doi: 10.3724/CBB-2026-0009

薏苡热激转录因子(*HSF*)基因家族的全基因组分析：进化、表达模式及对发育与非生物胁迫响应的研究

朱紫嫣[1†], 彭玲霞[1†], 李梦妮[2], 杨丽君[1], 李洪菊[1], 赵玉洋[3], 金燕[3], 陆若辉[4], 张晓丹[1*]

[1]浙江理工大学生命科学与医药学院, 浙江省植物次生代谢调控重点实验室, 杭州 310018; [2]浙江乐清市农业农村局, 乐清 325600; [3]中国中医科学院中药资源中心, 道地药材国家重点实验室培育基地, 北京 100700; [4]浙江省耕地质量与肥料管理总站, 乐清 325600

摘要 在影响生物体的众多非生物因子中, 热激转录因子(*HSF*)在干旱、高温、盐碱和低温等胁迫响应中发挥关键作用。薏苡(*Coix lacryma-jobi*)是一种全球性药食同源作物, 在中国、日本和韩国被广泛用于医药和食品领域。然而, 目前尚未见关于薏苡 *HSF* 基因家族(*ClHSF*)的系统鉴定及其在抗旱、耐盐、耐寒和耐热中功能的研究报道。本研究共鉴定出 20 个 *ClHSF* 基因家族成员。基于基因结构、蛋白保守基序及与拟南芥、水稻和玉米的系统进化关系, 这些成员可划分为 HSFA、HSFB 和 HSFC 三个亚家族。染色体定位分析表明, *ClHSF* 基因在薏苡 9 条染色体上呈不均匀分布。进一步通过 qRT-PCR 技术检测发现, 在干旱、高温、盐胁迫及低温等非生物胁迫条件下, 多个 *ClHSFs* 基因的表达水平呈现显著上调趋势, 表明该基因家族可能在薏苡逆境胁迫响应机制中发挥重要调控作用。该研究揭示了 *ClHSF* 基因在适应多种胁迫中的重要作用, 其中 *ClHSF13*、*ClHSF14* 和 *ClHSF17* 等特异性上调或下调基因可作为薏苡抗逆品种选育的潜在候选基因。

关键词 薏苡, *HSF* 基因家族, 非生物胁迫, 全基因组分析

当前, 全球地表平均温度上升主要源于人类活动引起的温室气体排放, 气候变化显著影响生物多样性, 并对其构成持续增强的威胁(Cahill et al., 2012; Friedrich et al., 2016; Jaureguiberry et al., 2022)。研究表明, 局域性物种灭绝事件已有多项例记录, 其中部分与气候变化存在直接关联(Cahill et al., 2012; Román et al., 2020)。热激蛋白(HSPs)作为响应高温胁迫的核心蛋白, 不仅参与环境应激反应, 还调控生长发育过程(Waters et al., 2013)。热激转录因子(*HSFs*)是 HSPs 基因表达的关键调控因子, 在正常生理和胁迫条件下均参与维持蛋白质稳态(Parker et al., 1984; Wu et al., 1984)。HSF 蛋白具有高度保守性, 包含 5 个特征结构域: DNA 结合域(DBD)、寡聚化域(OD)、核定位信号(NLS)、核输出信号(NES)和 C 端激活域(CTD) (Guo et al., 2016; Guo et al., 2013)。

研究人员最早从酵母和番茄中成功克隆获得 *HSF* 基因(Mishra et al., 2002)。目前已在

收稿日期: 2026-01-09; 接受日期 2026-06-01
基金项目: 国家自然科学基金区域创新发展联合基金重点项目(No.U25A20163)

† 共同第一作者

* 通讯作者。E-mail: zhangxiaodanzstu@163.com

© 植物学报 Chinese Bulletin of Botany

拟南芥(Wang et al., 2012)、马铃薯(Tang et al., 2016)、大豆(Li et al., 2014)、高粱(Nagaraju et al., 2015)、鹰嘴豆(Chidambaranathan et al., 2018)、猕猴桃(Ling et al., 2023)、紫花苜蓿(Ma et al., 2022)和黑麦(Ren et al., 2023)等多种植物中鉴定出 *HSF* 家族成员。研究表明, *HSF* 转录因子在植物响应热胁迫、氧化损伤及盐胁迫等非生物胁迫过程中发挥重要调控作用。在拟南芥中, *AtHSFA3* 的过量表达虽能提高耐热性, 但会降低盐胁迫耐受性(Zhu et al., 2016); *AtHSFA2* 可增强低氧耐受性并诱导特定基因表达(Baniwal et al., 2004); *AtHSFA6b* 在高盐和干旱胁迫下表达显著上调, 可能参与 ABA 信号途径(Zhang et al., 2020)。此外, *HSFs* 还可通过调控其他胁迫相关基因影响植物适应性: *ZmHSF11* 通过抑制氧化应激基因降低耐热性(Duan et al., 2019); *AtHSFA2* 通过激活 *APX2* 和 *HSP* 等基因增强拟南芥耐热性(Guotian et al., 2018); *SlHSFA3* 通过与 MAP 激酶互作参与热胁迫响应(Guo et al., 2015)。最新研究显示, *HSFs* 还可通过调控谷胱甘肽代谢(Huang et al., 2021)、ABA 信号(Chen et al., 2021)和次生代谢(Jeffares et al., 2008)等途径参与植物耐热性调控, 并在体细胞胚胎发育过程中发挥作用。

薏苡(*Coix lacryma-jobi*)作为东亚和东南亚地区重要的栽培作物(Arora et al., 1977), 其种子富含优质蛋白, 营养价值优于传统谷物(Zhu et al., 2016)。我国薏苡栽培历史悠久, 兼具粮食生产和酿酒用途(Yang et al., 2010; Wang et al., 2016)。现代医学研究表明, 薏苡仁提取物康莱特注射液在肿瘤临床治疗中具有重要应用价值(Zhang et al., 2019)。然而, 干旱和极端温度等非生物胁迫显著影响薏苡的生长发育及产量形成(Suzuki et al., 2014; Umezawa et al., 2006)。目前, 薏苡 *HSF* 基因的功能研究仍未见报道。

基于最新发布的薏苡基因组数据, 本研究对 *HSF* 基因家族进行了系统分析。共鉴定出 20 个薏苡 *HSF* 基因(*ClHSFs*), 通过系统进化分析对其进行分类, 并深入解析了基因结构特征、保守基序分布、基因复制事件与共线性关系、顺式调控元件组成及表达模式。研究结果为揭示 *ClHSFs* 的生物学功能提供了理论依据, 同时为深入理解薏苡响应非生物胁迫的分子机制提供了新的研究方向。

1 材料与方法

1.1 植物材料与胁迫处理

将薏苡(*Coix lacryma-jobi* L.)栽培于人工气候室(温度 25–28°C, 相对湿度 70–80%)。待幼苗生长至两周龄时, 分别施加以下胁迫处理: (1) 高温胁迫(40°C); (2) 低温胁迫(4°C); (3) 干旱胁迫(10%和 15% PEG6000, w/v); (4) 盐胁迫(100 mmol·L^{-1} 和 150 mmol·L^{-1} NaCl)。每种胁迫处理 12 小时后采集叶片样本, 立即液氮速冻后保存于–80°C 超低温冰箱。选择 12 h 作为取样时间点, 是基于前期预实验结果及相关文献报道, 该时间点能有效检测到 HSF 基因的早期表达响应(Duan et al., 2019)。每个处理设置 6 盆生物学重复, 每盆取样 3 次作为技术重复。

1.2 薏苡 *ClHSF* 基因的鉴定与染色体定位

基于国家基因组科学数据中心(BIG Data Center, https://ngdc.cncb.ac.cn/, 登录号: GWHAAYR00000000)提供的薏苡(*Coix lacryma-jobi*)全基因组数据(Hua et al., 2023), 我们构建了非冗余蛋白质数据库。为鉴定薏苡中潜在的 *HSF* 编码基因, 从 Pfam 数据库(https://pfam.xfam.org/, 2024 年 4 月 10 日检索)获取 *HSF* 特征性 DNA 结合域(DBD, PF00447)。采用隐马尔可夫模型(HMM)对薏苡全蛋白序列进行扫描, 设定 E 值阈值<10^{-5}筛选候选基因。为保证 *HSF* 家族成员的全面覆盖, 从 Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html)下载拟南芥 *HSF* 蛋白序列作为查询序列, 对薏苡参考基因组进行 BLAST 比对(E 值阈值 1×10^{-10})。通过 NCBI 保守结构域数据库(https://www.ncbi.nlm.nih.gov/cdd/, 2024 年 4 月 10 日访问)进一步验证候选基因的 *HSF* 结构域(E 值阈值<1×10^{-5})。使用 ExPASy 在线工具(https://web.expasy.org/protparam/)预测蛋白分子量与等电点, Cell-PLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/)进行亚细胞定位预测。最终基于 TBtools 软件将鉴定的 *ClHSF* 成员定位至参考基因组, 并按其染色体位置进行系统命名。为确保基因结构注释的准确性, 我们利用本研究中获得的 RNA-seq 数据, 使用 IGV 软件(Robinson et al., 2011)对所有初步鉴定的 *ClHSF* 基因的编码序列进行了人工检查与校正。校正后的基因模型用于后续所有分析。

1.3 系统进化树构建、基因结构与保守基序分析

首先对拟南芥 HSF 蛋白(AtHSF)与薏苡 ClHSF 蛋白序列进行多序列比对(MSA), 采用 MEGA 4.0 软件内置的 Muscle 程序识别比对结果中的特异性结构域(Tamura et al., 2021)。基于比对结果, 使用 MEGA 4.0 构建邻接法(Neighbor-Joining, NJ)系统进化树, 参数设置为: 成对缺失(pairwise deletion)、泊松模型(Poisson model), 并通过 1000 次重复计算获得 bootstrap 支持值(Kumar et al., 2018)。为进一步解析 *ClHSF* 的进化关系, 从 Ensembl 植物数据库(https://plants.ensembl.org/index.html, 访问日期: 2024 年 4 月 10 日)下载拟南芥(*A. thaliana*)、玉米(*Z. mays*)和水稻(*O. sativa*)的 HSF 蛋白序列, 构建多物种系统进化树。通过 NCBI 保守结构域数据库(NCBI-CDD, https://www.ncbi.nlm.nih.gov/cdd)鉴定 *ClHSFs* 的保守结构域, 并利用 MEME 在线工具(https://meme-suite.org/meme/tools/meme, 访问日期: 2024 年 4 月 10 日)预测 *ClHSFs* 所含的保守基序。最终采用 TBtools 软件(Larkin et al., 2007)对 *ClHSF* 基因的内含子-外显子结构、CDD 结构域及保守基序进行可视化呈现。

1.4 *ClHSF* 基因的复制与共线性分析

基于薏苡基因组数据, 采用 TBtools 软件对已定位到染色体的 *HSF* 基因进行共线性分析。从 Ensembl Plants 数据库(https://plants.ensembl.org/index.html)下载 7 个物种拟南芥(*Arabidopsis thaliana*)、二穗短柄草(*Brachypodium distachyon*)、大麦(*Hordeum vulgare*)、水稻(*Oryza sativa*)、高粱(*Sorghum bicolor*)、普通小麦(*Triticum aestivum*)和玉米(*Zea mays*)

的全基因组序列及基因注释文件。通过 MCScanX 算法，系统分析薏苡 *HSF* 家族成员与这些物种之间的共线性关系，以揭示薏苡与选定物种间 *HSF* 同源基因的进化关联。

1.5 *ClHSF* 基因启动子顺式作用元件分析

利用 PlantCARE 数据库 (Lescot et al., 2002) (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/，访问日期: 2024 年 4 月 10 日)对 *ClHSF* 基因启动子区(转录起始位点上游 2000 bp 序列)进行顺式调控元件预测。分析元件类型包括激素响应元件(如 ABRE、G-box)、胁迫响应元件(如 DRE、TC-rich repeats)以及光响应元件等。通过 TBtools 软件(Chen et al., 2020)对预测结果进行可视化展示，并统计各类元件的分布频率与位置特征。

1.6 RNA 提取与 qRT-PCR 验证

采用 TIANGEN 植物总 RNA 提取试剂盒提取样品 RNA，使用 PrimeScript™ RT Master Mix (TaKaRa)进行 cDNA 反转录。qRT-PCR 反应使用 TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMix 试剂盒，按照制造商推荐程序设置反应参数，每个实验组设置 3 个技术重复。以 UBQ5 基因为内参(Schmittgen et al., 2008)，采用 $2^{-\Delta\Delta Ct}$ 方法计算基因相对表达量。每个基因的 qRT-PCR 检测至少包含 3 个生物学重复的 cDNA 样本。基于 *ClHSFs* 基因序列，通过 Primer3Plus 软件设计特异性引物(附表 1)。在 20 个鉴定的 *ClHSF* 基因中，我们优先选择了 14 个具有完整 DBD 结构域、且在 RNA-seq 分析中表现出对至少一种非生物胁迫有显著响应(>2 倍变化)的基因进行 qRT-PCR 验证，以确保所选基因能代表该家族的主要胁迫响应模式。

表 1 通过 Primer3Plus 软件设计特异性引物

Table 1 Design specific primers using the Primer3Plus software

Gene	Gene ID	Forward primer (5'–3')	Reverse primer (5'– 3')
ClHSF01	Cl019516_T2	GTTGAAACAGGAACACCGCC	TTCTCCCTGTGCAGCTTCAG
ClHSF03	Cl020540_T1	CAGCGGGTTCTATAGCACCG	TCAGTAGCCGCTGTTGACG
ClHSF04	Cl000935_T1	GCCTGTCAAAACTGCACCTG	ATCACACGGAGGCGTTGAAT
ClHSF05	Cl003597_T1	CCAGCAGCAAGACAAGAGGA	AGCTCAGGCATTCCAGGTTC
ClHSF08	Cl027011_T1	CAGCCTACTCAACGCGATGA	AATCCCACCACCACTTCGTC
ClHSF09	Cl027124_T1	TTGTTTCAGGGTTCCGCAAG	TGCTGCGTCCTTCCTCTTG
ClHSF11	Cl008136_T1	AGCAGTAGCAAGGAACCGTC	TGGAATTGGACCCGTCTTGG
ClHSF12	Cl010934_T1	AGCAAGGTTCACCGATCCTG	CAAATCCTCCTGGGGCATGT
ClHSF13	Cl029438_T1	CCTCACAAGACGAGACCACC	CCTTCACATCCCCCTTTGCT
ClHSF14	Cl013647_T1	GCTATGCTCGCCAAGAAGGA	ACACTTGGCCGAGAACAGAG
ClHSF15	Cl014409_T1	ATCAGGACGAGGAGGAGGAC	CTGGCGAGGAAGGACATGAG

ClHSF17	Cl032883_T2	GTCAAGCAAGGACCAGCTCT	GCACGATCTTCCTGAACCCA
ClHSF19	Cl023145_T1	CGGACGAGACAATCTCCTGG	GGTGCTCACGTAGGTGTTGA
ClHSF20	Cl041956_T1	GCAGGAGACAACAGGCATCT	AAAGGCCAACGTCCAGAACA
UBQ5		TGCTGCAGTTCTACAAGGTG	GAAGTGGTTGGCCATGAAG

2 结果与分析

2.1 *ClHSF* 基因家族的全基因组鉴定

采用隐马尔可夫模型(HMM)对薏苡基因组进行 HSF 蛋白结构域(PF00447)扫描, 共鉴定出 20 个 *HSF* 基因。基于染色体定位信息, 将其系统命名为 *ClHSF01*–*ClHSF20* (表 1)。分析结果显示, 各 *ClHSF* 基因编码蛋白的理化性质存在明显差异: 氨基酸长度在 183 (*ClHSF14*)–453 aa (*ClHSF18*)之间; 分子量范围为 20.039 (*ClHSF14*)–48.030 kDa (*ClHSF18*); 等电点(pI)分布在 4.67 (*ClHSF11*)–10.15 (*ClHSF20*)之间。亚细胞定位预测表明, 所有 ClHSF 蛋白均定位于细胞核中。

表 2 薏苡中鉴定出的 20 个 *ClHSF* 基因信息

Table 2 Information of the 20 *ClHSF* genes identified in *Coix lacryma-jobi*

Gene name	Protein ID	Sequence ID	Chromosome	strand	start	End	Protein length(aa)	Mw (KDa)	pI	Predicted subcellular localization
ClHSF01	GWHPAAYR004191	Cl019516_T2	Chr1	–	54748837	54751675	395	43.80	4.92	Nuclear
ClHSF02	GWHPAAYR004557	Cl019723_T1	Chr1	–	74115291	74118095	416	46.96	5.52	Nuclear
ClHSF03	GWHPAAYR005972	Cl020540_T1	Chr1	–	16279262	16279532	217	23.06	6.33	Nuclear
ClHSF04	GWHPAAYR013980	Cl000935_T1	Chr2	+	13046453	13049036	297	33.93	4.79	Nuclear
ClHSF05	GWHPAAYR018266	Cl003597_T1	Chr2	+	12574020	12574369	281	32.136	4.90	Nuclear
ClHSF06	GWHPAAYR020662	Cl005150_T2	Chr2	–	16588902	16589497	236	25.478	9.54	Nuclear
ClHSF07	GWHPAAYR021040	Cl005368_T1	Chr2	–	17037182	17037921	259	28.634	5.15	Nuclear
ClHSF08	GWHPAAYR025252	Cl027011_T1	Chr3	–	53892509	53893782	279	30.388	4.84	Nuclear
ClHSF09	GWHPAAYR025428	Cl027124_T1	Chr3	+	62295748	62296771	214	22.179	5.30	Nuclear
ClHSF10	GWHPAAYR029476	Cl006812_T1	Chr4	+	62033536	62058156	230	25.376	8.05	Nuclear
ClHSF11	GWHPAAYR031686	Cl008136_T1	Chr4	+	27218284	27219608	318	33.226	4.67	Nuclear
ClHSF12	GWHPAAYR036612	Cl010934_T1	Chr4	–	15648440	15648799	261	29.957	4.74	Nuclear
ClHSF13	GWHPAAYR038668	Cl029438_T1	Chr5	–	30873204	30885514	386	42.661	4.92	Nuclear
ClHSF14	GWHPAAYR046011	Cl013647_T1	Chr6	–	27728667	27729540	183	20.039	9.66	Nuclear
ClHSF15	GWHPAAYR047254	Cl014409_T1	Chr6	–	42324139	42324996	285	30.739	6.06	Nuclear
ClHSF16	GWHPAAYR047809	Cl014764_T1	Chr6	–	51565154	51567802	292	32.348	5.43	Nuclear
ClHSF17	GWHPAAYR052209	Cl032883_T2	Chr7	–	11501553	11505114	378	39.514	8.29	Nuclear
ClHSF18	GWHPAAYR054538	Cl034430_T1	Chr7	–	12535124	12535287	453	48.030	8.43	Nuclear
ClHSF19	GWHPAAYR063030	Cl023145_T1	Chr9	–	18525430	18526422	330	35.581	6.29	Nuclear
ClHSF20	GWHPAAYR067071	Cl041956_T1	contig104	–	311039	314014	435	47.562	10.15	Nuclear

2.2 *ClHSF* 基因的染色体分布

染色体定位分析表明，本研究鉴定的 20 个 *ClHSF* 基因中，19 个不均匀地分布在薏苡 8 条染色体上(*ClHSF20* 未定位到特定染色体)。基于国家基因组科学中心(BIG Data Center)的基因组数据，我们确定了各基因的精确染色体位置(图 1)，并据此将其系统命名为 *ClHSF01–ClHSF20*。

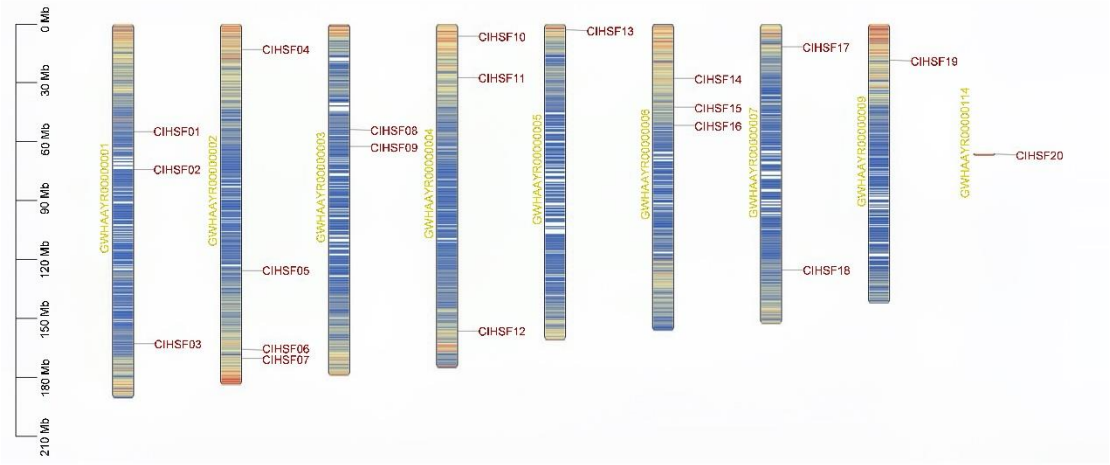


图 1　薏苡 *ClHSF* 基因成员的染色体分布图。左侧刻度尺标示染色体长度(单位: Mb)。染色体区域中，红色代表高基因密度区，蓝色代表低基因密度区。各 *ClHSF* 基因根据其系统进化分组采用不同颜色标记: HSFA 亚家族(红色)、HSFB 亚家族(蓝色)、HSFC 亚家族(绿色)。

Figure 1　Chromosomal distributions of *ClHSF* members in coix. The scale on the left represents the length of the chromosomes. In this chromosomal region, red denotes a higher gene density and blue, a lower gene density.

2.3 *ClHSF* 基因的系统进化分析与分类

为了研究 *ClHSF* 基因的分类情况，我们利用拟南芥、水稻、玉米和薏苡的 87 个 *HSF* 基因构建了系统进化树。HSF 蛋白的进化树分析显示可分为三类: HSFA、HSFB 和 HSFC (图 2)。其中 HSFA 类包含的 *ClHSF* 成员最多(59 个)，其次是 HSFB 类(24 个)，HSFC 类成员最少(4 个)。通过对 20 个 *ClHSF* 序列的分析，可将其划分为三个不同的 *HSF* 基因亚家族: HSFA 亚家族包含 *ClHSF1/2/3/4/5/7/8/9/12/13/14/15/16/20*; HSFB 亚家族包含 *ClHSF6/10/11/17/18/19*; HSFC 亚家族未发现薏苡成员。

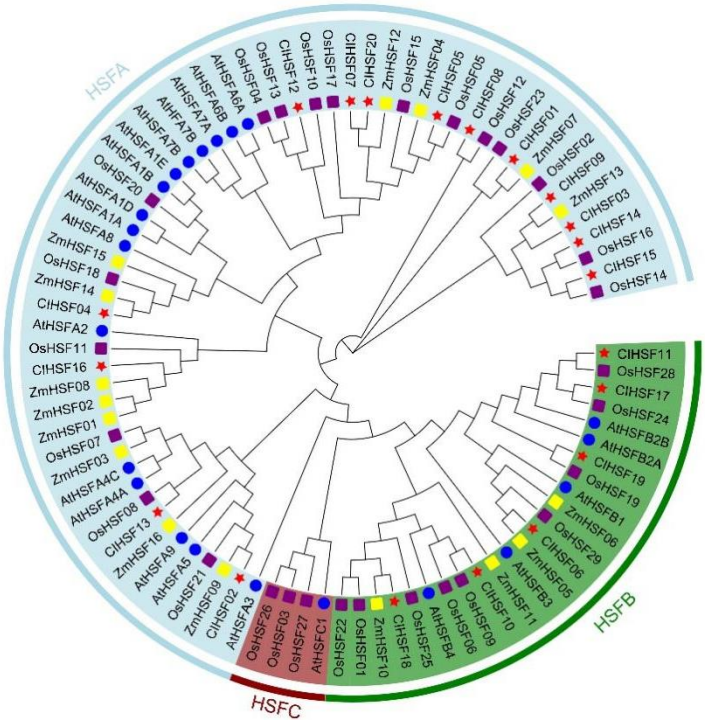


图 2 基于邻接法构建的薏苡、拟南芥、水稻和玉米 HSF 蛋白系统进化树。该进化树使用 MEGA 4.0 软件分析，采用 1000 次 bootstrap 重复检验。不同颜色的叶片背景代表不同的 HSF 蛋白分组：红色代表 HSFA 亚家族，蓝色代表 HSFB 亚家族，绿色代表 HSFC 亚家族。不同形状的节点代表不同物种：红色五角星代表薏苡，蓝色圆形代表拟南芥，紫色方块代表水稻，黄色方块代表玉米。分支上的数字表示 bootstrap 支持值(仅显示>50%的值)。标尺表示每个位点的氨基酸替换率。

Figure 2 Utilizing the neighbor-joining approach with 1000 bootstrap replicates, the Neighbor-Joining phylogenetic tree of Hsf proteins from *Coix*, *Arabidopsis*, *Oryza sativa*, and *Zea mays* was built with MEGA 4.0 software. Various colored leaf backgrounds stood for various groupings: red for the HSFA subfamily, blue for the HSFB subfamily, and green for the HSFC subfamily. Different node shapes represent different species: red five-pointed stars for Coix, blue circles for *Arabidopsis thaliana*, purple squares for Oryza sativa, and yellow squares for *Zea mays*. Numbers on branches indicate bootstrap support values (only values >50% are shown). The scale bar represents the amino acid substitution rate per site.

2.4 *ClHSF* 基因的结构与保守基序分析

为了阐明薏苡 *HSF* 基因家族的结构特征，我们对 20 个 *ClHSF* 基因成员进行了蛋白结构域和内含子-外显子结构的全面分析。通过 *ClHSF* 全序列比对构建的系统进化树(图 3A)清晰地展示了各成员间的进化关系。分析结果显示，*ClHSF* 基因平均含有 2 个外显子，数量变异范围为 1–4 个(图 3B)。使用 MEME 在线程序对 ClHSF 蛋白保守基序的分析鉴定出 15 个特征性 motif，其中 Motif 1、Motif 2 和 Motif 3 在几乎所有 ClHSF 蛋白中都存在(图 3C)。

保守结构域分析表明，所有 ClHSF 蛋白都含有高度保守的 HSF_DNA-bind 超家族结构域。此外，我们还发现 7 个 *ClHSF* 含有额外结构域：*ClHSF04* 含有 PRK10246 超家族结构域，*ClHSF16* 含有 ZIP_TSC22D-like 超家族结构域，*ClHSF12* 含有 MDN1 超家族结构域，*ClHSF08* 含有 DUF5401 超家族结构域，*ClHSF15* 含有 Smc 超家族结构域，*ClHSF19* 含有 bzip 和 UPF0182 超家族结构域，*ClHSF11* 含有 BRLZ 结构域(图 3D)。总体而言，外显子和内含子的序列与数量、基序模式以及保守结构域分布都支持 *ClHSF* 基因的进化关系。

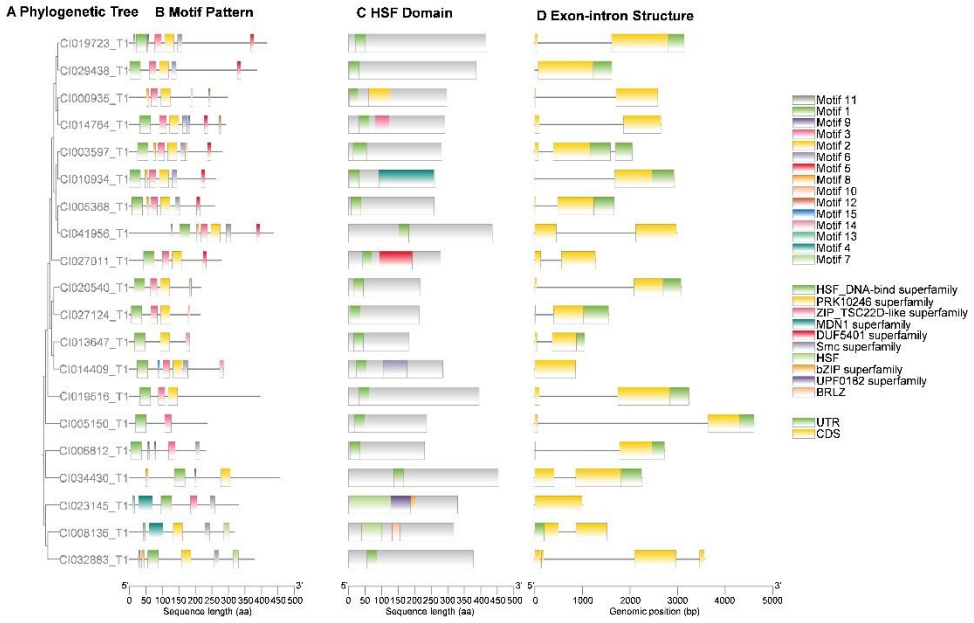


图 3 *ClHSF* 基因成员的结构特征展示

(A) 基于 *ClHSF* 蛋白全序列，通过 MEGA 4.0 软件构建的邻接法系统进化树；**(B)** ClHSF 蛋白保守基序分布图，共鉴定出 15 个特征性基序；**(C)** ClHSF 蛋白保守结构域预测结果，彩色方框代表不同保守结构域，灰色条形表示蛋白序列长度；**(D)** *ClHSF* 基因的外显子-内含子结构示意图，其中黄色方框表示编码区(CDS)，灰色方框标注 5'和 3'非翻译区，黑色线条代表内含子区域。

Figure 3 Members of the *ClHSF* gene are arranged as follows in their structure

(A) Using the full sequences of *ClHSF* proteins, MEGA 4.0 was used to create the Neighbor-Joining phylogenetic tree; **(B)** The distribution of conserved motifs in ClHSF proteins, with a total of 15 motifs identified; **(C)** The predicted conserved structural domains of ClHSF proteins, where colored boxes indicate conserved domains and grey bars denote protein sequence length. **(D)** The *ClHSF* genes' exon-intron structure, with gray boxes denoting the 5' and 3' untranslated regions and yellow boxes showing exons (CDS) and black lines indicating introns.

2.5 *ClHSF* 基因的全基因组复制事件与共线性分析

我们采用 MCScanX 软件对薏苡 *ClHSF* 基因成员的全基因组复制事件进行分析。研究

结果显示，薏苡基因组中的复制基因对呈现明显的共线性模式(图 4A)。其中红色线条标示 *ClHSF* 基因，灰色线条表示所有复制基因对。分析共鉴定出 8 个 *ClHSF* 基因，形成 4 对复制基因。值得注意的是，2 号染色体包含 4 个基因，1 号和 6 号染色体各含有 3 个基因，其余染色体通常分布 1–2 个基因，而 8 号和 10 号染色体未发现基因分布。这表明这些基因可能来源于全基因组复制事件，凸显了这些事件在推动 *ClHSF* 基因家族扩张中的关键作用。

为更深入研究 *ClHSF* 基因的复制情况，我们构建了 7 个比较共线性图谱(图 4B)。与 *ClHSF* 基因关联性最强的物种依次为：短柄草(23 个)、小麦(18 个)、高粱(24 个)、水稻(21 个)、玉米(18 个)和拟南芥(2 个)。比较基因组学分析表明，串联复制和片段复制事件在单子叶与双子叶植物进化过程中对基因家族扩张及多样性形成具有重要贡献。

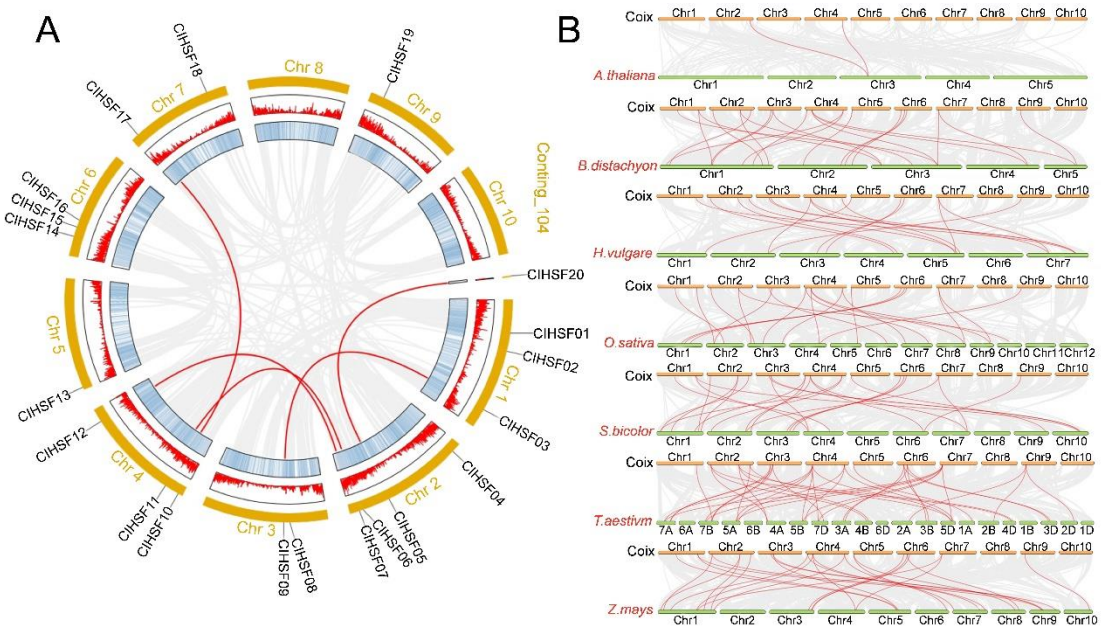


图 4 *ClHSF* 基因的染色体间关联示意图及与七种典型植物物种的共线性分析。

(A) *ClHSF* 基因的染色体间关联示意图，图中红色线条标示薏苡基因组内复制的 *HSF* 基因对，各染色体编号居中标注以便准确定位。灰色背景线条表示全基因组范围内的其他同源基因对，黑色刻度标示染色体物理距离; **(B)** *ClHSF* 基因与七种典型植物物种的共线性分析。

Figure 4 Diagrammatic representations of the *ClHSF* gene interchromosomal connections and Synteny investigation comparing seven typical plant species with the *ClHSF* genes.

(A) Diagrammatic representations of the *ClHSF* gene interchromosomal connections.Duplicated HSF gene pairs within *Coix* are delineated by red lines. Chromosome numbers are centrally annotated for precise identification; **(B)** Synteny investigation comparing seven typical plant species with the *ClHSF* genes.

2.6 *ClHSF* 基因的启动子顺式作用元件分析

为探究 *ClHSF* 基因响应非生物胁迫的调控机制，我们对 20 个 *ClHSF* 基因上游 2000 bp 序列的顺式作用元件进行了分析。研究发现，这些基因的启动子区含有大量与胁迫、环境及激素响应相关的顺式作用元件，表明不同 *ClHSF* 基因可能具有多样化的潜在功能。大多数 *ClHSF* 基因启动子含有赤霉素(gibberellin)、水杨酸(salicylic acid)、生长素(auxin)、茉莉酸甲酯(MEJA)和脱落酸(abscisic acid)等激素响应元件(图 5)。其中, 17 个 *ClHSF* 基因含有赤霉素响应元件，12 个基因含有茉莉酸甲酯(MEJA)响应元件。启动子区包含低温响应(LTR)、缺氧诱导、MYB 转录因子结合位点介导的胁迫响应，以及分生组织表达元件调控的生长发育相关元件。光响应元件(light response element)在几乎所有 *ClHSF* 基因启动子中均存在。此外，还发现与细胞周期调控、昼夜节律控制等相关的顺式作用元件。综合分析表明, *ClHSF* 基因可能通过复杂的顺式调控网络参与多种激素信号转导、生长发育调控及非生物胁迫响应过程。



图 5 *ClHSF* 基因上游 2000 bp 启动子区顺式作用元件的分布特征及其组成分析

Figure 5 The distribution of *cis*-acting elements found in the 2000 bp upstream promoter region of the *ClHSF* genes, the cis-acting element on the putative promoter of the *ClHSF* genes

2.7 *ClHSF* 基因在不同植物组织和多种非生物胁迫下的表达模式

为了全面了解 *ClHSF* 在不同组织中的表达模式，我们检测了种子、茎、根和叶的 RNA-seq 数据。使用 TBtools 软件，我们绘制了四种器官 log2FC 值的热图来检测 *ClHSF* 转录本的表达(图 6A)。不同组织中 *ClHSF* 的表达水平存在差异。几乎所有 *ClHSF* 基因至少在一种组织中显示出不同的表达水平。例如，叶片中 *ClHSF01*、*ClHSF05*、*ClHSF11*、*ClHSF13* 和 *ClHSF20* 的含量最高，其次是茎、种子和根。而另一个独特的 *ClHSF08* 仅在根中不表达，一个特定的、*ClHSF18* 在茎和种子中表达，在其他组织中不表达。所有检测的

组织都有一些 *ClHSF20* 基因表达，但某些器官的表达水平高于其他器官。例如，叶片中 *ClHSF01*、*ClHSF04*、*ClHSF11*、*ClHSF14* 和 *ClHSF20* 的表达水平最高，而根中 *ClHSF09*、*ClHSF15*、*ClHSF06* 和 *ClHSF03* 的表达水平最高。

我们利用 RNA-seq 数据检测了 *ClHSF* 基因在热、冷、干旱和盐胁迫处理下的表达水平，以探索这些基因在不同非生物胁迫下的潜在作用。在这些处理中，每个 *ClHSF* 基因都有表达，且至少被一种胁迫处理激活。图 6B 显示，其中 19 个被冷胁迫激活，20 个被热胁迫激活，18 个被干旱激活，18 个被盐胁迫激活；同时部分 *ClHSF* 基因可被相应胁迫处理抑制。综上，热胁迫对 *ClHSF* 基因的激活效应最为显著，而干旱与盐胁迫下被激活的 *ClHSF* 基因数量相对较少。

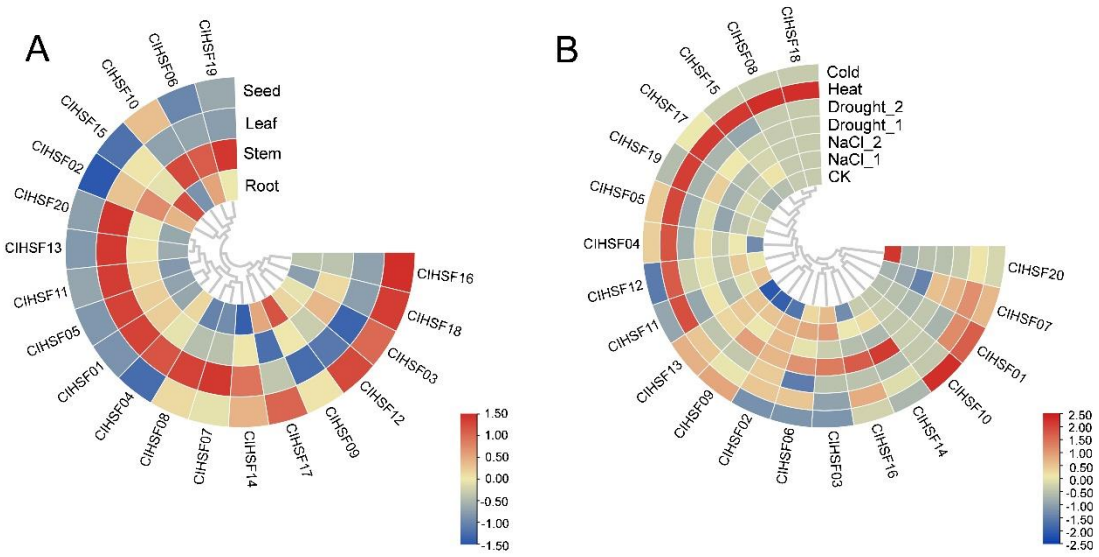


图 6 *ClHSF* 基因表达谱分析。

(A) *ClHSF* 基因在四种不同组织样本中的表达谱层次聚类分析。**(B)** 非生物胁迫处理下 HSF 基因的表达谱特征。

Figure 6 Profiles of *ClHSF* gene expression.

(A) Hierarchical clustering of *ClHSF* gene expression profiles across four samples, spanning several tissues. **(B)** HSF gene expression profiles under abiotic stress treatment.

2.8 *ClHSF* 基因在盐处理和干旱处理下的表达谱

采用 qRT-PCR 技术检测了 14 个 *ClHSF* 基因在盐胁迫和干旱胁迫下的表达模式(图 8)。盐胁迫分析显示，*ClHSF03*、*ClHSF04*、*ClHSF05*、*ClHSF09*、*ClHSF11*、*ClHSF14*、*ClHSF17*、*ClHSF19* 和 *ClHSF20* 表达显著下调(P<0.0001)，而 *ClHSF08*、*ClHSF13* 和 *ClHSF15* 则显著上调(图 7A)。干旱胁迫实验表明，随着 PEG6000 浓度增加，多数基因呈现先上调后下调的表达趋势(图 7B)。其中，PEG_1 处理下，*ClHSF03*、*ClHSF05*、*ClHSF08*、*ClHSF13* 和

ClHSF19 分别较对照上调 14.24、3.82、30.33、65.94 和 23.14 倍, *ClHSF20* 则下调 0.28 倍(P<0.0001)。这些结果表明 *ClHSF* 基因在薏苡响应盐胁迫和干旱胁迫过程中发挥重要调控作用。

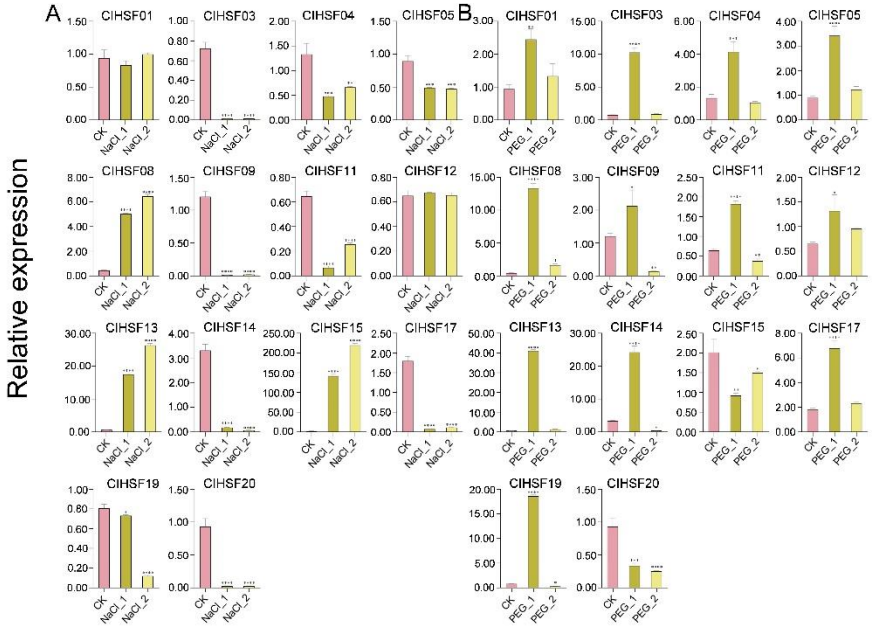


图 7 盐胁迫和干旱胁迫处理的薏苡叶片中 *ClHSF* 基因表达研究

(A) 盐处理实验设计: NaCl_1: 100 mmol·L^{-1} NaCl 处理; NaCl_2: 150 mmol·L^{-1} NaCl 处理。**(B)** 干旱处理实验设计: PEG_1: 10% PEG6000 处理; PEG_2: 15% PEG6000 处理。14 个选定 *ClHSF* 基因在盐胁迫处理下的相对表达水平。实验结果以均值±标准差表示。经 LSD 检验, 不同字母标记表示基因表达差异显著(P<0.05), 相同字母表示差异不显著(P>0.05)。

Figure 7 *ClHSF* gene expression study in *coix* leaves treated with salt and drought.

(A) The following were the ways of treatment: NaCl_1: *coix* treated with 100 mmol·L^{-1} NaCl; NaCl_2: *coix* treated with 150 mmol·L^{-1} NaCl. **(B)** The following were the methods of treatment: PEG_1: 10% and PEG_2: 15% of PEG6000, respectively, are treated. relative expression levels of 14 *ClHSF* genes that were chosen and treated with salt. The mean ± standard deviation represents the final results. Genes have been shown to differ significantly based on the LSD test; a significant difference (P<0.05) is indicated by the opposite letter, which indicates no significant difference (P>0.05).

2.9 *ClHSF* 基因在温度胁迫下的表达谱分析

为研究 *HSF* 基因对非生物胁迫的响应, 我们检测了薏苡基因表达对温度胁迫的反应。结果表明 *ClHSF* 基因表达水平具有温度依赖性。在不同温度条件下(图 8), 14 个基因表现出显著表达变化。具体而言: 高温胁迫下, *ClHSF01*、*ClHSF04*、*ClHSF05*、*ClHSF08*、*ClHSF12*、*ClHSF13*、*ClHSF15*、*ClHSF17* 和 *ClHSF19* 显著上调(P<0.0001), 其中 *ClHSF01* 表达量比对照组高 11.72 倍, *ClHSF08* 高 64.34 倍, *ClHSF12* 高 7.76 倍; 而 *ClHSF03*、*ClHSF09*、

ClHSF11 和 *ClHSF14* 显著下调(*P*<0.0001), *ClHSF09* 仅为对照组的 0.06 倍。低温胁迫下, *ClHSF01*、*ClHSF03*、*ClHSF04*、*ClHSF011*、*ClHSF14*、*ClHSF15* 和 *ClHSF19* 均显著上调(*P*<0.0001), 其中 *ClHSF01* 变化最显著(比对照组高 8.68 倍), *ClHSF15* 高 0.153 倍。由此推断这些基因在响应高低温胁迫中起重要调控作用, 对薏苡适应极端温度环境具有不可或缺的功能。

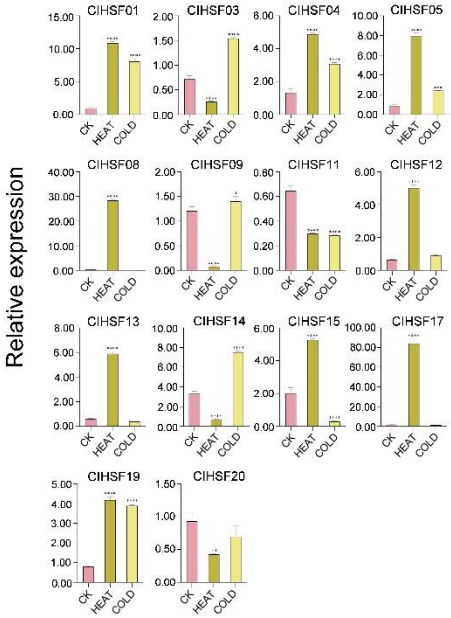


图 8 薏苡叶片 *ClHSF* 基因在不同温度处理下的表达分析。实验处理条件为: 高温组 40°C 处理, 低温组 4°C 处理。结果显示 14 个 *ClHSF* 基因在不同温度条件下的相对表达量(数据以均值±标准差表示)。经 LSD 检验, 不同小写字母表示组间差异显著(*P*<0.05), 相同字母表示差异不显著(*P*>0.05)。

Figure 8 *ClHSF* gene expression study in *coix* leaves treated at various temperatures. The following were the ways of treatment: Heat: *coix* when treated at 40°C; cold: *coix* when treated at 4°C. relative expression levels of 14 *ClHSF* genes that were chosen and treated with different temperatures. The mean ± standard deviation represents the final results. Genes have been shown to differ significantly based on the LSD test; a significant difference (*P*<0.05) is indicated by the opposite letter, which indicates no significant difference (*P*>0.05).

3 讨论

薏苡(*Coix lacryma-jobi* L.var.ma-yuen (Roman.) Stapf)是我国重要的药食同源作物, 兼具药用价值与食用价值。随着其基因组测序的完成, 为抗性基因挖掘及优良性状鉴定提供了基因组基础, 对培育抗逆性强、品质优良的薏苡新品种具有重要指导意义。

热激转录因子(*HSF*)是植物生长发育和防御过程中至关重要的转录调控因子, 通过调控热激蛋白(HSPs)及其他功能基因的表达(Baniwal et al., 2004), 在植物响应冷、热、旱、盐等非生物胁迫中发挥关键作用(Zhang et al., 2020)。尽管 HSF 基因家族已在黄瓜、小麦、

葡萄等重要经济作物中得到广泛研究(Duan et al., 2019; Guo et al., 2015; Huang et al., 2021), 但在薏苡中的研究尚未见报道。

本研究在薏苡基因组中鉴定出20个*HSF*基因家族成员，其染色体分布具有明显特异性，其中2号染色体呈现显著的基因富集特征。数量上，薏苡*HSF*基因少于玉米(25个) (Lin et al., 2011)和水稻(25 个) (Wan et al., 2010), 略少于拟南芥(21 个) (Guo et al., 2016)。这一差异可能反映了物种特异性基因复制或丢失事件，与禾本科植物在进化过程中经历的复杂基因组重组事件相关(Wang et al., 2015)。系统进化分析将 *ClHSF* 基因划分为 3 个亚家族，与拟南芥、大麦和玉米 HSF 蛋白的聚类模式一致。基因结构分析表明，所有 *ClHSF* 成员均具有保守的外显子结构和特征性 motif(1-3)。共线性比较显示，薏苡与高粱 *HSF* 基因具有高度同源性，进一步揭示了两者之间最强的共线性关系(图 4B)，这与两者同属禾本科黍亚科(Panicoideae)、亲缘关系较近的结论一致(Hua et al., 2023)，为阐明禾本科植物 *HSF* 基因的进化机制提供了重要依据。

启动子顺式作用元件分析表明，20 个 *ClHSF* 基因的启动子区含有参与植物生长发育、激素调控和逆境响应的多种顺式元件。值得注意的是，*ClHSF04*、*ClHSF05* 等基因启动子区含有脱落酸响应元件(盐胁迫)、MYB 结合位点(干旱诱导)和光响应元件(低温响应)等特异性调控元件。ABRE 和 MBS 元件分别在 ABA 信号介导的盐胁迫和干旱胁迫响应中起关键作用，这与水稻(Tao et al., 2024)和月季(Zhu et al., 2025)中的研究结果一致，暗示这些 *ClHSF* 可能通过类似的激素信号途径参与胁迫应答。组织表达谱显示，*ClHSF* 基因呈现器官表达特异性：*ClHSF01/04/05* 等在叶片中高表达，*ClHSF06/10/19* 等在茎中优势表达，*ClHSF03/15/17* 等在根中显著富集，*ClHSF03/12/16* 等在种子中活跃表达，这种表达模式差异提示不同 *ClHSF* 成员可能在器官发育与环境适应中发挥特异性调控作用。qRT-PCR 分析表明，*ClHSF* 基因对非生物胁迫呈现特异性响应模式。盐胁迫显著调控 *ClHSF11/13-15/19* 等基因表达(图 7A)，而温度胁迫主要影响 *ClHSF05/08-09/13-15/17* 等基因的表达谱(图 8)。值得注意的是，HSFA 亚家族的 *ClHSF13/14* 基因表现出广谱胁迫响应特性，可能与其代谢调控功能相关。该研究结果为阐明薏苡抗逆分子机制提供了新见解，并为禾本科作物抗逆遗传改良提供了潜在基因资源。

4 结论

本研究在薏苡基因组中鉴定出 20 个 *ClHSF* 基因家族成员，这些基因不均匀地分布在 9 条染色体上。根据系统发育分析，这些 *ClHSF* 成员可分为 HSFA、HSFB 和 HSFC 三个亚类。基因复制分析表明，串联重复和片段重复共同促进了 *ClHSF* 基因家族的扩张。表达分析显示，在热、冷、旱、盐等胁迫条件下，不同亚类的 *ClHSF* 基因均呈现显著差异表达。这些基因的表达模式不仅揭示了其在薏苡生长发育中的重要作用，更表明了其对温度和干旱胁迫的高度敏感性。本研究结果系统解析了 *ClHSF* 基因家族的进化特征及其在不同环境胁

迫下的潜在功能，为薏苡抗逆育种提供了重要的理论基础和基因资源。

作者贡献声明

朱紫嫣：研究方法、数据分析、实验操作、数据验证、论文初稿撰写。彭玲霞：研究方法、数据分析、实验操作、数据验证、论文初稿撰写。李梦妮：资源提供、实验操作、项目管理。杨丽君：资源提供、实验操作、项目管理。李洪菊：资源提供、实验操作、项目管理。赵玉阳和金燕：资源提供、实验操作、项目管理。陆若辉：资源提供、实验操作、项目管理。张晓丹：研究构思、论文修改、研究监督、资金获取、项目管理。

参考文献

Arora, R.K (1977). Job's-tears (*Coix lacryma-jobi*)—a minor food and fodder crop of northeastern India. Econ. Bot. 31:358–366. Doi: https://doi.org/10.1007/BF02866887.

Baniwal SK, Bharti K, Chan KY, Fauth M, Ganguli A, Kotak S, Mishra SK, Nover L, Port M, Scharf KD, Tripp J, Weber C, Zielinski D, von Koskull-Döring P (2004). Heat stress response in plants: a complex game with chaperones and more than twenty heat stress transcription factors. J Biosci. 29(4):471-87. doi: 10.1007/BF02712120.

Chidambaranathan P, Jagannadham PTK, Satheesh V, Kohli D, Basavarajappa SH, Chellapilla B, Kumar J, Jain PK, Srinivasan R (2018). Genome-wide analysis identifies chickpea (Cicer arietinum) heat stress transcription factors (Hsfs) responsive to heat stress at the pod development stage. J Plant Res.131(3):525-542. doi: 10.1007/s10265-017-0948-y.

Chen X. Wang Z. Tang R. Wang L. Chen C. Ren Z (2021). Genome-wide identification and expression analysis of Hsf and Hsp gene families in cucumber (Cucumis sativus L.). Plant Growth Regul. 95: 223–239. doi: 10.1007/s10725-021-00739-z.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 13(8):1194-1202. doi: 10.1016/j.molp.2020.06.009.

Cahill AE, Aiello-Lammens ME, Fisher-Reid MC, Hua X, Karanewsky CJ, Ryu HY, Sbeglia GC, Spagnolo F, Waldron JB, Warsi O, Wiens JJ (2013). How does climate change cause extinction? Proc Biol Sci. 280(1750):20121890. doi: 10.1098/rspb.2012.1890.

Duan S, Liu B, Zhang Y, Li G, Guo X (2019). Genome-wide identification and abiotic stress-responsive pattern of heat shock transcription factor family in Triticum aestivum L. BMC Genomics. 20(1):257. doi: 10.1186/s12864-019-5617-1.

Friedrich T, Timmermann A, Tigchelaar M, Elison Timm O, Ganopolski A (2016). Nonlinear climate sensitivity and its implications for future greenhouse warming. Sci Adv. 2(11): e1501923. doi: 10.1126/sciadv.1501923.

Guo M, Liu JH, Ma X, Luo DX, Gong ZH, Lu MH (2016). The Plant Heat Stress Transcription Factors (HSFs): Structure, Regulation, and Function in Response to Abiotic Stresses. Front Plant Sci. 7:114. doi: 10.3389/fpls.2016.00114.

Guo Q, Wei R, Xu M, Yao W, Jiang J, Ma X, Qu G, Jiang T (2023). Genome-wide analysis of HSF family and overexpression of PsnHSF21 confers salt tolerance in Populus simonii × P. nigra. Front Plant Sci. 14:1160102. doi: 10.3389/fpls.2023.1160102.

Guotian L, Fengmei C, Yi W, Jinzhu J, Wei D, Yuting W, Fangfang W, Shaohua L, Lijun W (2018).

Genome-wide Identification and Classification of HSF Family in Grape, and Their Transcriptional Analysis under Heat Acclimation and Heat Stress, Horticultural Plant Journal, 4 (4): 133-143. https://doi.org/10.1016/j.hpj.2018.06.001.

Guo M, Lu JP, Zhai YF, Chai WG, Gong ZH, Lu MH (2015). Genome-wide analysis, expression profile of heat shock factor gene family (CaHsfs) and characterisation of CaHsfA2 in pepper (Capsicum annuum L.). BMC Plant Biol. 15:151. doi: 10.1186/s12870-015-0512-7.

Guo M, Liu JH, Ma X, Luo DX, Gong ZH, Lu MH (2016). The Plant Heat Stress Transcription Factors (HSFs): Structure, Regulation, and Function in Response to Abiotic Stresses. Frontiers in Plant Science. 7:114. doi: 10.3389/fpls.2016.00114

Huang B, Huang Z, Ma R, Chen J, Zhang Z, Yrjälä K (2021). Genome-wide identification and analysis of the heat shock transcription factor family in moso bamboo (Phyllostachys edulis). Sci Rep. 11(1):16492. doi: 10.1038/s41598-021-95899-3.

Hua Y, Liu Q, Zhai Y, Zhao L, Zhu J, Zhang X, Jia Q, Liang Z, Wang D (2023). Genome-wide analysis of the HSP20 gene family and its response to heat and drought stress in *Coix* (*Coix lacryma-jobi* L.). BMC Genomics. 24(1):478. doi: 10.1186/s12864-023-09580-2.

Hua Y, Liu Q, Zhai Y, Zhao L, Zhu J, Zhang X, Jia Q, Liang Z, Wang D (2023). Genome-wide analysis of the HSP20 gene family and its response to heat and drought stress in Coix (Coix lacryma-jobi L.). BMC Genomics. 24(1):478. doi: 10.1186/s12864-023-09580-2

Jeffares DC, Penkett CJ, Bähler J (2008). Rapidly regulated genes are intron poor. Trends Genet. 24(8):375-8. doi: 10.1016/j.tig.2008.05.006. 2008. Erratum in: Trends Genet. 2008 Oct;24(10):488.

Jaureguiberry P, Titeux N, Wiemers M, Bowler DE, Coscieme L, Golden AS, Guerra CA, Jacob U, Takahashi Y, Settele J, Díaz S, Molnár Z, Purvis A (2022). The direct drivers of recent global anthropogenic biodiversity loss. Sci Adv. 8(45): eabm9982. doi: 10.1126/sciadv. abm9982.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 35(6):1547-1549. doi: 10.1093/molbev/msy096.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007). Clustal W and Clustal X version 2.0. Bioinformatics. 23(21):2947-8. doi: 10.1093/bioinformatics/btm404. Epub 2007 Sep 10. PMID: 17846036.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 30(1):325-7. doi: 10.1093/nar/30.1.325.

Ling C, Liu Y, Yang Z, Xu J, Ouyang Z, Yang J, Wang S (2023). Genome-Wide Identification of HSF Gene Family in Kiwifruit and the Function of AeHSFA2b in Salt Tolerance. Int J Mol Sci. Oct 27;24(21):15638. doi: 10.3390/ijms242115638.

Lin YX, Jiang HY, Chu ZX, Tang XL, Zhu SW, Cheng BJ (2011). Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. BMC Genomics. 12:76. doi: 10.1186/1471-2164-12-76

Ma J, Zhang G, Ye Y, Shang L, Hong S, Ma Q, Zhao Y, Gu C (2022). Genome-Wide Identification and Expression Analysis of HSF Transcription Factors in Alfalfa (Medicago sativa) under Abiotic Stress. Plants (Basel). 11(20):2763. doi: 10.3390/plants11202763.

Mishra SK, Tripp J, Winkelhaus S, Tschiersch B, Theres K, Nover L, Scharf KD (2002). In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of

thermotolerance in tomato. Genes Dev. 16(12):1555-67. doi: 10.1101/gad.228802.

Nagaraju M, Reddy PS, Kumar SA, Srivastava RK, Kishor PB, Rao DM (2015). Genome-wide Scanning and Characterization of Sorghum bicolor L. Heat Shock Transcription Factors. Curr Genomics. 16(4):279-91. doi: 10.2174/1389202916666150313230812.

Parker CS, Topol J (1984). A Drosophila RNA polymerase II transcription factor binds to the regulatory site of an hsp 70 gene. Cell. 37(1):273-83. doi: 10.1016/0092-8674(84)90323-4.

Román-Palacios C, Wiens JJ (2020). Recent responses to climate change reveal the drivers of species extinction and survival. Proc Natl Acad Sci U S A. 117(8):4211-4217. doi: 10.1073/pnas.1913007117.

Ren Y, Ma R, Xie M, Fan Y, Feng L, Chen L, Yang H, Wei X, Wang X, Liu K, Cheng P, Wang B (2023). Genome-wide identification, phylogenetic and expression pattern analysis of HSF family genes in the Rye (Secale cereale L.). BMC Plant Biol. 23(1):441. doi: 10.1186/s12870-023-04418-1.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011). Integrative genomics viewer. Nature Biotechnology, 29(1):24-6. doi: 10.1038/nbt.1754.

Schmittgen TD, Livak KJ (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc. 3(6):1101-8. doi: 10.1038/nprot.

Suzuki N, Rivero RM, Shulaev V, Blumwald E, Mittler R (2014). Abiotic and biotic stress combinations. New Phytol. 203(1):32-43. doi: 10.1111/nph.12797.

Scharf KD, Heider H, Hohfeld I, Lyck R, Schmidt E, Nover L (1998). The tomato hsf system: HsfA2 needs interaction with HsfA1 for efficient nuclear import and may be localized in cytoplasmic heat stress granules. Mol. Cell Biol. 18, 2240–2251. doi: 10.1128/MCB.18.4.2240.

Tang R, Zhu W, Song X, Lin X, Cai J, Wang M, Yang Q (2016). Genome-Wide Identification and Function Analyses of Heat Shock Transcription Factors in Potato. Front Plant Sci. 7:490. doi: 10.3389/fpls.2016.00490.

Tamura K, Stecher G, Kumar S (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 38(7):3022-3027. doi: 10.1093/molbev/msab120.

Tao H, Xu S, Shen H, Yang J, Xu Y, Huang G, Feng C, Wan W, Woldegiorgis ST, Liu W, He H (2024). Genome-wide characterization and functional analysis of heat shock transcription factors in wild and cultivated rice (*Oryza sativa* L.). Plant Stress. 12:100505. doi: 10.1016/j.stress.2024.100505.

Umezawa T, Fujita, M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K (2006). Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Curr. Opin. Biotechnol. 2006. 17, 113–22. doi: 10.1016/j.copbio.2006.02.002.

Waters ER (2013). The evolution, function, structure, and expression of the plant sHSPs. J Exp Bot. 64(2):391–403.

Wu C (1984). Activating protein factor binds in vitro to upstream control sequences in heat shock gene chromatin. Nature. 311(5981):81-4. doi: 10.1038/311081a0.

Wang F, Dong Q, Jiang H, Zhu S, Chen B, Xiang Y (2012). Genome-wide analysis of the heat shock transcription factors in Populus trichocarpa and Medicago truncatula. Mol Biol Rep. 39(2):1877-86. doi: 10.1007/s11033-011-0933-9.

Wang J, Liu L, Ball T, Yu L, Li Y, Xing F (2016). Revealing a 5,000-y-old beer recipe in China. Proc Natl Acad Sci U S A. 113(23):6444-8. doi: 10.1073/pnas.1601465113.

Wan BL, Zha ZP, Du XS (2010). Expression Profile Analysis of Rice Heat Shock Transcription Factor (HSF) Genes in Response to Plant Hormones and Abiotic Stresses. China Biotechnology. 30(10):22-32.

Wang X, Wang J, Jin D, Guo H, Lee TH, Liu T, Paterson AH (2015). Genome Alignment Spanning Major Poaceae Lineages Reveals Heterogeneous Evolutionary Rates and Alters Inferred Dates for Key Evolutionary Events. Molecular Plant. 8(6):885-898. doi: 10.1016/j.molp.2015.04.004

Yang X, and Jiang L (2010). Starch grain analysis reveals ancient diet at Kuahuqiao site, Zhejiang Province. Chin. Sci. Bull. 55:1150–1156. Doi: 10.1186/s12969-023-00879-8.

Zhang P, Meng X, Tang X, Ren L, Liang J (2019). The effect of a *coix* seed oil injection on cancer pain relief. Support Care Cancer. 27(2):461-465. doi: 10.1007/s00520-018-4313-z.

Zhu F (2017). *Coix*: chemical composition and health effects. Trends Food Sci. Technol. 61:160–175. doi: 10.1016/j.tifs.2016.12.003.

Zhang Q, Geng J, Du Y, Zhao Q, Zhang W, Fang Q, Yin Z, Li J, Yuan X, Fan Y, Cheng X, Du J (2022). Heat shock transcription factor (Hsf) gene family in common bean (Phaseolus vulgaris): genome-wide identification, phylogeny, evolutionary expansion and expression analyses at the sprout stage under abiotic stress. BMC Plant Biol. 22(1):33. doi: 10.1186/s12870-021-03417-4.

Zhu J, Fan S, Li R, Dong F, Liu Y, Wang C (2025). Genome-Wide Identification, Phylogenetic Classification, and Expression Profiling of the HSF Gene Family in *Rosa hybrida* Under Heat and Drought Stress. Plants. 14(20):3167. doi: 10.3390/plants14203167

Investigating the Heat Shock Transcription Factor (*HSF*) Gene Family in *Coix* (*Coix lacryma-jobi* L.): A Comprehensive Genome-Wide Analysis of Evolution, Expression Profiling, and Responses to Developmental and Abiotic Stresses

Ziyan Zhu[1†], Lingxia Peng[1†], Mengni Li[2], Lijun Yang[1], Hongju Li[1], Yuyang Zhao[3], Yan Jin[3], Ruohui Lu[4], Xiaodan Zhang[1*]

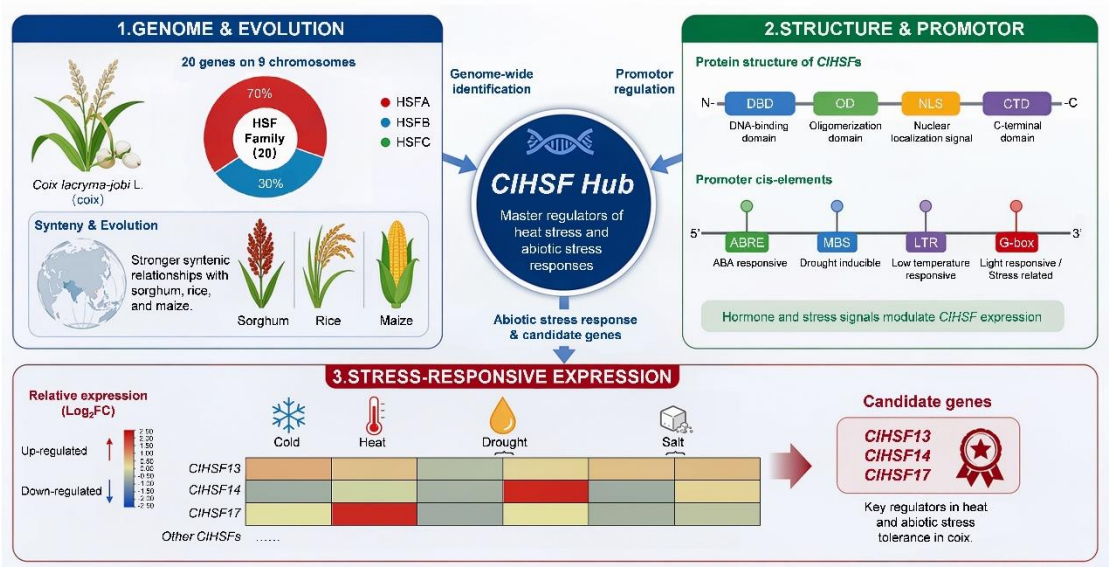
[1]Key Laboratory of Plant Secondary Metabolism and Regulation of Zhejiang Province, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou 310018, China; [2]Agricultural and Rural Bureau of Yueqing City, Yueqing 325600, Zhejiang, China; [3]State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China; [4]Cultivated Land Quality and Fertilizer Management Station of Zhejiang Province, Yueqing 325600, China

INTRODUCTION: Heat shock transcription factors (*HSFs*) are pivotal regulators in plant responses to a variety of abiotic stresses, including heat, drought, salinity, and cold. They function primarily by modulating the expression of heat shock proteins (HSPs) and other stress-responsive genes, thereby maintaining protein homeostasis under adverse environmental conditions. Although *HSF* gene families have been systematically characterized in several major crops and model plants, comprehensive information regarding *HSF* genes in *Coix lacryma-jobi* L., a widely cultivated medicine-food homologous crop in East and Southeast Asia, remains absent from the current literature.

RATIONALE: Given the increasing frequency of extreme climatic events that threaten crop productivity, it is imperative to explore the genetic resources that may confer enhanced stress tolerance in underutilized yet nutritionally and medicinally valuable crops such as coix. The recent release of a high-quality reference genome for coix provides an unprecedented opportunity to systematically identify and analyze the *HSF* gene family in this species. Elucidating the evolutionary relationships, structural features, and expression patterns of *ClHSF* genes under developmental and abiotic stress conditions will contribute to a better understanding of their functional roles and inform future molecular breeding efforts aimed at improving coix resilience.

RESULTS: In this study, a total of 20 *ClHSF* genes were identified at the whole-genome level and were found to be unevenly distributed across nine chromosomes. Phylogenetic analysis classified these genes into three subfamilies—HSFA, HSFB, and HSFC—consistent with the clustering patterns observed in Arabidopsis, rice, and maize. All identified ClHSF proteins contained the conserved HSF DNA-binding domain, and gene structure and conserved motif analyses further supported their evolutionary relationships. Collinearity analysis revealed that both tandem and segmental duplication events have contributed to the expansion of the *ClHSF* gene family, with a notably high degree of synteny observed between coix and sorghum. Promoter cis-acting element analysis identified multiple hormone- and stress-responsive elements, including ABRE, MBS, and LTR, suggesting that *ClHSF* genes are regulated through complex signaling networks. RNA-seq and qRT-PCR analyses demonstrated that *ClHSF* genes exhibited tissue-specific expression patterns and were differentially responsive to heat, cold, salt, and drought treatments. Notably, several genes, including *ClHSF13, ClHSF14, and ClHSF17*, showed significant and specific upregulation or downregulation under multiple stress conditions, highlighting their potential as key regulators in the stress adaptation machinery of coix.

CONCLUSION: This study provides the first systematic characterization of the *HSF* gene family in coix, revealing its evolutionary history, structural diversity, and dynamic expression in response to developmental cues and abiotic stresses. The identification of stress-responsive *ClHSF* members offers valuable candidate genes for future functional studies and genetic improvement programs. Collectively, these findings establish a foundational framework for understanding the molecular mechanisms through which *ClHSF* genes contribute to stress tolerance and adaptive growth in coix.



Coix *CIHSF* genes: Structure, promoter regulation, and abiotic stress responses

Keywords *Coix*, *HSF* gene family, abiotic stress, genome-wide

† These authors contributed equally to this paper

* Author for correspondence. E-mail: zhangxiaodanzstu@163.com

通讯作者简介

张晓丹, 浙江理工大学生命科学与医药学院副教授, 硕士研究生导师。研究方向为天然产物活性物质基础及生药资源开发与评价研究。