



宋蒙飞, 王星, 张开京, 等. 黄瓜 *DnaJ* 基因家族鉴定及对高温胁迫的表达响应 [J]. 南京农业大学学报, 2021, 44(2): 267-277.

SONG Mengfei, WANG Xing, ZHANG Kaijing, et al. Identification of *DnaJ* gene family in cucumber and its expression response to high temperature stress [J]. Journal of Nanjing Agricultural University, 2021, 44(2): 267-277.

黄瓜 *DnaJ* 基因家族鉴定及对高温胁迫的表达响应

宋蒙飞, 王星, 张开京, 陈劲枫, 娄群峰*

(南京农业大学作物遗传与种质创新国家重点实验室/园艺学院, 江苏 南京 210095)

摘要: [目的] 本文旨在研究黄瓜热激蛋白 *DnaJ* (*CsDnaJ*) 的家族基因鉴定, 并分析其在高温胁迫环境下的响应模式。[方法] 利用生物信息学软件及黄瓜的全基因组测序数据库、公共数据库对 *CsDnaJ* 进行基因家族成员、染色体位置、蛋白特性、结构特征、进化及复制模式和组织表达特性分析, 同时通过实时定量 PCR 技术分析其在高温胁迫环境下的响应模式。[结果] 黄瓜 *CsDnaJ* 基因家族共存在 81 个成员基因, 其中 Chr 3 含 *DnaJ* 基因最多, 为 16 个, Chr 4、Chr 7 最少, 均为 9 个; *CsDnaJ* 基因所编码的蛋白序列长度为 113~2 551 aa, 且至少含有 1 个完整的 J 结构域核心区序列; 亚细胞位置预测显示大部分蛋白定位于细胞核及叶绿体中。基因结构显示 *CsDnaJ* 基因均含有 1~22 个外显子。通过进化分析发现 *CsDnaJ* 蛋白可明显区分为 9 个亚簇, 每个亚簇均存在拟南芥同源蛋白。共线性分析发现, *CsDnaJ* 基因家族共存在 17 个共线基因, 组成 10 对复制基因对, 其中存在 5 对片段重复事件, 3 对串联重复事件, 2 对随机重复事件。组织表达特异性结果显示 *CsDnaJ* 基因在叶、茎、子房中的表达量略高于其在雄花、雌花、根及卷须中的表达量。荧光定量 PCR 结果显示, 高温环境(昼/夜温度为 42 °C/35 °C) 可以诱导 *CsDnaJ* 基因的高表达, 从而参与植物响应高温环境胁迫进程。[结论] 黄瓜 *CsDnaJ* 基因家族核心结构域高度保守, 在黄瓜响应高温胁迫中发挥重要作用。

关键词: 黄瓜; *DnaJ* 基因; 生物信息学; 高温胁迫; 表达分析

中图分类号: S642.2

文献标志码: A

文章编号: 1000-2030(2021)02-0267-11

Identification of *DnaJ* gene family in cucumber and its expression response to high temperature stress

SONG Mengfei, WANG Xing, ZHANG Kaijing, CHEN Jinfeng, LOU Qunfeng*

(State Key Laboratory of Crop Genetics and Germplasm Enhancement/College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: [Objectives] The paper aimed to identify the *DnaJ* protein family genes in cucumber and analyze its expression levels under high temperature stress. [Methods] Gene family members, chromosome position, protein characteristics, structural characteristics, evolution and replication patterns, tissue expression characteristics were identified using bioinformatics software, the cucumber reference genome database, and public databases, while their expression patterns under high temperature stress environment were analyzed by quantitative real-time PCR. [Results] A total of 81 *CsDnaJ* member genes of cucumber were identified. Chr 3 contained the largest number of *DnaJ* genes (16), and Chr 4 and Chr 7 contained the minimum (9). The length of the proteins sequence encoded by the *CsDnaJ* genes varied from 113 to 2 551 aa, which contained at least 1 complete core region sequence of the J-domain. The prediction of subcellular location showed that most of the proteins were located in the nucleus or chloroplast. The structure of *CsDnaJ* genes showed that each gene contained 1 to 22 exons. Evolutionary analysis showed that *CsDnaJ* proteins could be distinguished into 9 subclusters, each of which contained homologous proteins in *Arabidopsis thaliana*. Collinearity analysis showed that there were 17 collinear genes in the *CsDnaJ* gene family, forming 10 pairs of duplicating gene pairs, including 5 pairs of fragment repeating events, 3 pairs of tandem repeating events, and 2 pairs of random repeating events. The expression levels of *CsDnaJ* genes in leaf, stem and ovary were slightly higher than those in male flower, female flower, root and tendrils. Quantitative real-time PCR analysis showed that the high temperature environment (the temperature of day/night was 42 °C/35 °C) could induce the high expression of *CsDnaJ* genes to respond to high temperature stress environmental. [Conclusions] The core structure of the *CsDnaJ* gene family in cucumber is highly conserved and it plays an important role in response to high temperature stress in cucumber.

Keywords: cucumber; *DnaJ* gene; bioinformatics; high temperature stress; expression analysis

DnaJ 蛋白(又称 J 蛋白)是一类保守的辅伴蛋白家族,因包含 1 个标志性高保守的 J 结构域而得名,在生物体内作为热激蛋白(HSP70)的辅伴分子而存在^[1]。而 HSP70 是生物在高温环境下合成的一组蛋

收稿日期: 2020-06-24

基金项目: 江苏省农业科技自主创新基金 [CX(20)2019]; 江苏省农业重大新品种创制项目 (PZCZ201719); 江苏高校“青蓝工程”人才项目

作者简介: 宋蒙飞, 博士研究生。* 通信作者: 娄群峰, 教授, 从事葫芦科作物遗传育种与种质创新研究, E-mail: qlou@njau.edu.cn.

白质,存在于从细菌到人类的活生物体中,是生物响应高温胁迫的主要因子^[2]。作为 HSP70 的辅伴侣分子,DnaJ 蛋白在胁迫条件下能促进 HSP70 的 ATP 酶活性,并参与维持细胞内蛋白的折叠、复合物解聚等多种生命活动进程,在生命体抵抗和适应逆境胁迫方面发挥着重要作用^[3]。

DnaJ 蛋白通常情况下包含 4 个保守的结构域: N 端 J 结构域、甘氨酸/丙氨酸结构域、含 4 个 CxxCXGXC 重复基序的锌指结构域和 C 末端结构域,DnaJ 蛋白可根据其含有的结构域差异分为不同亚型(常见为 3 种类型)^[4]。DnaJ 蛋白广泛存在于生命体内不同部位,并因其在细胞内定位的不同而具有不同的功能,例如在拟南芥中已鉴定的 120 个 DnaJ 蛋白亚细胞分别定位于线粒体、叶绿体、内质网、细胞核和质膜上^[5]。目前研究最多的为定位于叶绿体的 DnaJ 蛋白,其在叶绿体中可参与 PS II 光合系统组装或修复、类囊体膜建成、叶绿体分裂发育、叶绿体转录、蛋白转运等过程^[6-8]。DnaJ 蛋白另一重要功能是能够参与植物胁迫响应进程,已有报道表明 DnaJ 蛋白在植物抗旱、抗高温、抗冷、抗盐以及抗病方面均发挥重要作用^[3]。

高温容易造成植物体内酶系统紊乱、细胞膜系统破坏、各组分改变等一系列生理生化变化,从而影响光合及呼吸作用,改变植物的源库关系,使作物的生长发育受到严重的影响,最终造成减产、减产等一系列问题^[9-10]。黄瓜(*Cucumis sativus* L.) 是重要的瓜类蔬菜作物之一,适宜生长温度为 16~26 °C,属喜温作物,但不耐热^[11]。目前在设施栽培或者南方地区露地栽培中,高温已成为设施黄瓜栽培生产的主要障碍^[12]。在黄瓜苗期、营养生长期、初花期、结果期遇到高温环境会影响黄瓜正常生长发育,降低黄瓜产量和品质,最终导致商品性降低^[13-15]。高温胁迫已经给黄瓜生产带来了巨大的损失,是我国黄瓜生产中急需解决的问题之一。目前关于黄瓜耐热性的研究主要集中在分子标记的开发上,研究者利用黄瓜耐热及不耐热材料通过连锁性分析鉴定了多个与耐高温相关的 QTL 位点及连锁标记^[16],但是关于黄瓜耐高温相关基因家族的分析尚未见报道。

目前在拟南芥、水稻、棉花中均有关于 *DnaJ* 基因家族的研究报道^[5,17-18]。但是,在黄瓜中关于 *DnaJ* 基因家族的研究尚未见报道。本研究基于已公布的黄瓜基因组信息对黄瓜 *DnaJ* 基因家族进行全基因组解析,注释该基因家族包括基因组成、结构、染色体分布和系统发育分析等信息,同时,通过公共数据库、RNA-seq 数据、实时定量 PCR 分析 *DnaJ* 基因的表达模式,并探讨其在植物响应高温环境胁迫中的功能,旨在为探明黄瓜响应高温胁迫应答机制提供理论参考。

1 材料与方法

1.1 黄瓜 *DnaJ* 基因家族全基因组鉴定

首先,从葫芦科作物基因组数据库(<http://cucurbitgenomics.org/>) 下载黄瓜‘9930’参考基因组相关信息文件(包括基因组序列、基因编码序列、蛋白序列)^[19],从 pfam 数据库数(<http://pfam.xfam.org/>) 下载 DnaJ 蛋白 HMM(隐马尔可夫模型)种子文件(编号:PF00226)^[20]。其次,参考崔梦杰等^[21]的方法,以种子文件为模板,通过 HMMER 3.1 软件的 hmmsearch 功能搜索蛋白质数据库,参数设为默认值。以初次筛选得到的黄瓜 DnaJ 蛋白序列为模板,通过 clusterW 软件将这些序列进行多序列对比,利用 HMMER 3.1 软件中 hmmbuild 功能将对对比结果建立黄瓜 DnaJ 蛋白特异 hmm 模型。最后,通过 hmmsearch 功能以黄瓜 DnaJ 蛋白特异 hmm 模型为模板再次搜索黄瓜蛋白数据库,所得结果通过 NCBI 的 CDD 数据库(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 检查蛋白质的保守结构域后^[22],手动筛选不含 DnaJ 特异 J 结构域核心区序列(组氨酸/脯氨酸/天冬氨酸,即 His/Pro/Asp,HPD 结构)的蛋白编号,最终得到黄瓜 DnaJ 蛋白家族基因。

1.2 黄瓜 *DnaJ* 基因家族染色体位置、蛋白特性及结构分析

黄瓜 *DnaJ* 基因家族染色体位置信息通过葫芦科基因组数据库 Browser 功能获得,并通过 Circos 软件(<http://circos.ca/>) 表现为图形模式;以基因名称搜索黄瓜基因组文件得到 DnaJ 基因家族蛋白全长序列,通过序列分析工具在线软件(http://www.detaibio.com/sms2/protein_mw.html) 计算蛋白相对分子质量、等电点(PI 值);利用在线软件 WOLF PRORT(<https://www.genscript.com/wolf-psort.html?src=leftbar>) 预测蛋白亚细胞定位;使用 MEME 在线软件(<http://meme-suite.org/tools/meme>) 分析蛋白 motif 特征^[23],motif 搜索个数为 10,其他参数默认,通过 TBtools 展示分析结果。

1.3 黄瓜 *DnaJ* 基因家族进化及复制模式分析

参考 Miernyk^[5] 的方法获得拟南芥 *DnaJ* 基因家族信息,并在拟南芥数据库下载拟南芥 *DnaJ* 基因家族蛋白序列。将黄瓜及拟南芥 *DnaJ* 基因家族 J 蛋白结构域序列通过 Neighbor-joining 进行聚类分析,使用 MEGA 7.0 软件构建系统发育树^[24], Bootstrap 值设置为 1 000。通过在线软件 EvolView(<https://www.evol-genius.info/evolview/>) 进行图形调整。以黄瓜 *DnaJ* 基因家族蛋白序列为模板,通过 MCSanX 软件进行基因组内复制事件鉴定^[25],并通过 Circos 软件表现为图形模式。

1.4 黄瓜 *DnaJ* 基因家族表达分析

在 NCBI 的 SRA 数据库中下载黄瓜基因在各组织的表达水平原始数据^[26],通过 salmon 软件计算黄瓜基因表达量^[27],根据基因名称获得黄瓜 *DnaJ* 基因家族在黄瓜各组织表达量,通过 Heml 软件绘制热图。

以华北密刺型黄瓜自交系品种‘长春密刺’为试验材料,于光照培养箱中培养黄瓜幼苗至 5~6 片真叶期,对其进行高温胁迫处理。高温处理:昼/夜时间为 16 h/8 h,温度为 42 °C/32 °C;对照:昼/夜时间为 16 h/8 h,温度为 28 °C/24 °C。分别于 0、6、12、24、48 h 时取黄瓜叶片,液氮速冻后保存于-80 °C,每个处理 3 个重复。利用 RNA 提取试剂盒(TaKaRa)提取植物叶片总 RNA,利用 Prime Script™ RT Reagent 试剂盒(TaKaRa)反转录合成 cDNA,使用 SYBR Premix Ex Taq™ 试剂盒(TaKaRa)进行荧光定量反应。考虑到具有高度相似性核心结构的基因可能具有相同的调控功能,因此分别在黄瓜 *DnaJ* 基因家族各亚簇组别中随机挑选 1 个基因(共 9 个基因),参考其编码 RNA 序列设计定量 PCR 引物,并检测基因在各处理中的相对表达量。定量 PCR 体系、反应程序及计算方法采用 Song 等^[28]的方法。

2 结果与分析

2.1 黄瓜 *DnaJ* 基因家族鉴定及蛋白序列信息分析

以 PF00226 的 HMM 文件为索引,对黄瓜基因组进行 *DnaJ* 基因筛选,在去除冗余基因后,共得到 81 个黄瓜 *DnaJ* 基因(表 1)。这些基因所编码的蛋白序列中至少含有 1 个完整的 J 结构域核心区序列(HPD 结构)。根据其在染色体上的位置信息,重新命名了这些基因名称(*CsDnaJ1*—*CsDnaJ81*)(表 1)。对 *DnaJ* 基因编码的蛋白序列进行分析,发现 *DnaJ* 基因编码的蛋白氨基酸长度为 113~2 551 aa,其中 66 个(80.2%)蛋白序列长度都大于 200 aa。另外,其蛋白相对分子质量为(12.10~279.60)×10³,等电点(PI)为 4.56~11.20。亚细胞位置预测显示,大部分蛋白定位于细胞核(42/81,51.8%)及叶绿体(20/81,24.7%)中,其余分别定位于细胞质(10)、细胞膜(3)、线粒体(2)、质体(2)、内质网(1)及液泡(1)中。

表 1 黄瓜 *DnaJ* 基因组基本特征

Table 1 Basic characteristics of *DnaJ* gene family in cucumber

基因序号 Gene ID	新命名 New name	染色体 Chromosome	位置 Position	分组 Group	氨基酸长度 The length of amino acid	相对分子质量/10 ³ Relative molecular mass	等电点 PI	亚细胞定位 Subcellular localization
Csa1G001510	<i>CsDnaJ1</i>	Chr 1	266 604~269 244	IV	842	94.27	9.33	细胞核 Nucleus
Csa1G013230	<i>CsDnaJ2</i>	Chr 1	1 782 802~1 785 241	V	589	67.46	5.11	细胞核 Nucleus
Csa1G031710	<i>CsDnaJ3</i>	Chr 1	3 284 188~3 303 460	IX	388	42.62	8.85	叶绿体 Chloroplast
Csa1G042670	<i>CsDnaJ4</i>	Chr1	4 430 048~4 431 545	IX	235	27.09	5.41	叶绿体 Chloroplast
Csa1G045610	<i>CsDnaJ5</i>	Chr 1	5 036 912~5 039 959	III	347	38.70	9.62	细胞质 Cytoplasm
Csa1G064710	<i>CsDnaJ6</i>	Chr 1	6 813 658~6 818 208	II	300	35.20	9.65	质体 Plastid
Csa1G320350	<i>CsDnaJ7</i>	Chr 1	13 738 487~13 741 140	IV	362	40.74	7.87	细胞核 Nucleus
Csa1G397110	<i>CsDnaJ8</i>	Chr 1	14 711 311~14 728 796	IX	2 551	279.60	6.12	细胞核 Nucleus
Csa1G535800	<i>CsDnaJ9</i>	Chr 1	19 204 081~19 207 106	IV	468	53.53	4.61	细胞核 Nucleus
Csa1G572990	<i>CsDnaJ10</i>	Chr 1	21 351 042~21 354 512	IV	848	94.57	8.85	细胞核 Nucleus
Csa1G637960	<i>CsDnaJ11</i>	Chr 1	25 490 096~25 498 071	IV	252	29.89	9.78	细胞核 Nucleus
Csa1G642540	<i>CsDnaJ12</i>	Chr 1	25 607 133~25 609 719	VII	414	45.83	6.16	细胞核 Nucleus
Csa2G022810	<i>CsDnaJ13</i>	Chr 2	2 736 153~2 738 850	IX	265	29.86	5.20	叶绿体 Chloroplast
Csa2G033330	<i>CsDnaJ14</i>	Chr 2	3 197 987~3 200 621	V	162	18.22	5.35	细胞核 Nucleus
Csa2G058620	<i>CsDnaJ15</i>	Chr 2	4 319 569~4 324 459	VII	334	37.39	5.53	叶绿体 Chloroplast
Csa2G249270	<i>CsDnaJ16</i>	Chr 2	12 158 601~12 162 192	VII	340	38.12	6.55	细胞质 Cytoplasm
Csa2G264570	<i>CsDnaJ17</i>	Chr 2	12 787 491~12 788 095	IX	146	16.06	8.46	叶绿体 Chloroplast
Csa2G301500	<i>CsDnaJ18</i>	Chr 2	14 466 098~14 470 314	VI	284	31.79	10.14	细胞核 Nucleus
Csa2G353470	<i>CsDnaJ19</i>	Chr 2	16 349 111~16 352 145	VI	245	29.25	9.81	细胞核 Nucleus
Csa2G360030	<i>CsDnaJ20</i>	Chr 2	16 976 993~16 979 050	IV	174	19.23	4.65	叶绿体 Chloroplast
Csa2G369750	<i>CsDnaJ21</i>	Chr 2	18 070 524~18 082 899	VIII	468	51.47	7.67	细胞核 Nucleus

续表1 Table 1 continued

基因序号 Gene ID	新命名 New name	染色体 Chromosome	位置 Position	分组 Group	氨基酸长度 The length of amino acid	相对分子质量/ 10^3 Relative molecular mass	等电点 PI	亚细胞定位 Subcellular localization
Csa2G372140	<i>CsDnaJ22</i>	Chr 2	18 428 099~18 431 824	V	135	15.91	4.79	细胞核 Nucleus
Csa2G408950	<i>CsDnaJ23</i>	Chr 2	21 233 928~21 236 619	III	343	38.05	9.46	细胞质 Cytoplasm
Csa3G020050	<i>CsDnaJ24</i>	Chr 3	2 083 620~2 090 586	V	540	60.09	9.40	细胞质 Cytoplasm
Csa3G129700	<i>CsDnaJ25</i>	Chr 3	8 368 761~8 370 568	IX	218	25.07	10.00	细胞核 Nucleus
Csa3G165080	<i>CsDnaJ26</i>	Chr 3	10 876 613~10 878 005	IV	268	30.78	9.20	细胞核 Nucleus
Csa3G184060	<i>CsDnaJ27</i>	Chr 3	12 941 508~12 946 289	II	686	76.82	5.40	叶绿体 Chloroplast
Csa3G223330	<i>CsDnaJ28</i>	Chr 3	14 791 608~14 794 730	V	648	74.17	8.23	细胞核 Nucleus
Csa3G710780	<i>CsDnaJ29</i>	Chr 3	26 888 164~26 889 931	IX	157	17.51	6.52	叶绿体 Chloroplast
Csa3G710790	<i>CsDnaJ30</i>	Chr 3	26 897 124~26 898 835	IX	157	17.64	5.26	叶绿体 Chloroplast
Csa3G730990	<i>CsDnaJ31</i>	Chr 3	27 445 080~27 449 407	VIII	573	62.81	9.95	液泡 Vacuole
Csa3G731180	<i>CsDnaJ32</i>	Chr 3	27 558 672~27 560 991	III	344	37.88	9.33	细胞膜 Cytoplasm
Csa3G782800	<i>CsDnaJ33</i>	Chr 3	30 422 428~30 425 931	I	276	31.20	4.97	叶绿体 Chloroplast
Csa3G806290	<i>CsDnaJ34</i>	Chr 3	30 865 483~30 869 971	VI	268	31.04	9.69	细胞核 Nucleus
Csa3G810550	<i>CsDnaJ35</i>	Chr 3	31 089 653~31 092 665	VI	283	32.56	8.70	细胞核 Nucleus
Csa3G843760	<i>CsDnaJ36</i>	Chr 3	34 066 279~34 069 137	IX	317	35.61	9.86	叶绿体 Chloroplast
Csa3G878950	<i>CsDnaJ37</i>	Chr 3	37 009 441~37 011 982	III	336	36.84	9.09	细胞质 Cytoplasm
Csa3G890060	<i>CsDnaJ38</i>	Chr 3	37 991 530~37 994 560	V	138	16.25	4.68	细胞核 Nucleus
Csa3G891630	<i>CsDnaJ39</i>	Chr 3	38 046 989~38 048 804	IV	336	37.63	8.19	细胞核 Nucleus
Csa4G006450	<i>CsDnaJ40</i>	Chr 4	1 129 977~1 137 749	VIII	506	54.50	9.12	叶绿体 Chloroplast
Csa4G011640	<i>CsDnaJ41</i>	Chr 4	1 683 695~1 688 164	IV	786	88.33	9.73	细胞核 Nucleus
Csa4G064050	<i>CsDnaJ42</i>	Chr 4	5 150 715~5 158 029	IX	558	63.67	8.38	细胞核 Nucleus
Csa4G095550	<i>CsDnaJ43</i>	Chr 4	6 246 049~6 250 698	VIII	446	47.93	9.43	叶绿体 Chloroplast
Csa4G292450	<i>CsDnaJ44</i>	Chr 4	11 482 044~11 483 605	IX	113	12.18	10.88	叶绿体 Chloroplast
Csa4G343090	<i>CsDnaJ45</i>	Chr 4	14 208 070~14 217 583	VII	1 369	150.61	6.22	细胞核 Nucleus
Csa4G496260	<i>CsDnaJ46</i>	Chr 4	17 293 149~17 303 290	IV	988	109.80	6.22	细胞核 Nucleus
Csa4G593910	<i>CsDnaJ47</i>	Chr 4	19 036 303~19 037 201	IX	199	22.36	9.46	细胞质 Cytoplasm
Csa4G665650	<i>CsDnaJ48</i>	Chr 4	23 260 271~23 282 496	VII	399	44.17	5.93	细胞质 Cytoplasm
Csa5G139530	<i>CsDnaJ49</i>	Chr 5	3 617 391~3 620 332	IV	760	85.67	9.54	细胞核 Nucleus
Csa5G175710	<i>CsDnaJ50</i>	Chr 5	7 327 362~7 329 584	IX	466	51.28	9.53	内质网 Endoplasmic reticulum
Csa5G467910	<i>CsDnaJ51</i>	Chr 5	16 420 635~16 426 713	VIII	408	46.07	8.74	细胞核 Nucleus
Csa5G523160	<i>CsDnaJ52</i>	Chr 5	18 508 684~18 512 685	IX	302	34.23	8.73	叶绿体 Chloroplast
Csa5G571570	<i>CsDnaJ53</i>	Chr 5	19 994 469~19 995 680	V	139	15.95	4.56	细胞核 Nucleus
Csa5G577340	<i>CsDnaJ54</i>	Chr 5	20 263 640~20 267 696	VIII	345	38.54	6.09	细胞核 Nucleus
Csa5G585980	<i>CsDnaJ55</i>	Chr 5	20 909 760~20 912 016	III	135	14.89	10.55	细胞核 Nucleus
Csa5G604140	<i>CsDnaJ56</i>	Chr 5	22 402 004~22 404 623	III	317	35.57	9.66	细胞质 Cytoplasm
Csa5G610320	<i>CsDnaJ57</i>	Chr 5	23 699 983~23 704 976	VIII	410	45.38	6.54	细胞核 Nucleus
Csa5G613470	<i>CsDnaJ58</i>	Chr 5	24 102 123~24 104 968	VII	414	46.07	6.44	细胞核 Nucleus
Csa5G623520	<i>CsDnaJ59</i>	Chr 5	24 863 348~24 866 050	IX	113	12.10	11.20	叶绿体 Chloroplast
Csa6G067950	<i>CsDnaJ60</i>	Chr 6	4 818 083~4 821 462	IV	941	107.28	5.89	细胞核 Nucleus
Csa6G190200	<i>CsDnaJ61</i>	Chr 6	12 300 916~12 303 900	VIII	274	31.51	5.14	细胞核 Nucleus
Csa6G306340	<i>CsDnaJ62</i>	Chr 6	14 649 575~14 651 979	IV	349	39.94	8.53	细胞核 Nucleus
Csa6G324860	<i>CsDnaJ63</i>	Chr 6	15 126 474~15 131 174	IX	731	81.36	8.31	质体 Plastid
Csa6G344270	<i>CsDnaJ64</i>	Chr 6	15 645 519~15 650 234	VII	386	43.46	5.51	细胞质 Cytoplasm
Csa6G367120	<i>CsDnaJ65</i>	Chr 6	16 836 152~16 842 153	I	394	45.03	6.67	细胞核 Nucleus
Csa6G399770	<i>CsDnaJ66</i>	Chr 6	17 734 003~17 735 932	IX	272	32.05	9.89	叶绿体 Chloroplast
Csa6G434360	<i>CsDnaJ67</i>	Chr 6	20 551 854~20 561 950	VII	1 342	148.43	5.46	细胞核 Nucleus
Csa6G445030	<i>CsDnaJ68</i>	Chr 6	20 678 413~20 683 581	VI	499	54.76	6.85	叶绿体 Chloroplast
Csa6G502820	<i>CsDnaJ69</i>	Chr 6	25 402 473~25 408 232	IV	1 434	160.40	4.97	细胞核 Nucleus
Csa6G505950	<i>CsDnaJ70</i>	Chr 6	25 811 706~25 812 690	IX	165	18.19	10.06	叶绿体 Chloroplast
Csa6G518110	<i>CsDnaJ71</i>	Chr 6	27 357 927~27 360 377	III	350	38.05	9.66	细胞膜 Cytoplasm
Csa6G524620	<i>CsDnaJ72</i>	Chr 6	28 237 753~28 241 178	IV	470	52.82	4.88	细胞核 Nucleus
Csa7G014590	<i>CsDnaJ73</i>	Chr 7	910 221~914 594	VI	291	32.60	9.81	细胞核 Nucleus
Csa7G032270	<i>CsDnaJ74</i>	Chr 7	1 819 347~1 823 445	IV	709	78.92	8.79	细胞核 Nucleus
Csa7G039200	<i>CsDnaJ75</i>	Chr 7	2 096 427~2 098 557	III	340	37.60	9.43	细胞膜 Cytoplasm
Csa7G061190	<i>CsDnaJ76</i>	Chr 7	3 563 962~3 564 459	VII	166	18.58	10.41	叶绿体 Chloroplast
Csa7G190690	<i>CsDnaJ77</i>	Chr 7	6 926 212~6 926 975	IX	121	13.69	9.55	线粒体 Chloroplast
Csa7G405850	<i>CsDnaJ78</i>	Chr 7	15 519 490~15 520 689	V	131	14.88	8.61	细胞核 Nucleus
Csa7G407530	<i>CsDnaJ79</i>	Chr 7	15 637 312~15 641 532	VIII	345	38.83	6.52	细胞质 Cytoplasm
Csa7G428930	<i>CsDnaJ80</i>	Chr 7	16 580 272~16 585 384	VIII	448	49.04	8.79	线粒体 Mitochondrion
Csa7G448690	<i>CsDnaJ81</i>	Chr 7	18 324 127~18 326 791	V	244	27.42	6.93	细胞核 Nucleus

注: I-IV为根据进化关系分析所产生的亚簇组别号。I-IV is the subcluster group number produced by evolutionary relationship analysis.

2.2 黄瓜 *DnaJ* 基因家族结构域及基因结构特征

利用 MEME 软件对黄瓜 *DnaJ* 基因家族蛋白序列进行分析, 结果发现所有蛋白都含有 motif 1, 其为 J 结构域核心序列, 含有稳定的 HPD 结构, 是 J 蛋白与 HSP70 蛋白相互作用的关键结构; 而 motif 2 为 J 结构的 C 端螺旋结构, motif 1 与 motif 2 结合形成完整的 J 结构域。由图 1 可知: 66 个黄瓜 *DnaJ* 蛋白都含有完整的 J 结构, 而其余 15 个蛋白只含有 HPD 结构及不完全的 α 螺旋结构所构成的 J 结构域; 而 motif 4 (81 个 *DnaJ* 蛋白中共 9 个蛋白含有该结构域) 为甘氨酸/苯丙氨酸富集结构域序列 (G/F 结构域), motif 5 (共 4 个) 为 CxxCXGXG 重复基序的锌指结构域序列, motif 3 (共 10 个) 或 motif 6 (共 12 个) 为 *DnaJ*-C 末端结构域序列。在其他作物中可以根据结构域特征将 *DnaJ* 蛋白分为 3 类, 而黄瓜中只含有其中 2 类, 即: I 只有 J 结构域 (共 72 个); II 含有 J 结构域、G/F 结构域及 C 末端结构域 (共 9 个)。另外, 一些 *DnaJ* 蛋白同样存在其他非 *DnaJ* 蛋白特有结构域。

通过对黄瓜 *DnaJ* 基因家族 mRNA 编码序列比对可知, 黄瓜 *DnaJ* 基因均含有 1~22 个外显子, 其中基因 *CsDnaJ8* 含有外显子最多, 为 22 个, 而 14 个 *CsDnaJ* 基因 (17.3%) 只含有 1 个外显子, 无内含子序列。

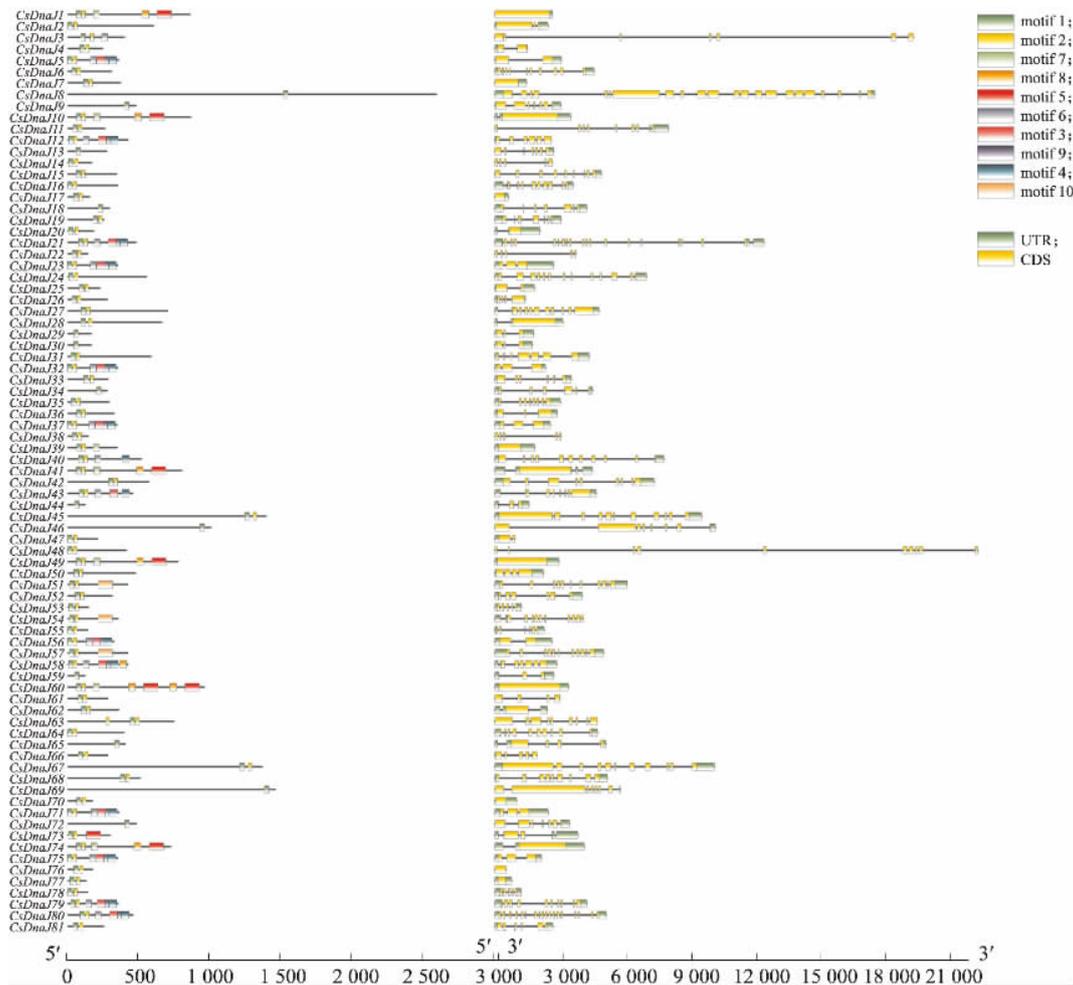


图 1 *CsDnaJ* 基因 motif 结构及基因结构分析

Fig. 1 Analysis of motif structure and gene structure of *CsDnaJ* gene

2.3 黄瓜 *DnaJ* 基因家族保守结构域及系统进化分析

对黄瓜 *DnaJ* 基因编码蛋白进行对比分析发现, 这些蛋白均含有保守的 J 结构域, 由 85 个氨基酸构成, 其中第 1~3 位 (DYY)、5~8 位 (ILGV)、17~18 位 (AS)、20~32 位 (EELKKAYRKLALK)、34~38 位 (HPDKN)、65 (F)、68 (I)、71~78 (AYEVLSDP)、80~81 (KR) 及 84~85 位 (YD) 氨基酸序列高保守。但核心结构域中其他序列仍存在部分差异, 可能造成其行使功能的不同 (图 2)。

利用黄瓜 *DnaJ* 基因家族及拟南芥 *DnaJ* 基因家族的核心结构域 (J 结构域) 序列进行聚类分析并构建系统进化树, 结果 (图 3) 显示, 黄瓜 *DnaJ* 蛋白与拟南芥 *DnaJ* 蛋白可明显区分为 9 个亚簇, 每个亚簇均存在拟南芥同源蛋白。其中 IX 组含有黄瓜 *DnaJ* 蛋白最多, 为 19 个 (23.5%), I 和 II 最少, 为 2 个 (2.7%)。

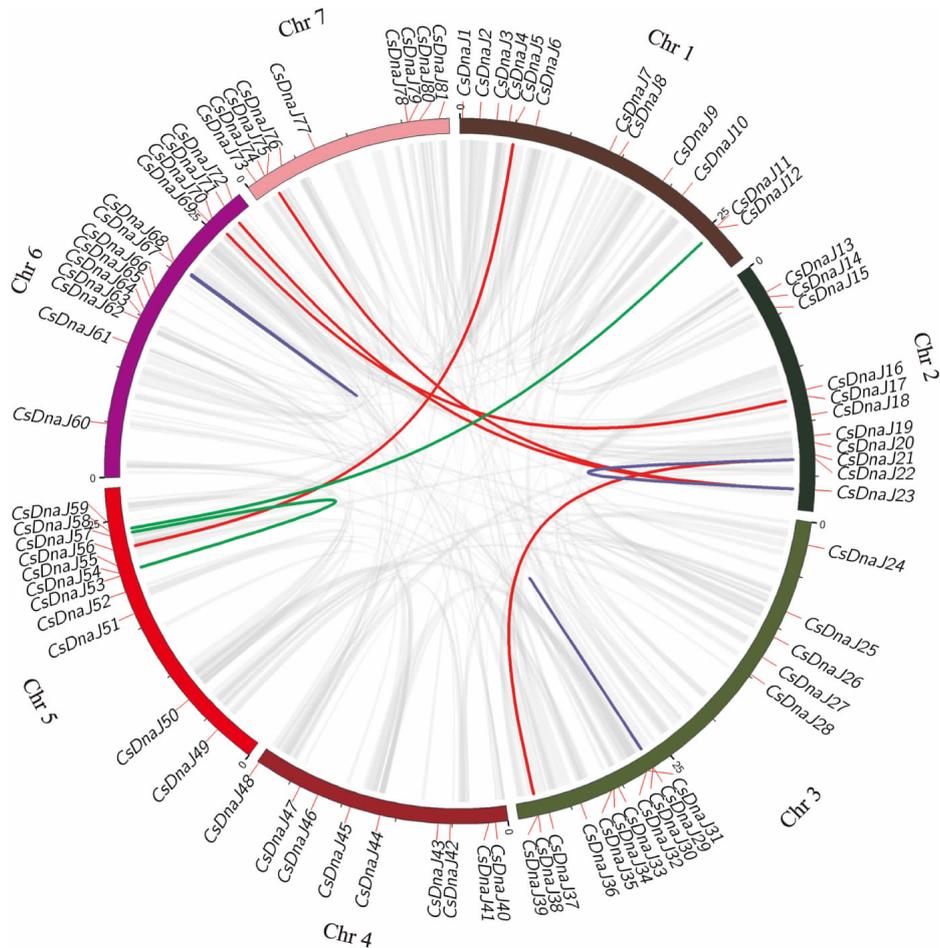


图 4 黄瓜 *DnaJ* 基因家族染色体定位及共线性分析

Fig. 4 Chromosome location and collinearity analysis of *DnaJ* gene family in cucumber

红线为片段重复事件,紫线为串联重复事件,绿线为随机重复事件。

The fragment repeat events are represented by red lines, the purple line are series repeat events, and the green lines are random repeat events.

2.5 黄瓜 *DnaJ* 基因家族组织表达特异性

由图 5 可知: 黄瓜 *DnaJ* 基因家族在各组织中均有表达, 其中, 在叶、茎、子房中的表达量略高于在雄花、雌花、根、雄花及卷须中的表达量。各基因间的表达量差异较大, 例如基因 *CsDnaJ20*、*CsDnaJ39*、*CsDnaJ71* 在各组织中表达量均较低(小于 1), 而基因 *CsDnaJ20*、*CsDnaJ35*、*CsDnaJ45* 在各组织中的表达量均较高(大于 10)。这些高表达基因分别属于第 IV、VI、VII 亚组, 且多数定位于叶绿体中。黄瓜 *DnaJ* 家族基因的组织表达特异性表明, 不同 *DnaJ* 家族基因成员在黄瓜各组织具有不同的调控作用, 且个别基因的高表达也表明其在 *DnaJ* 基因家族调控黄瓜生长发育过程中的重要性。

2.6 黄瓜 *DnaJ* 基因家族对高温胁迫应答模式

由图 6 可见: 在高温处理 48 h 后, 黄瓜植株出现明显萎蔫状态。由图 7 可见: 9 个 *DnaJ* 基因皆出现上调表达趋势, 且与对照组相比大部分基因表达量具有显著性差异。9 个 *DnaJ* 基因在不同处理时期表达模式有所差异: 5 个基因(*CsDnaJ33*、*CsDnaJ6*、*CsDnaJ37*、*CsDnaJ14*、*CsDnaJ19*) 在高温处理初期(6 h) 表达量达到最高值, 其后逐渐下降; *CsDnaJ72* 基因在高温处理 12 h 时表达量最高, 其他处理时间基本持平; *CsDnaJ16* 基因在不同处理时间表达量相近; *CsDnaJ31* 基因在处理前期表达量持续上升, 至 24 h 达到最高点, 其后下降; *CsDnaJ29* 基因在处理前期表达量变化不显著, 在 48 h 时上升达到最高值。*DnaJ* 基因家族在高温胁迫下的高表达现象表明了植物在高温环境下可以通过诱导 *DnaJ* 基因高表达来进一步响应胁迫环境, 维持植物体的正常发育状态。这也同样表明 *DnaJ* 蛋白作为应激蛋白在黄瓜响应胁迫环境下起重要调节作用。

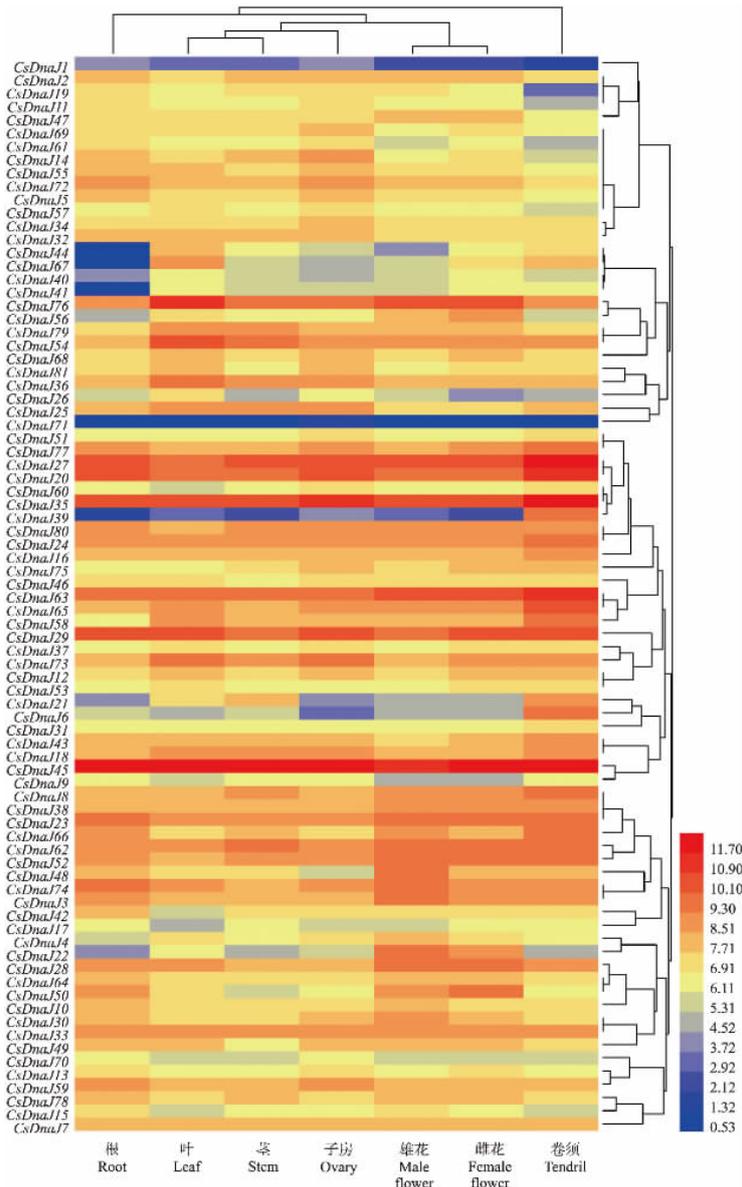


图5 黄瓜 *DnaJ* 基因家族组织表达分析

Fig. 5 Tissue expression analysis of *DnaJ* gene family in cucumber

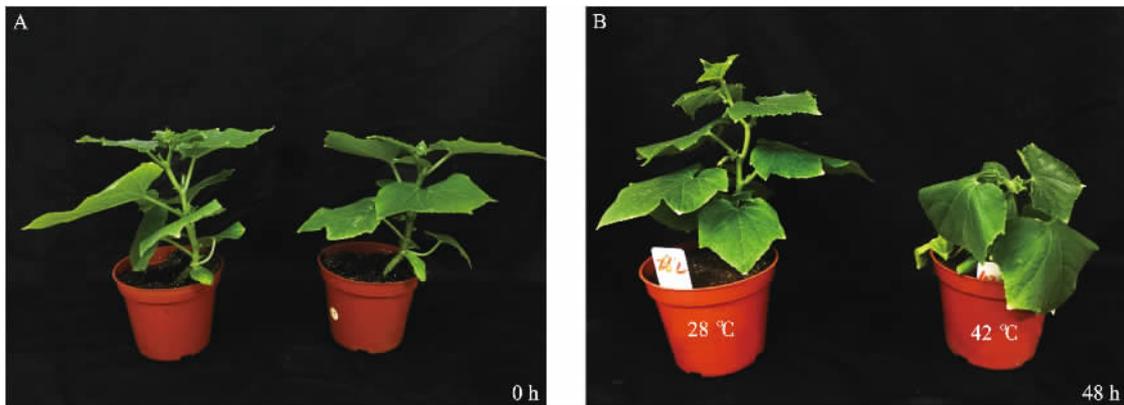
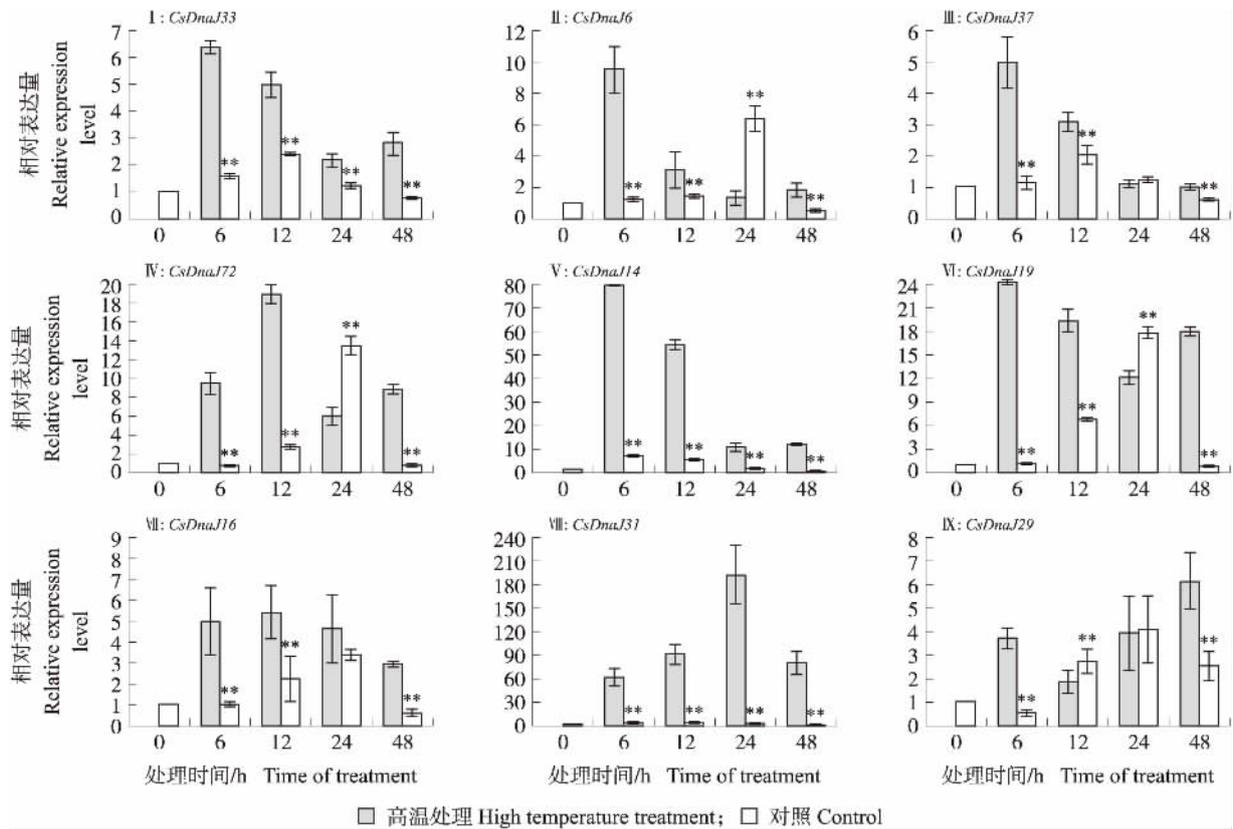


图6 高温胁迫下黄瓜表型观察

Fig. 6 Phenotype observation of cucumber under high temperature stress

A. 黄瓜植株高温处理前生长状态; B. 黄瓜植株经高温处理后生长状态对比。

A. Growth state of cucumber plants before high temperature treatment; B. Comparison of growth state of cucumber plants after high temperature.

图 7 黄瓜 *DnaJ* 基因家族在高温胁迫下表达分析Fig. 7 Expression analysis of *DnaJ* gene family in cucumber under high temperature stress** $P < 0.01$.

3 讨论

DnaJ 蛋白作为热激蛋白家族中的一类蛋白,大部分以胁迫响应因子而广泛存在于植物组织细胞中,在植物面临生物及非生物胁迫响应中扮演重要的角色^[1]。*DnaJ* 蛋白家族基因的表达能够显著提高植物耐热性,降低植物受高温影响造成的细胞损伤^[3]。对 *DnaJ* 蛋白家族基因的鉴定及分析有助于深入研究其在胁迫环境下的调控功能。

本试验利用生物信息学手段在作物黄瓜中鉴定了 81 个 *DnaJ* 基因家族成员,而在拟南芥、水稻、棉花中分别鉴定到 91、101、119 个 *DnaJ* 基因家族成员^[5,17-18],这几种植物在基因家族成员数量及蛋白特性上均存在高度相似性。这些植物的 *DnaJ* 家族成员蛋白序列中均存在由 HPD 三肽核心结构所构成的高保守 J 结构域,但该保守结构域同样存在部分氨基酸差异。由于黄瓜 *DnaJ* 家族蛋白序列长度差异较大,因此对黄瓜及拟南芥 *DnaJ* 家族核心 J 结构域蛋白序列信息进行进化分析,发现黄瓜 *DnaJ* 家族蛋白与拟南芥相似,同样可以分为 9 个亚簇,说明 *DnaJ* 家族基因间进化比 2 个物种间的进化早,且同一簇内蛋白可能行使相似的功能。

同样对黄瓜 *DnaJ* 家族蛋白中其他结构域进行分析,发现黄瓜中存在 G/F 结构、CxxCxCxG 锌指结构以及 *DnaJ*-C 末端结构,在其他植物中根据 *DnaJ* 基因含有结构域的不同可将其分为 3 种类型^[29],而黄瓜 *DnaJ* 蛋白只含有其中 2 种类型,缺乏类型 I 即 4 个结构域都含有的 *DnaJ* 蛋白,这可能会造成物种间对环境温度适应、胁迫反应的差异。*DnaJ* 基因的功能与其所在的亚细胞位置密切相关^[7-8]。本研究中通过对黄瓜 *DnaJ* 蛋白的亚细胞位置预测,发现大部分蛋白定位于细胞核与叶绿体内,目前在植物中对 *DnaJ* 蛋白的研究主要集中在定位于叶绿体的 *DnaJ* 蛋白,其能够参与叶绿体发育、光合作用、蛋白转运等多个进程^[6-8]。黄瓜中 *DnaJ* 蛋白的亚细胞定位预测结果同样说明 *DnaJ* 蛋白对黄瓜叶绿体发育至关重要。

基因复制在基因家族形成、物种适应中起重要作用,并可以为具有新功能的基因提供原始遗传材料^[30]。黄瓜 *DnaJ* 基因家族的共线性分析发现在黄瓜 *DnaJ* 蛋白家族进化过程中同样出现多次基因复制现象,包含串联复制、片段复制及随机复制事件。其中,片段复制出现最多(5 次),说明其在黄瓜 *DnaJ* 基

因家族形成过程中起重要作用,但总体复制现象仅占全部基因的21%(17/81),这说明存在着其他的进化方式。

本试验结果表明,黄瓜 *DnaJ* 基因在植物体内各组织均有表达,说明 *DnaJ* 基因在黄瓜的生长发育过程中发挥着不可缺少的作用。但基因间的表达模式并不相同,这可能是不同的 *DnaJ* 基因在黄瓜器官组织中可能发挥不同的作用模式所造成的。另外,一些基因在各组织中均呈低水平表达,可能是因为这些基因功能冗余或者发生了转录沉默/转录后沉默效应,这种现象及效应在家族基因行使功能中较为常见^[31]。

DnaJ 蛋白可以作为细胞内感应胁迫环境的主要因子,在诸如高温、低温、强光等逆境条件下会大量表达^[32-33],并以此参与植物响应环境胁迫进程。本试验通过定量 PCR 验证了 *DnaJ* 基因在高温环境的表达水平,结果表明高温环境可以诱导 *DnaJ* 基因高表达,因此黄瓜 *DnaJ* 被高温环境所激活,引起表达上调从而参与植物响应胁迫进程。另外,不同的 *DnaJ* 基因在随胁迫环境处理时间所表现出的表达模式、表达差异倍数具有明显差异,这反映了不同 *DnaJ* 基因在胁迫响应进程中调控功能的差异。下一步可通过基因转育或转基因方法将 *DnaJ* 基因响应胁迫功能应用于黄瓜抗胁迫品种的选育中。

参考文献 References:

- [1] 樊芳菲,杨暹,康云艳,等. 植物 DnaJ 蛋白的研究进展[J]. 分子植物育种,2018,16(6):2028-2034.
Fan F F, Yang X, Kang Y Y, et al. Research progress on DnaJ proteins in plants[J]. Molecular Plant Breeding, 2018, 16(6): 2028-2034 (in Chinese with English abstract).
- [2] 祁茂冬,谢鑫,魏凤菊. 禾本科植物 HSP70 研究进展[J]. 植物生理学报,2019,55(8):1054-1062.
Qi M D, Xie X, Wei F J. Research progress of HSP70s in Poaceae[J]. Plant Physiology Journal, 2019, 55(8): 1054-1062 (in Chinese with English abstract).
- [3] 张驰,刘丹丹,刘建中. 植物 J 蛋白的生物学功能及其作用机制[J]. 浙江大学学报(农业与生命科学版),2018,44(3):275-282.
Zhang C, Liu D D, Liu J Z. Biological functions and action mechanisms of J-domain proteins in plants[J]. Journal of Zhejiang University (Agriculture and Life Sciences), 2018, 44(3): 275-282 (in Chinese with English abstract).
- [4] Verma A K, Tamadaddi C, Tak Y, et al. The expanding world of plant J-domain proteins[J]. Critical Reviews in Plant Sciences, 2019, 38(5/6):382-400.
- [5] Miernyk J A. The J-domain proteins of *Arabidopsis thaliana*: an unexpectedly large and diverse family of chaperones[J]. Cell Stress & Chaperones, 2001, 6(3):209-218.
- [6] 孔凡英,邓永胜,孟庆伟. 叶绿体 J 蛋白研究进展[J]. 植物生理学报,2011,47(3):235-243.
Kong F Y, Deng Y S, Meng Q W. Research advancement of J proteins in chloroplast[J]. Plant Physiology Journal, 2011, 47(3): 235-243 (in Chinese with English abstract).
- [7] Wang G D, Kong F Y, Zhang S, et al. A tomato chloroplast-targeted DnaJ protein protects Rubisco activity under heat stress[J]. Journal of Experimental Botany, 2015, 66(11):3027-3040.
- [8] Kong F Y, Deng Y S, Zhou B, et al. A chloroplast-targeted DnaJ protein contributes to maintenance of photosystem II under chilling stress[J]. Journal of Experimental Botany, 2014, 65(1):143-158.
- [9] Wahid A, Gelani S, Ashraf M, et al. Heat tolerance in plants: an overview[J]. Environmental and Experimental Botany, 2007, 61(3):199-223.
- [10] 王改妮,白万鹏,王锁民. 植物响应高温胁迫的信号转导和转录调控机制研究进展[J]. 分子植物育种,2020,18(24):8109-8118.
Wang G N, Bai W P, Wang S M. Advances in research of signal transduction and transcriptional regulatory mechanism of plants in response to heat stress[J]. Molecular Plant Breeding, 2020, 18(24): 8109-8118 (in Chinese with English abstract).
- [11] 陈雪倩,韩静,任中海. 黄瓜响应高温胁迫转录组分析[J]. 分子植物育种. <http://kns.cnki.net/kcms/detail/46.1068.S.20200511.1658.013.html>.
Chen X Q, Han J, Ren Z H. Transcriptome analysis of cucumber in response to high temperature stress[J]. Molecular Plant Breeding. <http://kns.cnki.net/kcms/detail/46.1068.S.20200511.1658.013.html>.
- [12] 郝婷,丁小涛,金海军,等. 高温对黄瓜生长的影响及设施栽培对策[J]. 上海农业学报,2014,30(2):125-131.
Hao T, Ding X T, Jin H J, et al. Effects of high temperature on cucumber growth and countermeasures in protected culture[J]. Acta Agriculturae Shanghai, 2014, 30(2): 125-131 (in Chinese with English abstract).
- [13] 徐晓响. 高温胁迫对黄瓜叶片光合特性的影响[D]. 兰州:甘肃农业大学,2007.
Xu X Y. Effects of high temperature on photosynthetic characteristics in cucumber leaves[D]. Lanzhou: Gansu Agricultural University, 2007 (in Chinese with English abstract).
- [14] Baninasab B, Ghobadi C. Influence of paclobutrazol and application methods on high-temperature stress injury in cucumber seedlings[J]. Journal of Plant Growth Regulation, 2011, 30(2):213-219.
- [15] Ding X T, Jiang Y P, He L Z, et al. Exogenous glutathione improves high root-zone temperature tolerance by modulating photosynthesis, antioxidant and osmolytes systems in cucumber seedlings[J]. Scientific Reports, 2016, 6:35424.

- [16] 孙亚玲,臧传江,姚甜甜,等. 我国黄瓜耐热性研究进展[J]. 中国果菜,2018,38(10):57-62.
Sun Y L,Zang C J,Yao T T,et al. Research advances on heat tolerance of cucumber in China [J]. China Fruit & Vegetable,2018,38(10):57-62 (in Chinese with English abstract).
- [17] 赵旭,刘向国,王爱荣. 水稻 *DnaJ* 蛋白的生物信息学分析[J]. 亚热带农业研究,2011,7(3):206-211.
Zhao X,Liu X G,Wang A R. Bioinformatic analysis of the *DnaJ* protein family in rice [J]. Subtropical Agriculture Research,2011,7(3):206-211 (in Chinese with English abstract).
- [18] 徐珍珠,郭琪,刘静,等. 雷蒙德氏棉 *DnaJ* 蛋白的生物信息学预测及盐胁迫下的表达分析[J]. 棉花学报,2015,27(5):391-400.
Xu Z Z,Guo Q,Liu J,et al. Genome-wide identification and expression pattern analysis of the *DnaJ* protein family in *Gossypium raimondii* under salt stress using a bioinformatics method [J]. Cotton Science,2015,27(5):391-400(in Chinese with English abstract).
- [19] Huang S W,Li R Q,Zhang Z H,et al. The genome of the cucumber, *Cucumis sativus* L. [J]. Nature Genetics,2009,41(12):1275-1281.
- [20] Finn R D,Coggill P,Eberhardt R Y,et al. The Pfam protein families database: towards a more sustainable future [J]. Nucleic Acids Research, 2016,44(D1):D279-D285.
- [21] 崔梦杰,王晨,冷翔鹏,等. 葡萄 SBP-box 转录因子家族的生物信息学分析及其应答激素调控葡萄果实成熟的作用[J]. 南京农业大学学报,2018,41(3):429-439. DOI: 10.7685/jnau.201706045.
Cui M J,Wang C,Leng X P,et al. Bioinformatic analysis of the SBP-box transcription factor and function of their regulation modes on berry ripening process by responding to hormones [J]. Journal of Nanjing Agricultural University,2018,41(3):429-439(in Chinese with English abstract).
- [22] Marchler-Bauer A,Bo Y,Han L Y,et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures [J]. Nucleic Acids Research,2017,45(D1):D200-D203.
- [23] Bailey T L,Boden M,Buske F A,et al. MEME Suite: tools for motif discovery and searching [J]. Nucleic Acids Research,2009,37(Suppl 2): W202-W208.
- [24] Kumar S,Stecher G,Tamura K. MEGA 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets [J]. Molecular Biology and Evolution,2016,33(7):1870-1874.
- [25] Wang Y P,Tang H B,DeBarry J D,et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity [J]. Nucleic Acids Research,2012,40(7):e49.
- [26] Li Z,Zhang Z H,Yan P C,et al. RNA-Seq improves annotation of protein-coding genes in the cucumber genome [J]. BMC Genomics,2011,12(1):1-11.
- [27] Patro R,Duggal G,Kingsford C. Salmon: accurate, versatile and ultrafast quantification from RNA-seq data using lightweight-alignment [J]. bioRxiv,2015. DOI: 10.1101/021592.
- [28] Song M F,Zhang M R,Cheng F,et al. An irregularly striped rind mutant reveals new insight into the function of PG1 β in cucumber(*Cucumis sativus* L.) [J]. Theoretical and Applied Genetics,2020,133(2):371-382.
- [29] Fan F F,Yang X,Cheng Y,et al. The *DnaJ* gene family in pepper(*Capsicum annuum* L.): comprehensive identification, characterization and expression profiles [J]. Frontiers in Plant Science,2017,8:689.
- [30] Cannon S B,Mitra A,Baumgarten A,et al. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana* [J]. BMC Plant Biology,2004,4:10.
- [31] Jury J A,Frayne J,Hall L. Sequence analysis of a variety of primate fertilin alpha genes: evidence for non-functional genes in the *Gorilla* and man [J]. Molecular Reproduction and Development,1998,51(1):92-97.
- [32] Scarpeci T E,Zanor M I,Carrillo N,et al. Generation of superoxide anion in chloroplasts of *Arabidopsis thaliana* during active photosynthesis: a focus on rapidly induced genes [J]. Plant Molecular Biology,2008,66(4):361-378.
- [33] Piippo M,Allahverdiyeva Y,Paakkari V,et al. Chloroplast-mediated regulation of nuclear genes in *Arabidopsis thaliana* in the absence of light stress [J]. Physiological Genomics,2006,25(1):142-152.

责任编辑: 范雪梅