构和染色体定位分析发现，甜橙 *bHLH* 基因家族内含子数量差异较大，基因在染色体上的分布也不均匀 (图 1)，基因复制是基因重组以及扩增的最主要途径，其主要包括串联重复和大片段复制 2 种方式<sup>[36]</sup>。串联重复是指在同一条染色体的 2 个相邻基因，大片段复制事件发生在不同染色体上<sup>[37]</sup>。本研究中，126 个 *CsbHLH* 基因中的 22 个 (17.4%) 参与到串联重复事件中；虽然基因复制事件不多，但依然可以推测，基因复制是甜橙 *bHLH* 基因家族扩增的主要模式。串联复制的序列与保守区域非常的相似，进化树中的它们的亲缘关系也很近，因此推测，在进化的过程中他们的功能也类似，但还有待进一步的研究验证。通过在线软件 MEME 对甜橙 126 条 *bHLH* 基因的保守基序进行分析，从中找到了 10 个高度保守氨基酸基序，保守基序分析显

示,同一亚家族中的大多数保守基序相似,表明每个亚家族中编码蛋白的功能是稳定的, Motif1和 Motif2 几乎存在于所有的 CsbHLH 蛋白中,且总是相邻,共同构成 bHLH 结构域,综上,所有这些 bHLH 保守基序在同一亚家族的独特性及保守性也是对甜橙 *bHLH* 基因家族的进化分类的佐证,同时推测,除 bHLH 结构域外的其他保守基序也是每个亚家族发挥其相应功能的关键。

为了研究甜橙 *bHLH* 基因家族的进化关系,本研究把 126 个 bHLH 划分为 16 个亚族,同一亚家族内的成员大都具有相同的内含子模式和保守基序。bHLH TFs 家族成员数量巨大,对其进行分类有一定难度。bHLH TFs 参与植物多种调控代谢活动。

已发现植物开花过程是由 SPT 和 HEC 编码基因共同调控的<sup>[22]</sup>, Zhang 等<sup>[38]</sup>通过拟南芥突变株 DYT1 花药发育不良证实 DYT1 是花药发育的必须调控因子,可以与 bHLH010, bHLH089 和 bHLH090 通过互作来调节拟南芥的花药发育和转录表达<sup>[39]</sup>。预测分析发现 AMS、DYT1 和 AT2G31210 可能协同调控甜橙花器官的发生、形成及生长发育,而 MYC2 和 MYC4 协同调控甜橙的开花过程。ICE1 调控 *CBF/DRB1* 基因的冷诱导转录,与 CBF3/DREB1A 启动子中发现的 MYC 识别位点 (5'-CANNTG-3') 特异性结合,通过与 SPCH、MUTE 和 FAMA 等的相互作用控制气孔细胞分化及保卫细胞的形成<sup>[40]</sup>。本研究中 ICE1 与 SPCH、MUTE、FAM 和 bHLH093 相互作用,对甜橙气孔的发育有正调控作用。根据 Huq 和 Peter H.Quail 的研究结果, bHLH 转录因子 HFR1、PIF3 和 PIF4 在光敏色素信号传导中发挥调节作用<sup>[41]</sup>。甜橙中光敏色素互作蛋白 PIL5、PIL6、PIF3 和 PIF7 共同作用来调节光形态发生。参与拟南芥花色素苷合成调控的 bHLH 转录因子包括 TT8、EGL3、GL3 和 MYC1 TF 家族<sup>[11]</sup>, Zimmermann 等人发现它们都与 R2R3-MYB 转录因子相互作用,包括 PAP1 (AtMYB75) 和 PAP2 (AtMYB90), TT8 与 MYB75/PAP1 或 MYB90/PAP2 相互作用,参与调控黄酮类色素沉着,在调节无色花色色素还原酶 (BANYULS) 和二氢黄酮-4-还原酶 (DFR) 中起关键作用<sup>[42]</sup>。分析与甜橙发育相关的组织表达数据,为 *CsbHLH* 基因家族提供了初步的器官特异性表达数据,为进一步阐明这些基因在甜橙的进化发育中的作用奠定了基础。

#### [参 考 文 献]

- [1] Yanagisawa S. Transcription factors in plants: Physiological functions and regulation of expression [J]. *Journal of Plant Research*, 1998, 111(3): 363–371.
- [2] Riechmann J L, Heard J, Martin G, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes [J]. *Science*, 2000, 290(5499): 2105–2110.
- [3] 刘卫群,王永亮,郭红祥,等.烟草 DREBP 转录因子结合 DRE 元件的关键氨基酸 [J]. *中国生物化学与分子生物学报*, 2006, 22(2): 111–116.
- [4] Amoutzias G D, Veron A, Weiner J, et al. One billion years of bZIP transcription factor evolution: conservation and change in dimerization and DNA-binding site specificity [J]. *Molecular Biology and Evolution*, 2007, 24(3): 827–835.
- [5] Guo X J, Wang J R. Global identification, structural analysis and expression characterization of bHLH transcription factors in wheat [J]. *BMC Plant Biology*, 2017, 17(1): 1–12.
- [6] Yamasaki K, Kigawa T, Seki M, et al. DNA-binding domains of plant-specific transcription factors: structure, function, and evolution [J]. *Trends in Plant Science*, 2013, 18(5): 267–276.
- [7] Kosugi S, Ohashi Y. DNA binding and dimerization specificity and potential targets for the TCP protein family [J]. *The Plant Journal*, 2002, 30(3): 337–348.
- [8] Robinson K A, Koepke J I, Kharodawala M, et al. A network of yeast basic *Helix-loop-Helix* interactions [J]. *Nucleic Acids Research*, 2000, 28(22): 4460–4466.
- [9] Ledent V, Vervoort M. The basic *Helix-loop-Helix* protein family: Comparative genomics and phylogenetic analysis [J]. *Genome Research*, 2001, 11(5): 754–770.
- [10] Atchley W R, Fitch W M. A natural classification of the basic *Helix-loop-Helix* class of transcription factors [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1997, 94(10): 5172–5176.
- [11] 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 & Evolution*, 2003,



- 20(5): 735–747.
- [12] Jones S. An overview of the basic *Helix-loop-Helix* proteins [J]. *Genome Biology*, 2004, 5(6): 1–6.
- [13] Li X X, Duan X P, Jiang H X, 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.
- [14] 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.
- [15] Baudry A, Caboche M, Lepiniec L. TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell - specific accumulation of flavonoids in *Arabidopsis thaliana* [J]. *The Plant Journal*, 2006, 46(5): 768–779.
- [16] 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.
- [17] Simionato E, Ledent V, Richards G, et al. Origin and diversification of the basic *Helix-loop-Helix* gene family in metazoans: insights from comparative genomics [J]. *BMC Evolutionary Biology*, 2007, 7(1): 1–18.
- [18] Wang J Y, Hu Z Z, Zhao T M, et al. Genome-wide analysis of bHLH transcription factor and involvement in the infection by yellow leaf curl virus in tomato(*Solanum lycopersicum*) [J]. *BMC Genomics*, 2015, 16(1): 39.
- [19] Niu X, Guan Y X, Chen S K, et al. Genome-wide analysis of basic *Helix-loop-Helix* (bHLH) transcription factors in *Brachypodium distachyon* [J]. *BMC Genomics*, 2017, 18(1): 1–20.
- [20] Gao C, Sun J L, Wang C Q, et al. Genome-wide analysis of basic/*Helix-loop-Helix* gene family in peanut and assessment of its roles in pod development [J]. *PLoS One*, 2017, 12(7): e0181843.
- [21] Zhang X, Luo H M, Xu Z C, et al. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in *Salvia miltiorrhiza* [J]. *Scientific Reports*, 2015, 5: 11244.
- [22] Gremski K, Ditta G S, Yanofsky M F. The HECATE genes regulate female reproductive tract development in *Arabidopsis thaliana* [J]. *Development*, 2007, 134(20): 3593–3601.
- [23] Ni M, Tepperman J M, Quail P H. PIF<sub>3</sub>, a Phytochrome-Interacting factor necessary for normal photoinduced signal transduction, is a novel basic *Helix-loop-Helix* protein [J]. *Cell*, 1998, 95(5): 657–67.
- [24] 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]. *The Plant Cell*, 1997, 9(10): 1859–1868.
- [25] Abe H, Urao T, Ito T, et al. *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling [J]. *The Plant Cell*, 2003, 15(1): 63–78.
- [26] 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.
- [27] Goodrich J, Carpenter R, Coen E S. A common gene regulates pigmentation pattern in diverse plant species [J]. *Cell*, 1992, 68(5): 955–964.
- [28] 邓秀新, 彭抒昂. 柑橘学 [M]. 北京: 中国农业出版社, 2013: 15–84.
- [29] Bailey PC, Martin C, Toledo-Ortiz G, et al. Update on the basic *Helix-loop-Helix* transcription factor gene family in *Arabidopsis thaliana* [J]. *The Plant Cell*, 2003, 15(11): 2497–2502.
- [30] Chen Z, Chen X, Yan H W, et al. The lipoxygenase gene family in poplar: identification, classification, and expression in response to MeJA treatment [J]. *PLoS one*, 2015, 10(4): e0125526.
- [31] Guo C L, Guo R R, Xu X Z, et al. Evolution and expression analysis of the grape (*Vitis vinifera* L.) WRKY gene family [J]. *Journal of Experimental Botany*, 2014, 65(6): 1513–1528.
- [32] Vanstraelen M, Benková E. Hormonal interactions in the regulation of plant development [J]. *Annual Review of Cell and Developmental Biology*, 2012, 28: 463–487.
- [33] 杨金华. 苹果 bHLH 转录因子家族的鉴定及表达分析 [D]. 杨凌: 西北农林科技大学, 2017.
- [34] 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.
- [35] 耿晶晶. 甜橙 bHLH 家族转录因子发掘及 Cs bHLH18 抗寒功能鉴定与作用机制解析 [D]. 武汉: 华中农业大学. 2018.
- [36] Vision T J, Brown D G, Tanksley S D. The origins of

- genomic duplications in *Arabidopsis* [J]. *Science*, 2000, 290(5499): 2114–2117.
- [37] Liu Y, Jiang H Y, Chen W J, et al. Genome-wide analysis of the auxin response factor (ARF) gene family in maize (*Zea mays*) [J]. *Plant Growth Regulation*, 2011, 63: 225–234.
- [38] Zhang W, Sun Y J, Timofejeva L, et al. Regulation of *Arabidopsis* tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative b HLH transcription factor [J]. *Development*, 2006, 133(16): 3085–3095.
- [39] Zhu E G, You C J, Wang S S, et al. The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for *Arabidopsis* anther development and transcriptome [J]. *The Plant Journal*, 2015, 83(6): 976–990.
- [40] Macalister C A, Ohashi-Ito K, Bergmann D C. Transcription factor control of asymmetric cell divisions that establish the stomatal lineage [J]. *Nature*, 2007, 445(7127): 537–540.
- [41] Huq E, Quail P H. PIF4, a phytochrome-interacting bHLH factor, functions as a negative regulator of phytochrome B signaling in *Arabidopsis* [J]. *The EMBO Journal*, 2002, 21(10): 2441–2450.
- [42] Zimmermann I M, Heim M A, Weisshaar B, et al. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins [J]. *The Plant Journal*, 2004, 40(1): 22–34.

(责任编辑 张 坤)

