

# 绿豆 bHLH 转录因子家族的鉴定与生物信息学分析

陈红霖, 胡亮亮, 王丽侠, 王素华, 程须珍

(中国农业科学院作物科学研究所, 北京 100081)

**摘要:** bHLH 是真核生物中重要的一类转录因子, 其主要由碱性氨基酸区和螺旋-环-螺旋区组成。本研究利用生物信息学的方法鉴定到 122 个绿豆 bHLH 转录因子, 并对其理化性质、保守结构域、基因结构、在染色体上的分布、系统进化以及部分典型基因的组织表达差异等进行分析。结果表明, bHLH 转录因子理化性质差异较大; 含有 2 个保守结构域, 分别位于 N 端的碱性氨基酸区和 C 端的螺旋-环-螺旋区, 碱性氨基酸区含有 His5-Glu9-Arg13 保守序列, 与靶基因结合有关, HLH 区含有 Arg23 和 Arg55, 与形成二聚体有关, 同时含有 5 种保守元件; bHLH 基因在 11 条染色体上分布不均匀, 5 号、7 号和 8 号染色体上分布较多, 1 号、4 号和 10 号染色体上分布较少, 大部分基因含有 1~9 个不等的内含子, 在染色体上成簇状分布; 122 个 bHLH 转录因子可分为 11 个亚家族。多数 bHLH 基因在绿豆根、茎、叶、花和种子等组织中均有表达, 但具有组织表达特异性, 且不同基因表达量差异较大。本研究为进一步研究绿豆 bHLH 转录因子家族的生物学功能奠定基础。

**关键词:** 绿豆; bHLH; 转录因子; 生物信息学分析

## Genome-Wide Identification and Bioinformatics Analysis of bHLH Transcription Factor Family in Mung Bean (*Vigna radiata* L.)

CHEN Hong-lin, HU Liang-liang, WANG Li-xia, WANG Su-hua, CHENG Xu-zhen

(Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081)

**Abstract:** Basic helix-loop-helix protein (bHLH) is an important transcription factor in eukaryotes, and it is mainly composed of basic amino acid region and helix-loop-helix region. In this study, a total of 122 mung bean bHLH transcription factors were identified by bioinformatical methods, and their physicochemical properties, conserved domain, gene structure, chromosome distribution, family evolution and expression differences of some typical bHLH genes were analyzed. The results showed their physicochemical properties were different. The bHLH contained two conserved domains, which were located in the alkaline amino acid region of N and the helical ring and spiral region of C, respectively. There were five conserved motifs, and the basic amino acid region contained the conserved sequence: His5-Glu9-Arg13, and they were related to the binding with the target gene, while HLH region contained two conserved residues, Arg23 and Arg55, which were related to the formation of dimer. The bHLH genes were uneven distribution on the 11 chromosomes in mung bean. The number of genes on chromosome 5, 7 and 8 were larger, and the number of genes on chromosome 1, 4 and 10 were smaller. The number of introns ranging from 1 to 9, and the genes were clustered on some chromosomes. The phylogenetic analysis showed that 122 bHLH transcription factors were divided into 11 subgroups. Our results showed that the expression of bHLH is diverse and exhibit distinct patterns in different tissues of mung bean. This study lays a foundation for further study on the complex biological functions of bHLH transcription factor family in mung bean.

**Key words:** Mung bean (*Vigna radiata* (L.) R. Wilczek); basic helix-loop-helix; transcription factor; bioinformatics analysis

收稿日期: 2017-04-01 修回日期: 2017-05-23 网络出版日期: 2017-08-18

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20170818.1031.002.html>

基金项目: 国家自然科学基金项目(31401442); 中国农业科学院科技创新工程“杂粮特色育种材料创制和新品种培育”; 国家食用豆产业技术体系建设项目(CARS-08)

第一作者主要从事食用豆基因发掘研究。E-mail: [chenhonglin@caas.cn](mailto:chenhonglin@caas.cn)

通信作者: 程须珍, 主要从事食用豆遗传资源研究。E-mail: [chengxuzhen@caas.cn](mailto:chengxuzhen@caas.cn)

转录因子是真核生物中一类重要的调节因子,与顺式元件相互作用,调节特定基因的表达,以适应环境胁迫。bHLH(basic helix-loop-helix)是真核生物中成员数量较多的一类转录因子家族,在真核生物生长发育过程中发挥着重要的调控作用。在单细胞真核生物中,bHLH参与染色体分离、新陈代谢调节等过程<sup>[1]</sup>;在动物中,bHLH主要参与生长发育、调节细胞周期、组织分化以及抗逆胁迫等过程<sup>[2-4]</sup>;在植物中,bHLH在植物生长发育与形态建成<sup>[5]</sup>、生物合成与信号转导<sup>[6-7]</sup>、抗旱<sup>[8]</sup>、抗寒<sup>[9]</sup>、耐盐<sup>[10]</sup>等抗逆反应,以及一些次生代谢产物的合成<sup>[11]</sup>等均有重要作用。

bHLH转录因子是植物中家族成员最多的转录因子,研究者发现在拟南芥(*Arabidopsis thaliana* (L.) Heynh.)中含有162个成员<sup>[12]</sup>,在水稻(*Oryza sativa* L.)中含有167个成员<sup>[13]</sup>。bHLH转录因子结构域包含2个保守基序,分别位于N端的碱性氨基酸区和C端的螺旋-环-螺旋区(HLH区)。碱性氨基酸区具有DNA识别和结合位点,HLH区主要由疏水氨基酸组成,易于与同一bHLH或不同bHLH的HLH之间相互作用形成二聚体。由于家族成员庞大,造成了其功能的多样性和冗余性,对研究单个bHLH转录因子的功能带来了一定的困难。

绿豆(*Vigna radiata* (L.) Wilczek)是我国主要的食用豆类作物之一,营养价值丰富。目前国内外还没有关于绿豆bHLH转录因子方面的报道,本研究利用生物信息学方法对绿豆bHLH转录因子家族成员进行鉴定,对其理化性质、基因结构、基因在染色体上的位置分布以及系统发育进行了研究,旨在为揭示其复杂的生物学功能提供理论依据。

## 1 材料与方法

### 1.1 材料

基因的组织表达差异分析材料为绿豆品种中绿5号,种子来自中国农业科学院作物科学研究所国家种质资源库。幼苗期取根、茎、叶,开花期与成熟前期分别取花和幼嫩种子,液氮中速冻后-80℃保存备用。

### 1.2 绿豆bHLH基因的鉴定

在Pfam数据库(<http://pfam.xfam.org/>)中下载bHLH转录因子的隐马氏模型文件(Pfam号为:PF00010)<sup>[14]</sup>,从GenBank(<http://www.ncbi.nlm.nih.gov/genbank/>)数据库

中下载绿豆基因组数据和蛋白组数据,绿豆基因组大小约为459 Mb,分布于11条染色体<sup>[15]</sup>。利用HMMER 3.0软件(<http://www.hmmerr.org/>),以隐马氏模型文件PF00010对绿豆蛋白组数据进行搜索,获得绿豆bHLH转录因子家族蛋白序列,并根据蛋白序列从NCBI数据库中获得其基因及CDS序列。

### 1.3 bHLH蛋白理化特性分析

通过Expasy(<http://www.expasy.org/tools/>)对bHLH生理生化特征进行预测<sup>[16]</sup>,获得蛋白序列的氨基酸数目、分子量和等电点等理化性质。利用SMART(<http://smart.embl-heidelberg.de/>)对bHLH保守功能域进行预测<sup>[17]</sup>。

### 1.4 bHLH基因保守域和基因结构及染色体定位分析

利用在线软件WebLogo 3(<http://weblogo.threeplusone.com/>)分析bHLH保守结构域<sup>[18]</sup>,并通过MEME SUITE在线工具(<http://meme-suite.org/tools/meme>)预测bHLH蛋白序列的保守Motif,Motif的搜索数目设置为5,其他参数保持默认设置<sup>[19]</sup>。使用基因结构显示系统GSDS(<http://gsds.cbi.pku.edu.cn/index.php>)绘制基因结构示意图。从NCBI中获得各蛋白的基因序列以及在染色体上的位置信息,通过Mapinspect(<http://mapinspect.software.informer.com/>)进行染色体定位作图。

### 1.5 bHLH基因家族序列比对与系统进化分析

使用ClustalX<sup>[20]</sup>对所有蛋白序列进行比对,利用MEGA 5.0<sup>[21]</sup>的NJ(Neighbor-Joining)法构建系统发育树。

### 1.6 bHLH基因组织表达差异分析

利用Invitrogen公司的Trizol试剂盒提取绿豆总RNA,M-MLV反转录酶反转录合成cDNA,反应体系为:2 μL RNA(1 μg/μL),2 μL oligo(dT)<sub>18</sub>(600 ng/μL),DEPC ddH<sub>2</sub>O补足至15 μL。将上述混合物短暂离心,70℃温浴5 min,立即冰浴5 min。RT-PCR反应体系为25 μL,含2.5 μL 10×ExTaq缓冲液、2 μL dNTP(1 mmol/L)、各1 μL(10 mmol/L)上下游引物、0.25 μL ExTaq DNA聚合酶(5 U/μL)、1 μL cDNA、17.25 μL ddH<sub>2</sub>O。采用*VrACTIN*作为内参基因,PCR扩增程序为:94℃,5 min;94℃,30 s;55℃,30 s;72℃,1 min;目的基因为30个循环,内参基因为28个循环。每个样品重复3次。所用引物具体序列详见表1。

表 1 qRT-PCR 所用引物序列

Table 1 Primer sequences of quantitative RT-PCR

基因登录号(名称)	上游引物	下游引物	注释
Gene accession(name)	Forward primer(5'-3')	Reverse primer(5'-3')	Annotation
XP_014515687.1	GCAGGATGAGGACTTGAA	TTCTTGTCTTTGTTCTTGAT	bHLH19
XP_014515458.1	GAAGACATTGAAGTTGGA	TCATAGCAGACTCACTAG	bHLH25
XP_014498665.1	TGGAGTAATGCCTATAATC	AAGAGTATGTTGAATGGAA	bHLH62
XP_014510625.1	CGGCAGCAATAACCTCAAC	AGGCAGATAGTGATAGGATAGAAC	bHLH120
XP_014520018.1	CTCCTTCGTCTACATATCA	AGCACCATCTCATTACTT	bHLH121
XP_014494068.1	TTGCCTATATGAACAGTGA	GGAATGCTAACAGTCTCT	bHLH128
<i>VrACTIN</i>	TCCACGAGACAACATATAACT	TCCTTGCTCATCCTATCAG	

## 2 结果与分析

### 2.1 bHLH 转录因子家族成员的鉴定及其理化性质分析

利用 HMMER 3.0 软件搜索绿豆蛋白组数据,并通过 Pfam 和 SMART 对保守结构域进行验证,除去冗余,最后鉴定到具有典型 bHLH 结构域的转录因子家族成员共 122 个(表 2)。绿豆 bHLH 转录因

子家族成员蛋白序列含有氨基酸数目为 164 (XP\_014511795.1) ~ 795 (XP\_014512590.1),通过 Ex-pasy 预测其分子量大小为 18.17 (XP\_014510439.1、XP\_014496843.1) ~ 88.63kD (XP\_014512590.1、XP\_014502095.1),等电点为 4.67 (XP\_014500777.1) ~ 10.06 (XP\_014519722.1)。从以上理化性质来看,该转录因子家族成员之间特性差异较大,71.90% 的蛋白等电点小于 7,呈弱酸性。

表 2 绿豆 bHLH 转录因子家族的鉴定及特征

Table 2 The characteristics of bHLH transcription factor family in mung bean

登录号	氨基酸	分子量(kD)	等电点	登录号	氨基酸	分子量(kD)	等电点
Accession	AA	MW	pI	Accession	AA	MW	pI
XP_014516222.1	262	32.60	6.67	XP_014519522.1	289	32.60	5.10
XP_014515458.1	311	25.80	5.54	XP_014500100.1	232	25.80	6.46
XP_014498755.1	245	18.93	5.66	XP_014511795.1	164	18.93	7.80
XP_014506252.1	201	44.43	5.51	XP_014508430.1	402	44.43	5.59
XP_014511187.1	338	40.69	5.91	XP_014515222.1	365	40.69	6.72
XP_014511188.1	338	40.34	5.91	XP_014503403.1	359	40.34	4.84
XP_014507790.1	218	25.56	5.90	XP_014497134.1	232	25.56	6.24
XP_014503844.1	261	25.79	9.26	XP_014497133.1	234	25.79	6.06
XP_014510613.1	173	25.18	8.53	XP_014519722.1	221	25.18	10.06
XP_014507789.1	224	43.89	6.14	XP_014499846.1	396	43.89	6.17
XP_014507788.1	240	36.22	5.53	XP_014495077.1	323	36.22	9.05
XP_014494911.1	272	43.31	8.93	XP_014520595.1	394	43.31	7.61
XP_014520421.1	363	43.61	6.19	XP_014520587.1	397	43.61	7.59
XP_014505902.1	357	46.33	8.22	XP_014507909.1	338	37.84	8.58
XP_014505901.1	360	46.33	8.22	XP_014505002.1	429	44.33	6.32
XP_014504502.1	300	40.55	8.64	XP_014505003.1	429	46.33	6.02
XP_014518169.1	350	38.96	8.46	XP_014496023.1	368	40.55	5.92

表 2(续)

登录号 Accession	氨基酸 AA	分子量(kD) MW	等电点 pI	登录号 Accession	氨基酸 AA	分子量(kD) MW	等电点 pI
XP_014516125.1	302	32.11	8.65	XP_014511971.1	350	38.96	5.59
XP_014516133.1	302	52.08	8.65	XP_014503436.1	280	32.11	5.00
XP_014500414.1	319	24.63	5.59	XP_014506710.1	460	52.08	6.60
XP_014500412.1	321	37.97	5.70	XP_014511019.1	219	24.63	5.63
XP_014515687.1	262	38.07	6.55	XP_014514535.1	346	37.97	5.90
XP_014506449.1	517	57.84	5.34	XP_014514536.1	346	38.07	5.90
XP_014511443.1	211	38.57	8.35	XP_014506709.1	509	57.84	6.37
XP_014510279.1	270	60.80	7.51	XP_014498665.1	343	38.57	7.11
XP_014512557.1	379	63.34	7.70	XP_014496615.1	561	60.80	6.18
XP_014512556.1	381	36.71	7.71	XP_014516761.1	581	63.34	6.19
XP_014495363.1	398	58.93	9.49	XP_014511727.1	332	36.71	4.90
XP_014493470.1	435	58.93	8.45	XP_014521233.1	551	59.00	5.56
XP_014517676.1	321	58.93	5.66	XP_014521234.1	551	58.07	5.54
XP_014519977.1	325	27.92	4.80	XP_014521235.1	551	58.93	5.54
XP_014519976.1	331	30.95	4.88	XP_014515688.1	244	27.92	5.02
XP_014510625.1	191	30.13	6.79	XP_014518636.1	276	30.95	5.73
XP_014510633.1	191	36.73	6.79	XP_014511945.1	263	30.13	6.26
XP_014494205.1	417	36.85	5.55	XP_014507779.1	331	36.73	4.89
XP_014506568.1	396	36.85	6.89	XP_014507777.1	332	36.85	4.89
XP_014515684.1	323	48.34	5.08	XP_014507778.1	332	36.85	4.89
XP_014512100.1	337	48.34	5.38	XP_014500448.1	439	50.34	6.38
XP_014500777.1	317	56.55	4.67	XP_014500449.1	439	48.34	6.28
XP_014504156.1	331	27.27	4.82	XP_014515202.1	510	56.55	6.24
XP_014504157.1	329	49.01	4.85	XP_014500688.1	239	27.27	9.82
XP_014502121.1	316	49.19	6.47	XP_014517030.1	452	49.01	5.13
XP_014505734.1	397	24.60	6.49	XP_014517029.1	454	49.19	5.13
XP_014505733.1	399	26.15	6.49	XP_014495710.1	217	24.60	7.60
XP_014507581.1	381	25.68	4.99	XP_014495708.1	231	26.15	6.68
XP_014496843.1	255	18.17	7.12	XP_014513121.1	223	25.68	8.20
XP_014514707.1	504	18.40	5.50	XP_014510439.1	165	18.17	8.33
XP_014502095.1	481	88.63	6.03	XP_014510438.1	167	18.40	8.66
XP_014520389.1	337	33.78	7.06	XP_014512590.1	795	88.63	4.80
XP_014508404.1	346	55.34	5.84	XP_014500305.1	304	33.78	6.25
XP_014502421.1	439	55.56	5.70	XP_014512632.1	503	55.34	6.36
XP_014502422.1	439	52.13	5.70	XP_014512631.1	505	55.56	6.36
XP_014502423.1	439	25.82	5.70	XP_014507195.1	474	52.13	5.86
XP_014513050.1	245	35.38	7.13	XP_014495709.1	228	25.82	7.07

表 2(续)

登录号	氨基酸	分子量(kD)	等电点	登录号	氨基酸	分子量(kD)	等电点
Accession	AA	MW	pI	Accession	AA	MW	pI
XP_014502122.1	300	35.64	5.93	XP_014515781.1	322	35.38	6.59
XP_014494068.1	191	36.13	5.16	XP_014515780.1	324	35.64	6.59
XP_014520018.1	332	36.66	6.77	XP_014517608.1	333	36.13	6.00
XP_014515460.1	335	36.99	8.82	XP_014507018.1	335	36.66	6.24
XP_014513572.1	392	28.65	9.20	XP_014507017.1	338	36.99	6.24
XP_014511969.1	401	24.01	5.14	XP_014511973.1	264	28.65	8.51
XP_014507075.1	222	24.29	9.36	XP_014507599.1	219	24.01	5.74

## 2.2 bHLH 转录因子保守结构域和基因结构分析

通过 ClustalX 多序列比对和 SMART 分析,发现 122 个蛋白序列均含有 bHLH 结构域。它由约 60 个氨基酸组成,包含 N 端碱性区域和 C 端 HLH 2 个不同功能区,N 端的碱性氨基酸区由 10 ~ 15 个氨基酸组成,与 DNA 结合有关。bHLH 保守域氨基酸大部分成员在碱性氨基酸区含有高度保守的 H5-E9-R13 序列(His5-Glu9-Arg13),该序列对 bHLH 与靶基因结合是必需的<sup>[4,22]</sup>。另外,第 10 位和第 12 位上的精氨酸(Arg,R)也具有高度保守性。靶基因启动子上游

区域通常含有特殊的区域成为 E-box(CANNTG)<sup>[23]</sup>,较为常见的是 G-box(CACGTG),bHLH 转录因子的碱性区域就是与该特殊元件相互识别和结合。HLH 区位于 C 端,由约 50 个氨基酸组成,大部分为疏水性氨基酸,构成了螺旋-环-螺旋结构。同一个 bHLH 转录因子的 2 个  $\alpha$ -螺旋或不同 bHLH 转录因子的  $\alpha$ -螺旋之间可以相互作用形成二聚体,bHLH 转录因子正是通过同源或异源二聚体发挥作用。其中,第 23 位和第 55 位的亮氨酸(Leu,L)高度保守,对形成二聚体是必需的(图 1)。

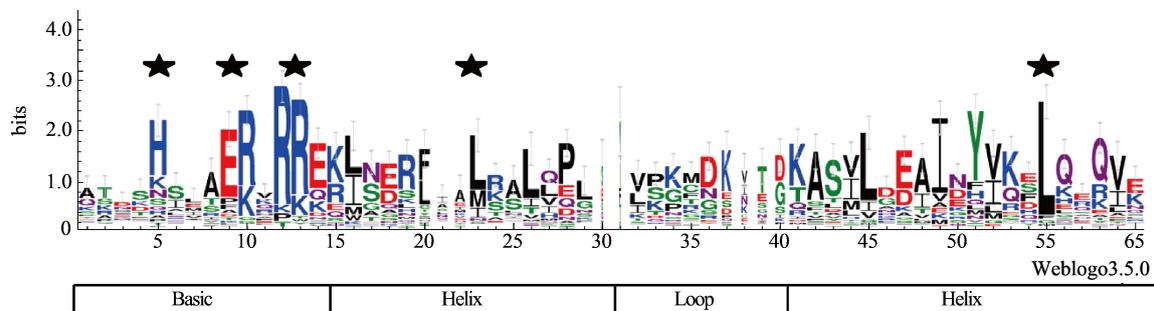


图 1 bHLH 结构域保守序列

Fig. 1 The conserved sequences and locations of bHLH

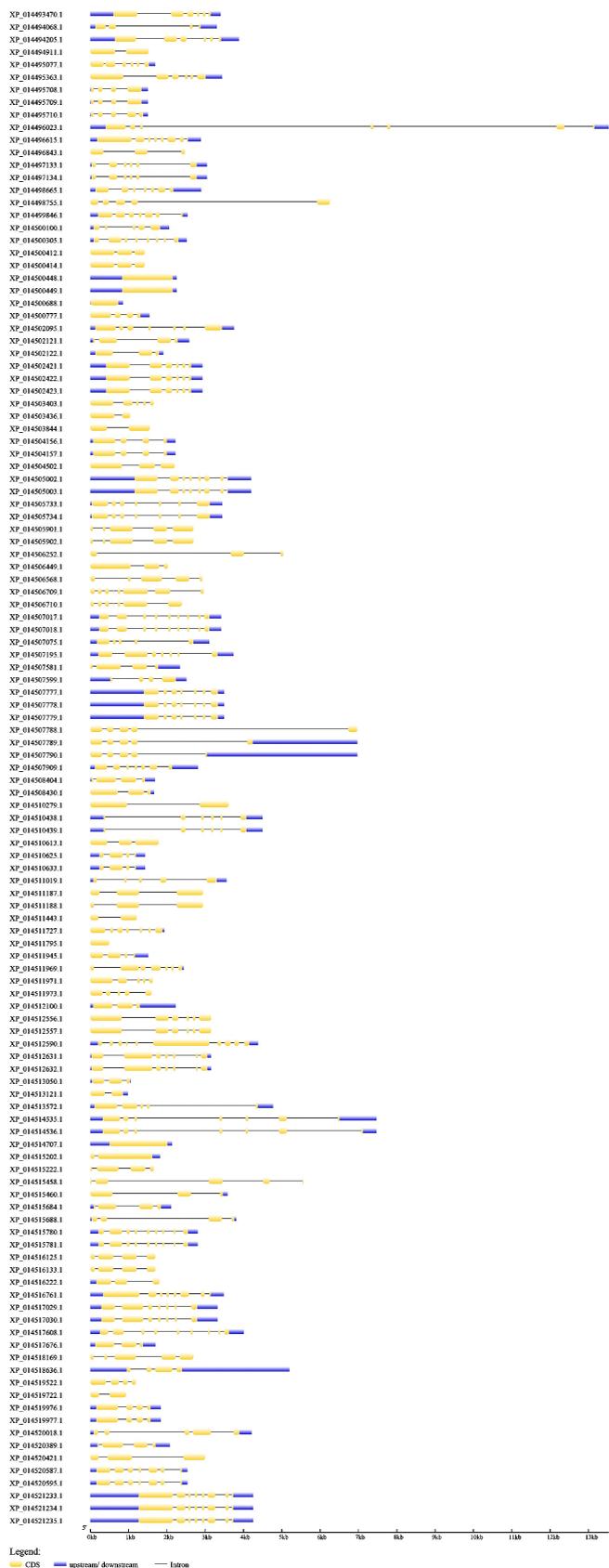
除了 bHLH 保守结构域外,bHLH 转录因子家族还含有其他保守型的结构域,由 MEME SUITE 在线工具分析得到该家族含有的 5 种保守元件(图 2)。其中元件 1 和元件 2 存在于所有序列中,而且总是相邻,两者共同组成了 bHLH 结构域。根据保守元件的不同,可将 bHLH 转录因子家族分为 4 组, I 组只含有元件 1、2; II 组除含元件 1、2 外,还含有元件 3; III 组除含有元件 1、2 外,还含有元件 5; IV 组除含元件 1、2 外,同时还含有元件 3 和元件 4。这些保守元件,对维持蛋白质的三维构象和与靶基因的结合至关重要,是转录因子发挥其调控作用的必要元件。

通过基因结构显示系统绘制了 122 个 bHLH 转录因子的基因结构图(图 2)。除 5 个基因(XP\_014500448.1、XP\_014500449.1、XP\_014500688.1、

XP\_014511795.1 和 XP\_014514707.1) 外,其他 bHLH 基因均含有内含子结构。内含子数目有 1 个(如 XP\_014503436.1、XP\_014503844.1 等)到 9 个(XP\_014512590.1)不等。从上述基因的结构特征可初步分析 bHLH 转录因子的基因复制方式。

## 2.3 bHLH 转录因子在染色体上的分布分析

根据 bHLH 基因家族在绿豆染色体上的位置信息,绘制了染色体位置分布图(图 3)。绿豆 122 个 bHLH 基因在 11 条染色体上均有分布,但分布并不均匀。在第 1 号、4 号和 10 号染色体上分布较少,分别有 5、4 和 4 个;5 号、7 号和 8 号染色体上分布较多,分别有 17、26 和 22 个;其他染色体上则存在 6 ~ 13 个不等。同一染色体上的 bHLH 基因的分布也不均匀,呈现出 2 个或多个基因聚集在一起形成基因簇。



蓝,黄表示不同的 motifs;黑线表示内含子

The motifs were represented in blue and yellow. The introns were represented by black gray lines

图2 bHLH 转录因子的基因结构

Fig. 2 Gene structure of bHLH transcription factor

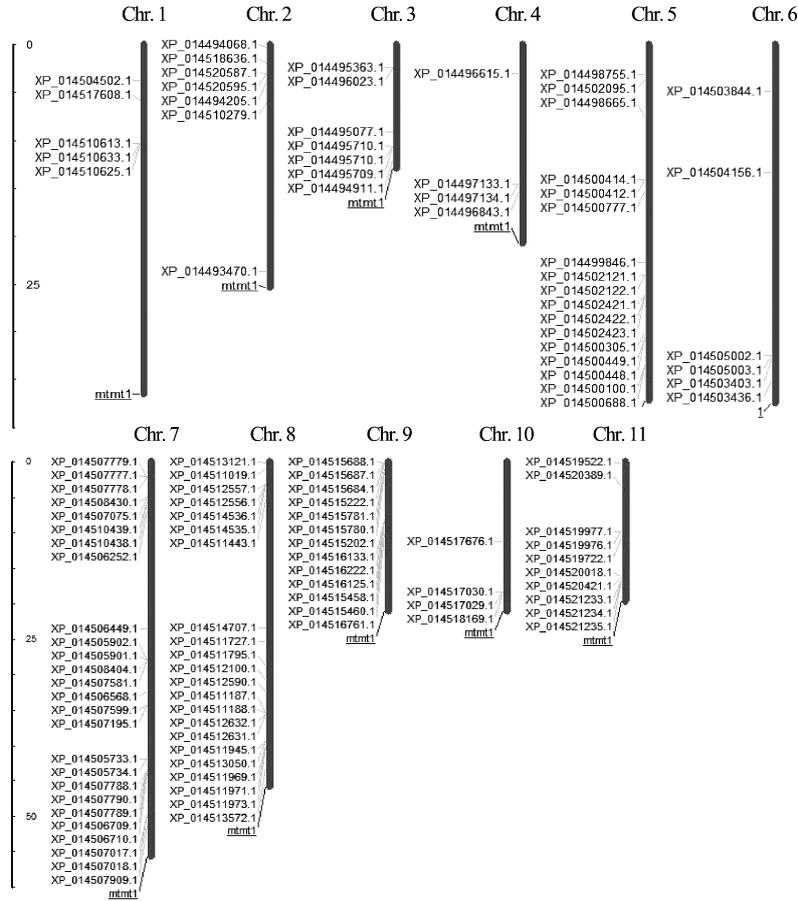


图 3 bHLH 基因家族染色体定位

Fig. 3 Chromosome locations of bHLH genes family

2.4 bHLH 转录因子家族进化分析

在 ClustalX 多序列联配的基础上,通过 MEGA 5.0 软件对上述 122 个 bHLH 转录因子家族氨基酸序列构建 NJ 系统进化树(图 4)。结果表明,122 个 bHLH 转录因子被划分在 4 个大进化组, a 组又分为 a1、a2 两个亚家族; b、c 和 d 组均又分为 3 个亚家族,最终分为 11 个亚家族,同一亚家族的成员趋向于拥有相同的保守模序,且趋向于拥有相同数量的内含子。

2.5 绿豆 bHLH 基因组织表达差异分析

根据 bHLH 序列差异,选取了 6 个典型的 bHLH 基因(图 5),分别对其在根、茎、叶、花和种子中的表达情况进行了定量 RT-PCR 表达差异分析。结果显示,6 个 bHLH 基因在 5 种不同的组织中具有不同的表达模式。*XP\_014498665.1* 和 *XP\_014520018.1* 在 5 种组织中均有表达,且表达量较高;*XP\_014515687.1* 在叶中未表达;*XP\_014515458.1* 在花中未表达;*XP\_014510625.1* 只在根和茎中表达,且表达量较低;*XP\_014494068.1* 在根、茎和花中表达,

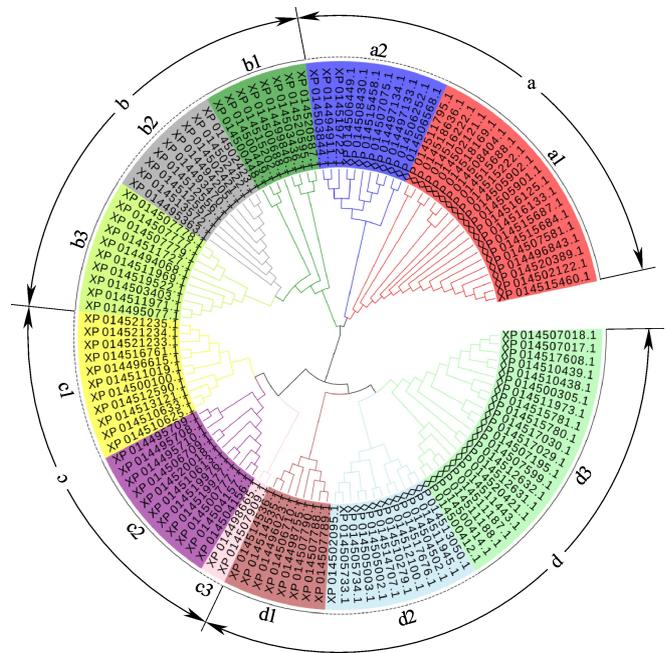


图 4 绿豆 bHLH 转录因子家族系统发育进化树

Fig. 4 Phylogenetic tree of bHLH transcription factor family in mung bean

表达量也较低。以上结果说明,bHLH 基因具有组织表达特异性,具有在特定发育阶段的特定组织中表达的特点。

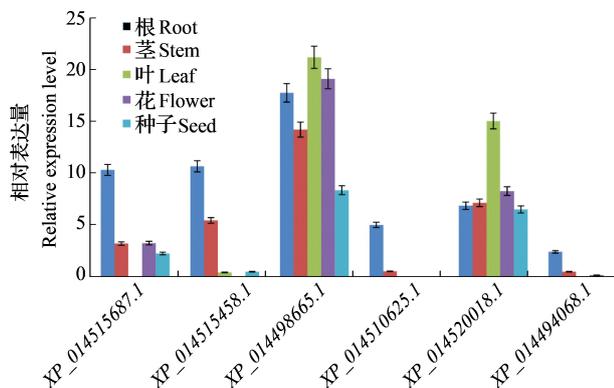


图5 绿豆 bHLH 基因在绿豆不同组织中的表达分析

Fig. 5 Expression analysis of bHLH genes in different tissues of mung bean

### 3 讨论

bHLH 转录因子家族是真核生物中一类重要的转录因子,其家族成员众多,功能多样<sup>[24-26]</sup>,本课题组通过转录组测序也发现绿豆含有大量的 bHLH 转录因子<sup>[27]</sup>。目前,对动物和模式植物拟南芥以及水稻的 bHLH 转录因子家族的研究较为成熟,但豆科植物 bHLH 家族的研究还不深入。本研究利用生物信息学的手段鉴定了绿豆 bHLH 转录因子家族,并对其保守结构域、基因结构和在染色体上的位置分布以及系统进化等方面进行了分析。

通过验证,去除冗余,共鉴定到 122 个绿豆 bHLH 转录因子。bHLH 保守结构域由两部分组成,碱性氨基酸区和 HLH 区。碱性氨基酸区位于 N 端,HLH 区位于 C 端。碱性氨基酸区含有高度保守的 His5-Glu9-Arg13 序列,第 10 位和第 12 位的 Arg 也高度保守,而 HLH 区第 23 位和第 55 位的 Leu 高度保守。研究者发现在番茄 (*Lycopersicon esculentum* Mill) 中有 152 个 bHLH 转录因子家族中同样存在 His5、Glu9 和 Arg13,以及 Arg10、Arg12 等高度保守的氨基酸残基,而在 HLH 区则存在保守残基 Leu23 和 Leu64<sup>[28]</sup>。碱性氨基酸区的保守氨基酸残基与靶基因结合有关,而 HLH 区保守氨基酸残基与形成二聚体密切相关。

通过基因结构和染色体定位分析,发现绿豆 bHLH 基因家族含有数量不等的内含子,且基因在染色体上的分布不均匀,并形成了很多基因簇(图 3),相互之间物理距离很近说明它们可能属于串联

复制,且它们的序列和保守区域非常相似,在进化树中亲缘关系很近,因此推测他们在进化过程中功能类似,但有待进一步研究确证。据报道,拟南芥的 NBS-LRR 抗病基因也发生了基因的大规模复制<sup>[29]</sup>,有 146 个 NBS-LRR 基因呈大小不等的簇存在,占基因总数的 71.57%。同一染色体上的基因簇,可能是染色体内基因的串联复制的结果,如果某基因产物的需求量很大时,该基因可以产生大量的串联重复;不同染色体上的基因簇同系物则是染色体间片段复制的结果<sup>[29-30]</sup>。同一基因簇内的基因,其功能往往相似,这是为满足生命活动某一特定产物的大量需求所导致的。

聚类分析表明,122 个 bHLH 被划分为 11 个亚家族,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。由于 bHLH 转录因子家族数量庞大,对其分类带来了一定难度。研究者依据进化关系、与 DNA 结合模式,将动物 bHLH 转录因子家族划分为 A~F 6 个家族<sup>[31]</sup>;但在植物中还未有系统的划分,在模式植物拟南芥 (*Arabidopsis thaliana*) 中,研究者将 162 个 bHLH 转录因子划分为 21 个亚家族,而在 167 个水稻 (*Oryza sativa*) 和 230 个大白菜 (*Brassica rapa*) bHLH 转录因子家族中,分别被划分为 22 和 24 个亚家族。A 类成员大都可以与 CAGCTG 或 CACCTG 类型的 E-box 结合,B 类则与 G-box (CACCTG) 结合,且大都有 His5-Glu9-Arg13 序列。C 类除 bHLH 结构域外,C 端还含有 PAS 结构域,而 D 类只含有 HLH 区,缺乏碱性氨基酸区,故其不能与靶 DNA 结合。E 类第 6 位氨基酸为保守的脯氨酸,与 CACGCG 或 CACGAG 结合。F 类含有 COE 结构域。研究者发现在 167 个水稻 bHLH 转录因子中,有 141 个成员具有与 DNA 结合的能力,其中 95 个是与 G-box 结合,19 个与其他种类 E-box 结合<sup>[13]</sup>。植物中大部分 bHLH 转录因子属于 B 类,8% 与 A 类结构相似,小部分不能与 DNA 结合,还未发现植物 bHLH 与 C 类、F 类相同<sup>[32-33]</sup>。

定量 RT-PCR 表达分析表明,绿豆 6 个 bHLH 基因在不同组织中具有不同的表达模式,这为研究其功能提供参考。在苜蓿<sup>[23]</sup>、小麦<sup>[33]</sup> 和 水稻<sup>[34]</sup> 等多种植物中也进行了 bHLH 基因的组织表达研究,并得到了相似的结果。看家基因以组成型方式在所有细胞中表达,而奢侈基因只在特定组织细胞中表达,使得细胞生理生化活动得以正常进行,又对庞大的 bHLH 家族成员进行了合理的分工。

目前,绿豆研究起步较晚,其分子生物学方面的

研究相对滞后,对 bHLH 转录因子的功能研究远没有拟南芥、水稻和大豆等植物研究得清楚和深入。本研究通过对绿豆 bHLH 家族的保守结构域、基因结构、染色体分布、进化关系等多种生物信息学和组织表达情况进行了分析,为将来深入研究该基因家族的表达调控、结构和功能等提供参考,为揭示该基因家族参与植物生长、发育的调控机制提供了理论依据。

#### 参考文献

- [1] Robinson K A, Lopes J M. SURVEY AND SUMMARY: Saccharomyces cerevisiae basic helix-loop-helix proteins regulate diverse biological processes [J]. Nucleic Acids Res, 2000, 28 (7): 1499-1505
- [2] Jiang H, Guo R, Powell-Coffman J A. The Caenorhabditis elegans *hif-1* gene encodes a bHLH-PAS protein that is required for adaptation to hypoxia [J]. Proc Natl Acad Sci USA, 2001, 98 (14): 7916-7921
- [3] Lu J R, Webb R, Richardson J A, et al. MyoR: a muscle-restricted basic helix-loop-helix transcription factor that antagonizes the actions of MyoD [J]. Proc Natl Acad Sci USA, 1999, 96 (2): 552-557
- [4] Stevens J D, Roalson E H, Skinner M K. Phylogenetic and expression analysis of the basic helix-loop-helix transcription factor gene family: genomic approach to cellular differentiation [J]. Differentiation, 2008, 76 (9): 1006-1022
- [5] Penfield S, Josse E M, Kamangara R, et al. Cold and light control seed germination through the bHLH transcription factor SPATULA [J]. Curr Biol, 2005, 15 (22): 1998-2006
- [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] Friedrichsen D M, Nemhauser J, Muramitsu T, et al. Three redundant brassinosteroid early response genes encode putative bHLH transcription factors required for normal growth [J]. Genetics, 2002, 162 (3): 1445-1456
- [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] Chinnusamy V, Ohta M, Kanrar S, et al. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis* [J]. Gene Dev, 2003, 17 (8): 1043-1054
- [10] Jiang Y, Deyholos M K. Comprehensive transcriptional profiling of NaCl-stressed *Arabidopsis* roots reveals novel classes of responsive genes [J]. BMC Plant Biol, 2006, 6: 25
- [11] Albert N W, Lewis D H, Zhang H, et al. Light-induced vegetative anthocyanin pigmentation in *Petunia* [J]. J Exp Bot, 2009, 60 (7): 2191-2202
- [12] Bailey P C, Martin C, Toledo-Ortiz G, et al. Update on the basic helix-loop-helix transcription factor gene family in *Arabidopsis thaliana* [J]. Plant Cell, 2003, 15 (11): 2497-2502
- [13] 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 Physiol, 2006, 141 (4): 1167-1184
- [14] Finn R D, Tate J, Mistry J, et al. The Pfam protein families database [J]. Nucleic Acids Res, 2008, 36: 281-288
- [15] Kang Y J, Kim S K, Kim M Y, et al. Genome sequence of mungbean and insights into evolution within *Vigna* species [J]. Nat Commun, 2014, 5: 5443
- [16] Gasteiger E, Gattiker A, Hoogland C, et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis [J]. Nucleic Acids Res, 2003, 31 (13): 3784-3788
- [17] Schultz J, Milpetz F, Bork P, et al. SMART, a simple modular architecture research tool: Identification of signaling domains [J]. Proc Natl Acad Sci USA, 1998, 95 (11): 5857-5864
- [18] Crooks G E, Hon G, Chandonia J M, et al. WebLogo: a sequence logo generator [J]. Genome Res, 2004, 14 (6): 1188-1190
- [19] Bailey T L, Boden M, Buske F A, et al. MEME SUITE: tools for motif discovery and searching [J]. Nucleic Acids Res, 2009, 37: 202-208
- [20] Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. Nucleic Acids Res, 1997, 25 (24): 4876-4882
- [21] Tamura K, Peterson D, Peterson N, et al. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. Mol Biol Evol, 2011, 28 (10): 2731-2739
- [22] Pires N, Dolan L. Origin and diversification of basic-helix-loop-helix proteins in plants [J]. Mol Biol Evol, 2010, 27 (4): 862-874
- [23] 刘晓月, 王文生, 傅彬英. 植物 bHLH 转录因子家族的功能研究进展 [J]. 生物技术进展, 2011, 1 (6): 391-397
- [24] Godiard L, Lepage A, Moreau S, et al. MtBHLH1, a bHLH transcription factor involved in *Medicago truncatula* nodule vascular patterning and nodule to plant metabolic exchanges [J]. New Phytol, 2011, 191 (2): 391-404
- [25] Hao Y, Oh E, Choi G, et al. Interactions between HLH and bHLH factors modulate light-regulated plant development [J]. Mol Plant, 2012, 5 (3): 688-697
- [26] Koini M A, Alvey L, Allen T, et al. High temperature-mediated adaptations in plant architecture require the bHLH transcription factor PIF4 [J]. Curr Biol, 2009, 19 (5): 408-413
- [27] Chen H, Wang L, Wang S, et al. Transcriptome sequencing of mung bean (*Vigna radiate* L.) genes and the identification of EST-SSR markers [J]. PLoS ONE, 2015, 10: e0120273
- [28] 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: 39
- [29] 王岩, 李兆阳, 唐心龙, 等. 拟南芥基因组 NBS-LRR 类基因家族的生物信息学分析 [J]. 中国农学通报, 2009, 25 (15): 40-45
- [30] Shatters R G, Boykin L M, Lapointe S L, et al. Phylogenetic and structural relationships of the *PR5* gene family reveal an ancient multigene family conserved in plants and select animal taxa [J]. J Mol Evol, 2006, 63 (1): 12-29
- [31] Atchley W R, Fitch W M. A natural classification of the basic helix-loop-helix class of transcription factors [J]. Proc Natl Acad Sci USA, 1997, 94 (10): 5172-5176
- [32] Pires N, Dolan L. Origin and diversification of basic-helix-loop-helix proteins in plants [J]. Mol Biol Evol, 2010, 27 (4): 862-874
- [33] 刘文学, 李文学. 植物 bHLH 转录因子研究进展 [J]. 生物技术进展, 2013, 3 (1): 7-11
- [34] 张子佳, 王迪, 傅彬英. 水稻转录因子 bHLH 家族基因响应环境胁迫表达谱分析 [J]. 分子植物育种, 2008, 6 (3): 425-431