

· 环境生物技术 ·

## 紫鸭跖草热激转录因子基因家族鉴定及在 $\text{Cu}^{2+}$ 胁迫下表达模式分析

彭国颖, 卢山, 杨坤, 万玮, 黄长干

江西农业大学 南昌市植物资源化学利用重点实验室, 江西 南昌 330045

彭国颖, 卢山, 杨坤, 万玮, 黄长干. 紫鸭跖草热激转录因子基因家族鉴定及在  $\text{Cu}^{2+}$  胁迫下表达模式分析. 生物工程学报, 2022, 38(1): 238-251.

PENG GY, LU S, YANG K, WAN W, HUANG CG. Identification of heat stress transcription factors gene family in *Setcreasea purpurea* and analysis of its expression pattern under  $\text{Cu}^{2+}$  stress. Chin J Biotech, 2022, 38(1): 238-251.

**摘要:** 热激转录因子 (Hsf) 是植物中最重要的转录因子家族之一, 在植物遭遇热、干旱和重金属等多种非生物胁迫中具有重要的作用。为明确紫鸭跖草 *Hsf* 基因家族的结构及功能, 从紫鸭跖草全长转录组数据库筛选鉴定出 20 个 *SpbHsf* 基因, 并运用生物信息学工具和 qRT-PCR 进行分析。结果显示: *SpbHsf* 蛋白均为亲水性蛋白, 定位于细胞核的 *SpbHsf* 蛋白有 12 个, 所有 *SpbHsf* 蛋白二级结构的  $\alpha$  螺旋和无规卷曲的含量较高。*SpbHsf* 基因分为 3 个亚族, 各亚族都包含独特的保守基序。全部 *SpbHsf* 蛋白含有 DBD 和 HR-A/B 结构域。系统进化分析显示, 与 *SpbHsf* 蛋白同源性最高的是水稻 *OsHsf* 蛋白。20 个 *SpbHsf* 基因在紫鸭跖草根组织中均有表达。相比于对照组, 根组织在  $\text{Cu}^{2+}$  胁迫下, 8 个 *SpbHsf* 基因的表达量显著上调, 8 个 *SpbHsf* 基因的表达量显著下调。研究结果为紫鸭跖草 *Hsf* 基因家族的功能验证及表达模式提供了一定的理论参考。

**关键词:** 紫鸭跖草; 热激转录因子; 生物信息学; 表达分析

Received: January 6, 2021; Accepted: March 19, 2021

Supported by: National Natural Science Foundation of China (31760068)

Corresponding author: HUANG Changgan. Tel: +86-791-83828063; E-mail: windhcg@163.com

基金项目: 国家自然科学基金 (31760068)

# Identification of heat stress transcription factors gene family in *Setcreasea purpurea* and analysis of its expression pattern under Cu<sup>2+</sup> stress

PENG Guoying, LU Shan, YANG Kun, WAN Wei, HUANG Changgan

Nanchang Key Laboratory of Chemical Utilization of Plant Resources, Jiangxi Agricultural University, Nanchang 330045, Jiangxi, China

**Abstract:** Heat stress transcription factors (Hsf) family is one of the most important transcription factor families in plants, and plays an important role in the growth and development of plants when encountering abiotic stresses such as heat, drought, and heavy metals. In this study, 20 *SpbHsf* genes were identified from the full-length transcriptome database of *Setcreasea purpurea*, and the structure and function of the *Hsf* gene family were analyzed using bioinformatics tools and qRT-PCR. The results showed that all *SpbHsf* proteins were hydrophilic. There were 12 *SpbHsf* proteins located in the nucleus, and the content of  $\alpha$ -helix and random coil in the secondary structure of all *SpbHsf* proteins was high. The *SpbHsf* genes are divided into three subfamilies, each of which contains unique conserved motifs. All *SpbHsf* proteins contain DBD and HR-A/B domains. Phylogenetic analysis showed that OsHsf in *Oryza sativa* protein had the highest homology with *SpbHsf* protein. All the 20 *SpbHsf* genes were expressed in the root tissues of *S. purpurea*. Among them, 8 were significantly up-regulated while 8 were significantly down-regulated under Cu<sup>2+</sup> stress. This study may help better understand the function and expression pattern of the *S. purpurea Hsf* gene family.

**Keywords:** *Setcreasea purpurea*; Hsf; bioinformatics; expression analysis

热激转录因子 (heat stress transcription factors, Hsf), 是介导对热或其他应激反应基因激活的信号转导链的末端元件, 被认为是调控热激蛋白 (heat shock proteins, Hsps) 表达的主要转录调节因子<sup>[1]</sup>。在正常生长条件下, 热休克蛋白大多数充当分子伴侣, *Hsf* 基因通过与 Hsp70 等分子伴侣的结合保持在惰性单体状态。在发生应激反应时, *Hsf* 基因能调控热休克蛋白表达, Hsps 能够阻止蛋白质解体和聚集, 从而维持细胞内蛋白质的稳态、保护细胞结构和调节应激反应的功能。与许多其他转录因子类似, *Hsf* 基因家族在进化上十分保守。Hsf 蛋

白由多个结构和功能保守域组成, 包括 DNA 结合域 (DNA binding domain, DBD)、核输出信号 (nuclear export signal, NES)、寡聚化结构域 (oligomerization domain, OD)、C 端转录激活结构域 (carboxy-terminal domain, CTD)、核定位信号 (nuclear localization signal, NLS) 和抑制蛋白结构域 (repression domain, RD)<sup>[2]</sup>。在这些保守结构域中, N 末端的高度结构化的 DNA 结合域是 Hsf 蛋白最保守的基序, 由 4 个  $\beta$ -折叠 ( $\beta 1$ 、 $\beta 2$ 、 $\beta 3$ 、 $\beta 4$ ) 和 3 个  $\alpha$ -螺旋 ( $\alpha 1$ 、 $\alpha 2$ 、 $\alpha 3$ ) 组成。该结合域的疏水核心形成一个特定螺旋-转角-螺旋 ( $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\alpha 3$ - $\beta 3$ - $\beta 4$ ) 结构, 能

特异性地识别并结合靶基因的热休克元件的保守基序,进而激活抗逆基因的转录和表达<sup>[3]</sup>。OD结构域是2个以疏水性氨基酸残基(七肽重复区域)结合的HR-A/B部位构成的,后者通过可变长度柔性连接体(约15–80个氨基酸)连接到DBD<sup>[4]</sup>。根据DBD距离OD的长度以及HR-A/B的残基数量,可将Hsf基因家族分为3组:A、B和C。A类可细分为9个子类:A1–A9;B类分为4个子类:B1–B4;而C类只有2个子类:C1和C2<sup>[5–7]</sup>。

尽管Hsf基因家族在进化上十分保守,但Hsf基因家族在数量上却存在差异。现今已从许多物种中鉴定出Hsf基因家族成员,例如拟南芥*Arabidopsis thaliana*中鉴定出22个、水稻*Oryza sativa*中鉴定出25个、紫花苜蓿*Medicago truncatula*中鉴定出16个、毛果杨*Populus trichocarpa*中鉴定出28个、陆地棉*Gossypium hirsutum*和小麦*Triticum aestivum*中分别鉴定出40个和56个成员等<sup>[8]</sup>。迄今为止,在甘蓝型油菜*Populus trichocarpa*中鉴定出Hsf基因家族包含的基因最多,有64个Hsf基因成员<sup>[9]</sup>。近年研究发现,许多Hsf基因成员在不同物种中应对各种胁迫反应的生物学功能不断被证实。例如,在对拟南芥HsfA2亚类基因敲除突变和过表达分析中,发现该亚家族基因参与了热激、强光、氧化应激和缺氧等多种胁迫反应<sup>[10]</sup>,HsfA3亚类也参与了受热、干旱胁迫反应<sup>[11]</sup>。向日葵*Helianthus annuus* HsfA4a和HsfA9基因在烟草中的共表达可以提高幼苗对严重脱水和氧化胁迫的抗性<sup>[12]</sup>;镉超积累型东南景天*Sedum alfredii*基因SaHsfA4ah和SaHsfA4ch的异源表达可以增强酵母对Cd的耐受性<sup>[13]</sup>。

紫鸭跖草*Setcreasea purpurea* B.K.Boom又名紫锦草,是一种多年生草本植物,具有较高

的观赏价值。紫鸭跖草是一种超积累铜植物<sup>[14]</sup>,耐受铜离子的能力强于其他普通植物,在修复铜污染的土地具有一定的潜力。目前尚未有对紫鸭跖草的Hsf基因的报道,本研究从紫鸭跖草转录组测序的GO富集和KEGG通路中,通过基因的注释和阈值 $P \leq 0.001$ 的过滤,筛选与铜诱导相关且在300  $\mu\text{mol/L}$   $\text{Cu}^{2+}$ 胁迫组和1000  $\mu\text{mol/L}$   $\text{Cu}^{2+}$ 胁迫组均有显著性上调表达的基因。发现*SpbHsfA3*基因和*SpbHsfB2b*基因显著上调表达,说明Hsf基因家族成员在受铜胁迫时可能发挥关键作用。本研究从紫鸭跖草的转录组数据库中筛选鉴定出Hsf基因家族成员,并对其进行生物信息学分析和 $\text{Cu}^{2+}$ 胁迫下的表达模式分析,从而获得潜在的铜离子胁迫表达基因。

## 1 材料与方法

### 1.1 数据库与试验材料

本研究以紫鸭跖草根组织在 $\text{Cu}^{2+}$  (0.25、100、300  $\mu\text{mol/L}$ )胁迫下的全长转录组测序所得的数据(NCBI的SRA数据库,登录号为SAMN11265427),筛选和鉴定出紫鸭跖草的Hsf基因家族成员。

剪取紫鸭跖草4–6 cm的嫩茎,置于培养箱中的Hoagland's(霍格兰氏)营养液中培养,培养箱昼夜温度为(25±1) °C/(20±1) °C,光周期为14 h光照/10 h黑暗,光照强度为300  $\mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 。待培养30 d后嫩根长成,依次在0.25(对照组)、100、300、1000  $\mu\text{mol/L}$   $\text{Cu}^{2+}$ 的Hoagland's营养液中培养4 d,以各 $\text{Cu}^{2+}$ 营养液培养紫鸭跖草的根为qRT-PCR分析试验材料,取样后迅速用液氮冷冻并放置–80 °C冰箱保存。

### 1.2 紫鸭跖草Hsf基因家族成员的鉴定

通过本地BLAST程序和Hmmer程序共同

检索紫鸭跖草全长转录组数据库 (本课题组转录组数据库) 中的 *Hsf* 基因: (1) 利用本地 BLAST 程序, 以已知的 *Hsf* 基因 DBD 保守结构域的氨基酸序列为查询序列, 运用 BLASTp 搜索本地紫鸭跖草转录组数据库的蛋白序列库中的同源基因, 筛选阈值:  $E < 10^{-5}$ , 得到潜在的 *SpbHsf* 基因序列。(2) 利用 Hmmer 程序, 以 *Hsf* 基因的隐马尔可夫模型 (Pfam: PF0047) 在紫鸭跖草转录组测序的蛋白质序列库中进行搜索, 筛选  $E < 0.0001$  且编码氨基酸长度大于 100 的序列作为潜在的 *Hsf* 基因。合并两个程序搜索的序列数据, 再通过 Pfam 数据库和在线 SMART 程序验证所有候选 Hsf 蛋白序列是否包含 Hsf 型 DBD 结构域。由于转录组测序获得的部分 unigene 所包含的编码区信息不完整, 因此利用在线 BLASTP 分析和 SMART 分析对初步筛选的紫鸭跖草 *Hsf* 基因家族成员进行鉴定分析。对于编码序列不完整的基因成员, 委托生工生物工程 (上海) 股份有限公司利用同源克隆的方法进行补足。

### 1.3 SpbHsf 蛋白理化性质及结构的分析

运用 ExPASy 在线软件 (<https://web.expasy.org/protparam/>) 分析 SpbHsf 蛋白的分子质量、等电点、平均疏水性和不稳定指数等理化性质; 运用 Prabi 在线软件 ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)) 预测 SpbHsf 蛋白的二级结构; 运用 Swiss-Model 在线软件 (<https://swissmodel.expasy.org/interactive>) 预测 SpbHsf 蛋白的三级结构; 运用 PSOR 在线软件 (<http://psort1.hgc.jp/form.html>) 分析 SpbHsf 蛋白的亚细胞定位。

### 1.4 SpbHsf 基因保守基序的分析

对紫鸭跖草 Hsf 基因家族各成员, 利用 Clustal X 进行同源性比对, 比对结果由

GeneDoc 编辑。利用在线程序 MEME 对 *Hsf* 基因家族成员进行 Motif 提取, 参数: 搜索 motif 为 20 个、其他参数不变。利用 MEGA7 (neighbour joining 法) 软件和在线程序 EvolView 绘制系统进化树与保守基序组成图。

### 1.5 SpbHsf 蛋白结构域分析

通过在线数据库 Heatster, 以已鉴定的植物 Hsf 蛋白结构域为基础分析 *SpbHsf* 基因家族成员的保守结构域。再运用 MEGA7、GeneDoc 软件对 SpbHsf 蛋白的 DBD 保守结构域进行聚类分析。

### 1.6 多物种进化树的构建

在 NCBI 中检索出 22 个拟南芥 *AtHsf* 蛋白、25 个水稻 *OsHsf* 蛋白、28 个毛果杨 *PtHsf* 蛋白和 22 个东南景天 (*S. alfredii*) *SaHsf* 蛋白, 再运用 MEGA7 软件 (neighbour joining 法) 将检索出的 Hsf 蛋白与 *SpbHsf* 蛋白共同构建系统进化树, 再通过 EvolView 在线软件对其进行美化。

### 1.7 qRT-PCR 分析 SpbHsf 基因的表达

运用 Primer Premier 5 软件设计 *SpbHsf* 基因的特异性引物 (表 1), 选取泛素连接酶 (UBI) 基因 (UBI-F: 5'-AGCACTCTTCACCTCGTCC-3'; UBI-R: 5'-CTTCGCCTTCACATTCTC-3') 作为荧光定量 qRT-PCR 分析的内参基因, 运用 qRT-PCR 分析紫鸭跖草根组织中 20 个 *SpbHsf* 基因的表达模式。运用 qRT-PCR 分析不同  $\text{Cu}^{2+}$  浓度胁迫下紫鸭跖草根各 *Hsf* 基因表达不同组织 (根、茎、叶) 的 *SpbHsf* 基因表达量, 利用实时荧光定量仪 LightCycler480 System 进行反应, 每个反应进行 3 次生物学重复, 运用  $2^{-\Delta\Delta C_t}$  法计算基因的相对表达量。采用 SPSS 25.0 软件对数据进行统计分析, 均值间比较采用 Duncan's 法 ( $P < 0.05$ )。

表 1 *SpbHsfs* 基因家族的引物设计Table 1 Primers used for amplification of *SpbHsfs* gene family

Names	Primer sequences		$T_m$ (°C)	Length (bp)
	Former primers	Reverse primers		
SpbHsfA1a	CGAGGTTGGTAGTGGCAGTG	ACGCCCCGAGCCTCAAGT	58/60	186
SpbHsfA1b	AGAAACTCAACAAGGCGTAA	CGGAAGGCTCGTAGGTAG	57/60	156
SpbHsfA1d	AGAACGCACGGCTCACC	CATCGCCAGGCTTTACG	60/60	178
SpbHsfA2b	AAAGGCTCAACCACATAA	TTCCCAAGGTCTTCAACT	60/59	134
SpbHsfA3	AATTTCTCCAGTTTCGTC	TTATGTGGTTGAGCCTTC	60/60	169
SpbHsfA5	CTGGTGGCAATGCTGAT	GCAAACCTCCCTGACTT	60/57	202
SpbHsfA6	GTGATGCCTGACCTGTGG	CAATGCTGGTGATGAAGT	60/60	142
SpbHsfA4c	CCACCTGCATTCCTTTA	TCGTTGTCTTCTGCGTTG	56/60	106
SpbHsfA4b	ATTAGAAGTGTTAGCCAAGC	GTAGCCAACCTCCGTGATT	58/58	159
SpbHsfA4a	CCATTATCCAGCAAAGAA	CTGCCTGAACAAAGACATTA	60/60	123
SpbHsfB1b	CCCGATAGGTGGGAGTTT	ATGTGGTGAGCCGTTTGT	58/59	225
SpbHsfB1c	CACCAATGCCACAGTCCC	CCAAGCAGTTTCGTCTAG	60/60	101
SpbHsfB1a	GGGGAACAACAGGGACTT	CACTTTGCGGAAACCATA	59/59	146
SpbHsfB2b	CAGACCACAGCAACCCAC	AACCTCCAGCAGAGCAGA	60/60	134
SpbHsfB2a	TCGCCAGCTCAACACTTA	TTCTTTCCATTCCTCCA	60/58	156
SpbHsfB4b	CGAGTTACAAACGCATACA	AAGGAGGACTTCTGAACCAT	60/61	152
SpbHsfC1b	TCACTGCGTTTACAAGAAT	AAGTAGTCGGCGGATAAG	60/60	112
SpbHsfC1a	TTATGGATTTAGGAAGGTGGAT	CGCCGAACAATGAGATGA	60/60	108
SpbHsfC2b	GCCACCTGCTCCTCCATA	CTGCCATCAATACTTGCTTA	60/60	189
SpbHsfC2c	TCACTGCGTTTACAAGAAT	AAGTAGACGGCGGATAAG	60/60	201

## 2 结果与分析

### 2.1 *SpbHsf* 基因鉴定结果与理化性质

运用本地 BLAST 程序和 Hmmer 程序检索、Pfam 数据库和在线 SMART 程序验证、同源克隆,最后获得 20 个紫鸭跖草 *Hsf* 基因家族成员,将 20 个 *SpbHsf* 基因在 NCBI 上进行注释,发现大部分与水稻 *Hsf* 基因有较高的同源性,故将各基因成员根据水稻直系同源基因进行命名(表 2)。

通过理化性质分析可知:紫鸭跖草 20 个 *Hsf* 基因的编码序列长度在 630–1 524 bp 之间,编码的蛋白质含 209–507 个氨基酸,分子量在 23.84–55.40 kDa 之间;蛋白等电点介于 4.59–8.94 之间,其中 14 个 *SpbHsf* 蛋白为酸性

蛋白,6 个 *SpbHsf* 蛋白为碱性蛋白,说明 *SpbHsf* 蛋白以酸性蛋白为主;蛋白平均疏水性在  $-0.778$  至  $-0.154$  之间,均为负值,说明 *SpbHsf* 蛋白均为亲水性蛋白;蛋白的不稳定指数在 37.62–65.25 之间,除了 *SpbHsfA2b* 蛋白的不稳定指数小于 40,为稳定蛋白外,其余 *SpbHsf* 蛋白的不稳定指数均大于 40,为不稳定蛋白。

### 2.2 *SpbHsf* 蛋白的结构

Prabi 在线软件能预测出 *SpbHsf* 蛋白二级结构的  $\alpha$  螺旋、延伸链、 $\beta$  转角和无规卷曲 4 种元件(表 1), $\alpha$  螺旋 (27.58%–55.90%) 和无规卷曲 (27.95%–55.26%) 2 种结构元件含量显著高于延伸链 (7.01%–15.04%) 和  $\beta$  转角 (27.95%–55.26%)。紫鸭跖草 *SpbHsf* 蛋白三级结构的预测如图 1 所示。亚细胞定位分析表明:

表2 *SpbHsf* 基因的基本信息  
Table 2 Basic information of *SpbHsf* gene

Names	ORF length (bp)	No. of AA	Mol Wt	pI	GRAVY	Instability index	$\alpha$ -Helix (%)	Extended strand (%)	$\beta$ -Turn (%)	Random coil (%)	Subcellular location (%)
SpbHsfA1a	1 452	483	53.50	4.59	-0.496	56.97	33.95	7.45	4.55	54.04	Endoplasmic reticulum
SpbHsfA1b	1 464	487	53.84	4.87	-0.547	47.23	31.42	9.24	5.13	54.21	Endoplasmic reticulum
SpbHsfA1d	1 524	507	55.32	4.65	-0.487	59.47	35.11	9.07	5.52	50.30	Extracellular
SpbHsfA2b	786	261	29.47	4.91	-0.778	37.62	45.59	8.81	8.43	37.16	Nucleus
SpbHsfA3	1 491	496	55.36	8.00	-0.513	59.19	37.30	12.50	6.85	43.35	Peroxisome
SpbHsfA5	1 371	456	55.40	5.16	-0.662	56.08	35.53	7.02	4.17	53.29	Nucleus
SpbHsfA6	969	322	36.82	5.03	-0.681	59.89	55.90	9.94	6.21	27.95	Nucleus
SpbHsfA4c	1 026	341	38.40	5.11	-0.730	54.31	46.33	10.56	5.28	37.83	Chloroplast
SpbHsfA4b	1 305	434	49.00	5.43	-0.715	57.59	39.17	8.76	3.00	49.08	Nucleus
SpbHsfA4a	1 203	400	45.56	5.33	-0.764	53.34	40.00	11.25	3.25	45.50	Nucleus
SpbHsfB1b	942	313	35.17	8.56	-0.680	59.73	43.45	9.90	4.15	42.49	Nucleus
SpbHsfB1c	903	300	33.52	7.71	-0.584	50.71	42.33	10.33	6.67	40.67	Nucleus
SpbHsfB1a	1 020	339	38.31	8.21	-0.615	53.61	46.61	7.96	5.01	40.41	Nucleus
SpbHsfB2b	1 065	354	38.57	6.49	0.419	54.80	34.18	9.32	4.24	52.26	Nucleus
SpbHsfB2a	10 249	342	37.07	5.36	-0.547	65.25	33.04	8.19	3.51	55.26	Endoplasmic reticulum
SpbHsfB4b	1 080	359	39.75	8.21	-0.439	56.16	27.58	15.04	5.29	52.09	Nucleus
SpbHsfC1b	645	214	24.70	8.25	-0.489	51.19	49.53	7.01	7.01	36.45	Nucleus
SpbHsfC1a	795	265	27.84	4.98	-0.154	47.23	35.85	12.83	9.06	42.26	Endoplasmic reticulum
SpbHsfC2b	630	209	23.84	8.94	-0.594	53.89	52.63	9.57	8.61	29.19	Nucleus
SpbHsfC2c	705	234	26.26	6.47	-0.525	48.72	51.71	11.54	7.69	29.06	Peroxisome



图 1 SpbHsf 蛋白的三级结构

Figure 1 The tertiary structure of SpbHsf proteins.

SpbHsfA3 和 SpbHsfC2c 蛋白定位于过氧化物酶体, SpbHsfA4c 定位于叶绿体, SpbHsfA1d 蛋白定位于细胞外, SpbHsfA1a、SpbHsfA1b、SpbHsfB2a 和 SpbHsfC1a 蛋白定位于内质网, 其余 12 个 SpbHsf 蛋白均定位于细胞核, 这表明紫鸭跖草的 *Hsf* 基因可能在多种细胞结构中具有生物学功能。

### 2.3 *SpbHsf* 基因的保守基序

*SpbHsf* 基因家族共鉴定出 20 个 motif。通过构建 *SpbHsf* 基因的保守基序组成分析图 (图 2), 显示出 *SpbHsf* 基因被分为 A、B、C 三个亚族, 从 *SpbHsf* 基因的保守序列可以看出, 所有的 *SpbHsf* 基因均含有 motif 1 和 motif 3, 从单一亚族来看, motif 8、13、11 和 19 仅存在 A 类

亚族中, motif 10、16 仅存在 B1 亚族中, motif 4 仅存在 C2 亚族中, 说明 *SpbHsf* 基因在相同亚族中具有一定的保守性。

### 2.4 SpbHsf 蛋白的结构域

通过在线数据库 Heatster, 以已鉴定的植物热激转录因子结构域为基础分析 SpbHsf 蛋白保守结构域 (表 3)。所有的 SpbHsf 蛋白中均含有 DBD 和 HR-A/B 两种保守结构域, AHA 和 NES 结构域仅存在于 A 类亚家族, RD 仅存在于 B 类亚家族, 而 C 类亚家族三者都不包含。运用 MEGA7 和 GeneDoc 2 个软件对这 20 个 SpbHsf 蛋白成员的结构域进行聚类分析 (图 3), 可知 SpbHsf 蛋白含有由 3 个  $\alpha$ -螺旋和 4 个  $\beta$ -折叠构成的 DBD 结构域<sup>[15]</sup>, SpbHsf 蛋白的

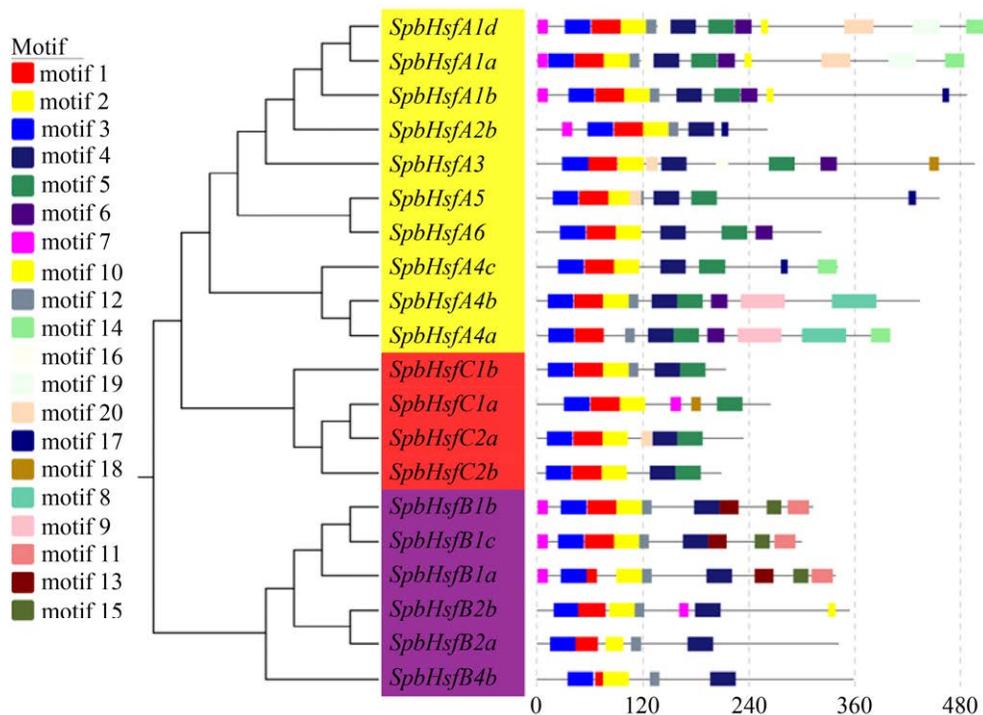


图 2 *SpbHsf* 基因的保守基序组成分析

Figure 2 The conserved motif composition of *SpbHsf* genes.

表 3 基于 Heaster 数据库分析的 *SpbHsf* 蛋白的结构域定位

Table 3 Domain location of *SpbHsf* proteins based on Heaster database analysis

Names	E-value	DBD	HR-A/B	NLS	AHA	RD	NES
SpbHsfA1d	7.8e-185	33-128	132-224	226-243	-	-	489-503
SpbHsfA1a	3.7e-184	14-109	113-205	207-224	(AHA2) 420-437	-	465-479
SpbHsfA1b	2.9e-206	37-132	139-213	233-250	(AHA2) 455-472	-	475-484
SpbHsfA2b	3.3e-96	53-150	159-173	179-199	(AHA) 224-236	-	-
SpbHsfA3	7.6e-35	27-112	225-292	361-378	(AHA3) 388-389	-	-
SpbHsfA5	1.2e-138	17-111	126-205	206-223	(AHA) 415-432	-	-
SpbHsfA6	4.7e-49	22-117	143-208	255-265	-	-	-
SpbHsfA4c	4.3e-93	21-118	134-213	214-224	(AHA1) 276-288	-	324-338
SpbHsfA4b	2.6e-144	9-104	124-203	207-226	(AHA1) 376-393	-	-
SpbHsfA4a	3.2e-77	10-105	131-198	203-222	(AHA2) 367-381	-	384-398
SpbHsfB1b	4.2e-96	23-118	160-215	284-289	-	261-269	-
SpbHsfB1c	5.2e-95	20-119	166-214	254-263	-	248-256	-
SpbHsfB1a	4.7e-86	23-118	191-237	297-306	-	290-299	-
SpbHsfB2b	2.0e-144	14-109	171-220	292-300	-	280-288	-
SpbHsfB2a	9.0e-126	10-105	163-212	282-289	-	270-278	-
SpbHsfB4b	3.7e-141	34-129	195-241	320-326	-	207-316	-
SpbHsfC1b	2.1e-96	11-106	129-191	-	-	-	-
SpbHsfC1a	6.7e-75	30-125	171-234	-	-	-	-
SpbHsfC2a	1.7e-104	13-107	129-185	-	-	-	-
SpbHsfC2b	4.9e-111	12-106	127-183	202-209	-	-	-

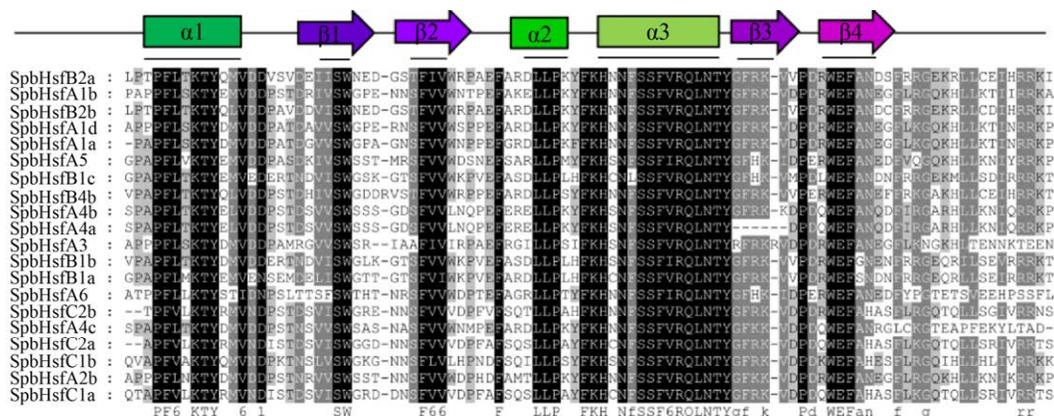


图3 SpbHsf 蛋白 DBD 结构域多序列比对

Figure 3 Multiple sequence alignment of the DBD domain of SpbHsf proteins.

DBD 结构域由 94 个氨基酸构成,其中 SpbHsf4a 蛋白的 DBD 结构域中缺失  $\beta 3$  折叠,可能是本身遗传进化的特性所导致。

## 2.5 多物种 Hsf 蛋白系统进化树分析

为了分析紫鸭跖草 *Hsf* 基因家族成员的进化特征,构建了 SpbHsf 蛋白与拟南芥、水稻、毛茛和东南景天 *Hsf* 蛋白的系统进化树(图 4),可以看出 SpbHsf 蛋白总体上被分为 A、B、C 三类。其中 A 类包含 10 个 SpbHsf 蛋白, B 类包含 6 个 SpbHsf 蛋白, C 类包含 4 个 SpbHsf 蛋白。从系统进化树可以看出紫鸭跖草 SpbHsf 蛋白与水稻 OsbHsf 蛋白的亲缘距离最近,说明紫鸭跖草与水稻的 *Hsf* 基因的亲缘性最高,这与 NCBI 上的功能注释结果一致。进化树中 *SpbHsf* 基因家族缺失 A7、A8、A9 和 B3 亚类,部分亚类缺失可能是由于转录组测序拼接的紫鸭跖草基因序列不包含所有基因组信息所导致。

## 2.6 *SpbHsf* 基因在不同铜浓度胁迫下紫鸭跖草根组织的表达模式

紫鸭跖草根组织中的 *SpbHsf* 基因在营养液组 (0.25  $\mu\text{mol/L}$   $\text{Cu}^{2+}$ )、胁迫组 (100  $\mu\text{mol/L}$ 、300  $\mu\text{mol/L}$ 、1 000  $\mu\text{mol/L}$   $\text{Cu}^{2+}$ ) 表达模式如图 5 所示,不同 *SpbHsf* 基因在不同  $\text{Cu}^{2+}$  溶液中

的表达模式可分为 3 类,第一类:有 8 个 *SpbHsf* 基因在  $\text{Cu}^{2+}$  胁迫下的表达量显著上调,其中 5 个表达量显著上调的 *SpbHsf* 基因 (*SpbHsfA4c*、*SpbHsfB2b*、*SpbHsfC1a*、*SpbHsfC1b*、*SpbHsfB2b*) 在 100  $\mu\text{mol/L}$   $\text{Cu}^{2+}$  溶液中的表达量最高,3 个表达量显著上调的 *SpbHsf* 基因 (*SpbHsfA1b*、*SpbHsfA3*、*SpbHsfA4b*) 在 300  $\mu\text{mol/L}$   $\text{Cu}^{2+}$  溶液中的表达量最高,说明这 8 个 *SpbHsf* 基因在不同  $\text{Cu}^{2+}$  溶液中始终参与胁迫应答反应,不同 *SpbHsf* 基因在不同  $\text{Cu}^{2+}$  溶液中的应答反应程度有所不同;第二类:有 4 个 A 类基因 (*SpbHsfA1a*、*SpbHsfA1d*、*SpbHsfA4a*、*SpbHsfA5*)、2 个 B 类基因 (*SpbHsfB1a*、*SpbHsfB4b*) 和 2 个 C 类基因 (*SpbHsfC2a*、*SpbHsfC2b*) 在  $\text{Cu}^{2+}$  胁迫下的表达量显著下调,说明紫鸭跖草的根组织中,  $\text{Cu}^{2+}$  的胁迫在一定程度上会抑制这 8 个 *SpbHsf* 基因的表达;第三类:有 4 个 *SpbHsf* 基因的表达量随着  $\text{Cu}^{2+}$  浓度增加先显著上调后显著下调,其中 *SpbHsfA2b*、*SpbHsfB1a*、*SpbHsfB2a* 基因在 100  $\mu\text{mol/L}$   $\text{Cu}^{2+}$  溶液中表达量最高, *SpbHsfB1b* 基因在 300  $\mu\text{mol/L}$   $\text{Cu}^{2+}$  溶液中表达量最高,说明这 4 个 *SpbHsf* 基因在低浓度  $\text{Cu}^{2+}$  胁迫中参与应答反应,在高浓度  $\text{Cu}^{2+}$

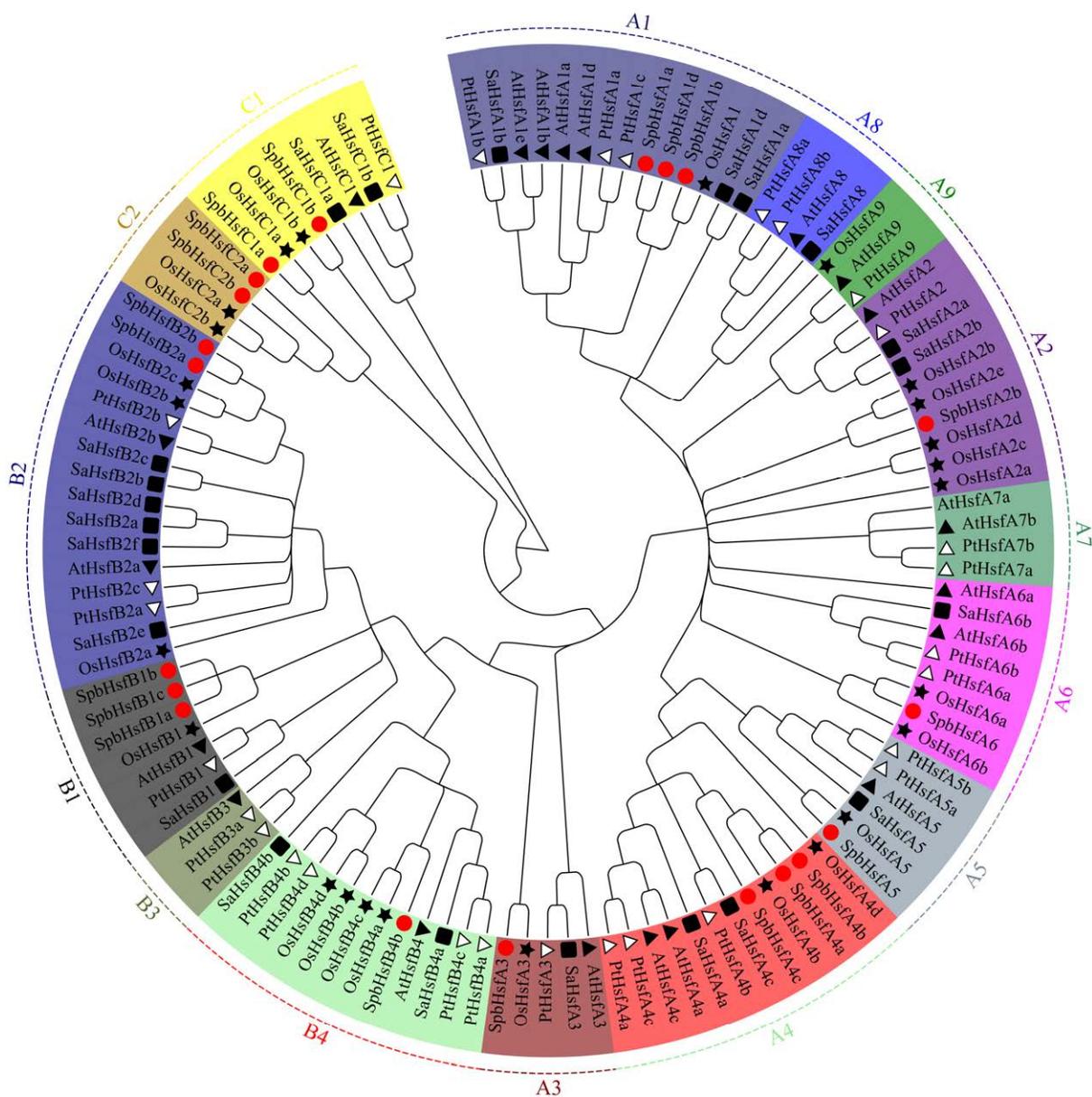


图 4 5 种物种 *Hsf* 基因的系统进化树

Figure 4 Phylogenetic trees of *Hsf* genes from 5 species.

胁迫中的表达受到抑制从而不参与应答反应。通过铜胁迫下的表达模式分析，我们发现紫鸭跖草中的 *SpbHsfA4b*、*SpbHsfB2b* 基因的显著上调表达与镉胁迫下东南景天的 *SaHsfA4*、*SaHsfB2* 的表达结果一致<sup>[13]</sup>，这一结果表明植物的 *SpbHsfA4* 亚类和 *SpbHsfB2* 亚类可能在参与调节重金属胁迫反应过程中具有重要的作用。

### 3 讨论

植物的生长发育和生产受到许多外界生物和非生物胁迫的干扰，如热、冷、潮和重金属污染等<sup>[16]</sup>。在应对外界的各种胁迫中，植物都具有自己独特的防御系统或应对机制，在此过程中会涉及不同类型的转录因子。除了 Hsf，还

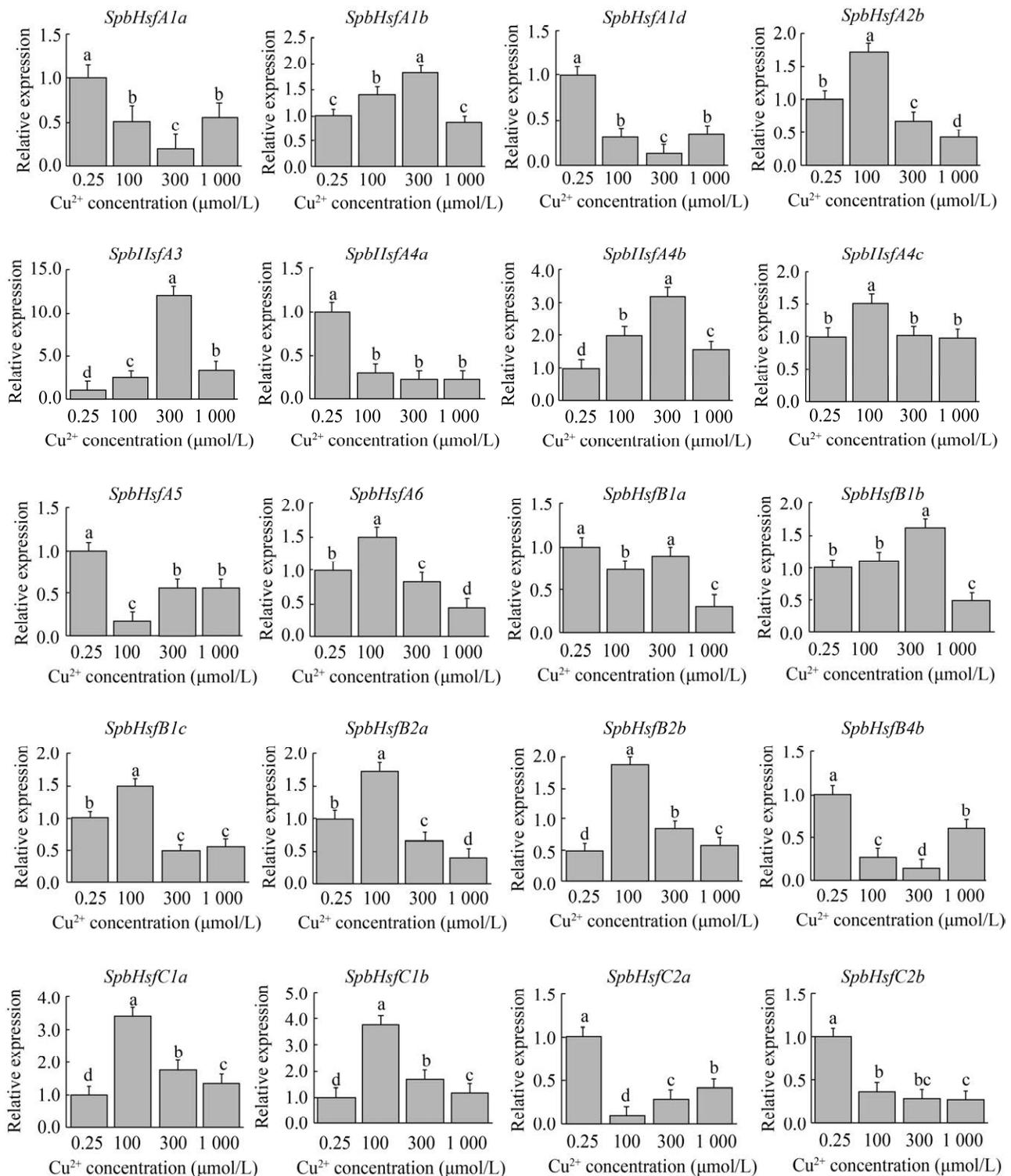


图 5 根组织中的 *SpbHsf* 基因家族成员在不同铜浓度胁迫下的表达模式分析

Figure 5 Expression pattern of *SpbHsf* genes in root tissue under different copper concentrations. Different lowercase letters indicate significant difference ( $P < 0.05$ ).

有 WRKY、MYB、AP2/ERF、NAC 等, 它们在不同的应激条件下调控数千个基因的表达来抵御外界刺激<sup>[17-19]</sup>。在植物中, *Hsf* 基因家族是植物中最重要的转录因子家族之一, 其不仅受高温诱导, 而且在抵抗非生物胁迫时也起到一定的作用<sup>[20]</sup>。研究植物对热、冷、潮和重金属等条件下的分子机制是提高植物抗逆性和生产力的前提, 其基因产物起着为植物提供抗逆性的作用<sup>[21]</sup>。植物 *Hsf* 基因是调节各种应激反应基因表达的重要转录因子, 在提供对多种非生物胁迫的耐受性方面起着关键作用<sup>[22]</sup>。

本研究通过挖掘紫鸭跖草全长转录组数据库数据, 筛选鉴定出 20 个 *SpbHsf* 基因。并对 *SpbHsf* 基因的理化性质与结构、系统进化树、保守基序、结构域及在不同 Cu<sup>2+</sup>浓度胁迫下表达模式进行分析。保守基序分析表明, 各亚类的 *SpbHsf* 基因均含有一类或几类相同的保守基序, 表明 *SpbHsf* 基因在进化过程中相对比较保守。结构域分析表明, *SpbHsf* 蛋白含有 3 个  $\alpha$ -螺旋和 4 个  $\beta$ -折叠的特征的 DBD 结构域, 表明鉴定出的 20 个 *SpbHsf* 基因具有 *Hsf* 基因的结构特征。在与 5 个物种的系统进化树构建中, 不同物种间的各个亚类进化保守、层次分明, 表明 *SpbHsf* 基因与水稻的 *OsHsf* 基因的同源性最高。

*Hsf* 基因会通过调节活性氧 (reactive oxygen species, ROS) 的浓度和调控下游基因/蛋白的表达来增强非生物胁迫的耐受性。例如, *HsfA1a* 基因可通过调节咖啡酸-O-甲基转移酶 (COMT) 基因的转录和诱导褪黑激素的积累来增强番茄的镉耐受性, 褪黑素可与 ROS 发生氧化还原反应, 从而维持胁迫条件下植物体内的氧化还原平衡。番茄的 *HsfA1a* 基因沉默会降低褪黑激素的含量, 同时番茄的镉耐受性也随之减低; *HsfA1a* 基因的过表达能刺激 *COMT* 基因

的表达及褪黑激素的积累, 并增强番茄的镉耐受性<sup>[23-24]</sup>。陈双双<sup>[13]</sup>发现东南景天 *SaHsfA4ah* 和 *SaHsfA4ch* 基因参与调控 Cd 胁迫应答, 野生型拟南芥受到镉胁迫后, 其叶片会发黄甚至枯萎, 但过表达 *SaHsfA4ah* 或 *SaHsfA4ch* 基因的拟南芥生长状况较好, 镉吸收速率高于野生型, 且能通过调控下游 *AtNramp6* 基因和 *AtHsps* 蛋白的表达水平来增强拟南芥的耐镉性。已有研究表明, DBD 结构域是 *Hsf* 蛋白参与胁迫应答的关键结构, Chen 等发现含有 DBD 结构域的 *SaHsfA4c* 蛋白具有镉抗性, 而包含不完整 DBD 结构域的 *InSaHsfA4c* 蛋白不具有镉抗性<sup>[25]</sup>。我们发现除了 *SpbHsf4a* 蛋白, 其他 *SpbHsf* 蛋白都含有 motif 2, 这可能是由于 *SpbHsf4a* 蛋白的 DBD 结构域缺失  $\beta$ 3 折叠所导致, 且从 *SpbHsf4a* 基因的表达模式中可以发现, *SpbHsf4a* 基因在铜胁迫下的表达量始终显著下调, 说明不含完整 DBD 结构域的 *SpbHsf4a* 蛋白不具有抗铜性。*Hsf* 基因是一类研究相对比较深入的基因家族, 目前对植物 *Hsf* 基因的功能在重金属的研究多集中于镉胁迫, 目前关于 *Hsf* 基因在铜胁迫响应的相关研究不够深入, 例如桑建<sup>[26]</sup>发现东南景天的 3 个 *SaHsf* 基因在铜胁迫下的表达显著上调, 董雪芳<sup>[27]</sup>发现菠菜的 *SoHsf* 基因在酵母中异源表达能提高酵母对铜的耐受性, 但他们均未对 *Hsf* 基因的耐铜机制进行进一步研究及分析。基因的表达模式与它们所行使的功能息息相关, 分析表达模式有利于找出潜在的胁迫应答基因。本文通过分析不同铜浓度胁迫下紫鸭跖草根组织的 20 个 *SpbHsf* 基因的表达模式, 并从中鉴定出 8 个未报道过的潜在铜胁迫应答基因, 为后续研究紫鸭跖草耐铜性机制及 *Hsf* 基因在铜胁迫的应答机制提供了一定的理论依据。

## 4 结论

本研究共筛选鉴定出 20 个 *SpbHsf* 基因家族成员, 预测分析表明, *SpbHsf* 基因作用于 5 个细胞结构中, *SpbHsf* 蛋白均含有 DBD 及 HR-A/B 保守结构域。20 个 *SpbHsf* 基因在不同  $\text{Cu}^{2+}$  浓度中表达模式不同, 发现 8 个 *SpbHsf* 基因会受不同铜浓度胁迫诱导显著上调表达, 其中 *SpbHsfA4b* 和 *SpbHsfB2b* 基因的表达模式与镉胁迫下东南景天 *SaHsf* 基因的表达模式相似, 表明 *SpbHsf* 基因在防御  $\text{Cu}^{2+}$  毒害上起一定作用。本研究为研究紫鸭跖草的耐铜性机制和 *Hsf* 基因应对  $\text{Cu}^{2+}$  胁迫的分子机制提供了初步依据。

## REFERENCES

- [1] Nover L, Bharti K, Döring P, et al. *Arabidopsis* and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell Stress Chaperones*, 2001, 6(3): 177-189.
- [2] 黄小云, 陶鹏, 李必元, 等. 植物热激转录因子基因家族的研究进展. *浙江农业科学*, 2014(9): 1323-1332, 1336.  
Huang XY, Tao P, Li BY, et al. Research progress of plant heat shock transcription factor gene family. *J Zhejiang Agric Sci*, 2014(9): 1323-1332, 1336 (in Chinese).
- [3] Von Koskull-Döring P, Scharf KD, Nover L. The diversity of plant heat stress transcription factors. *Trends Plant Sci*, 2007, 12(10): 452-457.
- [4] 瞿光航. 梭梭热激转录因子 *HaHsfA3*、*HaHsfB3* 的克隆、亚细胞定位和表达模式分析[D]. 南京: 南京农业大学, 2016.  
Qu GH. The cloning subcellular localization and expression patten analysis of heat shock transcription factor *HaHsfA3*, *HaHsfB3* in *Haloxylon ammodendron*[D]. Nanjing: Nanjing Agricultural University, 2016 (in Chinese).
- [5] Singh A, Mittal D, Lavania D, et al. OsHsfA2c and OsHsfB4b are involved in the transcriptional regulation of cytoplasmic *OsClpB* (*Hsp100*) gene in rice (*Oryza sativa* L.). *Cell Stress Chaperones*, 2012, 17(2): 243-254.
- [6] Cheng Q, Zhou Y, Liu Z, et al. An alternatively spliced heat shock transcription factor, *OsHSFA2dI*, functions in the heat stress-induced unfolded protein response in rice. *Plant Biol*, 2015, 17(2): 419-429.
- [7] Giesguth M, Sahm A, Simon S, et al. Redox-dependent translocation of the heat shock transcription factor AtHSFA8 from the cytosol to the nucleus in *Arabidopsis thaliana*. *FEBS Lett*, 2015, 589(6): 718-725.
- [8] Lin YX, Jiang HY, Chu ZX, et al. Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. *BMC Genomics*, 2011, 12: 76.
- [9] Li W, Wan XL, Yu JY, et al. Genome-wide identification, classification, and expression analysis of the *Hsf* gene family in carnation (*Dianthus caryophyllus*). *Int J Mol Sci*, 2019, 20(20): 5233.
- [10] Nishizawa A, Yabuta Y, Yoshida E, et al. *Arabidopsis* heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant J*, 2006, 48(4): 535-547.
- [11] Ogawa D, Yamaguchi K, Nishiuchi T. High-level overexpression of the *Arabidopsis HsfA2* gene confers not only increased thermotolerance but also salt/osmotic stress tolerance and enhanced callus growth. *J Exp Bot*, 2007, 58(12): 3373-3383.
- [12] Personat JM, Tejedor-Cano J, Prieto-Dapena P, et al. Co-overexpression of two heat shock factors results in enhanced seed longevity and in synergistic effects on seedling tolerance to severe dehydration and oxidative stress. *BMC Plant Biol*, 2014, 14: 56.
- [13] 陈双双. 超积累型东南景天 *SaHsf* 基因家族分析及 *SaHsfA4a/c* 耐镉功能研究[D]. 北京: 中国林业科学研究院, 2018.  
Chen SS. Analysis of *SaHsf* family in hyperaccumulator *Sedum alfredii* Hance and cadmium tolerance function of *SaHsfA4a/c*[D]. Beijing: Chinese Academy of Forestry Sciences, 2018 (in Chinese).
- [14] 黄长干. 紫鸭跖草对铜的积累规律及在铜胁迫下的生理反应研究[D]. 长沙: 湖南农业大学, 2007.  
Huang CG. Study on copper accumulation rule and physiological responses to copper stress in *Setcreasea purpurea* boom[D]. Changsha: Hunan Agricultural University, 2007 (in Chinese).
- [15] 徐坚, 王燕, 陈先知, 等. 大白菜热激转录因子基因家族鉴定及表达分析. *核农学报*, 2014, 28(4): 586-596.  
Xu J, Wang Y, Chen XZ, et al. Genome-wide identification and analysis of heat shock factors family in Chinese cabbage. *J Nuclear Agric Sci*, 2014, 28(4):

- 586-596 (in Chinese).
- [16] 崔岩伟. 拟南芥受体激酶 CIKs 决定早期花药的细胞命运[D]. 兰州: 兰州大学, 2018.  
Cui YW. CIK receptor kinases determine cell fate specification during early anther development in *Arabidopsis*[D]. Lanzhou: Lanzhou University, 2018 (in Chinese).
- [17] Guo M, Liu JH, Ma X, et al. The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. *Front Plant Sci*, 2016, 7: 114.
- [18] 成舒飞, 端木慧子, 陈超, 等. 大豆 MYB 转录因子的全基因组鉴定及生物信息学分析. *大豆科学*, 2016, 35(1): 52-57.  
Cheng SF, Duanmu HZ, Chen C, et al. Whole genome identification of soybean MYB transcription factors and bioinformatics analysis. *Soyb Sci*, 2016, 35(1): 52-57 (in Chinese).
- [19] 张书玲, 刘莹, 张惠敏, 等. 棉花 WRKY 转录因子家族成员的鉴定和生物信息学分析. *中国农业科技导报*, 2017, 19(8): 9-15.  
Zhang SL, Liu Y, Zhang HM, et al. Genome-wide identification and analysis of WRKY transcription factors in *G. arboreum*. *J Agric Sci Technol*, 2017, 19(8): 9-15 (in Chinese).
- [20] Zhang J, Liu BB, Li JB, et al. *Hsf* and *Hsp* gene families in *Populus*: genome-wide identification, organization and correlated expression during development and in stress responses. *BMC Genomics*, 2015, 16: 181.
- [21] 李丽, 刘双清, 杨远航, 等. 热激转录因子在植物抗非生物胁迫中的功能研究进展. *生物技术进展*, 2018, 8(3): 214-220.  
Li L, Liu SQ, Yang YH, et al. Progress on the function of heat shock transcription factors in plant abiotic stress tolerance. *Curr Biotechnol*, 2018, 8(3): 214-220 (in Chinese).
- [22] Wang XM, Shi X, Chen SY, et al. Evolutionary origin, gradual accumulation and functional divergence of heat shock factor gene family with plant evolution. *Front Plant Sci*, 2018, 9: 71.
- [23] 孙莎莎, 韩亚萍, 闫燕燕, 等. 过表达咖啡酸-O-甲基转移酶基因(*COMT1*)调控番茄幼苗对干旱胁迫生理响应. *植物生理学报*, 2019, 55(8): 1109-1122.  
Sun SS, Han YP, Yan YY, et al. Overexpression of caffeic acid-O-methyltransferase gene (*COMT1*) regulates physiological response of tomato seedlings to drought stress. *Plant Physiol J*, 2019, 55(8): 1109-1122 (in Chinese).
- [24] Cai SY, Zhang Y, Xu YP, et al. HsfA1a upregulates melatonin biosynthesis to confer cadmium tolerance in tomato plants. *J Pineal Res*, 2017, 62(2): e12387.
- [25] Chen SS, Yu M, Li H, et al. *SaHsfA4c* from *Sedum alfredii* Hance enhances cadmium tolerance by regulating ROS-scavenger activities and heat shock proteins expression. *Front Plant Sci*, 2020, 11: 142.
- [26] 桑健. 基于转录组测序的超积累型东南景天内参基因筛选与 HSF 基因家族初步分析[D]. 北京: 中国林业科学研究院, 2014.  
Sang J. Validation of reference genes and analysis of *HSF* gene family based on RNA-Seq in hyper-accumulating *Sedum alfredii* Hance[D]. Beijing: Chinese Academy of Forestry, 2014 (in Chinese).
- [27] 董雪芳. 菠菜 *HsfA1* 和 *HsfBs* 基因的功能研究[D]. 上海: 上海师范大学, 2019.  
Dong XF. Functional studies of *HsfA1* and *HsfBs* genes in *spinacia oleracea*[D]. Shanghai: Shanghai Normal University, 2019 (in Chinese).

(本文责编 郝丽芳)