

蓖麻 RcHSF 基因家族鉴定与冷胁迫下的表达模式分析

刘栩铭^{1,2}, 李敏³, 段琼^{1,2}, 张洪雨^{1,2}, 孟迪^{1,2}, 张丽雪^{1,2},
张继星^{1,2,4}, 王晓宇^{1,2,4}

(1. 内蒙古民族大学 生命科学与食品学院, 内蒙古 通辽 028000; 2. 内蒙古民族大学 科尔沁植物逆境生物学研究所, 内蒙古 通辽 028000;
3. 内蒙古民族大学 农学院, 内蒙古 通辽 028000; 4. 内蒙古自治区高校蓖麻产业工程技术研究中心, 内蒙古自治区蓖麻育种重点实验室,
内蒙古自治区蓖麻产业协同创新培育中心, 内蒙古 通辽 028000)

摘要:为明确蓖麻 HSF 基因家族的结构特征与冷胁迫下的表达模式, 利用序列比对法实现蓖麻 HSF 基因家族成员的全基因组鉴定, 通过生物信息学手段对基因家族成员的序列结构、理化性质、染色体定位、启动子顺式作用元件、蛋白保守结构域、系统进化关系以及聚类方式进行分析, 并结合 RNA-Seq 与 qRT-PCR 对 RcHSFs 的表达水平进行验证。结果表明, 蓖麻基因组中含有 19 个 RcHSF 基因家族成员; 基因分析发现 RcHSFs 分别含有 1~3 个外显子, 定位在 17 个染色体片段上且呈不均匀分布, 其中大多数 RcHSFs 均含有 AAGAA-motif、MYB 与 MYC 元件, 暗示其响应植物激素与逆境胁迫; 蛋白分析表明, RcHSFs 按照基序类型与排布方式被聚为 A、B、C 3 类, 推测不同类型 HSF 的蛋白功能存在差异, 尽管所有成员均存在不同程度的 motif 缺失, 但结构域高度保守; RNA-Seq 分析发现仅 *RcHSF4* 与 *RcHSF10* 受冷胁迫诱导表达, 与 qRT-PCR 验证结果一致。结果为 RcHSFs 基因的克隆与功能研究提供重要的理论基础, 为蓖麻抗冷新种质资源的建立提供候选基因。

关键词: 蓖麻; HSF 基因家族; 冷胁迫; 表达模式

中图分类号: Q78; S563.03 文献标识码: A 文章编号: 1000-7091(2020)05-0062-010

doi: 10.7668/hbxb.20191312



Identification of RcHSF Gene Family in Castor and Analysis of Expression Pattern under Cold Stress

LIU Xuming^{1,2}, LI Min³, DUAN Qiong^{1,2}, ZHANG Hongyu^{1,2}, MENG Di^{1,2},
ZHANG Lixue^{1,2}, ZHANG Jixing^{1,2,4}, WANG Xiaoyu^{1,2,4}

(1. College of Life Sciences and Food Engineering, Inner Mongolia University for Nationalities, Tongliao 028000, China; 2. Horqin Plant Stress Biology Research Institute of Inner Mongolia University for Nationalities, Tongliao 028000, China; 3. Agricultural College, Inner Mongolia University for Nationalities, Tongliao 028000, China; 4. Inner Mongolia Industrial Engineering Research Center of Universities for Castor, Inner Mongolia Key Laboratory of Castor Breeding, Inner Mongolia Collaborative Innovation Cultivate Center for Castor, Tongliao 028000, China)

Abstract: The purpose of this study was to illustrate the structure traits of HSF genes family in castor and its expression pattern in different tissues. The bioinformatics method was used to analyze the sequence structure, physical and chemical properties, chromosome positioning, and cis-acting elements of promoters, the conserved domains, phylogenetic relationships, and clustering methods of the encoded protein. RNA-Seq and qRT-PCR was

收稿日期: 2020-06-11

基金项目: 国家自然科学基金项目(31701467; 31760399); 内蒙古自治区高等学校“青年科技英才”支持计划项目(NJYT-20-B25); 国家服务特殊需求蒙药学博士点建设专项开放基金项目(MDMYBJ2019008); 内蒙古民族大学博士科研启动基金项目(BS398); 内蒙古民族大学自治区科技储备项目(2018NDCB05-3)

作者简介: 刘栩铭(1994-), 男, 内蒙古阿拉善人, 硕士, 主要从事植物抗逆分子生物学研究。

通讯作者: 王晓宇(1987-), 男, 内蒙古通辽人, 副教授, 博士, 主要从事植物抗逆分子生物学研究。

张继星(1971-), 男, 内蒙古呼和浩特人, 教授, 博士, 主要从事植物抗逆分子生物学研究。

used to assess the expression of RcHSFs. The results showed that the castor genome contained 19 members of the RcHSF gene family. Gene structure analysis showed that RcHSFs contained 1 – 3 exons, and were located on 17 chromosome fragments with uneven distribution; most RcHSFs contained AAGAA-motif, MYB and MYC elements, which indicated that these genes were responsive to plant hormone and stress; Protein analysis showed that RcHSFs were grouped into A, B, and C categories according to motif type and arrangement, which indicated that the protein functions of different types of HSF are different. Although all members have different degrees of motif deletion, the domain is highly conserved; RNA-Seq analysis showed that only *RcHSF4* and *RcHSF10* were induced by cold stress, which was consistent with qRT-PCR verification results. This result provides an important theoretical basis for the cloning and functional research of RcHSFs, and also provides candidate genes for the establishment of new germplasm resources of castor cold resistance.

Key words: Castor; HSF gene family; Cold stress; Expression pattern

热激转录因子 (Heat shock transcription factor, HSF) 是生物体参与热胁迫信号转导与耐热性调控的关键转录因子^[1]。研究表明, HSF 在接收和传递信号、调控下游基因表达、激发植物抗逆反应以及提高植物耐热性的过程中发挥重要作用^[2]。热激转录因子能够通过特异性识别并结合热激元件 (Heat shock element, HSE) 激活热激蛋白 (Heat shock protein, HSP)。热激蛋白作为分子伴侣, 协助受损蛋白重新折叠、组装、分配和降解, 从而提高植物热抵抗能力^[3-4]。研究表明, 在真核生物中典型的 HSF 具有 5 个保守结构域, 即 DNA 结合域 (DNA-binding domain, DBD)、寡聚化结构域 (Oligomerization domain, OD)、转录激活结构域 (C-terminal activation domain, CTAD)、核定位信号 (Nuclear localization signals, NLS) 以及核输出信号 (Nuclear export signals, NES)。其中 DBD 与 OD 结构域最为保守, 分别用于精准识别并结合 HSE 与 HSP 的启动子; NLS 和 NES 结构域用于保障 HSF 的跨核运输, 维持了基因在细胞核与细胞质中的自由分布; CTAD 具有转录激活功能, 可以直接调控热激过程中的相关基因表达^[5-8]。

HSF 基因最早在酵母中被克隆, 随后多种哺乳动物 HSF 相继被挖掘。然而植物的第 1 个 HSF 基因是在番茄中克隆得到的^[5], 随着基因组测序技术的不断完善, HSF 又相继在拟南芥、水稻、大豆、玉米等植物中被报道。不同植物的 HSF 基因家族成员数量差异较大, 小麦中数量最多含有 56 个^[8], 大豆^[9]、玉米^[10]、拟南芥^[11] 中分别含有 52, 30, 21 个。目前有关植物 HSF 基因功能的研究主要集中在高温、干旱响应与生长发育等方面^[2, 7, 12-16]。此外, 其他研究表明 HSF 在冷胁迫应答中发挥着重要作用。朱华生^[17] 指出玉米 *ZmHsf04* 基因受低温胁迫诱导表达; 余文才等^[18] 发现低温条件下巴西橡胶树

HbHsfA4a 基因的表达量显著上调。

蓖麻 (*Ricinus communis* L.) 是一种油料作物, 被广泛应用于医药、化工、航空航天等领域, 其含油量丰富, 稳居世界十大油料作物之首, 有“可再生绿色石油资源”的美称^[19]。作为一种热带作物, 蓖麻具有喜温不耐冷的特点, 但目前关于蓖麻 HSF 基因的研究还未见报道。本研究对蓖麻 HSF 基因家族进行全基因组鉴定, 利用生物信息学手段分析该基因家族成员的序列特征、染色体定位及基因组复制事件, 结合冷胁迫下蓖麻叶片的转录组数据, 分析蓖麻 RcHSF 基因在冷胁迫下的表达模式, 以期为更深入研究蓖麻 HSF 基因家族成员的冷应答机制与生物学功能奠定基础。

1 材料和方法

1.1 试验材料

本研究选用通蓖 5 号为试验材料, 由内蒙古通辽市农业科学研究院提供。将混合的草炭土与蛭石 (混合比例 3:1) 作为栽培土壤, 设置室内培养温度 24 °C, 相对湿度 70%~75%, 光照强度 450 $\mu\text{mol}/(\text{m}^2 \cdot \text{s})$ 。当植株生长到四叶期时, 取生长状态相同的植株分别进行 0, 2, 6, 8 h 冷处理 (4 °C)。每株幼苗取两片真叶作为样本, 将 3 株幼苗混样作为 1 次生物学重复, 共设置 3 次生物学重复。液氮速冻样本, 用于转录组测序与基因表达试验。

1.2 试验方法

1.2.1 总 RNA 提取、cDNA 合成及转录组测序

通过 TRIzol 法提取叶片的总 RNA, 利用 DNase I 去除 gDNA 污染, Agilent2100 Bioanalyzer 和 NanoDrop 分别检测 RNA 的完整性和纯度, 质检合格后进行文库构建, 随后使用 Illumina HiSeq X-ten 测序仪进行转录组测序。UEIris RT-PCR System (US Everbright, 苏州, 中国) 反转录合成 cDNA 第 1 条链, 用

于后续的 qRT-PCR 分析。

1.2.2 蓖麻 HSF 基因的鉴定与序列分析 蓖麻基因组数据与结构注释信息下载于 Phytozome 数据库 (<https://phytozome.jgi.doe.gov/pz/portal.html>), 21 个拟南芥 HSF 基因家族成员序列下载于 Plant Tair 数据库 (<https://www.arabidopsis.org/>), 24 个水稻 HSF 基因家族序列信息下载于 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/>)。利用基因组结构注释信息提取蓖麻全基因组 CDS, 成功构建蓖麻蛋白数据库。以拟南芥 HSF 基因家族成员为探针检索蓖麻蛋白数据库, 去除冗余的蓖麻蛋白, 预筛选 HSF 同源蛋白, E-value 值小于 0.01。提交预筛选蛋白于 NCBI-swissprot 数据库检索同源蛋白信息, 去除蓖麻基因组中 HSF 家族成员的近源序列, 最终将含有 HSF 保守结构域的蛋白序列保留。

1.2.3 蓖麻 HSF 家族成员的理化性质 利用 Prot Param 工具 (<https://web.expasy.org/protparam/>) 分析蓖麻 HSF 蛋白的氨基酸数量、等电点、分子量。利用 CELLO 工具 (<http://cello.life.nctu.edu.tw/>) 在线预测蓖麻 HSF 家族成员的亚细胞定位信息。利用 TMHMM 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) 工具统计蓖麻 HSF 家族成员的跨膜结构域。

1.2.4 蓖麻 HSF 保守基序与结构域分析 利用 MEME 在线工具 (<http://meme.nbcr.net/meme/tools/meme>) 预测蓖麻 HSF 蛋白的保守 motif; 参数设置: 设置位点分布情况为任意重复次数, 设置保守基序上限为 10, 设置基序大小为 10 ~ 60 bp, 其他参数均为默认。利用 Batch CD-search 在线工具 (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) 分析蛋白的保守结构域。

1.2.5 蓖麻 HSF 家族成员进化分析和组分鉴定

利用 MEGA 5.1 工具对拟南芥、水稻与蓖麻 HSF 蛋白序列进行比对, 最大迭代次数设置为 100, 其他参数默认; 参考比对结果, 采用邻接法 (Neighbor-Joining, NJ) 构建系统进化树, 选择 JTT 模型, 采用 Partial deletion 处理遗漏数据, 设置 Bootstrap 检验次数为 1 000。

1.2.6 蓖麻 HSF 染色体定位信息 从蓖麻基因组注释信息中获取 RcHSF 基因家族成员的染色体定位信息, 使用 TBtools 绘制基因染色体分布图, 通过微共线性分析, 探讨基因复制事件。

1.2.7 蓖麻 HSF 基因家族成员顺式作用元件预测 提取 RcHSF 基因起始位点上游 2 000 bp 的启动子区序列, 利用 Plant CARE 在线工具预测该家族成员与环境胁迫、激素调控等有关的顺式作用元件, 以此进一步推测基因家族成员具有的潜在功能。

1.2.8 蓖麻 HSF 在冷胁迫下的转录组分析 结合蓖麻冷胁迫 (4 °C) 下的转录组数据, 对 FoldChange 标准化数值进行可视化分析, 分析 RcHSF 基因在冷胁迫下的转录水平。

1.2.9 蓖麻 HSF 在冷胁迫下的表达验证 利用 qRT-PCR 验证差异表达基因在冷胁迫下的表达水平。参考王晓宇等^[20]的研究方法, 将 *RcADP* (XP_002515227) 与 *RcEF1β* (XP_002511200) 作为内参基因。利用 Primer premier 5 设计目的基因与内参基因的特异性引物 (表 1)。qRT-PCR 所用仪器为 qTOWER 2.2 (Analytik Jena, 德国), 其反应体系与程序设置参考 2 × SYBR Green qPCR Master Mix (USEverbright, 中国, 苏州) 试剂盒说明书。数据采用平均数 ± 标准差表示, 并利用 2^{-ΔΔCt} 法进行基因相对表达量的计算, 两组间比较采用 Students' t-test, 当 P < 0.05 时认为统计学上具有显著差异。

表 1 qRT-PCR 引物序列信息

Tab. 1 Primers used for qRT-PCR

| 引物名称 Primer name | 正向引物(5' - 3') Forward primer | 反向引物(5' - 3') Reverse primers | 退火温度/°C T _m |
|---------------------|---------------------------------|----------------------------------|---------------------------|
| <i>RcADP</i> | GAGAGATGCTGTGCTGCTTG | TGAAGGCCGAGCTTATCACT | 55 |
| <i>RcEF1β</i> | GCTCGAGGAAGCAGTTAGGA | TTCCATATCCAACCTGGCACA | 55 |
| <i>RcHSF4</i> | TTTCGCATAGTATCACCG | CGAACACGAGGACACCTA | 55 |
| <i>RcHSF10</i> | AGCAAGCCTTTGACTGTC | TGTGCCTTCTATTGGGTC | 55 |

2 结果与分析

2.1 蓖麻 HSF 基因家族鉴定与结构分析

本研究从 NCBI 数据库获取蓖麻全基因组信息, 通过拟南芥 HSF 基因家族比对和 Swissprot 数据库验证, 筛选出 20 个蓖麻 HSF 基因家族成员。利

用 NCBI-Batch CD-search 保守结构域预测去除 1 个错误比对 (30190. m011154. TIGRR0. 1), 最终成功鉴定到 19 个蓖麻 HSF 基因家族成员, 并根据染色体定位顺序对所有成员依次命名 (表 2)。CELLO 工具预测显示所有蛋白均定位于细胞核; Prot Param 工具分析表明, 19 个 RcHSF 蛋白的氨基酸个数在

84 (RcHSF16) ~ 521 (RcHSF2), 平均氨基酸个数为 366; 分子量 (ku) 大小在 9.605 8 (RcHSF16) ~ 58.650 6 (RcHSF2); 等电点 (pI) 为 4.68 (RcHSF11) ~ 9.76 (RcHSF7), 其中 RcHSF11 最小, 显示出较弱的溶解度与更强的沉淀能力, 而 RcHSF7 分子间相对作用力较大, 推测沉降系数低。尽管 RcHSF16 的氨基酸数量明显少于 HSF 家族其他成员, 但基于 Domains 保守结构域分析, 该蛋白存在明显的 HSF 活性位点, 推测能发挥完整的 HSF 蛋白功能。

RcHSFs 基因结构分析显示, 所有成员均含有长度与数量不等的 UTR 与 CDS, 基因 *RcHSF1*、*RcHSF2*、*RcHSF4*、*RcHSF5*、*RcHSF6*、*RcHSF7*、*RcHSF8*、*RcHSF9*、*RcHSF11*、*RcHSF14*、*RcHSF15*、*RcHSF16*、*RcHSF18* 均呈现以外显子开头和结尾的特殊结构, 其中 *RcHSF15* 仅含有 1 个外显子, *RcHSF1* 与 *RcHSF4* 含有 3 个外显子, 其他成员均含有 2 个外显子 (图 1)。

表 2 蓖麻 RcHSF 基因家族成员信息

Tab. 2 The information of RcHSF gene family in *Ricinus communis* L.

| 基因命名 Gene names | 基因组注释 ID Genome annotation ID | 编码氨基酸数量/个 Amino acids number | 分子量 /ku Molecule weight | 等电点 pI |
|--------------------|----------------------------------|---------------------------------|----------------------------|-----------|
| <i>RcHSF1</i> | 28312. m000133. TIGRR0.1 | 191 | 22.144 0 | 6.80 |
| <i>RcHSF2</i> | 29092. m000454. TIGRR0.1 | 521 | 58.650 6 | 4.89 |
| <i>RcHSF3</i> | 29629. m001435. TIGRR0.1 | 491 | 54.951 9 | 5.72 |
| <i>RcHSF4</i> | 29636. m000750. TIGRR0.1 | 464 | 52.555 7 | 5.99 |
| <i>RcHSF5</i> | 29739. m003565. TIGRR0.1 | 371 | 41.674 2 | 4.76 |
| <i>RcHSF6</i> | 29844. m003222. TIGRR0.1 | 359 | 41.213 3 | 5.24 |
| <i>RcHSF7</i> | 29851. m002407. TIGRR0.1 | 218 | 24.931 4 | 9.76 |
| <i>RcHSF8</i> | 29883. m001955. TIGRR0.1 | 360 | 41.024 0 | 5.73 |
| <i>RcHSF9</i> | 29912. m005397. TIGRR0.1 | 478 | 54.185 3 | 4.92 |
| <i>RcHSF10</i> | 29912. m005526. TIGRR0.1 | 337 | 37.322 8 | 5.16 |
| <i>RcHSF11</i> | 29968. m000635. TIGRR0.1 | 409 | 46.934 7 | 4.68 |
| <i>RcHSF12</i> | 30006. m000287. TIGRR0.1 | 248 | 28.738 2 | 7.01 |
| <i>RcHSF13</i> | 30026. m001484. TIGRR0.1 | 402 | 46.226 4 | 5.15 |
| <i>RcHSF14</i> | 30054. m000802. TIGRR0.1 | 510 | 55.981 2 | 4.83 |
| <i>RcHSF15</i> | 30073. m002275. TIGRR0.1 | 494 | 55.326 3 | 5.29 |
| <i>RcHSF16</i> | 30073. m002276. TIGRR0.1 | 84 | 9.605 8 | 5.59 |
| <i>RcHSF17</i> | 30147. m014282. TIGRR0.1 | 323 | 35.794 9 | 5.25 |
| <i>RcHSF18</i> | 30170. m013859. TIGRR0.1 | 362 | 40.499 6 | 7.72 |
| <i>RcHSF19</i> | 30190. m011156. TIGRR0.1 | 337 | 37.054 1 | 5.31 |

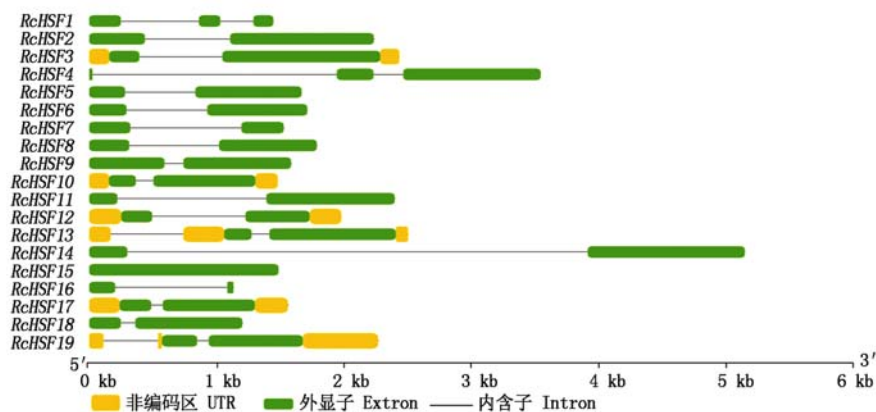


图 1 蓖麻 RcHSF 基因结构分布

Fig. 1 Structure of RcHSF gene in *Ricinus communis* L.

2.2 RcHSF 蛋白保守基序及结构域鉴定

利用 MEME 进一步分析蓖麻 HSF 蛋白家族成员的保守基序, 结果显示, 所有 RcHSFs 的基序类型与排布方式基本相似 (图 2-A), 但无一例外都存在

不同程度的 motif 缺失, 其中 RcHSF16 仅含有 2 个保守基序, 缺失最为严重。DNA 结合域 (DBD) 与寡聚化结构域 (OD) 是保守度最高的结构域, 确保了热激转录元件 HES 的精准定位, 尽管本研究中

RcHSF15 没有 DBD 基序,但 Domains 分析显示该基因存在 HSF 活性,推测可以发挥完整的基因功能;核定位信号 NLS 是由 Lys、Pro、Arg 等组成的微小片段,参与并调控转录因子向核内运输的过程,但由于不同物种间 NLS 序列类型多样,导致转录因子的核运输方式同样存在差异^[21],本研究中 RcHSF2、RcHSF16、RcHSF17、RcHSF19 蛋白没有检测到 NLS 信号区,故推测热激转录因子可能存在其他 NLS 非依赖性核运输方式。

Batch CD-search 保守结构预测表明,所有 RcHSF 中均存在有关热激转录因子的活性区(图 2-

B)。由于 RcHSF15 缺少 DBD 基序,导致其结构域比对率相对较低,该现象同样存在于龙眼 HSF 家族成员鉴定过程中^[22],但在番茄与拟南芥中没有相似案例^[11]。

通过氨基酸序列的同源比对,将所有 RcHSFs 聚为 I、II、III 3 个组,其中 RcHSF2、RcHSF15 与其他氨基酸序列的同源性相对较低,不具有代表性,不作分组处理。又根据基序类型与数量又将 I 组成员细化为 I-1、I-2、I-3 3 个亚组。推测在同组的 RcHSFs 同源性较高,且存在功能冗余。

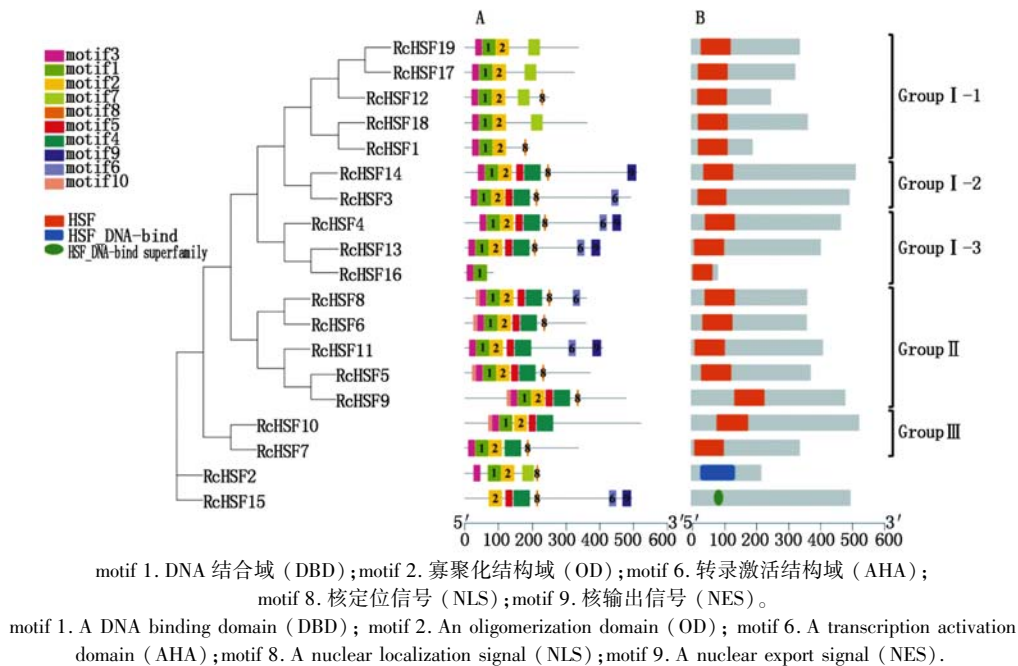


图 2 蓖麻 RcHSF 蛋白的保守基序 (A) 与结构域 (B) 分布

Fig. 2 Distribution of conserved motif (A) and conserved domain (B) of castor RcHSF proteins

2.3 RcHSF 基因家族系统发育关系

通过比对不同植物的 HSF 基因,可以更准确分析基因家族成员间的进化关系,也为蛋白的精准分类提供依据。本研究对蓖麻、拟南芥与番茄的 HSF 氨基酸序列进行比对,并利用 MEGA 工具构建系统进化树(图 3)。根据序列的同源性将所有 HSF 蛋白聚为 I、II、III 3 个主分支与多个亚分支,除了 RcHSF4 与 RcHSF16 为旁系同源基因,其他蓖麻 HSF 家族成员均与拟南芥 AtHSF 蛋白聚合为末端分支,故推测蓖麻与拟南芥 HSF 蛋白的同源性高于番茄 HSF 蛋白。

参考拟南芥 HSF 聚类方式,将所有 HSF 聚为 A、B、C 3 个组,其中只有 A 组蛋白含有 C 端启动基序(AHA),B 组蛋白的寡聚化基序(HR-A/B)比 C 组蛋白少 7 个碱基^[23]。组内成员数按照 A > B > C 进行排序。根据基序的类型与数量,又将 A 组蛋白

分为 3 个亚组 (A1、A2、A3)。推测在同组中,异源 HSF 可能具有相似功能。

2.4 RcHSF 基因家族成员染色体定位信息

根据基因家族成员的染色体定位信息,使用 TBtools 工具绘制基因染色体分布图,由于蓖麻基因组拼接与染色体组装不完整,只能将 RcHSFs 基因定位在未完整拼接的染色体片段上。最终分析结果(图 4)显示,所有 RcHSF 基因被定位在 17 个染色体片段上,呈不均匀分布。Gu 等^[24]将 2 个同源性均大于 75% 且比对长度大于较长序列 80% 的基因定义为串联重复基因。本研究中,由于 RcHSF9 与 RcHSF10 基因没有相连,RcHSF15 与 RcHSF16 基因的同源性较低,且所有染色体片段之间没有表现出微共线性,充分证明在基因组复制过程中 RcHSF 基因家族成员并未发生串联重复与片段重复事件。

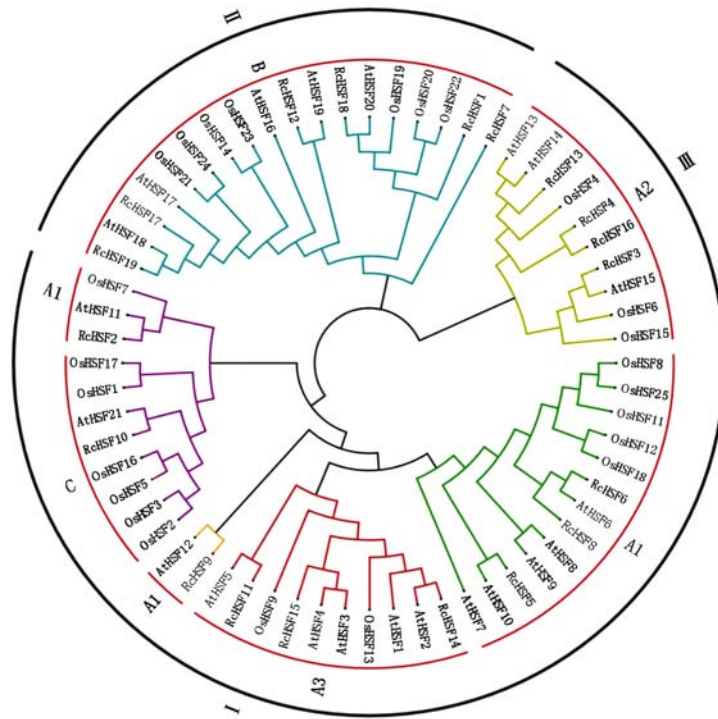


图 3 蓖麻、拟南芥与番茄 HSF 的系统进化关系

Fig. 3 Phylogenetic relationship between castor, *Arabidopsis* and tomato HSF

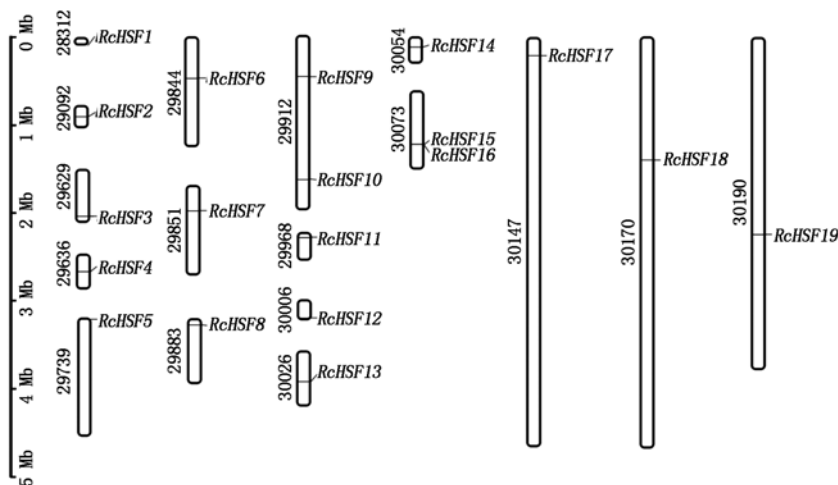


图 4 蓖麻 RcHSF 基因的染色体分布

Fig. 4 Chromosome distribution map of RcHSF gene in *Ricinus communis* L.

2.5 启动子顺式作用元件的预测分析

为挖掘蓖麻 RcHSFs 基因的潜在功能,利用 Plant CARE 工具在线预测处于基因起始位点上游 2 000 bp 的多种顺势作用元件,基于顺势作用元件的多样性,推测不同基因可能具有多种潜在功能。预测结果(表 3)显示,该基因家族含有多种顺势作用元件,其中含有 AAGAA-motif、MYB 与 MYC 元件的基因数量最多,证明 RcHSFs 可能响应植物激素与逆境胁迫,如盐、干旱、低温胁迫与 ABA 诱导等。但基因表达受多种元件共同调控,复杂的调控类型

与调控方式使基因功能多元化,基因功能的专一性仍需进一步验证。

2.6 RcHSF 基因在冷胁迫下的表达模式分析

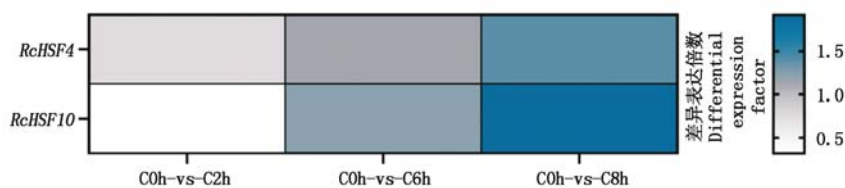
本研究采用 DEseq2 进行差异表达基因的筛选 ($\log_2(\text{FoldChange}) > 1$; $q \text{ value} < 0.001$),发现蓖麻中仅有 *RcHSF4* 与 *RcHSF10* 能够响应低温诱导(图 5)。其中,*RcHSF4* 在冷处理 2,6 h 时与对照组差异表达不显著,但冷处理 8 h 的表达量显著上调 1.51 倍;*RcHSF10* 在冷处理 2 h 时与对照组差异表达不显著,但在冷处理 6,8 h 后差异表达水平均显著上

调,上调倍数分别为 1.05,1.61 倍。该结果表明, 要作用。
RcHSF4 与 RcHSF10 可能在蓖麻抗冷方面发挥着重

表 3 蓖麻 RcHSF 基因家族成员启动子顺式作用元件分析

Tab.3 The cis acting element analysis of promoter region of castor RcHSF gene family members

| 元件名称 Element name | 基序序列 Motif sequence | 基序作用 The function of the motif | RcHSF 基因 RcHSF genes |
|----------------------|---|-----------------------------------|---|
| AAGAA-motif | GGTAAAGAAA GAAAGAA | 逆境响应元件 | 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 16, 17, 18, 19 |
| ABRE | ACGTG TACGGTC GACACGTGGC GCCGCGTGGC CACGTG GCAACGTGTC AACCCGG | 脱落酸响应元件 | 2, 3, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19 |
| ARE | AAACCA | 厌氧感应元件 | 1, 2, 3, 5, 6, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19 |
| CGTCA-motif | CGTCA | 茉莉酸甲酯响应元件 | 2, 5, 7, 8, 9, 13, 14, 15, 17, 18, 19 |
| ERE | ATTTTAAA ATTTCATA | 乙烯响应元件 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19 |
| GARE-motif | TCTGTTG | 赤霉素响应元件 | 2, 3, 15 |
| G-box | CCACGTAA TACGTG GCCACGTGGA TGACACGTGGCA CACGTG(T) CACGA(T)C TAACACGTAG ACACGTGGC | 盐胁迫, 脱落酸, 伤害 响应元件 | 2, 3, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19 |
| LTR | CCGAAA | 低温响应元件 | 2, 5, 9, 10, 13, 14, 17 |
| MBS | CAACTG | 干旱诱导下的 MYB 转 录因子结合位点 | 2, 3, 6, 7, 9, 13, 14, 16, 18 |
| MYB | CAACA(T)G C(T)AACCA T(C)AACTG | 干旱、高盐、低温响应 元件 | 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 |
| MYC | CAATTG CATG(T)TG TCTCTTA | 抗脱落酸、抗旱、抗冻、 抗寒元件 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 19 |
| TCA-element | CCATCTTTTT TCAGAAGAGG | 水杨酸响应元件 | 2, 5, 7, 9, 11, 13, 14, 16, 18 |
| TGACG-motif | TGACG | 茉莉酸甲酯响应元件 | 2, 5, 7, 8, 9, 13, 14, 15, 17, 18, 19 |
| W-box | TTGACC | 生物胁迫响应元件 | 9, 10, 14, 16, 18 |
| WUN-motif | AAATTTCTT AAATTACTA AAATTACT CAATTACAT TTATTACAT | 伤害相关响应元件 | 1, 2, 4, 5, 7, 9, 11, 12, 14, 16, 17 |
| TC-rich repeats | ATTCTCTAAC GTTTTCTTAC | 防御和应激反应元件 | 3, 4, 5, 6, 7, 8, 9, 12, 14, 18 |
| TGA-element | AACGAC | 生长素响应元件 | 9, 10, 11, 13, 17 |



COh-vs-C2h. 冷处理 0, 2 h 的差异表达倍数; COh-vs-C6h. 冷处理 0, 6 h 的差异表达倍数; COh-vs-C8h. 冷处理 0, 8 h 的差异表达倍数。
COh-vs-C2h. The differential expression multiple of cold treatment 0 h and 2 h; COh-vs-C6h. The differential expression multiple of cold treatment 0 h and 6 h; COh-vs-C8h. The differential expression multiple of cold treatment 0 h and 8 h.

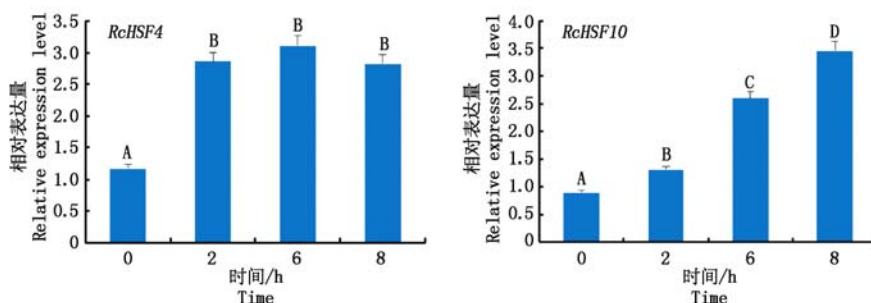
图 5 冷胁迫下 RcHSF 差异表达基因的转录水平

Fig. 5 Transcription level of differentially expressed genes of RcHSF under cold stress

2.7 RcHSF 基因受冷胁迫的表达水平验证

为了验证转录组测序结果的可重复性,本研究对蓖麻 *RcHSF4* 与 *RcHSF10* 进行 qRT-PCR 分析(图 6)。结果显示,这 2 个基因在不同持续时间冷处理

下的表达水平与 RNA-seq 的数据基本一致,表明测序结果可靠。该结果进一步证明, *RcHSF4* 与 *RcHSF10* 可能对于调节蓖麻冷耐受性具有重要意义,其作用途径有待进一步探究。



大写字母 A、B、C、D 代表基因受不同时间冷处理差异表达极显著 ($P < 0.01$)。

The capital letters A, B, C and D represent that the differential expression of genes is extremely significant after cold treatment at different times ($P < 0.01$).

图 6 蓖麻 *RcHSF4* 与 *RcHSF10* 在冷胁迫下的表达水平

Fig. 6 Expression levels of *RcHSF4* and *RcHSF10* of castor under cold stress

3 讨论与结论

蓖麻全基因组测序的实现为 RcHSF 基因家族的鉴定提供了便利,本研究在蓖麻全基因组范围内挖掘 RcHSF 基因家族成员,最终成功鉴定到 19 个热激转录因子,其数量约占所有编码基因(31 221)的 0.000 6%。与拟南芥(21)、番茄(25)、玉米(30)等植物相比,蓖麻 HSF 基因数量相对较少,这可能与蓖麻基因组注释不完全或染色体组装不完整有关。利用同源比对与多物种比对,将 RcHSF 基因进一步聚类为 A、B、C 3 个组,其中 A 组基因数量最多,C 组数量最少,该聚类方式与拟南芥 HSF 基因家族成员的聚类方式相同^[11],推测基序类型与排布顺序相似的同源基因,可能存在功能冗余,异源基因可能具有相同功能。雷娟利等^[25]发现当拟南芥受到热激胁迫时,A 类 HSF 的表达量上升最为明显,过表达 *AtHSFA2* 基因能够显著提高植物的基础耐热性;B 类 HSF 也能够识别并结合热启动元件 HSE,但受到羧基末端特征结构域(-LFGV-)的阻遏作用,导致既无法正常启动热响应基因的表达^[21],又无法调节植物的热耐受性。因此,不同类

型蓖麻 RcHSF 的蛋白功能与作用方式是否存在差异有待进一步考证。

进化分析表明,大部分蓖麻 HSF 基因与拟南芥 HSF 基因被聚为同一分支末端,属于直系同源基因,而在蓖麻基因组中仅存在 *RcHSF4* 与 *RcHSF16* 一对旁系同源基因,表明在植物漫长的进化过程中,大部分 HSF 基因均按照物种特异性的方式进行扩展,该现象在植物其他基因家族的研究中也得到了广泛验证^[26-28]。

大多数 RcHSFs 启动子区均含有 AAGAA-motif、MYB 与 MYC 元件,暗示其响应植物激素与逆境胁迫。目前,植物 HSF 基因的功能研究主要集中在干旱与盐胁迫,而在调节植物冷适应性方面鲜有报道。郭丽红等^[29]研究发现,低温胁迫下拟南芥 *AtHS-FAIA* 可以通过负调控 Caspase-3 蛋白酶活性,进而抑制细胞程序性死亡。作物抗冷性是由多基因控制的复杂性状,因此,耐冷数量性状基因座的定位是抗冷基因克隆的前提。本研究结合蓖麻基因组与转录组数据对 RcHSF 进行分析,发现冷胁迫下仅 *RcHSF4* 与 *RcHSF10* 差异表达显著;其中 *RcHSF4* 在冷处理早期差异表达倍数没有明显变化,而

RcHSF10 的差异表达倍数随冷处理时间的延长逐渐提高。该表达模式与茶树 *CsHSFA4*、*CsHSFA5* 的表达模式相似^[30], 推测 *RcHSF4* 与 *RcHSF10* 的上调表达对于蓖麻适应低温胁迫、触发冷响应机制具有重要意义。本研究为探索蓖麻 HSF 基因家族成员的冷应答机制提供重要数据, 也为蓖麻抗冷新种质资源的建立提供候选基因。

参考文献:

- [1] 翁锦周, 洪月云. 植物热激转录因子在非生物逆境中的作用[J]. 分子植物育种, 2006, 4(1): 88 - 94. doi:10.3969/j.issn.1672-416X.2006.01.016.
Weng J Z, Hong Y Y. The roles of plant heat shock transcription factors in abiotic stress [J]. *Molecular Plant Breeding*, 2006, 4(1): 88 - 94.
- [2] Ohama N, Sato H, Shinozaki K, Yamaguchi-Shinozaki K. Transcriptional regulatory network of plant heat stress response [J]. *Trends in Plant Science*, 2017, 22(1): 53 - 65. doi:10.1016/j.tplants.2016.08.015.
- [3] Guo M, Liu J H, Ma X, Luo D X, Gong Z H, Lu M H. The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses [J]. *Frontiers in Plant Science*, 2016, 7: 114. doi:10.3389/fpls.2016.00114.
- [4] Bharti K, Koskull-Döring P V, Bharti S, Kumar P, Tintschl-Körbitzer A, Treuter E, Nover L. Tomato heat stress transcription factor HsfB1 represents a novel type of general transcription coactivator with a histone-like motif interacting with the plant CREB binding protein ortholog HAC1 [J]. *The Plant Cell*, 2004, 16(6): 1521 - 1535. doi:10.1105/tpc.019927.
- [5] Scharf K D, Rose S, Zott W, Schöff F, Nover L, Schöff F. Three tomato genes code for heat stress transcription factors with a region of remarkable homology to the DNA-binding domain of the yeast HSF [J]. *The EMBO Journal*, 1990, 9(13): 4495 - 4501. doi:10.1002/j.1460-2075.1990.tb07900.x.
- [6] Li P S, Yu T F, He G H, Chen M, Zhou Y B, Chai S C, Xu Z S, Ma Y Z. Genome-wide analysis of the Hsf family in soybean and functional identification of *GmHsf-34* involvement in drought and heat stresses [J]. *BMC Genomics*, 2014, 15(1): 1009. doi:10.1186/1471-2164-15-1009.
- [7] Zhu X Y, Huang C Q, Zhang L, Liu H F, Yu J H, Hu Z Y, Hua W. Systematic analysis of Hsf family genes in the *Brassica napus* genome reveals novel responses to heat, drought and high CO₂ stresses [J]. *Frontiers in Plant Science*, 2017, 8: 1174. doi:10.3389/fpls.2017.01174.
- [8] Scharf K D, Berberich T, Ebersberger I, Nover L. The plant heat stress transcription factor (Hsf) family: structure, function and evolution [J]. *Biochimica et Biophysica Acta*, 2012, 1819(2): 104 - 119. doi:10.1016/j.bbagr.2011.10.002.
- [9] Chung E, Kim K M, Lee J H. Genome-wide analysis and molecular characterization of heat shock transcription factor family in *Glycine max* [J]. *Journal of Genetics and Genomics*, 2013, 40(3): 127 - 135. doi:10.1016/j.jgg.2012.12.002.
- [10] Lin Y X, Jiang H Y, Chu Z X, Tang X L, Zhu S W, Cheng B J. Genome-wide identification, classification and analysis of heat shock transcription factor family in maize [J]. *BMC Genomics*, 2011, 12(1): 76 - 89. doi:10.1186/1471-2164-12-76.
- [11] Guo J K, Wu J, Ji Q, Wang C, Luo L, Yuan Y, Wang Y H, Wang J. Genome-wide analysis of heat shock transcription factor families in rice and *Arabidopsis* [J]. *Journal of Genetics and Genomics*, 2008, 35(2): 105 - 118. doi:10.1016/s1673-8527(08)60016-8.
- [12] Prasinos C, Krampis K, Samakovli D, Hatzopoulos P. Tight regulation of expression of two *Arabidopsis* cytosolic *Hsp90* genes during embryo development [J]. *Journal of Experimental Botany*, 2005, 56(412): 633 - 644. doi:10.1093/jxb/eri035.
- [13] Swindell W R, Huebner M, Weber A P. Transcriptional profiling of *Arabidopsis* heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways [J]. *BMC Genomics*, 2007, 8(1): 125. doi:10.1186/1471-2164-8-125.
- [14] Hu W H, Hu G C, Han B. Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed overlapped and stress specific response under abiotic stresses in rice [J]. *Plant Science*, 2009, 176(4): 583 - 590. doi:10.1016/j.plantsci.2009.01.016.
- [15] Giorno F, Wolters-Arts M, Grillo S, Scharf K D, Vriezen W H, Mariani C. Developmental and heat stress regulated expression of HsfA2 and small heat shock proteins in tomato anthers [J]. *Journal of Experimental Botany*, 2010, 61(2): 453 - 462. doi:10.1093/jxb/erp316.
- [16] Ishka M R, Brown E, Weigand C, Tillett R L, Schlauch K A, Miller G, Harper J F. A comparison of heat-stress transcriptome changes between wild-type *Arabidopsis* pollen and a heat-sensitive mutant harboring a knockout of *cyclic nucleotide-gated cation channel 16 (cngc16)* [J]. *BMC Genomics*, 2018, 19(1): 549. doi:10.1186/s12864-018-4930-4.
- [17] 朱华生. 玉米热激转录因子 ZmHsf04 的克隆及抗逆功能研究 [D]. 合肥: 安徽农业大学, 2016.
Zhu H S. Cloning and stress-resistance study of heat shock transcription factor ZmHsf04 from *Zea mays* [D]. Hefei: Anhui Agricultural University, 2016.
- [18] 余文才, 李言, 吴绍华, 晁金全, 张世鑫, 杨曙光, 田维敏. 巴西橡胶树 *HbHsfA4a* 的克隆及其对低温胁迫的响应 [J]. 热带作物学报, 2019, 40(5): 898 - 904. doi:10.3969/j.issn.1000-2561.2019.05.010.
Yu W C, Li Y, Wu S H, Chao J Q, Zhang S X, Yang S G, Tian W M. Cloning of *HbHsfA4a* and its transcriptional responses to cold stress in *Hevea brasiliensis* [J]. *Chinese Journal of Tropical Crops*, 2019, 40(5): 898 - 904.
- [19] 李金琴, 朱国立, 吴国林, 何智彪, 李靖霞, 张春华, 田福东, 贾娟霞. 蓖麻种子含油量与主要数量性状的相关及通径分析 [J]. 中国油料作物学报, 2004, 26(2): 43 - 46.
Li J Q, Zhu G L, Wu G L, He Z B, Li J X, Zhang C

- H, Tian F D, Jia J X. Correlation and path analysis between seed oil content and major quantitative characters in castor[J]. *Chinese Journal of Oil Crop Sciences*, 2004, 26(2): 43–46.
- [20] 王晓宇, 李敏, 刘栩铭, 卜祥琪, 丁雪, 张继星. 蓖麻 *RcSEC14p* 基因的克隆及低温胁迫下的表达分析[J]. *分子植物育种*, 2019, 17(13): 4204–4209. doi:10.13271/j. mpb. 017. 004204.
- Wang X Y, Li M, Liu X M, Bu X Q, Ding X, Zhang J X. Cloning of *RcSEC14p* in *Ricinus communis* L. and expression analysis under cold stress[J]. *Molecular Plant Breeding*, 2019, 17(13): 4204–4209.
- [21] 李家朋. 玉米热激转录因子 ZmHsf-17 的功能研究[D]. 合肥:安徽农业大学, 2015.
- Li J P. Function analysis of heat shock transcription factor ZmHsf-17 of *Zea mays* [D]. Hefei: Anhui Agricultural University, 2015.
- [22] 王云, 彭丽云, 孙雪丽, 高玉莹, 陈晓慧, 陈裕坤, 林玉玲, 赖钟雄. 龙眼 Hsf 基因家族全基因组鉴定及体胚发生过程中的表达分析[J]. *应用与环境生物学报*, 2019, 25(2): 420–431. doi:10.19675/j. cnki. 1006-687x. 2018. 06004.
- Wang Y, Peng L Y, Sun X L, Gao Y Y, Chen X H, Chen Y K, Lin Y L, Lai Z X. Genome-wide identification of longan Hsf family members and their functional analysis during somatic embryogenesis in longan [J]. *Chinese Journal of Applied and Environmental Biology*, 2019, 25(2): 420–431.
- [23] 陈先知, 王燕, 史建磊, 朱隆静, 王克磊, 徐坚. 黄瓜全基因组热激转录因子(HSFs)的鉴定与表达分析[J]. *遗传*, 2014, 36(4): 376–386. doi:10.3724/SP. J. 1005. 2014. 0376.
- Chen X Z, Wang Y, Shi J L, Zhu L J, Wang K L, Xu J. Genome-wide identification, sequence characteristic and expression analysis of heat shock factors (HSFs) in cucumber[J]. *Hereditas*, 2014, 36(4): 376–386.
- [24] Gu Z L, Cavalcanti A, Chen F C, Bouman P, Li W H. Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast[J]. *Molecular Biology and Evolution*, 2002, 19(3): 256–262. doi:10.1093/oxfordjournals. molbev. a004079.
- [25] 雷娟利, 徐志豪. 转座子在植物功能基因组学中的应用[J]. *浙江农业学报*, 2002, 14(5): 291–296. doi:10.3969/j. issn. 1004-1524. 2002. 05. 011.
- Lei J L, Xu Z H. Application of transposons in the plant functional genomics[J]. *Acta Agriculturae Zhejiangensis*, 2002, 14(5): 291–296.
- [26] Bai J F, Pennill L A, Ning J C, Lee S W, Ramalingam J, Webb C A, Zhao B Y, Sun Q, Nelson J C, Leach J E, Hulbert S H. Diversity in nucleotide binding site-leucine-rich repeat genes in cereals[J]. *Genome Research*, 2002, 12(12): 1871–1884. doi: 10.1101/gr.454902.
- [27] Zhang S B, Chen C, Li L, Meng L, Singh J, Jiang N, Deng X W, He Z H, Lemaux P G. Evolutionary expansion, gene structure, and expression of the rice wall-associated kinase gene family[J]. *Plant Physiology*, 2005, 139(3): 1107–1124. doi:10.1104/pp. 105. 069005.
- [28] Jain M, Tyagi A K, Khurana J P. Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*) [J]. *Genomics*, 2006, 88(3): 360–371. doi: 10.1016/j. ygeno. 2006. 04. 008.
- [29] 郭丽红, 徐娅, 邵秋霞, 李念, 檀文涛, 张学兰. 拟南芥热激因子 AtHsfA1a 在低温胁迫下对细胞程序性死亡中 Caspase-3 活性的影响[J]. *江苏农业科学*, 2017, 45(21): 24–27. doi:10.15889/j. issn. 1002-1302. 2017. 21. 006.
- Guo L H, Xu Y, Xi Q X, Li N, Tan W T, Zhang X L. Effect of *Arabidopsis* heat shock factor AtHsfA1a on Caspase-3 activity in programmed cell death under low temperature stress[J]. *Jiangsu Agricultural Sciences*, 2017, 45(21): 24–27.
- [30] 熊洋洋, 顾雅琦, 刘志薇, 李彤, 吴致君, 庄静. 茶树中两个 HSF 转录因子基因分离与温度胁迫响应的比较分析[J]. *植物生理学报*, 2015, 51(12): 2143–2151. doi:10.13592/j. cnki. ppj. 2015. 0532.
- Xiong Y Y, Gu Y Q, Liu Z W, Li T, Wu Z J, Zhuang J. Isolation and expression profiles analysis of two HSF transcription factor genes under temperature stress in *Camellia sinensis* [J]. *Plant Physiology Journal*, 2015, 51(12): 2143–2151.