

# 拟南芥 *Hsp20* 基因家族的特征及其在干旱和盐胁迫下的表达分析

黄思沛, 黄德娅, 付连郭, 秦余, 唐婕, 吕思慧, 廖海\*, 周嘉裕\*

(西南交通大学 生命科学与工程学院, 中国四川 成都 610031)

**摘要:** 热激蛋白 20 (heat shock protein 20, Hsp20) 是植物响应非生物胁迫时广泛合成的一类功能蛋白质, 参与植物的抗逆过程。借助拟南芥基因组数据库, 利用生物信息学方法分析 *Hsp20* 基因家族。HMMER 搜索和蛋白质的理化性质分析表明, 拟南芥至少含有 30 个 *Hsp20* 基因, 其编码蛋白的分子质量为 14.6~41.4 kD, 且均含有  $\alpha$ -晶状体蛋白结构域。系统发育分析表明, 在 30 个拟南芥 *Hsp20* 成员中, 22 个归属于 12 个不同的亚家族, 其余 8 个归于未知分类, 可能为 *Hsp20* 类蛋白质(Hsp20-like)。共线性分析表明, 片段复制与串联复制是 *Hsp20* 成员的主要扩增事件, 大部分 *Hsp20* 不含或只含 1 个内含子。MEME 分析结果表明, 基序 1、2、9 是 *Hsp20* 共有的保守基序。转录组分析提示, 15 个 *Hsp20* 的表达水平在干旱和(或)盐胁迫后出现了上调, 该结果得到了 qRT-PCR 实验的验证。以上结果为人们进一步探究 *Hsp20* 基因的生物学功能提供了理论依据。

**关键词:** 拟南芥; 热激蛋白 20 (Hsp20); 基因家族; 生物信息学分析; 非生物胁迫; 表达分析

中图分类号: Q943.2, Q945.78

文献标志码: A

文章编号: 1007-7847(2023)02-0162-08

## Characterization of *Arabidopsis thaliana* *Hsp20* Gene Family and Its Expression Analysis Under Drought and Salt Stress

HUANG Sipei, HUANG Deya, FU Lianguo, QIN Yu, TANG Jie, LÜ Sihui, LIAO Hai\*, ZHOU Jiayu\*

(School of Life Science and Engineering, Southwest Jiaotong University, Chengdu 610031, Sichuan, China)

**Abstract:** Heat shock protein 20 (Hsp20) family contains a kind of functional proteins that are widely synthesized in plants in response to abiotic stresses. Herein, the *Hsp20* gene family was analyzed by bioinformatics using *Arabidopsis thaliana* genome database. HMMER search and protein physicochemical property analysis showed that there are at least 30 *Hsp20* genes in *A. thaliana*, encoding 14.6~41.4 kD molecular weight proteins, all of which contain an  $\alpha$ -crystallin domain (ACD). Phylogenetic analysis showed that 22 of the 30 *A. thaliana* *Hsp20* members belong to 12 different subfamilies, and the remaining 8 belong to unknown taxa, which may be Hsp20-like proteins. Collinearity analysis showed that fragment replication and tandem replication are the main amplification events of *Hsp20* members, and most *Hsp20* members contain no or only one intron. Analysis using MEME showed that motifs 1, 2, and 9 are common conserved motifs in *Hsp20* family. Transcriptome analysis indicated that the expression levels of 15 *Hsp20* genes were up-regulated after drought and (or) salt stress, which was verified by qRT-PCR assay. These results provide theoretical basis for further investigation of the biological functions of *Hsp20* genes.

收稿日期: 2022-02-28; 修回日期: 2022-07-01; 网络首发日期: 2023-03-28

基金项目: 国家自然科学基金青年科学基金项目(31500276); 四川省中医药管理局项目(2020JJC0128, 2021MS116); 西南交通大学医工结合项目(2682021ZTPY017); 西南交通大学个性化实验(GX2022160170)

作者简介: 黄思沛(1998—), 男, 陕西汉中, 硕士研究生; \*通信作者: 廖海(1974—), 男, 四川遂宁人, 博士, 副教授, 主要从事生物化学和分子生物学研究, E-mail: ddliaohai@home.swjtu.edu.cn; 周嘉裕(1976—), 女, 四川宜宾人, 博士, 副教授, 主要从事植物学和分子生物学研究, E-mail: spinezhou@home.swjtu.edu.cn.

**Key words:** *Arabidopsis thaliana*; heat shock protein 20 (Hsp20); gene family; bioinformatics analysis; abiotic stress; expression analysis

(*Life Science Research*, 2023, 27(2): 162–169)

热激蛋白(heat shock protein, Hsp)是原核和真核生物受到非生物胁迫时合成的一类应激蛋白质,其作为分子伴侣,负责细胞内蛋白质折叠、组装、易位和降解,维持新生多肽的天然构象或将变性蛋白质重组为天然构象,在稳定蛋白质结构和维持细胞稳态方面发挥着重要作用<sup>[1]</sup>。根据质量大小和序列同源性,Hsp 被分为 5 个家族:Hsp100(约 100 kD)、Hsp90(83~90 kD)、Hsp70(66~78 kD)、Hsp60(约 60 kD)和 Hsp20(14~42 kD)<sup>[2]</sup>。相较于其他类型的 Hsp,Hsp20 有两个重要特征:1) 其与靶蛋白的结合不依赖于 ATP; 2) 其与靶蛋白的结合能力较强。

Hsp20 的结构可分为 3 个结构域,即 N 端的可变结构域、长度为 80~100 个氨基酸残基的  $\alpha$ -晶状体蛋白结构域( $\alpha$ -crystallin domain, ACD)和短的 C 端延伸。Hsp20 的 3 个不同区域分别行使不同的功能,其中 N 端可变结构域参与靶蛋白的结合,ACD 能够稳定靶蛋白的构象,C 端延伸则负责同源寡聚化<sup>[3]</sup>。根据亚细胞定位、序列同源性和功能差异,Hsp20 又可细分为不同的亚家族:C I~C VII、MT I、MT II、ER、P 和 Po,C I~C VII 定位于细胞质/细胞核,MT I 和 MT II 定位于线粒体,ER、P 和 Po 分别定位于内质网、叶绿体和过氧化物酶体<sup>[4]</sup>。迄今为止,*Hsp20* 基因已在拟南芥<sup>[5]</sup>、水稻<sup>[6]</sup>、小麦<sup>[7]</sup>与大豆<sup>[8]</sup>等多种植物中被报道,它们在干旱、盐与低温等非生物胁迫响应中发挥重要作用<sup>[9]</sup>。例如:拟南芥 *Hsp17.8* 在生菜中异源表达可增强转基因植物对脱水和高盐胁迫的抗性<sup>[10]</sup>; *Hsp17.7* 过表达的转基因水稻在干旱复水后能够复苏<sup>[11]</sup>; 过表达小麦 *Hsp23.9* 的转基因拟南芥对热和盐胁迫表现出更强的耐受性<sup>[12]</sup>。

已有研究发现,*Hsp20* 在植物基因组中常以基因家族形式存在,不同成员显示出广泛的序列变异性、进化差异与功能的多样化<sup>[13]</sup>。作为模式植物,拟南芥在基因家族功能及进化研究中具有突出的优势。2001 年,Scharf 等<sup>[14]</sup>在拟南芥基因组中首次发现 13 个 *Hsp20* 核心成员;2008 年,Siddique 等<sup>[5]</sup>对 5 个拟南芥 *Hsp20* 新成员完成了功能验证;2012 年,Bondino 等<sup>[15]</sup>发现部分具有 ACD 结构域的蛋白质并不具有分子伴侣功能,将它们归

为 Hsp20 类蛋白质(Hsp20-like)。为了后续更好地开展拟南芥 *Hsp20* 基因家族的功能研究与利用,本文从基因组水平对拟南芥 *Hsp20* 基因家族开展了整体性分析。首先,使用 HMMER 软件搜索拟南芥 *Hsp20* 基因家族成员,并结合系统发育分析确定了 22 个经典 Hsp20 和 8 个 Hsp20-like;其次,对其理化性质、共线性、顺式作用元件以及在非生物胁迫下的表达特征进行了分析。上述工作可为人们充分认识拟南芥 *Hsp20* 基因家族的进化关系及其可能的生物学功能提供基础资料,同时也可对其他物种 *Hsp* 家族的相关研究提供理论依据。

## 1 材料与方法

### 1.1 基因家族成员的确定

拟南芥蛋白质序列来源于 Ensembl 数据库(<http://plants.ensembl.org/index.html>)。首先,参考 Yu 等<sup>[9]</sup>的方法,采用 HMMER 3.0 软件,从 Pfam 数据库(<http://pfam.xfam.org/>)下载 Hsp20 结构域(编号为 PF00011)的隐马尔可夫模型配置文件,用于搜索拟南芥基因组中的 Hsp20 成员,其中 *E* 值<0.01;其次,将分子质量大小超出 14~42 kD 的序列剔除;再次,以 Hsp20 为关键词检索 TAIR (The Arabidopsis Information Resource; <https://www.arabidopsis.org/>)、Phytozome (<https://phytozome-next.jgi.doe.gov/>)等数据库;最后,将去冗余的蛋白质序列提交至 Pfam (<http://pfam.xfam.org/>)、SMART (<http://smart.embl-heidelberg.de/>)和 InterPro (<https://www.ebi.ac.uk/interpro/>)等在线工具,再次检验并确认其是否含 Hsp20 结构域,将不符合要求的序列剔除。以上数据获取及工具使用在 2021 年 8 月进行,后续生物信息学分析在 2021 年 9 月完成。

### 1.2 蛋白质理化性质分析

将拟南芥 Hsp20 氨基酸序列提交至 ExPASy 的 ProtParam 工具(<http://web.expasy.org/protparam/>),以计算其氨基酸数量、分子量。

### 1.3 系统发育分析

从 Phytozome (<https://phytozome-next.jgi.doe.gov/>)数据库下载来自大豆和水稻的 Hsp20 氨基酸

序列, 结合拟南芥 Hsp20 氨基酸序列进行系统发育分析。使用 MEGA-X 软件的 ClustalW 工具比对后构建邻接系统发育树, 替换模型选择 Poisson model, 空位处理选择 pairwise deletion, 自举检验次数为 1 000。

#### 1.4 保守基序分析

利用 MEME 5.5.1 (<http://meme-suite.org/tools/meme>)程序识别拟南芥 Hsp20 家族蛋白质序列中的保守基序, 识别基序数量的上限设为 20, 其他值均为系统默认, 筛选出  $E$  值 $<0.05$  的保守基序。

#### 1.5 共线性分析

使用 MCSanX 软件分析拟南芥基因组内的共线性区块<sup>[16]</sup>。

#### 1.6 启动子中与胁迫相关的顺式作用元件分析

使用 TBtools 工具提取拟南芥 Hsp20 编码序列的上游 2 000 bp 序列, 提交至 PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>), 以筛选参与胁迫响应的调控元件。

#### 1.7 拟南芥转录组数据

拟南芥 Columbia 株系(Col-0)在 1/2MS (PhytoTechnology Laboratories 公司, 美国)平板上培养 10 d 后移苗, 分为对照组、干旱胁迫组和盐胁迫组。对照组为 1/2MS 平板, 干旱胁迫组为含 2% PEG6000 (北京兰杰柯科技有限公司)的 1/2MS 平板, 盐胁迫组为含 100 mmol/L NaCl 的 1/2MS 平板。胁迫 7 d 后收集拟南芥根系样品, 送至北京百迈客生物科技有限公司进行转录组测序, 该数据为实验室前期获得<sup>[17]</sup>。对 Hsp20 基因表达水平进行标准化  $\log_2$ (FPKM)后 (FPKM: fragments per kilobase of exon model per million mapped reads), 使用 TBtools 工具绘制基因差异表达热图。

#### 1.8 qRT-PCR 验证

参照 1.7 的处理方法获得拟南芥根系材料, 利用植物 RNA 提取试剂盒(天根生化科技有限公司)提取总 RNA, 使用 Goldenstar™ RT6 cDNA Synthesis Mix 试剂盒(北京擎科新业生物技术有限公司)进行逆转录, 合成 cDNA。随机选取 8 个差异表达的 Hsp20 基因, 以 18S rRNA 为内参, 采用 SYBR Green I (北京擎科新业生物技术有限公司)荧光法进行 qRT-PCR 验证, 实施 3 次生物学重复, 采用  $2^{-\Delta\Delta C_t}$  法计算相对表达量。使用 SPSS 软件进行统计学分析, 3 组数据采用单因素方差分析, \* $P<0.05$  和 \*\* $P<0.01$  分别表示显著差异和极显著差异, 数据结果以平均值 $\pm$ 标准差表示。

## 2 结果

### 2.1 拟南芥 Hsp20 基因的筛选

针对拟南芥的全基因组分析共发现 30 个可能的 Hsp20 基因(表 1), 其编码蛋白的分子质量介于 14.6~41.4 kD。

表 1 拟南芥 Hsp20 成员的基本信息  
Table 1 Basic information on Hsp20 members in *Arabidopsis thaliana*

Serial number	Gene	Molecular weight/kD	Number of amino acid
1	AT1G06460.1	31.2	285
2	AT1G07400.1	17.8	157
3	AT1G52560.1	26.5	232
4	AT1G53540.1	17.6	157
5	AT1G54050.1	17.4	155
6	AT1G54400.1	20.4	183
7	AT1G54840.1	39.4	349
8	AT1G54850.1	22.3	206
9	AT1G59860.1	17.6	155
10	AT1G76770.1	28.7	224
11	AT2G03020.1	27.7	255
12	AT2G19310.1	18.5	162
13	AT2G27140.1	25.1	224
14	AT2G29500.1	17.6	153
15	AT3G46230.1	17.4	156
16	AT4G10250.1	22.0	195
17	AT4G16540.1	25.4	232
18	AT4G16545.1	29.1	262
19	AT4G21870.1	15.4	134
20	AT4G25200.1	23.6	210
21	AT4G27670.1	25.3	227
22	AT5G04890.1	41.3	366
23	AT5G12020.1	17.6	155
24	AT5G12030.1	17.7	156
25	AT5G20970.1	28.1	249
26	AT5G37670.1	15.7	137
27	AT5G47600.1	14.7	131
28	AT5G51440.1	23.5	210
29	AT5G54660.1	21.7	192
30	AT5G59720.1	18.1	161

### 2.2 拟南芥 Hsp20 成员的系统发育分析

利用 MEGA-X 软件对 30 个拟南芥 Hsp20 成员及已报道的水稻和大豆的 Hsp20 成员进行系统发育树构建(图 1)。结果表明, 30 个拟南芥 Hsp20 成员可以分为 7 类, 包括 16 个 C I~C VII、2 个 M I、1 个 M II、1 个 P、1 个 ER 和 1 个 Po 成员; 另有 8 个 Hsp20 成员形成两个单系群(图 1, 红色), 它们与质核类亚家族(C V 与 C VII)邻近且具有较近的亲缘关系, 由于它们与已知功能的 Hsp20 未成功聚类, 被定义为 Hsp20-like, 推测它们可能具有非典型 Hsp20 功能。从图 1 可知, 不同物种

来源的 *Hsp20* 成员并未按照植物种属关系严格聚类,而是按照亚家族类别分开聚类,表明 *Hsp20* 的进化历史较早。例如: C I、C II、C III 等大部分分支含有拟南芥、大豆和水稻的 *Hsp20* 成员,表明以上亚家族的形成早于单双子叶植物的分化时期; C IV 亚家族分支只含有双子叶植物拟南芥和大豆的 *Hsp20* 成员,表明该亚家族为单双子叶植物分化后形成;而 C VII 亚家族分支只含有拟南芥的 *Hsp20* 成员,推测该分支的形成是拟南芥物种分化后的近期事件。

### 2.3 拟南芥 *Hsp20/Hsp20-like* 的保守基序及基因结构

拟南芥 *Hsp20/Hsp20-like* 成员共被鉴定出 14 个保守基序(motifs 1~14),保守基序的数量、分布与系统发育关系有密切联系(图 2),其中基序 1、2 和 9 均位于 ACD 结构域,在所有 *Hsp20* 成员中均高度保守,推测其可能与 *Hsp20* 的分子伴侣功

能有关。同时,分析结果还显示,C I、C II、C IV、ER 和 Po 亚家族成员不含有内含子。

### 2.4 拟南芥 *Hsp20/Hsp20-like* 的共线性分析

*Hsp20/Hsp20-like* 基因分布在拟南芥的 5 条染色体上,其中 1 号和 5 号染色体最为集中(图 3)。有 3 对基因(*AT1G07400/AT2G29500*、*AT3G46230/AT5G59720* 和 *AT4G25200/AT5G51440*)发生了片段复制事件,前两对属于 C I 类 *Hsp20*,第 3 对属于 M I 类 *Hsp20*。另外,有 3 对基因(*AT1G54840/AT1G54850*、*AT4G16540/AT4G16545* 和 *AT5G12020/AT5G12030*)发生了串联重复事件,其中第 1 对为 *Hsp20-like* 基因,后两对分别属于 C VII 类和 C II 类 *Hsp20*。

### 2.5 拟南芥 *Hsp20/Hsp20-like* 启动子的顺式作用元件分析

启动子序列分析显示,*Hsp20/Hsp20-like* 含有干旱响应元件 DRE core、脱落酸响应元件 ABRE、

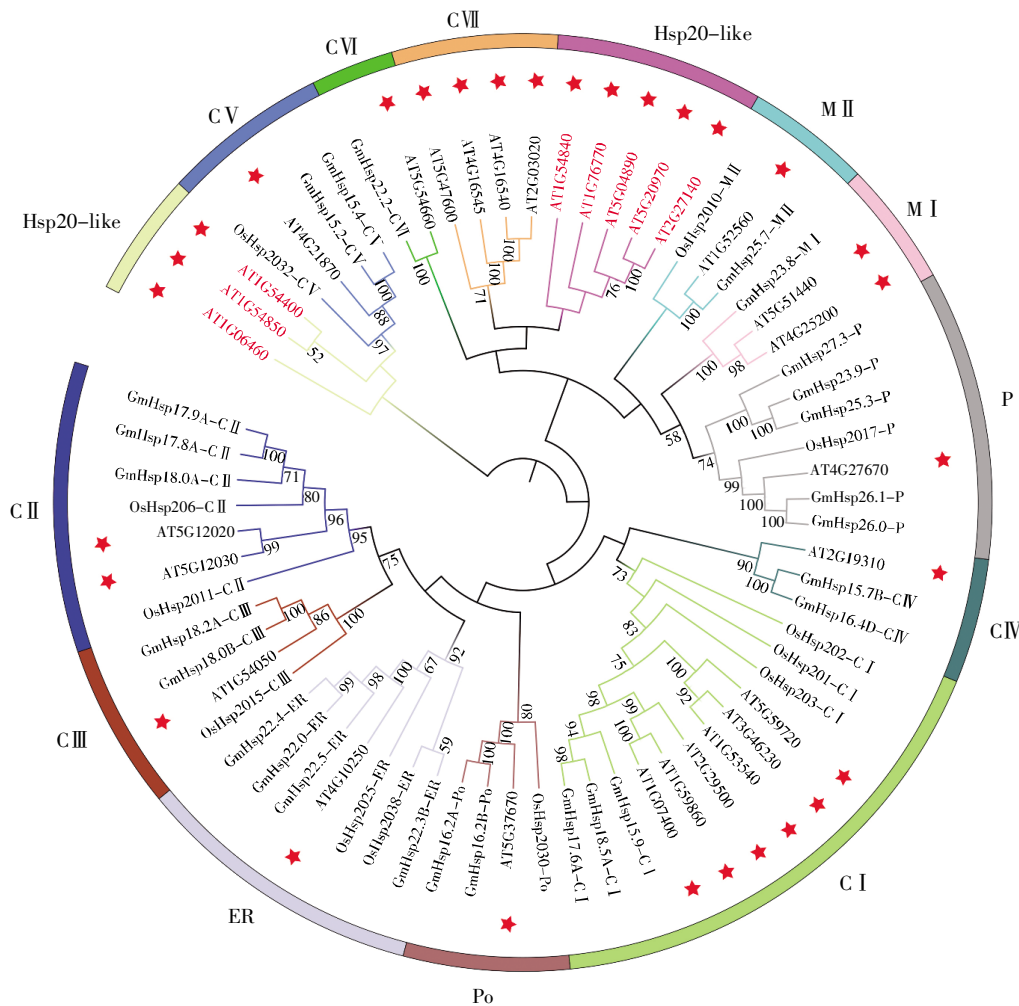


图 1 拟南芥、大豆和水稻 *Hsp20/Hsp20-like* 氨基酸序列的系统发育分析  
 Fig.1 Phylogenetic analysis of *Hsp20/Hsp20-like* amino acid sequences in *A. thaliana*, *Glycine max* and *Oryza sativa*

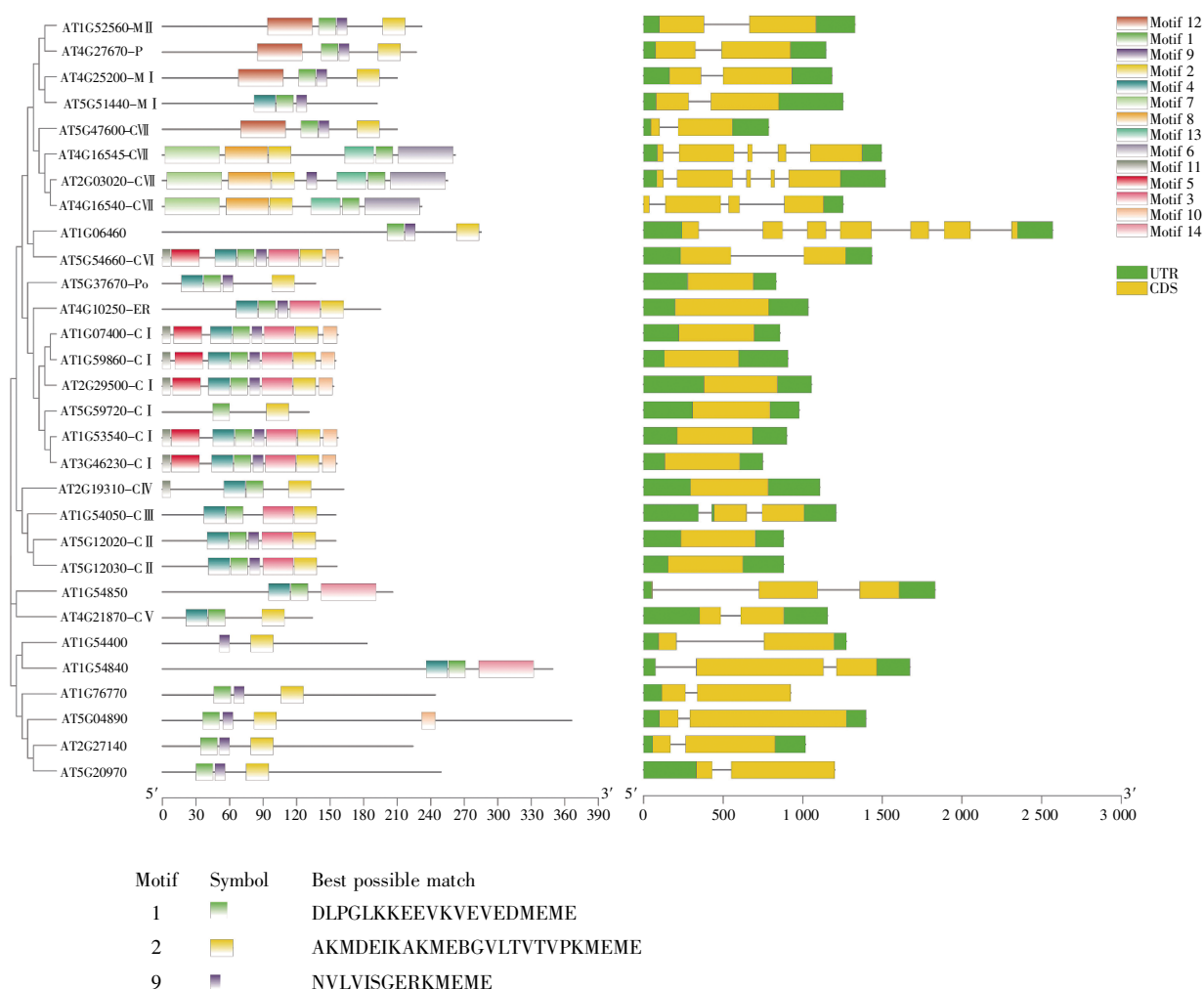


图2 拟南芥 *Hsp20/Hsp20-like* 的保守基序与基因结构

Fig.2 Conserved motifs and gene structures of *Hsp20/Hsp20-like* in *A. thaliana*

光调控元件 G-box 等 9 种响应非生物胁迫的顺式作用元件<sup>[18]</sup>(图 4)。

## 2.6 拟南芥 *Hsp20/Hsp20-like* 的差异表达分析

转录组数据显示, 9 个 *Hsp20* 基因(*AT1G07400*、*AT1G53540*、*AT2G29500*、*AT3G46230*、*AT4G10250*、*AT4G25200*、*AT5G12020*、*AT5G12030*、*AT5G59720*)在干旱与盐胁迫条件下均显著上调; 4 个 *Hsp20* 基因(*AT1G52560*、*AT1G54050*、*AT1G59860*、*AT5G37670*)在干旱胁迫下显著上调; 1 个 *Hsp20-like* 基因(*AT1G06460*)和 1 个 *Hsp20* 基因(*AT4G16540*)在盐胁迫下显著上调(图 5A)。另外, *AT4G16545* 和 *AT4G27670* 两个 *Hsp20* 基因无转录组数据。进一步采用 qRT-PCR 分析 *AT4G16545* 和 *AT4G27670* 的表达情况, 结果显示: *AT4G16545* 无扩增信号, 可能为假基因, 而 *AT4G27670* 虽然有扩增信号, 但表达强度极低, 并且在干旱及盐胁迫下无差异表达。

随机选择 8 个 *Hsp20* 基因进行 qRT-PCR 验证(图 5B), 结果显示: 所选基因的相对表达水平与转录组数据的趋势基本一致, 表明转录组数据可信度较高。

## 3 讨论

经全基因组分析发现, 拟南芥有 30 个 *Hsp20* 成员(表 1), 其数量少于水稻<sup>[6]</sup>、小麦<sup>[7]</sup>和大豆<sup>[8]</sup>, 这可能与拟南芥基因组小有关系。尽管拟南芥 *Hsp20* 成员所编码氨基酸序列的长度变化较大(131 到 366), 但它们均含有 ACD 与十二聚化的结构域。除 8 个 *Hsp20-like* 成员外, 其余 22 个 *Hsp20* 成员归属于 12 个亚家族(图 1), 该结果与 Siddique 等<sup>[5]</sup>的报道结果一致, 不同的是本研究发现有 3 个新成员(*AT4G16540*、*AT4G16545* 和 *AT2G03020*)被聚为 C VII 类, 这可能是因为采取的鉴定方法不同或近期注释了更多的 *Hsp20* 基因。细胞质是蛋白

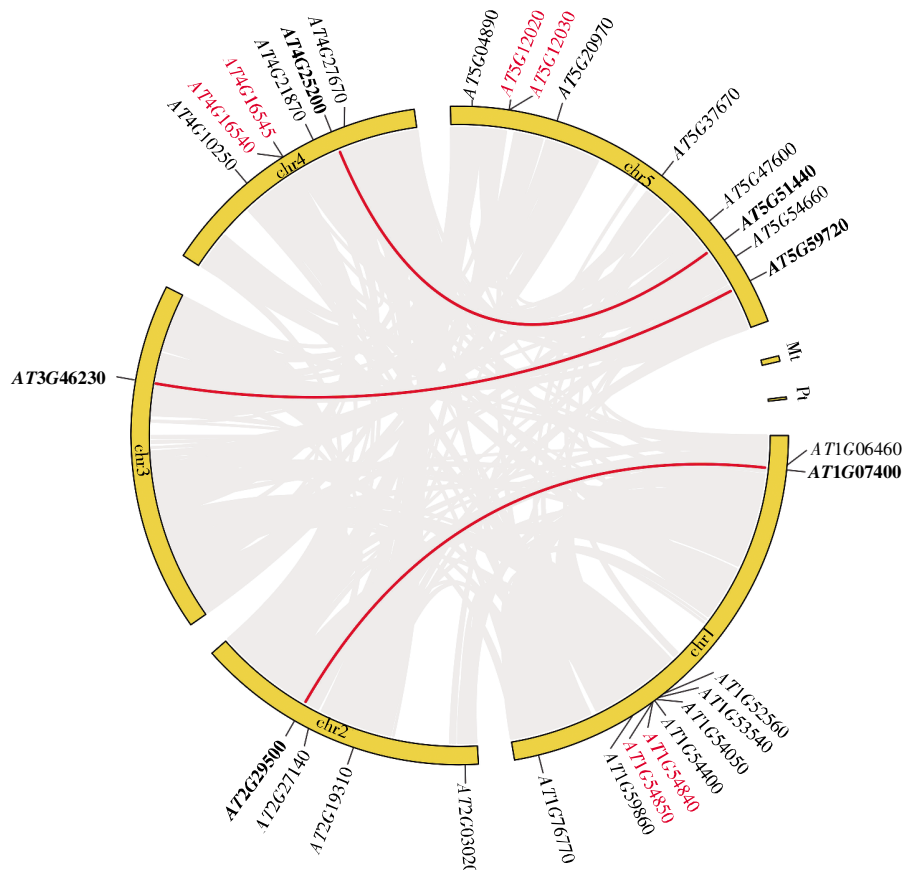


图 3 拟南芥 *Hsp20/Hsp20-like* 基因的共线性分析

红线连接基因对为片段复制, 标红基因对为串联复制。

Fig.3 Analysis of colinear data of *A. thaliana Hsp20/Hsp20-like* genes

Fragment duplication occurs in the gene pairs connected by a red line, and tandem duplication occurs in the gene pairs marked in red.

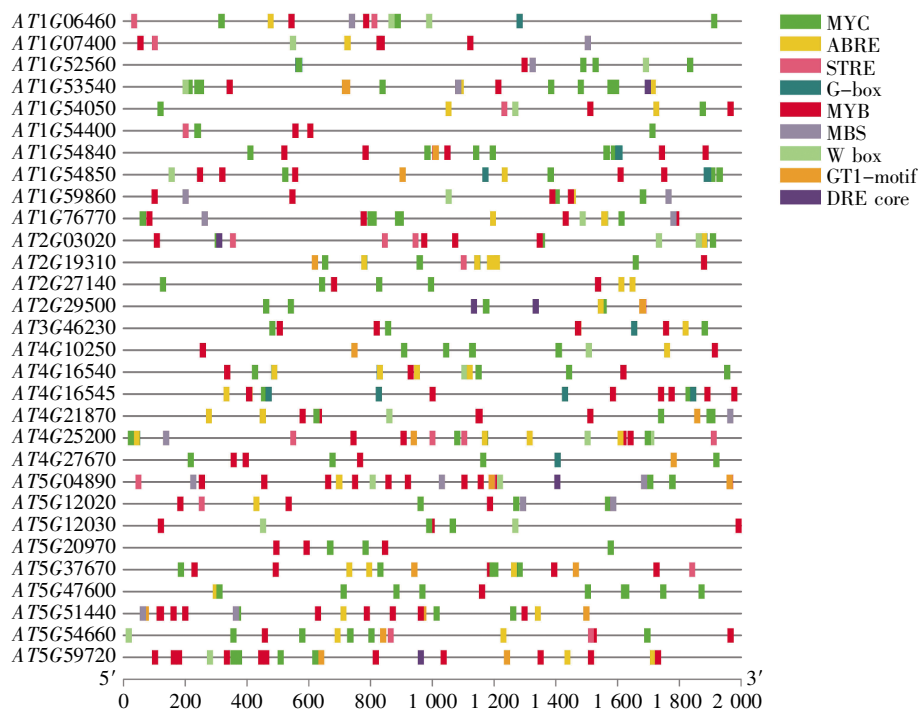


图 4 拟南芥 *Hsp20/Hsp20-like* 基因的顺式作用元件

Fig.4 The *cis*-acting elements of *A. thaliana Hsp20/Hsp20-like* genes

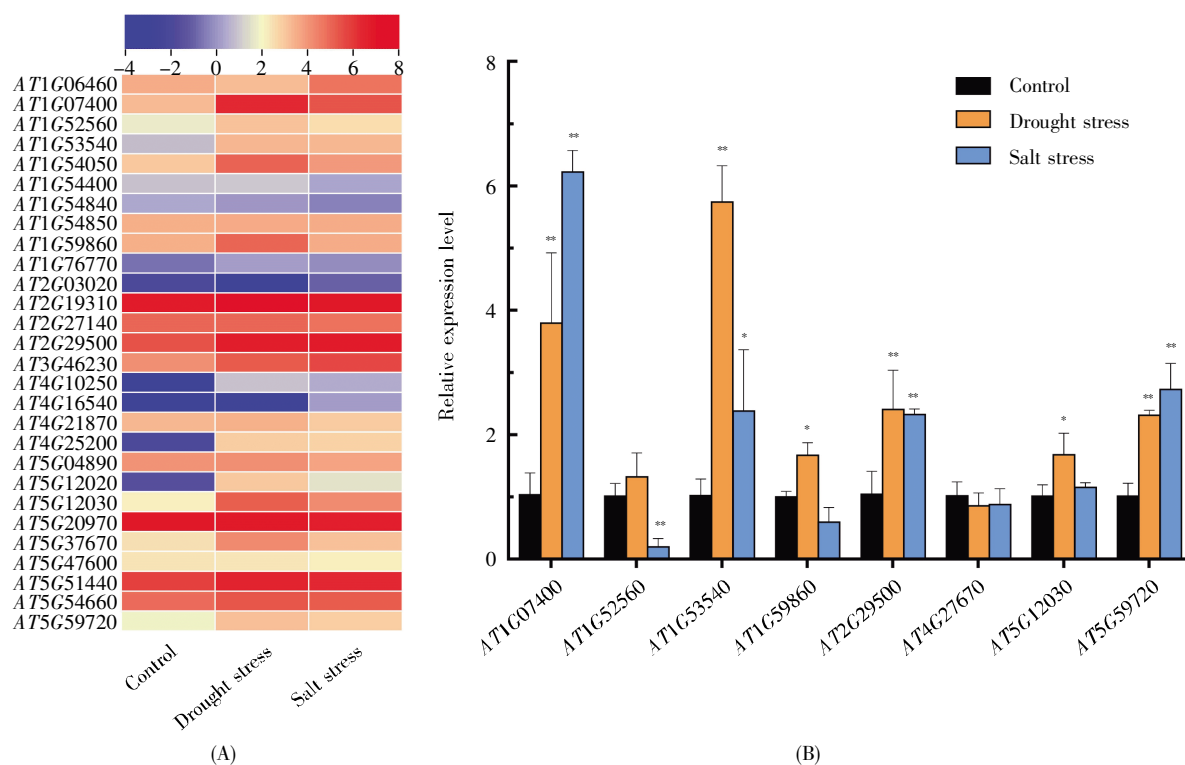


图5 *Hsp20/Hsp20-like* 基因在干旱和盐胁迫下的表达

(A) *Hsp20/Hsp20-like* 基因差异表达的热图。转录组 mRNA 表达水平通过  $\log_2(\text{FPKM})$  标准化; (B) qRT-PCR 验证部分 *Hsp20/Hsp20-like* 基因的相对表达水平。数据以平均值  $\pm$  标准差表示; 与对照组比较, \*:  $P < 0.05$ , \*\*:  $P < 0.01$ 。

**Fig.5 Expression of *Hsp20/Hsp20-like* genes under drought and salt stress**

(A) Heatmap of differential expression of *Hsp20/Hsp20-like* genes. The transcriptome mRNA expression levels were normalized by  $\log_2(\text{FPKM})$ ; (B) Verification of the relative expression levels of some *Hsp20/Hsp20-like* genes by qRT-PCR. The data are expressed as mean  $\pm$  standard deviation. \*:  $P < 0.05$  and \*\*:  $P < 0.01$ , compared with the control group.

质合成的主要场所,也是 *Hsp20* 行使分子伴侣功能的主要场所,这可能是大多数 *Hsp20* 成员被归属于质核亚家族的原因。需要指出的是,8个 *Hsp20-like* 蛋白可能具有其他功能,如 *AT5G04890* 可限制烟草蚀纹病毒的传播<sup>[9]</sup>,但目前人们对大部分 *Hsp20-like* 蛋白的了解非常有限。

C I、C II、C IV、ER 和 Po 亚家族成员不含内含子(图 2),推测它们能够快速响应非生物胁迫,这与先前的研究报道相符,内含子很少或没有内含子的基因结构有利于基因的快速激活,及时提升基因表达水平,以应对胁迫<sup>[20-22]</sup>。基因家族的扩展进化主要依赖于基因复制事件,本研究发现片段重复更有助于 *Hsp20* 基因的扩增(图 3),类似现象也出现在番茄等农作物中<sup>[9]</sup>。

同一亚家族成员具有相似的表达趋势(图 5),推测它们在功能上具有冗余性。相较于水稻 *Hsp20* 家族成员的表达模式<sup>[9]</sup>,我们发现来自 C I、C II、ER 和 M I 类亚家族的部分成员在干旱和盐胁迫下均表达上调,推测这 3 类成员在响应干旱和盐胁迫

过程中发挥重要作用。启动子分析表明,以上 3 类成员均存在响应干旱和盐胁迫的顺式作用元件(图 4)。此外,在 8 个 *Hsp20-like* 基因中,除 *AT1G06460* 在盐胁迫下表达上调外,其余基因均对干旱及盐胁迫无响应,这进一步说明它们可能不具有分子伴侣的功能。另一方面,在胁迫条件下表达上调的 *Hsp20* 基因极有可能参与胁迫响应过程,例如: *AT1G07400* (C I 类)的表达水平在干旱与盐胁迫条件下均出现上调,将其在生菜中过表达时,转基因植物对脱落酸表现为超级敏感,并获得了对于干旱和高盐胁迫的抗性<sup>[10]</sup>; *AT5G12030* (C II 类)与 *AT1G07400* (C I 类)具有相似的表达情况,其过表达也能够提高植株的耐盐性和抗旱性<sup>[23]</sup>。

上述有关拟南芥 *Hsp20* 基因家族的分析为今后相关基因家族成员的功能研究提供了理论基础。而且,不同亚家族成员应对不同胁迫时的响应有差异,对其分别进行功能研究有助于认识其功能分化以及抗逆反应机制,对今后其他转基因作物的改良研究也具有积极意义。

## 参考文献(References):

- [1] WANG W X, VINOCUR B, SHOSEYOV O, *et al.* Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response[J]. *Trends in Plant Science*, 2004, 9(5): 244–252.
- [2] WATERS E R, LEE G J, VIERLING E. Evolution, structure and function of the small heat shock proteins in plants[J]. *Journal of Experimental Botany*, 1996, 47(3): 325–338.
- [3] JAYA N, GARCIA V, VIERLING E. Substrate binding site flexibility of the small heat shock protein molecular chaperones[J]. *Proceedings of the National Academy of Sciences of USA*, 2009, 106(37): 15604–15609.
- [4] WATERS E R. The evolution, function, structure, and expression of the plant sHSPs[J]. *Journal of Experimental Botany*, 2013, 64(2): 391–403.
- [5] SIDDIQUE M, GERNHARD S, VON KOSKULL-DÖRING P, *et al.* The plant *sHSP* superfamily: five new members in *Arabidopsis thaliana* with unexpected properties[J]. *Cell Stress & Chaperones*, 2008, 13(2): 183–197.
- [6] SARKAR N K, KIM Y K, GROVER A. Rice *sHsp* genes: genomic organization and expression profiling under stress and development[J]. *BMC Genomics*, 2009, 10: 393.
- [7] MUTHUSAMY S K, DALAL M, CHINNUSAMY V, *et al.* Genome-wide identification and analysis of biotic and abiotic stress regulation of small heat shock protein (*HSP20*) family genes in bread wheat[J]. *Journal of Plant Physiology*, 2017, 211: 100–113.
- [8] LOPES-CAITAR V S, DE CARVALHO M C C G, DARBEN L M, *et al.* Genome-wide analysis of the *Hsp20* gene family in soybean: comprehensive sequence, genomic organization and expression profile analysis under abiotic and biotic stresses[J]. *BMC Genomics*, 2013, 14: 577.
- [9] YU J H, CHENG Y, FENG K, *et al.* Genome-wide identification and expression profiling of tomato *Hsp20* gene family in response to biotic and abiotic stresses[J]. *Frontiers in Plant Science*, 2016, 7: 1215.
- [10] KIM D H, XU Z Y, HWANG I. *AtHSP17.8* overexpression in transgenic lettuce gives rise to dehydration and salt stress resistance phenotypes through modulation of ABA-mediated signaling[J]. *Plant Cell Reports*, 2013, 32(12): 1953–1963.
- [11] SATO Y, YOKOYA S. Enhanced tolerance to drought stress in transgenic rice plants overexpressing a small heat-shock protein, *sHSP17.7*[J]. *Plant Cell Reports*, 2008, 27(2): 329–334.
- [12] WANG J, GAO X, DONG J, *et al.* Over-expression of the heat-responsive wheat gene *TaHSP23.9* in transgenic *Arabidopsis* conferred tolerance to heat and salt stress[J]. *Frontiers in Plant Science*, 2020, 11: 243.
- [13] BASHA E, O'NEILL H, VIERLING E. Small heat shock proteins and  $\alpha$ -crystallins: dynamic proteins with flexible functions[J]. *Trends in Biochemical Sciences*, 2012, 37(3): 106–117.
- [14] SCHARF K D, SIDDIQUE M, VIERLING E. The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing alpha-crystallin domains (Ac proteins)[J]. *Cell Stress & Chaperones*, 2001, 6(3): 225–237.
- [15] BONDINO H G, VALLE E M, TEN HAVE A. Evolution and functional diversification of the small heat shock protein/ $\alpha$ -crystallin family in higher plants[J]. *Planta*, 2012, 235(6): 1299–1313.
- [16] WANG Y P, TANG H B, DEBARRY J D, *et al.* MCSanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity[J]. *Nucleic Acids Research*, 2012, 40(7): e49.
- [17] 陈安琪, 宋思敏, 邓楷煜, 等. 拟南芥 COMT 基因家族的全基因组分析[J]. 分子植物育种(CHEN Anqi, SONG Simin, DENG Kaiyu, *et al.* Genome-wide analysis of COMT gene family in *Arabidopsis thaliana*[J]. *Molecular Plant Breeding*), 2022, 20(19): 6388–6396.
- [18] 郭晋艳, 郑晓瑜, 邹翠霞, 等. 植物非生物胁迫诱导启动子顺式元件及转录因子研究进展[J]. 生物技术通报(GUO Jinyan, ZHENG Xiaoyu, ZOU Cuixia, *et al.* Research progress of cis-elements of abiotic stress inducible promoters and associated transcription factors[J]. *Biotechnology Bulletin*), 2011(4): 16–20, 30.
- [19] WHITHAM S A, ANDERBERG R J, CHISHOLM S T, *et al.* *Arabidopsis RTM2* gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein[J]. *The Plant Cell*, 2000, 12(4): 569–582.
- [20] REN X Y, VORST O, FIERS M W E J, *et al.* In plants, highly expressed genes are the least compact[J]. *Trends in Genetics*, 2006, 22(10): 528–532.
- [21] JEFFARES D C, PENKETT C J, BÄHLER J. Rapidly regulated genes are intron poor[J]. *Trends in Genetics*, 2008, 24(8): 375–378.
- [22] CHUNG B Y W, SIMONS C, FIRTH A E, *et al.* Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana*[J]. *BMC Genomics*, 2006, 7: 120.
- [23] SUN W, BERNARD C, VAN DE COTTE B, *et al.* *At-HSP17.6A*, encoding a small heat-shock protein in *Arabidopsis*, can enhance osmotolerance upon overexpression[J]. *The Plant Journal*, 2001, 27(5): 407–415.