

王占军, 汪虹妍, 杨妍萍, 等. 油桐 *HSP70* 基因家族的全基因组鉴定与表达分析[J]. 江苏农业学报, 2024, 40(5): 806-816.
doi: 10.3969/j.issn.1000-4440.2024.05.005

油桐 *HSP70* 基因家族的全基因组鉴定与表达分析

王占军^{1,2}, 汪虹妍¹, 杨妍萍¹, 叶青芳¹, 王淑雯¹, 王朝霞¹, 张翼², 张颢²

(1. 合肥师范学院生物与食品工程学院, 安徽 合肥 230601; 2. 湖南省林业科学院省部共建木本油料资源利用国家重点实验室, 湖南 长沙 410004)

摘要: 为探究油桐 *HSP70* 基因的功能和进化关系, 本研究对 *HSP70* 基因家族成员进行了生物信息学鉴定和分析, 包括理化性质、进化关系、基因结构、保守基序及染色体定位。同时分析了 *HSP70* 基因家族成员在不同组织中的表达差异。结果表明, 油桐 *HSP70* 基因家族包含 24 个家族成员, 24 个 *HSP70* 蛋白的理论等电点为 4.57~8.70, 且大部分属于稳定的、亲水性蛋白质, 系统进化树分析结果表明 *HSP70* 家族成员分为 5 个聚类, 油桐与毛果杨亲缘关系较近。24 个 *VfHSP70* 蛋白氨基酸序列都包含 *HSP70* 保守结构域。22 条 *HSP70* 基因分布在 10 条染色体上。分析不同时空的 *VfHSP70* 基因的表达量发现, *VfHSP70-5* 在开花前 30 d、开花前 25 d、开花前 10 d 的花蕾中表达量均较高。*VfHSP70-4* 在油桐根、茎、叶和种子中均有较高表达量。以上结果表明, *VfHSP70* 基因可能在调控油桐生长发育中具有重要作用。

关键词: 油桐; *HSP70*; 基因家族; 生物信息学

中图分类号: S794.3

文献标识码: A

文章编号: 1000-4440(2024)05-0806-11

Genome-wide identification and expression analysis of *HSP70* gene family in *Vernicia fordii*

WANG Zhanjun^{1,2}, WANG Hongyan¹, YANG Yanping¹, YE Qingfang¹, WANG Shuwen¹,
WANG Zhaoxia¹, ZHANG Yi², ZHANG Xie²

(1. School of Biology and Food Engineering, Hefei Normal University, Hefei 230601, China; 2. State Key Laboratory of Utilization of Woody Oil Resource, Hunan Academy of Forestry, Changsha 410004, China)

Abstract: In order to investigate the function and evolution of *HSP70* gene in *Vernicia fordii*, the members of *HSP70* gene family were identified and analyzed by bioinformatics, including physical and chemical properties, evolutionary relationships, gene structure, conserved motifs and chromosomal localization. The expression differences of *HSP70* gene family members in different tissues were also analyzed. The results showed that the *HSP70* gene family of *Vernicia fordii* contained

收稿日期: 2024-02-05

基金项目: 省部共建木本油料资源利用国家重点实验室开放课题 (GZKF202103, GZKF202203); 安徽省高校优秀青年人才支持计划项目 (gxyq2020040); 安徽省省级大学生创新创业训练计划项目 (S202414098204, S202414098205); 国家级大学生创新创业训练计划项目 (202414098077)

作者简介: 王占军 (1984-), 男, 安徽寿县人, 博士研究生, 主要从事植物细胞分子生物学研究。(E-mail) wangzhanjunhxj@163.com。汪虹妍为共同第一作者。

通讯作者: 张颢, (E-mail) zhangxie3@163.com

24 members, and the theoretical isoelectric points of 24 *HSP70* proteins ranged from 4.57 to 8.70. Most of *HSP70* proteins were stable and hydrophilic proteins. Phylogenetic tree analysis showed that *HSP70* family members were divided into five clusters, and *Vernicia fordii* was closely related to *Populus trichocarpa*. The amino acid sequences of 24 *VfHSP70* proteins contained the conserved domain of *HSP70*. Twenty-two *HSP70* genes were distributed on 10 chromosomes. By analyzing the expression of *VfHSP70*

gene in different time and space, it was found that the expression levels of *VfHSP70-5* were higher in buds at 30 d before anthesis, 25 d before anthesis and 10 d before anthesis. *VfHSP70-4* was highly expressed in roots, stems, leaves and seeds of *Vernicia fordii*. These results suggest that *VfHSP70* gene may play an important role in regulating the growth and development of *Vernicia fordii*.

Key words: *Vernicia fordii*; *HSP70*; gene family; bioinformatics

油桐 (*Vernicia fordii*) 属多年生落叶乔木, 原产中国, 广泛分布于亚热带地区^[1], 为大戟科 (Euphorbiaceae) 油桐属植物, 是中国四大木本油料树种之一^[2]。油桐喜光, 耐阴性差, 畏寒怕冷, 适宜生长在温暖湿润的环境, 低温是限制油桐生长发育的重要环境因素。油桐为雌雄同株异花植物, 生长速度快, 单位面积产量高, 种子含油率高^[3], 而且抗逆性较强, 可以生长在贫瘠的土壤上, 因此适合作为造林树种。油桐品种资源丰富, 广泛分布在中国的 16 个省份和多个国家, 其中重庆、湖南、湖北等地的油桐栽培面积最大^[4]。油桐是中国研究最早最完善的油料树种, 中国有较为完整的油桐栽培技术, 能够实现油桐的早产、高产、优产^[5]。

油桐是重要的能源树种, 其经济价值很高, 在制造油漆、合成树脂、船舶制造、药品生产等方面均有广泛的应用^[6]。油桐种子的含油量高达 50% ~ 70%, 出油率为 50%, 油脂含量高于油菜等油料作物, 是植物油中最优质的干性油^[7]。油桐籽的油脂具有干燥速度快, 附着能力强, 防水性能好, 耐酸碱高温, 抗冻裂抗腐蚀等优点^[8]。油桐籽榨油后剩余的桐饼和桐麸是优良的有机肥料, 不仅能提高土壤肥力, 还能为农作物提供养分, 促进其生长发育, 提高产量; 并且桐饼和桐麸可供家畜食用^[9]。桐油还能用于制作生物柴油替代石油, 在国防工业、材料工业、生物质能源产业等多个领域均有很大的发展前景。

目前, 油桐的生物学研究主要包括已完成全基因组测序^[10], 油桐花发育与性别决定的深入探究^[11], 建立植物组织培养体系^[12]等。这些研究成果为进一步开展油桐遗传改良提供了有力支撑。油桐对于发展中国工业油料产业起着战略性作用, 加强油桐基础研究尤其是分子生物学和林业生物技术研究, 鉴定及克隆其生长发育关键基因, 并挖掘各基因关键生物学功能, 对于油桐产业发展有重要意义。

热激蛋白 (Heat shock protein, HSP) 最早见于地中海果蝇 (*Drosophila melanogaster*)^[13], 在动物、植

物和人体内广泛存在, 在原核生物和真核生物中也普遍存在^[14]。*HSP70* 基因家族成员作为分子伴侣^[15], 能维持细胞内稳态, 当细胞受到外部环境的不利刺激时, *HSP70* 基因会被激活并诱导表达。*HSP70* 蛋白的相对分子量为 $6.8 \times 10^4 \sim 7.8 \times 10^4$, 是一类高度序列保守的蛋白质家族^[14]。*HSP70* 蛋白包括位于 N 端的高度保守的腺苷三磷酸结合区 (AT-Pase domain) 和位于 C 端的底物结合结构域 (Substrate binding domain, SBD)^[16]。

在植物的生长发育过程中, *HSP70* 基因扮演着重要的角色, 同时 *HSP70* 蛋白在植物细胞响应生物胁迫 (如病原体) 与非生物胁迫 (如冷、热、干旱等) 过程中起到重要的作用; 例如, 在拟南芥 (*Arabidopsis thaliana*) 中, 将 *HSP70* 基因过表达、敲除或表达下调拟南芥表型无明显变化, 但双突变体 *hsp70-1 hsp70-4* 和三突变体 *hsp70-2 hsp70-4 hsp70-5* 表型改变, 且这两个突变体对高温、低温以及高糖、高盐等渗透胁迫敏感^[17]。研究表明, 经过 48 h 的 12 °C 低温锻炼后, 大戟科的麻疯树 (*Jatropha curcas*) 幼苗在 1 °C 强冷胁迫下的耐冷性显著增强^[18-19]。Wang 等^[20]对其转录组数据分析表明, 在低温锻炼阶段, *HSP70* 是对 12 °C 低温表现出高响应的基因。另外 Wimmer 等^[21]发现, 在番茄 (*Solanum lycopersicum*) 和菠菜 (*Spinacia oleracea*) 中, 冷诱导时 *HSP70* 在 12 h 和 48 h 时出现表达峰值。

拟南芥^[22]、水稻 (*Oryza sativa*)^[23]、毛果杨 (*Populus trichocarpa*)^[24] 等植物的 *HSP70* 基因家族成员已经完成鉴定, 其中拟南芥中共鉴定出 18 个 *HSP70* 基因, 水稻中共鉴定出 32 个 *HSP70* 基因, 毛果杨中鉴定出 34 个 *HSP70* 基因。近年来, 有关油桐的全基因组测序也已完成^[10], 且有关生物胁迫下油桐基因家族的表达模式已有报道^[25], 但目前关于油桐抗寒性基因家族的研究却鲜有报道^[26], 因此本研究基于油桐全基因组对 *HSP70* 基因家族进行全面分析, 揭示油桐生长发育和非生物胁迫响应的分子基础, 为进一步培育或改良油桐抗寒品种提供理

论依据。

1 材料与方法

1.1 序列数据下载

本研究所用的油桐全基因组数据由中南林业科技大学经济林培育与保护教育部重点实验室提供 (<https://bigd.big.ac.cn/gsa/>)^[27], 拟南芥基因组^[22]和毛果杨基因组^[24]来源于 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/>), 水稻基因组^[23]来源于 Phytosome 数据库 (<https://phytosome-next.jgi.doe.gov/>)。

1.2 油桐 HSP70 基因家族成员的鉴定

为筛选出油桐 HSP70 候选基因序列, 在 NCBI 数据库中下载拟南芥已鉴定出的 18 条 HSP70 蛋白氨基酸序列, 利用 TBtools 软件^[28]将其与油桐基因组蛋白质氨基酸序列进行同源比对 (E 值 $< 1 \times 10^{-10}$)^[29], 获取油桐 HSP70 基因家族成员编码的蛋白质氨基酸序列, 共检测出 24 条待定的 HSP70 基因序列, 同时通过 BLASTP 同源搜索, 从 Pfam 数据库 (<http://pfam.xfam.org/>) 中下载 HSP70 蛋白结构域 (PF00010) 的隐马尔科夫模型 (HMM) 文件^[30], 使用 HMMER v3.0 软件查找油桐中 VjHSP70 蛋白氨基酸序列。使用 SMART 工具 (<http://smart.embl-heidelberg.de/>) 进行结构域筛选, 确定这 24 个 HSP70 蛋白氨基酸序列均含有 HSP70 结构域, 以此确认油桐有 24 条 HSP70 蛋白氨基酸序列。

1.3 油桐 HSP70 家族蛋白质的理化性质分析

利用在线工具 ExPASy (<http://web.expasy.org/protparam/>)^[31]分析油桐 HSP70 蛋白的理化性质, 包括蛋白质长度 (Predicted protein length)、相对分子量 (Relative molecular weight, RMW)、等电点 (pI)、不稳定性系数 (Instability index, II)、亲水性指数 (Grand average of hydropathicity, GRAVY) 等。借助在线工具 WOLF PSOR (<https://wolfsort.hgc.jp/>)^[32]进行亚细胞定位分析。

1.4 HSP70 家族的系统进化树构建

从数据库中下载拟南芥、水稻、毛果杨的 HSP70 蛋白氨基酸序列构建种间进化树, 根据文献^[33]下载的喙叶泥炭藓 (*Sphagnum recurvum*) 和北美乔柏 (*Thuja plicata*) HSP70 蛋白氨基酸序列均来自 Phytosome (<https://phytosome-next.jgi.doe.gov/>) 数据库。利用 MAGE 11^[34]软件中的 Clustal W 对拟南芥、毛

杨、水稻、油桐、喙叶泥炭藓和北美乔柏的 HSP70 蛋白氨基酸序列进行多序列比对, 对比结果采用最大似然法 (Maximum likelihood) 构建系统进化树, Bootstrap 值设置为 1 000, 对其结果使用在线软件 Evolview 与 Adobe Illustrator 进行美化。同时结合 BLASTP 与进化树对油桐的 24 个 HSP70 基因进行命名。

1.5 油桐 HSP70 基因家族的基因结构分析与保守基序分析

使用 GSDS 2.0 在线工具 (<http://gsds.cbi.pku.edu.cn/index.php>)^[35]绘制油桐 HSP70 基因家族的基因结构图。采用在线工具 MEME (<http://meme-suite.org/tools/meme>)^[36]进行 HSP70 基因家族成员保守基序 (Motif) 分析, 基序最大数设为 10, 其他参数使用网站默认值, 获得 xml 格式文件, 再在 NCBI 数据库中下载 24 条油桐 HSP70 蛋白的 hitdata 文件, 利用准备的 meme.xml 文件、hitdata 文件、油桐的进化树 Newick Export 文件和 24 条油桐 HSP70 蛋白氨基酸序列, 通过 TBtools 软件分析油桐 HSP70 家族蛋白质保守序列和蛋白质结构域。

1.6 油桐 HSP70 基因家族成员的染色体定位

使用油桐基因组 GFF 注释文件, 对储存基因的 ID 文件进行提取, 获取的基因 ID 应与 GFF 注释文件相对应。通过 TBtools 软件定位油桐 HSP70 基因, 获取其染色体长度及位置信息。

1.7 油桐 HSP70 基因家族的表达分析

为探究油桐中 HSP70 基因的表达情况, 利用来自中南林业科技大学油桐研究团队的油桐根、茎、叶、种子 (开花后 10 周、开花后 15 周、开花后 20 周、开花后 25 周、开花后 30 周)、花蕾 (开花前 30 d、开花前 25 d、开花前 10 d、开花前 1 d) 的基因表达数据^[37], 使用 TBtools 进行作图分析, 让试验结果可视化。

2 结果与分析

2.1 油桐 HSP70 基因家族鉴定结果

根据参考文献中的基因序列号在数据库中下载拟南芥的 HSP70 蛋白氨基酸序列为比对序列, 运用 BLASTP 对油桐基因组数据库进行比对, 检索得到 24 个油桐 HSP70 候选基因。使用在线 SMART 数据库检验这 24 个 HSP70 蛋白氨基酸序列, 结果显示均含有 HSP70 结构域, 结合多个验证结果综合考量, 24 个油桐 HSP70 基因家族成员最终被确定。依据 24 个 VjHSP70 基因在 NCBI 数据库 BLASTP 中所得结果和系统进化

树综合分析,为 24 个 VfHSP70 基因命名。

2.2 油桐 HSP70 家族成员的理化性质

24 个 HSP70 蛋白的理化性质如表 1 所示,蛋白质氨基酸序列长度为 88~2 008 aa,其中 VfHSP70-10 蛋白最长,为 2 008 aa, VfHSP70-34 蛋白最短,为 88 aa; VfHSP70-15 的相对分子量最大,为 9.872×10^4 , VfHSP70-20 的相对分子量最小,为 1.374×10^4 ;理论等电点为 4.57~8.70,其中 VfHSP70-1 的等电点是 8.70,为碱性蛋白,其余 23 个 HSP70 蛋白的等电点均在 7 以下,为酸性蛋白。亲水性指数是用来衡量蛋白质亲疏水性的大小,亲水性指数越大,蛋白质的疏水性就越强,HSP70 蛋白亲水性指数为 -1.069~0.095, VfHSP70-8 亲水性指数为 0.095, VfHSP70-34

亲水性指数为 0.066,其余 HSP70 蛋白亲水性指数均<0,为负值。从亲水性来看,大多数 HSP70 家族成员亲水性多数为负值,可见油桐 HSP70 蛋白大部分属于亲水性蛋白。不稳定系数是用来衡量蛋白质在体外的稳定性,不稳定系数<40 的为稳定蛋白。结果显示,HSP70 家族成员不稳定系数为 18.36~50.66,其中不稳定系数>40 的有 9 个 HSP70 蛋白。其中 10 个油桐 HSP70 蛋白定位于细胞质,4 个油桐 HSP70 蛋白定位于叶绿体,3 个油桐 HSP70 蛋白定位于细胞核,3 个油桐 HSP70 蛋白定位于线粒体,3 个油桐 HSP70 蛋白定位于内质网,1 个油桐 HSP70 蛋白定位于细胞外。不同的亚细胞定位意味着功能区别和系统发育分化。

表 1 油桐 HSP70 蛋白家族成员基本信息

Table 1 Basic information of HSP70 protein family members of *Vernicia fordii*

蛋白质	基因 ID	氨基酸长度(aa)	相对分子量	等电点	亲水性指数	不稳定系数	亚细胞定位
VfHSP70-1	Vf03G2087	182	2.102×10^4	8.70	-0.672	39.49	细胞质
VfHSP70-2	Vf03G2089	209	2.294×10^4	4.57	-0.568	41.35	细胞质
VfHSP70-3	Vf02G0481	233	2.515×10^4	4.67	-0.571	38.15	细胞质
VfHSP70-4	Vf03G2096	203	2.235×10^4	4.79	-0.941	41.85	细胞核
VfHSP70-5	Vf05G2015	665	7.334×10^4	5.13	-0.449	26.76	内质网
VfHSP70-6	Vf03G0479	1 858	2.020×10^4	6.85	-0.109	36.97	叶绿体
VfHSP70-7	Vf09G0690	282	3.039×10^4	4.74	-0.345	18.36	细胞质
VfHSP70-8	Vf06G0942	514	5.587×10^4	5.73	0.095	33.89	细胞质
VfHSP70-9	Vf01G0036	679	7.321×10^4	5.74	-0.323	37.83	线粒体
VfHSP70-10	Vf03G0752	2 008	2.224×10^4	5.80	-0.336	42.31	叶绿体
VfHSP70-13	Vf02G2688	493	5.542×10^4	5.09	-0.408	40.96	细胞质
VfHSP70-14	Vf02G2178	851	9.390×10^4	5.14	-0.446	44.93	细胞质
VfHSP70-15	Vf04G2080	889	9.872×10^4	5.33	-0.451	39.20	内质网
VfHSP70-16	Vf11G0994	771	8.608×10^4	5.56	-0.399	42.47	细胞核
VfHSP70-18	Vf03G2085	237	2.652×10^4	5.54	-0.206	48.06	细胞质
VfHSP70-19	Vf02G2681	163	1.862×10^4	6.43	-1.069	38.52	叶绿体
VfHSP70-20	Vf07G2428	1 217	1.374×10^4	6.22	-0.505	50.66	细胞核
VfHSP70-23	Vf02G1850	133	1.465×10^4	4.66	-0.887	35.63	线粒体
VfHSP70-24	Vf02G1851	198	2.152×10^4	4.73	-0.688	32.59	细胞质
VfHSP70-25	Vf08G0248	124	1.378×10^4	4.57	-0.909	23.80	线粒体
VfHSP70-26	Vf00G1966	208	2.276×10^4	4.76	-0.657	38.09	细胞质
VfHSP70-28	Vf10G1894	666	7.339×10^4	5.26	-0.417	30.78	内质网
VfHSP70-29	Vf07G2429	662	7.300×10^4	5.15	-0.446	29.27	细胞外
VfHSP70-34	Vf09G0687	88	9.592×10^4	5.02	0.066	45.26	叶绿体

2.3 HSP70 家族进化关系

利用最大似然法构建 6 个物种的 HSP70 系统发育树。如图 1 所示,其中包括 17 个拟南芥 HSP70 蛋白氨基酸序列,34 个毛果杨 HSP70 蛋白氨基酸序列,16 个

北美乔柏 HSP70 蛋白氨基酸序列,24 个水稻 HSP70 蛋白氨基酸序列和 1 个喙叶泥炭藓 HSP70 蛋白氨基酸序列。根据同源关系将 116 个 HSP70 蛋白分为 5 个亚族(I~V),每个亚族都含有油桐 HSP70 蛋白。I 亚族占比

最高,共 59 个 HSP70 蛋白,包括 6 个拟南芥 HSP70 蛋白,14 个毛果杨 HSP70 蛋白,13 个水稻 HSP70 蛋白,10 个油桐 HSP70 蛋白。Ⅱ亚族有 8 个 HSP70 蛋白,包括 1 个拟南芥 HSP70 蛋白,3 个毛果杨 HSP70 蛋白,3 个油桐 HSP70 蛋白。Ⅲ亚族含有 17 个 HSP70 蛋白,Ⅳ亚族

含有 12 个 HSP70 蛋白,Ⅴ亚族含有 20 个 HSP70 蛋白,其中北美乔柏 HSP70 蛋白聚集于Ⅰ亚族,喙叶泥炭藓 HSP70 蛋白聚集在Ⅱ亚族,生物进化距离表明,油桐与喙叶泥炭藓的亲缘关系较远。

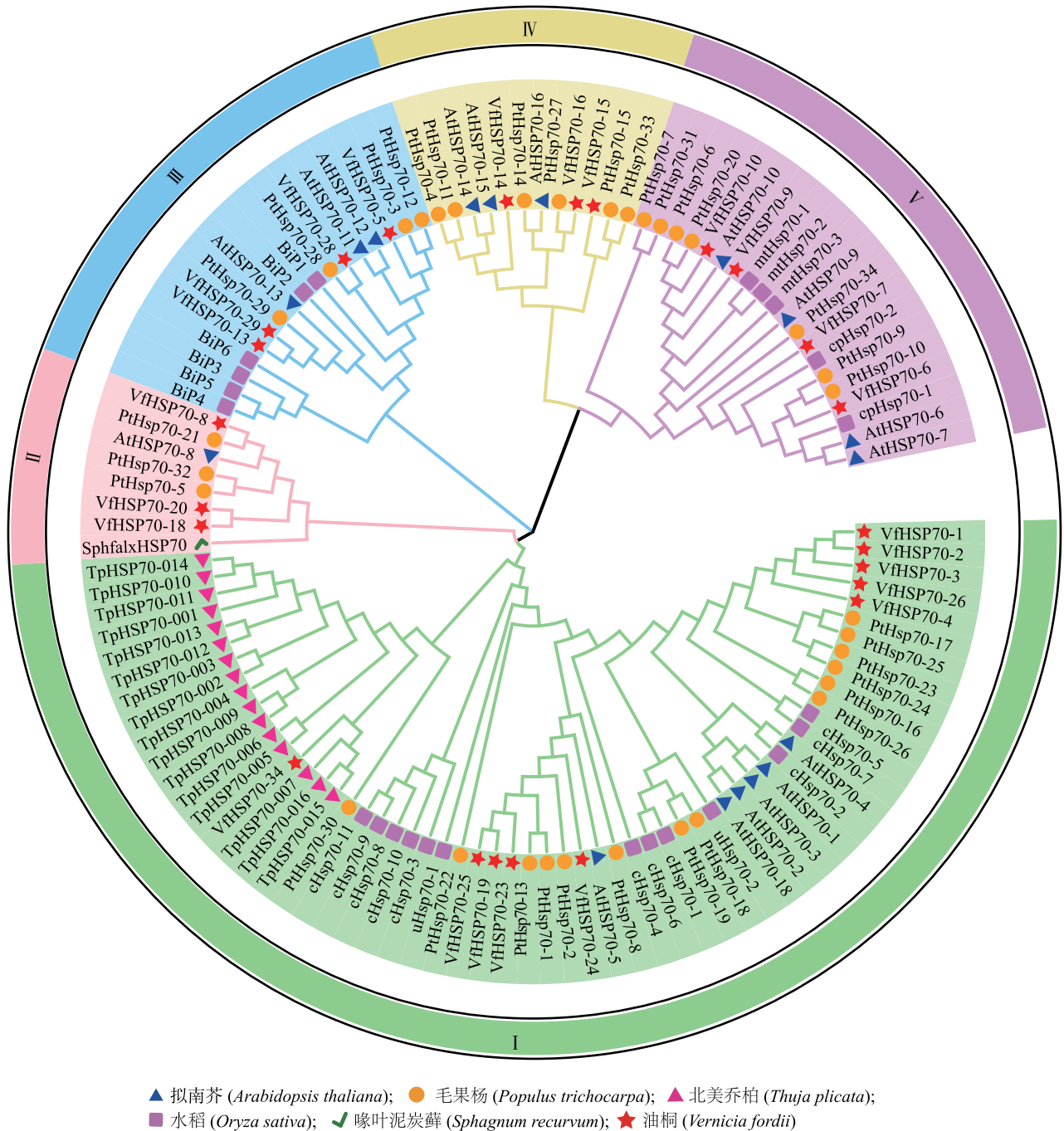


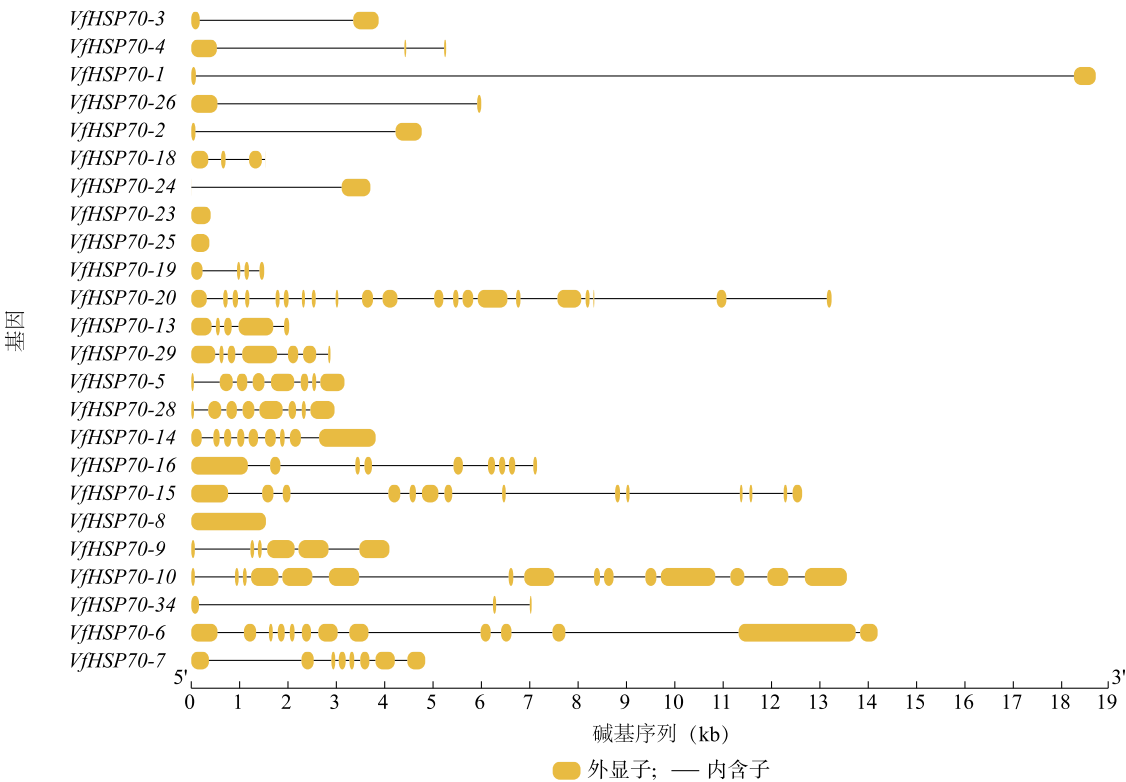
图 1 油桐、拟南芥、毛果杨、北美乔柏、水稻和喙叶泥炭藓中 HSP70 家族系统发育进化树

Fig.1 Phylogenetic tree of HSP70 family in *Vernicia fordii*, *Arabidopsis thaliana*, *Populus trichocarpa*, *Thuja plicata*, *Oryza sativa* and *Sphagnum recurvum*

2.4 油桐 *HSP70* 基因家族的基因结构

根据油桐 *HSP70* 基因注释文件,利用 GSDS 2.0 在线分析工具,获得油桐 *HSP70* 基因家族成员基因结构。如图 2 所示, *VfHSP70-8*、*VfHSP70-23* 与 *VfHSP70-25* 仅含有外显子,其他 21 个 *HSP70* 基因都含有内含子和外显子。在相同的进化树亚族中的 *VfHSP70* 基因具有相似的外显子-内含子分布模式,如

VfHSP70-5 与 *VfHSP70-28*、*VfHSP70-23* 与 *VfHSP70-25* 外显子-内含子分布模式相似。总体来说,不同亚族 *VfHSP70* 基因家族成员的内含子和外显子数量差异较大,且内含子位置和长度也存在很大差异,但同一亚族成员大部分基因具有相似的外显子-内含子分布模式。



黄色的矩形表示外显子结构,黑色线条表示内含子。

图 2 油桐 *HSP70* 基因家族成员基因结构

Fig.2 Gene structure of *HSP70* gene family members in *Vernicia fordii*

2.5 油桐 *HSP70* 家族的保守基序

为了解 *HSP70* 蛋白的结构特征,使用 MEME 在线工具分析油桐 *HSP70* 家族成员蛋白氨基酸序列的保守基序,最大 Motif 检测数设置为 10。由图 3 可知,有 5 个 *VfHSP70* 蛋白氨基酸序列包含全部 10 个保守基序,分别为 *VfHSP70-5*、*VfHSP70-6*、*VfHSP70-10*、*VfHSP70-28*、*VfHSP70-29*。同一进化分支上的蛋白质具有相同的保守序列,如 *VfHSP70-9* 与 *VfHSP70-10*,*VfHSP70-5* 与 *VfHSP70-28* 等。*VfHSP70-34* 仅含有 1 个保守基序。其中 17 个 *VfHSP70* 蛋白包含 Motif 1,15 个 *VfHSP70* 蛋白包含 Motif 2,Motif 1 和 Motif 2 可能在 *HSP70* 蛋白的进化中表现出特异性。各个 *HSP70*

蛋白的 Motif 数量以及分布上的不同可能会导致 *HSP70* 蛋白功能上的差异,保守基序种类及分布可能与蛋白质功能多样性相关。基序结构域分析结果表明,24 个蛋白质氨基酸序列共含有 13 个结构域,其中都含有 *HSP70* 结构域。结构域的分布存在差异,但是在同一亚群的家族成员具有类似的结构域,如 *VfHSP70-5* 与 *VfHSP70-28*。

2.6 油桐 *HSP70* 基因家族成员的染色体定位

为了明确油桐 *HSP70* 基因在染色体上的分布情况,使用 TBtools 软件对 24 个油桐 *HSP70* 基因家族成员进行染色体定位分析。在染色体 1 上有 6 个基因,分别为 *VfHSP70-3*、*VfHSP70-23*、*VfHSP70-24*、

VfHSP70-14、*VfHSP70-19*、*VfHSP70-13*。染色体 2 上也有 6 个 *VfHSP70* 基因,分别为 *VfHSP70-6*、*VfHSP70-10*、*VfHSP70-18*、*VfHSP70-1*、*VfHSP70-2*、*VfHSP70-4*。染色体 6 上有 2 个 *VfHSP70* 基因,分别为 *VfHSP70-20*、*VfHSP70-29*,染色体 8 有 2 个 *VfHSP70* 基因,分别为 *VfHSP70-34*、*VfHSP70-7*。*VfHSP70-15*

基因分布在染色体 3 上,*VfHSP70-5* 分布在染色体 4 上,*VfHSP70-8* 分布在染色体 5 上,*VfHSP70-25* 基因分布在染色体 7 上,*VfHSP70-28* 基因分布在染色体 9 上,*VfHSP70-16* 基因分布在染色体 10 上。*VfHSP70-26* 基因和 *VfHSP70-9* 未定位到染色体上。

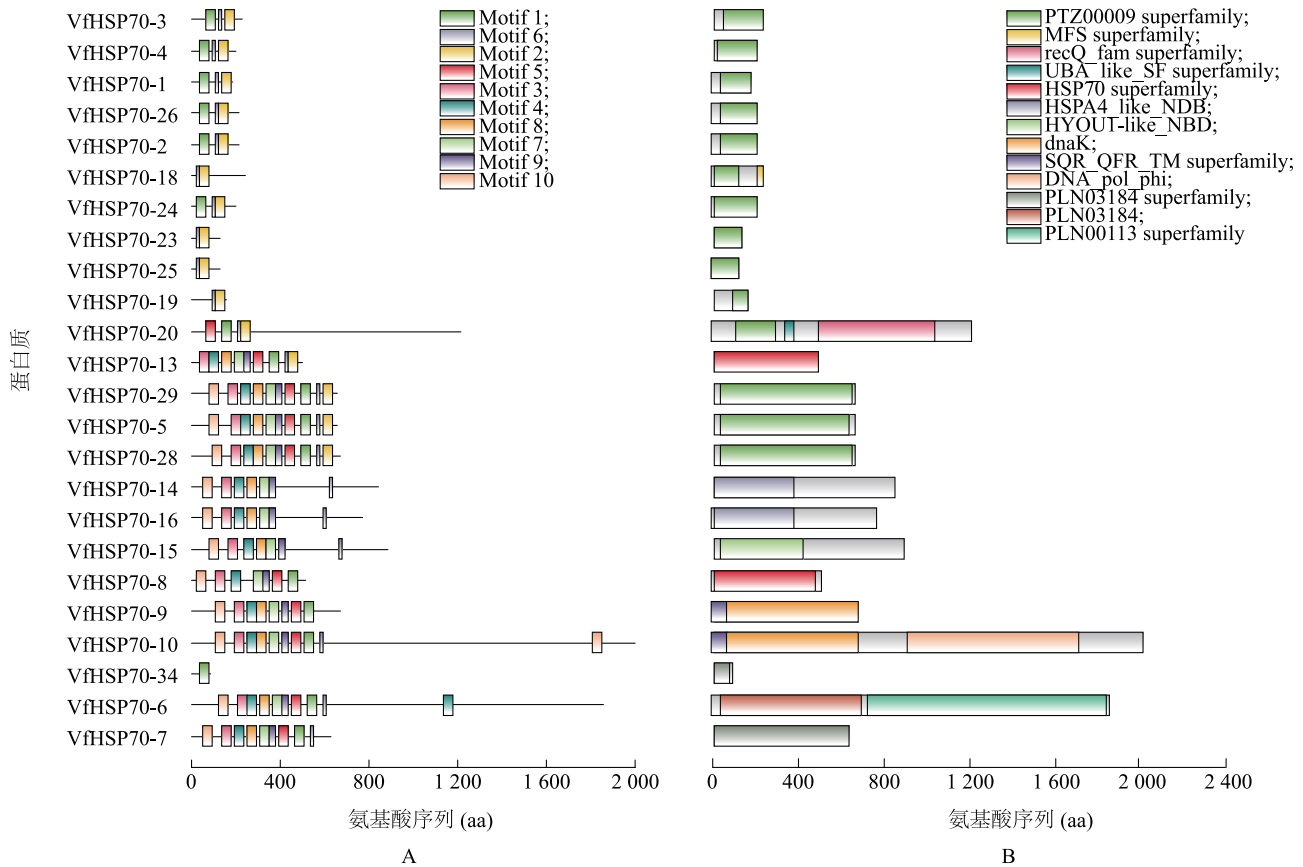


图3 油桐 HSP70 蛋白质保守序列 (A) 和蛋白质结构域 (B)

Fig.3 The conserved motifs (A) and domains (B) of HSP70 in *Vernicia fordii*

2.7 油桐 HSP70 基因家族的表达特性

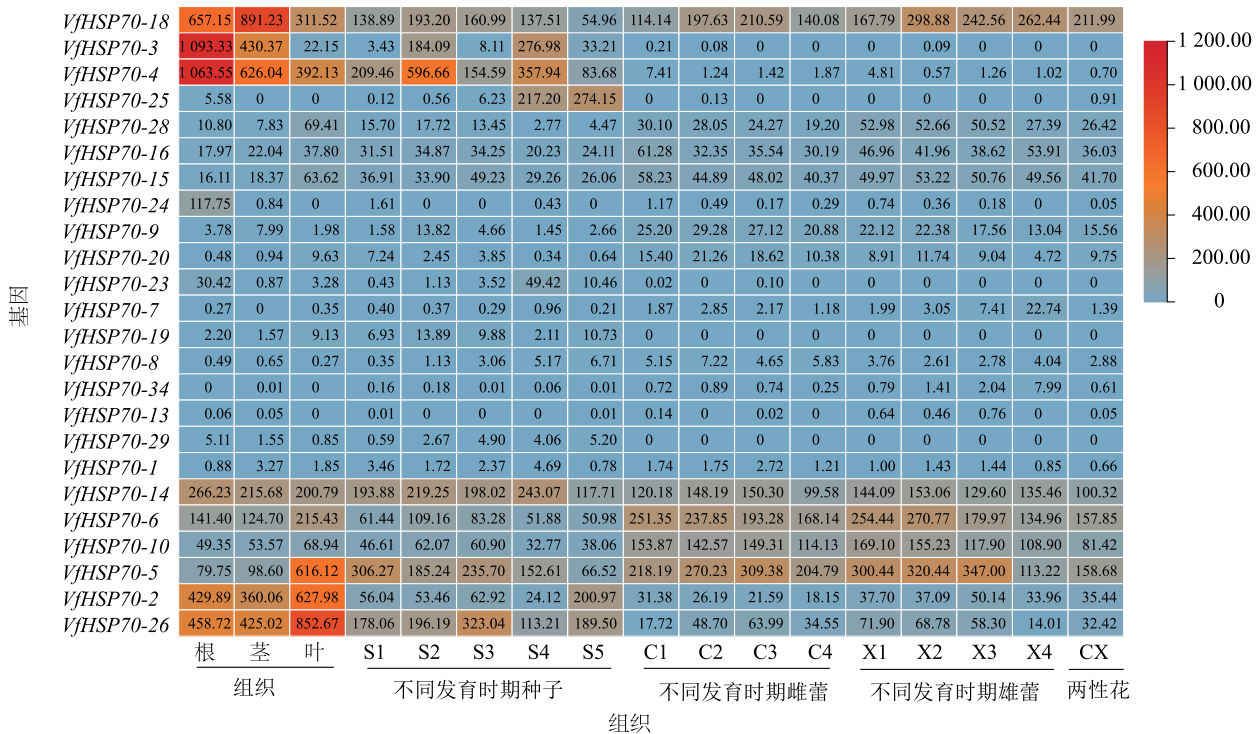
收集油桐 HSP70 基因在不同组织中的表达数据,利用转录组数据获取 FPKM 值,使用 TBtools 制作热图使数据可视化,对 *VfHSP70* 基因家族成员在根、茎、叶、种子和花中的表达量进行研究。

对于 *VfHSP70* 基因家族成员在根、茎、叶中的表达,结果如图 4 所示,*VfHSP70-3* 与 *VfHSP70-4* 在根中的表达量较高,在茎和叶中的表达量明显低于根,*VfHSP70-18* 在茎中表达量较高;*VfHSP70-26* 在叶中表达量较高,说明这些基因对油桐营养组织的生长发育过程起重要作用,且同一个基因在不同的组织中发挥不同的功能。关于 *VfHSP70* 在油桐种子中的

表达模式,分析了以下 5 个具有代表性的种子发育阶段:开花后 10 周 (S1)、开花后 15 周 (S2)、开花后 20 周 (S3)、开花后 25 周 (S4)、开花后 30 周 (S5)。在开花后 10 周,油桐种子中 *VfHSP70-5* 表达量较高;在开花后 15 周,油桐种子中 *VfHSP70-4* 的表达量较高;在开花后 20 周,油桐种子中 *VfHSP70-26* 表达量较高;开花后 25 周,油桐种子中 *VfHSP70-4* 表达量较高;在开花后 30 周,油桐种子中 *VfHSP70-25* 表达量较高。这些较高表达的基因在油桐种子发育过程中发挥重要的作用,不同发育时期种子的高表达基因也有所不同。Feng 等^[38] 将在某一阶段表达量是其余阶段表达量 2 倍以上的 *VfHSP70* 基因定义

为“阶段特异性”基因,由图 4 可见,*VfHSP70-2* 基因在开花后 30 周是“阶段特异性”基因。在油桐发育

阶段,这些特异性表达的基因起着重要的作用。



S1:开花后 10 周的种子;S2:开花后 15 周的种子;S3:开花后 20 周的种子;S4:开花后 25 周的种子;S5 开花后 30 周的种子;C1:开花前 30 d 的雌蕾;C2:开花前 25 d 的雌蕾;C3:开花前 10 d 的雌蕾;C4:开花前 1 d 的雌蕾;X1:开花前 30 d 的雄蕾;X2:开花前 25 d 的雄蕾;X3:开花前 10 d 的雄蕾;X4:开花前 1 d 的雄蕾;CX:两性花。

图 4 油桐 HSP70 基因家族成员在根、茎、叶、花、种子组织中的表达量

Fig.4 Expression levels of HSP70 gene family members in roots, stems, leaves, flowers, and seeds of *Vernicia fordii*

对 *VfHSP70* 基因在花中表达量进行可视化,分析了以下 4 个具有代表性的雌蕾和雄蕾发育阶段:开花前 30 d 的雌蕾(C1)和雄蕾(X1);开花前 25 d 的雌蕾(C2)和雄蕾(X2);开花前 10 d 的雌蕾(C3)和雄蕾(X3);开花前 1 d 的雌蕾(C4)和雄蕾(X4)。在 4 个阶段中,*VfHSP70-3*、*VfHSP70-13*、*VfHSP70-19*、*VfHSP70-23*、*VfHSP70-25* 和 *VfHSP70-29* 基因 $FPKM < 0.2$,几乎不表达;*VfHSP70-5*、*VfHSP70-6*、*VfHSP70-10*、*VfHSP70-14*、*VfHSP70-18* 基因在 4 个阶段表达量较高,*VfHSP70-5* 在开花前 30 d 到开花前 10 d 均有较高表达量;在两性花发育阶段,*VfHSP70-18* 有较高表达量。这些结果表明,不同的基因可能阶段性参与油桐花的发育。

油桐 HSP70 基因家族成员在根、茎、叶、种子、花 5 种组织中各有不同的表达模式,*VfHSP70-4*、*VfHSP70-5*、*VfHSP70-6*、*VfHSP70-14*、*VfHSP70-18*、*VfH-*

SP70-26 在 5 种组织中均有较高表达,表明这些基因在油桐的整个生长发育中发挥着较大作用,且不同的基因在这 5 个组织中的表达具有特异性;*VfHSP70-4* 在根、茎、叶、种子均有较高表达,说明 *VfHSP70-4* 在油桐根、茎、叶和种子的生长和发育过程中可能起着重要作用;*VfHSP70-5* 从开花前 30 d 至开花前 10 d 均表现出较高表达水平。因此,这些基因可能在油桐繁殖阶段起着重要调控作用。

3 讨论

油桐是一种经济价值较高的能源树种,能够提供十分优质的生物质能源,如干性油料。油桐属典型的中亚热带树种,为阳性树种,喜温、喜光、喜水,但又不耐水湿,适应暖湿气候,增强油桐的抗逆性将对油桐的种植范围与生产产量有很大的帮助。HSP70 是一种普遍存在的分子伴侣,在植物生长发

育过程中发挥重要作用,植物中 *HSP70* 基因家族成员广泛参与植物的逆境胁迫反应和代谢调控。目前,*HSP70* 基因的生物学功能已在许多植物中得到了研究,例如拟南芥、水稻、毛果杨和小麦 (*Triticum aestivum*)^[39],但油桐的 *HSP70* 基因鲜有报道。因此本研究对油桐 *HSP70* 基因家族进行了全面的全基因组分析,以期为今后油桐 *HSP70* 基因功能研究提供基础。

HSP70 是一种重要的热休克蛋白^[40-41],在植物的细胞、组织、器官以及整个生命周期中发挥着广泛的作用。近年来,随着 *HSP70* 基因家族成员的深入研究,人们对 *HSP70* 基因在植物油脂代谢中的作用有了更深入的理解。然而,关于油桐油脂代谢中 *HSP70* 基因的具体作用仍待探秘。研究发现,随着环境温度的升高,拟南芥萼片与花朵早上随之开放,然而对于其潜在的分子机制知之甚少^[42]。轻度热应激会使脂质代谢减慢,特别是表皮脂质代谢减慢,萼片中的相关基因表达量变高,*HSP70-16* 的突变同样使脂质代谢减慢,萼片中的相关基因表达量变高^[43-44]。另外,研究发现大豆 (*Glycine max*) 中的一些 miRNA 通过靶向负调控 *HSP70* 基因的表达,从而对大豆籽粒的蛋白质合成起到促进作用。磷酸烯醇式丙酮酸羧化酶 (*PEPC*) 可能在大豆籽粒成熟过程中油脂和蛋白的转化中扮演着重要角色^[45]。1,5 二磷酸核酮糖羧化酶/加氧酶有 4 个大亚基 (*RbcL*) 和 4 个小亚基 (*RbcS*),*RbcS2* 基因表达量高且相对稳定,被用于微藻重组蛋白表达载体的构建。*HSP70-RbcS2* 嵌合启动子是促进莱茵衣藻 (*Chlamydomonas reinhardtii*) 外源基因表达的较好启动子^[46]。含有 *HSP70* 和 *RBCS* 双启动子的蛋白质表达载体可以降低小球藻的转基因沉默效应^[47]。随着研究的不断深入,期待 *HSP70* 基因在油桐油脂代谢中的作用机制能被明确阐释。这些研究结果不仅有助于人们理解植物在环境变化下的适应机制,也为改良油桐品种提供了新思路。

4 结 论

本研究从油桐全基因组数据中鉴定出 *HSP70* 家族基因,分析其理化性质、系统发育关系、染色体定位、基因结构、保守基序及表达特性等。与其他植物相比,油桐的 *HSP70* 基因家族成员数量适中(拟南芥 18 个,水稻 32 个,大豆 61 个^[48],北美乔柏 16

个,喙叶泥炭藓 1 个),油桐的 *HSP70* 蛋白等电点 (4.57~8.70) 与大戟科的麻疯树 (5.06~6.06) 相近,油桐 *HSP70* 蛋白主要为酸性的、亲水性的、稳定的蛋白质。有研究表明,*HSP70* 蛋白的亲水性能增强植物对非生物胁迫的抵御能力^[49]。*HSP70* 基因编码的蛋白质多定位于细胞质、叶绿体、细胞核和线粒体中,表明 *HSP70* 在油桐的整个生命活动中都发挥作用,其中较多的 *HSP70* 基因编码的蛋白质定位于细胞质中,这可能与它所承担的功能有关,即位于胞质的 *HSP70* 成员在信息传递到细胞核的过程中起着至关重要的作用,从而对基因在细胞核内的表达进行调节。由油桐系统发育进化树、基因结构与保守基序结果综合分析来看,同一聚类的蛋白质常具有相同或相似的结构,该结果与拟南芥、水稻、毛果杨、大豆的研究结果一致,可能是由于在进化过程中油桐基因组的串联重复和随机重复,串联重复和随机重复在基因家族的进化和扩张中发挥重要作用^[50]。其物种间进化树分析结果表明,有 5 对直系同源基因,其中 3 对是毛果杨与油桐,说明油桐与毛果杨的亲缘关系较为接近。基因结构分析结果表明,属同一进化分支的家族成员,具有明显相同或相似的内含子-外显子数量结构特征;油桐基序分析结果表明,其中 5 个 *VfHSP70* 蛋白氨基酸序列包含全部 10 个保守基序,其余 19 个 *VfHSP70* 蛋白氨基酸序列均缺少部分 Motif。以上结果表明,*VfHSP70* 家族不同聚类在进化关系、基因结构、保守基序上均具有明显的差异性。*VfHSP70* 基因在油桐不同组织和器官中表达,这些结果表明它们可能参与油桐生长发育的各个过程,其中 *VfHSP70-5* 在开花前 30 d 到开花前 10 d 的花蕾中均有较高表达量,推测可能促进油桐花的发育;*VfHSP70-4* 在根、茎、叶、种子中均有较高表达量,推测可能与油桐的生长发育有较强关联。

低温是影响油桐生长发育的重要环境因子,本研究对油桐 *HSP70* 基因家族成员进行鉴定和生物信息学分析,系统揭示了油桐 *HSP70* 基因的理化性质、进化特点和表达特性等,为将来利用分子手段培育抗寒油桐新品种提供了理论基础。

致谢: 感谢合肥师范学院生物与食品工程学院陶瑞松副教授在本文进化树研究工作中的支持与帮助!

参考文献:

- [1] 赵 斌. 油桐的嫁接繁殖技术[J]. 种子, 2005(8): 118-120.
- [2] 谭晓风. 油桐的生产现状及其发展建议[J]. 经济林研究, 2006, 24(3): 62-64.
- [3] 贺舍予, 钟海雁, 金 超, 等. 桐油主要脂肪酸定量分析方法的构建及其应用[J]. 中南林业科技大学学报, 2014, 34(5): 83-87.
- [4] 张玲玲, 彭俊华. 油桐资源价值及其开发利用前景[J]. 经济林研究, 2011, 29(2): 130-136.
- [5] 潘 登. 两步法催化桐油制备生物柴油工艺及动力学研究[D]. 重庆: 重庆大学, 2009.
- [6] 李永梅, 魏远新, 周大林, 等. 油桐的价值及其发展途径[J]. 现代农业科技, 2008(16): 113.
- [7] 谭晓风, 蒋桂雄, 谭方友, 等. 我国油桐产业化发展战略调查研究报告[J]. 经济林研究, 2011, 29(3): 1-7.
- [8] 张庆伟, 温智婷, 罗克明. 油桐生物学研究进展[J]. 经济林研究, 2020, 38(4): 238-245.
- [9] 黄健生. 油桐发展前景分析及丰产栽培技术[J]. 安徽农学通报, 2018, 24(24): 71-72, 95.
- [10] 周 静, 韩杰铨, 涂洲溢, 等. 我国油桐籽的生产及发展对策[J]. 中国油脂, 2021, 46(4): 118-122.
- [11] 毛颖基. 油桐花发育及其性别决定机制研究[D]. 合肥: 中国科学技术大学, 2017.
- [12] 冯邦朝, 黄 艳, 马 博, 等. 大戟科主要木本油料植物组织培养快繁技术研究进展[J]. 南方农业学报, 2012, 43(11): 1650-1654.
- [13] RITOSSA F. A new puffing pattern induced by temperature shock and DNP in drosophila[J]. Experientia, 1962, 18(12): 571-573.
- [14] VIERLING E. The roles of heat shock proteins in plants[J]. Annual Review of Plant Biology, 1991, 42: 579-620.
- [15] MIEMYK J. The 70 kDa stress-related proteins as molecular chaperones[J]. Trends in Plant Science, 1997, 2(5): 180-187.
- [16] DRAGOVIC Z, BROADLEY S A, SHOMURA Y, et al. Molecular chaperones of the *Hsp110* family act as nucleotide exchange factors of *Hsp70s*[J]. EMBO Journal, 2006, 25(11): 2519-2528.
- [17] LENG L, LIANG Q, JIANG J, et al. A subclass of *HSP70s* regulate development and abiotic stress responses in *Arabidopsis thaliana*[J]. Journal of Plant Research, 2017, 130(2): 349-363.
- [18] AO P X, LI Z G, FAN D M, et al. Involvement of antioxidant defense system in chill hardening-induced chilling tolerance in *Jatropha curcas* seedlings[J]. Acta Physiol Plant, 2013, 35: 153-160.
- [19] AO P X, LI Z G, GONG M. Involvement of compatible solutes in chill hardening-induced chilling tolerance in *Jatropha curcas* seedlings[J]. Acta Physiol Plant, 2013, 35: 3457-3464.
- [20] WANG H, ZOU Z, WANG S, et al. Global analysis of transcriptome responses and gene expression profiles to cold stress of *Jatropha curcas* L.[J]. PLoS One, 2013, 8(12): e82817.
- [21] WIMMER B, LOTTSPEICH F, VAN DER KLEI I, et al. The glyoxysomal and plastid molecular chaperones (70-kDa heat shock protein) of watermelon cotyledons are encoded by a single gene[J]. PNAS, 1997, 94(25): 13624-13629.
- [22] LIN B, WANG J, LIU H, et al. Genomic analysis of the *HSP70* superfamily in *Arabidopsis thaliana*[J]. Cell Stress Chaperones, 2001, 6(3): 201-208.
- [23] SARKAR N K, KUNDNANI P, GROVER A. Functional analysis of *HSP70* superfamily proteins of rice (*Oryza sativa*) [J]. Cell Stress Chaperones, 2013, 18: 427-437.
- [24] YER E N, BALOGLU M C, ZIPLAR U T, et al. Drought-responsive *HSP70* gene analysis in populus at genome-wide level[J]. Plant Molecular Biology Reporter, 2016, 34(2): 483-500.
- [25] 江慎秀, 王静雅, 庄婧怡, 等. 油桐 *NPR1* 家族全基因组鉴定及表达模式分析[J]. 植物遗传资源学报, 2021, 22(2): 521-531.
- [26] 张庆伟, 温智婷, 罗克明. 油桐生物学研究进展[J]. 经济林研究, 2020, 38(4): 238-245.
- [27] ZHANG L, LIU M, LONG H, et al. Tung tree (*Vernicia fordii*) genome provides a resource for understanding genome evolution and improved oil production[J]. Genomics Proteomics Bioinformatics, 2019, 17(6): 558-575.
- [28] CHEN J, CHEN H, ZHANG Y, et al. TBtools an integrative toolkit developed for interactive analyses of big biological data[J]. Molecular Plant, 2020, 13(8): 1194-1202.
- [29] GUAN Y, LIU S, WU W, et al. Genome-wide identification and cold stress-induced expression analysis of the *CBF* gene family in *Liriodendron chinense*[J]. Journal of Forestry Research, 2021, 32(6): 2531-2543.
- [30] 宋晋辉, 马海莲, 翁巧云, 等. 玉米 *HSP70* 基因家族的全基因组鉴定与分析[J]. 核农学报, 2017, 31(7): 1245-1254.
- [31] WILKINS M R, GASTEIGER E, BAIROCH A, et al. Protein identification and analysis tools in the ExPASy server[J]. Methods in Molecular Biology, 1999, 112: 531-552.
- [32] HORTON P, PARK J, OBAYASHI T, et al. WoLF PSORT: protein localization predictor[J]. Nucleic Acids Research, 2007, 35(2): 585-587.
- [33] CHEN Y J, CHENG S Y, LIU C H, et al. Exploration of the truncated cytosolic Hsp70 in plants-unveiling the diverse T1 lineage and the conserved T2 lineage[J]. Frontiers in Plant Science, 2023, 14: 1279540.
- [34] TAMURA K, STECHER G, PETERSON D, et al. MEGA6: molecular evolutionary genetics analysis version 6.0[J]. Molecular Biology And Evolution, 2013, 30(12): 2725-2729.
- [35] HU B, JIN J, GUO Y, et al. GSDS 2.0: an upgraded gene feature visualization server[J]. Bioinformatics, 2015, 31(8): 1296-1297.
- [36] BAILEY T L, BODEN M, BUSKE F A, et al. MEME SUITE: tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37(2): 202-208.
- [37] LIU W, YI Y, ZHUANG J, et al. Genome-wide identification and transcriptional profiling of the basic helix-loop-helix gene family in tung tree (*Vernicia fordii*) [J]. Peer J, 2022, 10: e13981.

- [38] FENG N, SONG G, GUAN J, et al. Transcriptome profiling of wheat inflorescence development from spikelet initiation to floral patterning identified stage-specific regulatory genes [J]. *Plant Physiology*, 2017, 174(3):1779-1794.
- [39] GUO H, ZHANG H, WANG G, et al. Identification and expression analysis of heat-shock proteins in wheat infected with powdery mildew and stripe rust [J]. *Plant Genome*, 2021, 14(2):e20092.
- [40] 孙旋辉, 邴旭文, 丁炜东, 等. 高温应激对鳊鱼血清生化指标及肝脏 *sod* 基因和热休克蛋白基因表达的影响 [J]. *南方农业学报*, 2022, 53(12):3539-3547.
- [41] 楚宗丽, 李亮杰, 姬虹, 等. 小麦 *Hsp70* 基因家族鉴定及蛋白互作网络分析 [J]. *江苏农业科学*, 2022, 50(10):37-44.
- [42] CHEN Z, JASINSKA W, ASHRAF M, et al. Lipidomic insights into the response of *Arabidopsis* sepals to mild heat stress [J]. *aBIOTECH*, 2023, 4(3):224-237.
- [43] CHEN X, SHI L, CHEN Y, et al. *Arabidopsis HSP70-16* is required for flower opening under normal or mild heat stress temperatures [J]. *Plant Cell & Environment*, 2019, 42:1190-1204.
- [44] RAN X, CHEN X, SHI L, et al. Transcriptomic insights into the roles of HSP70-16 in sepal's responses to developmental and mild heat stress signals [J]. *Environmental and Experimental Botany*, 2020, 179:104225.
- [45] 陈锦玲. 大豆油脂和蛋白质合成与累积相关基因的表达及调控研究 [D]. 桂林: 广西师范大学, 2020.
- [46] 朱振. 莱茵衣藻核编码叶绿体定位蛋白表达及其应用 [D]. 大连: 大连工业大学, 2018.
- [47] TOKUNAGA S, SANDA S, URAGUCHI Y, et al. Overexpression of the DOF-Type transcription factor enhances lipid synthesis in *Chlorella vulgaris* [J]. *Applied Biochemistry and Biotechnology*, 2019, 189(1):116-128.
- [48] ZHANG L, ZHAO H K, DONG Q L, et al. Genome-wide analysis and expression profiling under heat and drought treatments of *HSP70* gene family in soybean (*Glycine max* L.) [J]. *Frontiers in Plant Science*, 2015, 6:773.
- [49] 周辉, 宋莉, 赵德刚. 大肠杆菌 HspQ 基因的生物信息学分析 [J]. *山地农业生物学报*, 2018, 37(5):57-61.
- [50] 徐剑文, 赵君, 刘剑光, 等. 芝麻全基因组 *Hsp70* 基因的鉴定与生物信息学分析 [J]. *江苏农业学报*, 2017, 33(3):493-502.

(责任编辑: 成纾寒)