

研究报告

Research Report

香蕉广谱胁迫蛋白(USP)基因的鉴定与表达分析

王霞^{1,2} 林秋妹² 赵东方² 贾彩红² 王静毅² 刘菊华² 黄绵佳¹ 从心黎^{1*} 王卓^{2*}

1 海南大学生命科学与药学院, 海口, 570228; 2 中国热带农业科学院热带生物技术研究所, 农业部热带作物生物学与遗传资源利用重点实验室, 海口, 571101

* 通信作者, wangzhuocd@126.com, cong0890@163.com

摘要 广谱胁迫蛋白(Universal Stress Protein, USP)是生物体中一类保守的应激反应蛋白, 参与响应生物和非生物胁迫, 但在植物中 USP 的功能研究较少。本研究在香蕉 A 基因组中鉴定 53 个 *MaUSPs* 基因, 系统发育分析结果表明 *MaUSPs* 可划分为 A、B、C、D 四个亚家族, 并且都具有 USP 保守结构域。染色体定位结果表明 52 个 *MaUSPs* 定位在 11 条染色体上, 将其命名为 *MaUSP01* 至 *MaUSP52*, *Ma00_t00840.1* 未锚定在染色体上, 命名为 *MaUSP53*。基因复制事件表明片段重复是 *MaUSPs* 家族成员扩张的主要途径。RNA 测序(RNA-seq)数据表明, 大多数 *MaUSPs* 在香蕉果实发育和采后成熟过程中组成型表达; *MaUSP08*、*MaUSP47* 和 *MaUSP49* 在低温、干旱、盐胁迫下显著差异表达; 香蕉接种 Foc TR4 后, *MaUSP03*、*MaUSP25*、*MaUSP39*、*MaUSP06*、*MaUSP08*、*MaUSP24*、*MaUSP17*、*MaUSP48* 和 *MaUSP49* 基因在感病品种巴西蕉中均显著下调, 而在抗病品种 GCTCV-119 中均显著上调, 上述结果说明这些 *MaUSPs* 可能是参与香蕉果实品质形成和响应逆境胁迫的过程中重要的候选基因。本研究结果为进一步研究 *MaUSPs* 基因家族重要成员在香蕉果实品质形成和响应逆境胁迫过程中的功能解析提供基础。

关键词 香蕉, 广谱胁迫蛋白(Universal Stress Protein), 非生物胁迫, 生物胁迫, 基因表达

Identification and Expression of Universal Stress Protein (USP) Genes in Banana (*Musa acuminata* L. AAA group, cv. Cavendish)

Wang Xia^{1,2} Lin Qiumei² Zhao Dongfang² Jia Caihong² Wang Jingyi² Liu Juhua² Huang Mianjia¹ Cong Xinli^{1*} Wang Zhuo^{2*}

1 College of Life Science and Pharmacy, Hainan University, Haikou, 570228; 2 Key Laboratory of Tropical Crop Bioscience and Biotechnology, Ministry of Agriculture, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou, 571101

* Corresponding author, wangzhuocd@126.com, cong0890@163.com

DOI: 10.5376/mpb.cn.2020.18.0019

Abstract The universal stress proteins (Universal Stress Protein, USP) are conserved stress-responsive proteins in various organisms that are involved in response to biotic and abiotic stresses, but the function of most USP in plants is unknown. In this study, there are identified 53 genes in the *Musa acuminata* genome. The phylogenetic analysis indicated that these *MaUSPs* were divided into four subfamilies which the subfamily A, B, C and D. All *MaUSPs* have USP conserved domains. The distribution mapping results showed that a total of 52 *MaUSPs* were located on the 11 chromosomes in *Musa acuminata* genome, which were named *MaUSP01* to *MaUSP52*, based on their chromosomal locations. *Ma00_t00840.1* was not anchored on the chromosome, and named *MaUSP53*. The gene duplication events showed that six pairs of *MaUSPs* were located in collinear fragments, and segment

收稿日期: 2020 年 11 月 22 日; 接受日期: 2021 年 1 月 5 日; 发表日期: 2021 年 1 月 12 日

引用格式: 王霞, 林秋妹, 赵东方, 贾彩红, 王静毅, 刘菊华, 黄绵佳, 从心黎, 王卓, 2020, 香蕉广谱胁迫蛋白(USP)基因的鉴定与表达分析, 分子植物育种(网络版), 19(1): 1-10 (doi: 10.5376/mpb.cn.2021.19.0001) (Wang X., Lin Q.M., Zhao D.F., Jia C.H., Wang J.Y., Liu J.H., Huang M.J., Cong X.L., and Wang Z., 2020, Identification and expression of universal stress protein (USP) genes in banana (*Musa acuminata* L. AAA group, cv. cavendish), *Fengzi Zhiwu Yuzhong* (Molecular Plant Breeding (online)), 19(1): 1-10 (doi: 10.5376/mpb.cn.2021.19.0001))

duplications might be the only way for the expansion of MaUSPs. RNA sequencing (RNA-seq) data indicate that most MaUSPs showed constitutive expression during banana fruit development and ripening-specification. Under drought, low temperature and salt stress, *MaUSP08*, *MaUSP47* and *MaUSP49* were significantly expressed that may be involved in banana response to abiotic stress. Under Foc TR4 stress, *MaUSP03*, *MaUSP25*, *MaUSP39*, *MaUSP06*, *MaUSP08*, *MaUSP24*, *MaUSP17*, *MaUSP48* and *MaUSP49* genes were significantly down-regulated in Cavendish and up-regulated in GCTCV-119, indicating that *MaUSPs* may play an important role in banana resistance to banana wilt. These results suggest that these *MaUSPs* may be important candidate genes in banana fruit quality formation and response to stresses. This study providing a basis for the function analysis the roles of *MaUSPs* gene in banana during fruit quality formation and response to stresses.

Keywords Banana, Universal stress proteins, Abiotic stress, Biotic stress, Gene expression

香蕉(*Musa spp.*)是热带和亚热带地区的粮食和经济作物之一。香蕉果实富含蛋白质和碳水化合物,具有重要的经济价值,对世界各地人民的收入有巨大贡献(Aurore et al., 2009)。香蕉在生长发育过程中会受到许多非生物胁迫和生物胁迫的影响,如干旱(Mahajan and Tuteja, 2005)、低温(Kang et al., 2013)、盐(Munns, 2005)和香蕉枯萎病(D'Hont et al., 2012)等,从而影响香蕉果实的产量和品质。香蕉枯萎病是香蕉生产中的瓶颈问题,是由尖孢镰刀菌古巴专化型(*Fusarium oxysporum* f. sp. *Cubense*, Foc)引起的毁灭性土传维管束病害,其中4号生理小种(Foc TR4)对香蕉主栽品种巴西蕉(*M. acuminata* AAA group cv. Cavendish)的危害最大(Ploetz, 2006; Visser et al., 2010)。尽管已经通过筛选体细胞组培变异获得抗病性较强的突变体 GCTCV-119 (*M. acuminata* L. AAA group, cv. GCTCV-119) (Hwang and Ko, 2004),但任然无法满足当前香蕉产业的需求,且目前还没有有效的预防和防控香蕉枯萎病的措施(Ploetz, 2015),因此挖掘重要的抗胁迫相关基因是当前重要的工作。

广谱胁迫蛋白 (Universal Stress Protein, USP)是一种古老而保守的蛋白质。USP 不仅存在于细菌的基因组中,也存在于古生菌、真菌、原生物、植物和动物的基因组中,其表达受到多种胁迫的影响(Vanbogelen et al., 1990; Sauter et al., 2002; Kvint et al., 2003; Aravind et al., 2010; Foret et al., 2011; Vollmer and Bark, 2018)。USP 最早在大肠杆菌(*Escherichia coli*) W3110 K-12 菌株中被发现(Vanbogelen et al., 1990)。USP 属于丝氨酸和苏氨酸的自磷酸化蛋白,是GTP和ATP的磷酸盐供体(Freestone et al., 1998)。在大肠杆菌中,USP 基因被分为6组:*uspA*、*uspC*、*uspD*、*uspE*、*uspF*和*uspG* (Gustavsson and Nyström, 2002)。根据序列的相似性,USP 基因可分为两个亚家族:*uspA*亚家族包括 *uspA*、*uspC* 和 *uspD*; *uspFG* 亚家族包括

*uspF*和*uspG*。在植物中,将拥有与细菌相似的USPA结构域(大肠杆菌USPA)的基因定义为USP基因(Kerk et al., 2003),在植物中,USP的功能研究较少。

广义上USPs可分为两组:第一组USPs与没有ATP结合活性的流感嗜血杆菌(*Haemophilus influenzae*)的*uspA*结构相似;第二组USPs与詹氏甲烷球菌(*Methanococcus jannaschii*)中具有ATP结合活性的*uspFG*型蛋白结构相似。USP蛋白最小的只有一个USP结构域;也有包含两个USP结构域串联重复或与其他功能结构域(如Na⁺/H⁺交换剂,氨基酸渗透酶和蛋白激酶)一起存在,形成较大的蛋白(Kvint et al., 2003)。

迄今为止,在不同植物中已有多个USP基因被研究,如拟南芥(*Arabidopsis thaliana*)、番茄(*Lycopersicon esculentum*)、水稻(*Oryza sativa*)、棉花(*Gossypium spp.*)、黄芪(*Astragalus sinicus*)、大麦(*Hordeum vulgare*)、丹参(*Salvia miltiorrhiza*)、木豆(*Cajanus cajan*)和烟草(*Nicotiana tabacum*) (Zegzouti et al., 1999; Sauter et al., 2002; Kerk et al., 2003; Chou et al., 2007; Li et al., 2010; Merkourouopoulos and Tsiftaris, 2013; Sinha et al., 2016; Udawat et al., 2016; Gutierrez-Beltran et al., 2017; Wang et al., 2017)。研究表明,USP基因可提高在胁迫下的细胞存活率,从而提高植物抗逆性(Zahur et al., 2009)。在番茄中,ER6 (Ethylene-regulated gene)与USP高度相似,可以被乙烯诱导(Zegzouti et al., 1999)。水稻*OsUSP1*在水分胁迫和乙烯处理中上调表达(Sauter et al., 2002)。在紫云英中,USP基因在感染根瘤菌的根中被诱导表达(Chou et al., 2007)。亚洲棉的GUSP1和GUSP2可以响应水分胁迫(Zahur et al., 2009)。在木豆(*Cajanus cajan* L.)中发现了4个类似*uspA*的基因响应干旱胁迫(Sinha et al., 2016)。拟南芥AtUSP重组蛋白具有抗真菌活性,能够抑制多种病原真菌的生长(Jung et al., 2015)。在烟草中过

表达苜蓿 *MfUSP1* 可增强其对低温、干旱、盐等非生物胁迫的耐受性(Gou et al., 2020)。过表达 *SbUSP* (*Salicornia brachiata*) 可增强烟草的耐盐性和耐旱性(Udawat et al., 2016)。到目前为止,香蕉的 USP 基因还未见报道。

香蕉 A 基因组已经完成测序,为 *MaUSPs* 基因全基因组分析提供了数据支持(France et al., 2012)。由于 *USP* 基因已被证明与许多植物物种的抗逆性有关,本研究对 *MaUSPs* 基因家族进行系统鉴定,分析其在果实发育和成熟以及胁迫响应中的作用,为进一步针对这些基因作为遗传改良靶点提高香蕉果实品质和抗逆能力提供了遗传资源。

1 结果与分析

1.1 *MaUSP* 基因的鉴定和分类

利用香蕉 A 基因组网站获得全部 *MaUSPs* 成员的蛋白序列,使用 SMART 对其进行蛋白结构域预测,共鉴定出 53 个香蕉 USP 基因(*MaUSPs*)。根据其在染色体上的位置将其命名为 *MaUSP01* 至 *MaUSP52*, *Ma00_t00840.1* 未锚定在染色体上,命名为 *MaUSP53*。详细的 *MaUSP* 蛋白序列位置(表 1)。*MaUSP* 蛋白的长度包含 158~817 个氨基酸不等;等电点分析表明 *MaUSP* 蛋白序列等电点差异明显,分布在 4.88~10.95 之间;分子量分布在 16.9~90.89 kD 之间(表 1)。亚细胞预测结果表明 *MaUSPs* 分别定位于细胞核(20 个)、细胞质(14 个)、叶绿体(9 个)、线粒体(6 个)、细胞外(3 个)、细胞质膜(1 个),说明不同 *MaUSPs* 基因成员间可能具有不同的功能(表 1)。

1.2 *MaUSPs* 系统发育分析

为研究 *MaUSPs* 的进化关系,以拟南芥的 26 个 USP 蛋白序列与香蕉 53 个 USP 蛋白序列构建系统进化树。利用 MEGA6.0 软件中的比邻法构建系统进化树,采用 poisson model 计算进化距离。结果显示, *MaUSPs* 可分为 4 个亚家族:其中 A 亚家族包含 24 个成员,是最大的亚家族;B 亚家族最小,只包含一个成员;C 亚家族和 D 亚家族,分别有 16 和 12 个 *MaUSPs* (图 1)。

1.3 USP 结构域中的保守氨基酸残基

使用 SMART 预测发现 53 个 *MaUSPs* 蛋白序列都含有 USP 结构域。根据 MJ-0577 的晶体结构,通过 ClustalW align 和 weblog visualization 发现 6 个保守结构域,如保守的 ATP 腺嘌呤(A)、ATP 磷酸(P)或

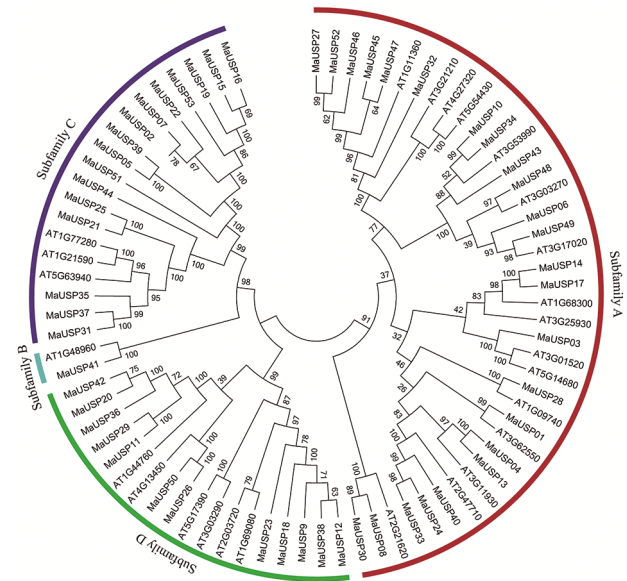


图 1 *MaUSPs* 的系统进化树

Figure 1 Phylogenetic tree of the *MaUSPs*

ATP 核糖(R)或位于二聚结构域(D)(图 2A;图 2B;图 2C;图 2D)。在这些 USP 结构域中, *MaUSPs* 的保守程度不高。根据蛋白结构, *UspA* 蛋白家族主要分为两个亚群:一个含有 ATP 结合位点结构域,另一个无 ATP 结合位点结构域(Tkaczuk et al., 2013)。此外,并非所有 *MaUSPs* 基因都含有甘氨酸(G),在 53 个 *MaUSPs* 的 USP 结构域的 168 个氨基酸位点上,只有 23 个位点具有完整的 ATP 结合保守结构域的特异结合序列(G-2x-G-9x-G[S/T])(图 2E)。

1.4 *MaUSPs* 的染色体定位和基因重复

根据 *MaUSPs* 在香蕉 A 基因组数据库中的位置确定其在染色体上的相对分布。染色体定位分析表明,共有 52 个 *MaUSPs* 分布在 11 条染色体上。*MaUSPs* 分布在 11 条香蕉染色体上(图 3),但其分布是不均匀的;其中 2 号染色体上分布的 *MaUSPs* 数量最多,3 号和 6 号染色体上有 8 个 *MaUSPs*,5 号、9 号和 11 号染色体上分别有 4 个 *MaUSPs*,1 号和 2 号染色体上有 3 个 *MaUSPs*,还有 2 个分别位于 10 号和 7 号染色体上。

片段重复、串联重复和反转录是基因家族扩增的关键因素(Hurles, 2004; Freeling, 2009)。根据香蕉 A 基因组的节段性片段信息,在本研究中发现 12 个共线的 *MaUSPs* 位于共线区,属于片段重复(图 3);在系统进化树中,共线区大部分共线 *MaUSPs* 属于同一类群; *MaUSP48* 和 *MaUSP28* 在系统进化树中分布在不同的位置(图 1),有可能是因为 *MaUSP48* 和

表 1 MaUSPs 基因家族成员的基本信息

Table 1 Basic information of the MaUSPs genes

基因 ID	蛋白 ID	染色体编号	起始位置	终止位置	大小(aa)	等电点	分子量(kD)	亚细胞定位
Gene ID	Protein ID	ChromosomeNumber	Gene start	Gene end	Size (aa)	pI	MW (kD)	Subcellular localization
MaUSP01	Ma01_p17560.1	chr01	12865234	12866097	192	6.21	21.02	Cytoplasmic
MaUSP02	Ma01_p18690.1	chr01	13993019	13997192	737	6.95	81.38	Nuclear
MaUSP03	Ma01_p20470.1	chr01	18756079	18766107	183	5.88	20.56	Nuclear
MaUSP04	Ma02_p10680.1	chr02	20167925	20168710	187	6.44	20.46	Extracellular
MaUSP05	Ma02_p17530.1	chr02	24389651	24394839	755	8.43	84.96	Nuclear
MaUSP06	Ma02_p23920.1	chr02	28726254	28727072	159	8.57	17.34	Extracellular
MaUSP07	Ma03_p13520.1	chr03	10653936	10656958	741	7.88	81.98	Nuclear
MaUSP08	Ma03_p15600.1	chr03	16525280	16526124	171	6.06	19.04	Cytoplasmic
MaUSP09	Ma03_p17280.1	chr03	22500965	22501781	217	10.14	23.8	Mitochondrial
MaUSP10	Ma03_p18710.1	chr03	24245233	24250039	166	5.77	18.29	Cytoplasmic
MaUSP11	Ma03_p19910.1	chr03	25174234	25174896	169	8.72	18.62	Nuclear
MaUSP12	Ma03_p26460.1	chr03	30176872	30178022	223	10.38	24.91	Mitochondrial
MaUSP13	Ma03_p26810.1	chr03	30404820	30405593	179	6.96	19.34	Mitochondrial
MaUSP14	Ma03_p30690.1	chr03	33211461	33214328	158	5.93	16.9	Cytoplasmic
MaUSP15	Ma04_p04660.1	chr04	3571046	3574467	752	8.23	83.31	Nuclear
MaUSP16	Ma04_p05780.1	chr04	4315798	4319003	688	8.96	76.01	Nuclear
MaUSP17	Ma04_p06800.1	chr04	4957383	4963299	161	5.95	17.36	Chloroplast
MaUSP18	Ma04_p10790.1	chr04	7651049	7652071	281	10.06	31.29	Nuclear
MaUSP19	Ma04_p11510.1	chr04	8121247	8125592	756	6.58	84.32	Nuclear
MaUSP20	Ma04_p15910.1	chr04	14537255	14538119	224	9.76	24.71	Nuclear
MaUSP21	Ma04_p22180.1	chr04	24542381	24545047	621	5.35	69.26	Cytoplasmic
MaUSP22	Ma04_p27320.1	chr04	28516538	28520468	734	8.58	81.99	Nuclear
MaUSP23	Ma04_p30030.1	chr04	30789361	30790265	228	10.95	25.05	Nuclear
MaUSP24	Ma05_p00170.1	chr05	122344	124687	170	6.35	18.09	Cytoplasmic
MaUSP25	Ma05_p21310.1	chr05	32950978	32957757	817	9.41	90.89	Nuclear
MaUSP26	Ma05_p22960.1	chr05	34731415	34732775	199	9.58	21.7	Chloroplast
MaUSP27	Ma05_p28710.1	chr05	39626842	39631219	235	4.91	25.48	Cytoplasmic
MaUSP28	Ma06_p03910.1	chr06	2832207	2835571	174	5.45	18.1	Chloroplast
MaUSP29	Ma06_p08940.1	chr06	6265563	6266507	200	9.79	22.16	Nuclear
MaUSP30	Ma06_p10620.1	chr06	7329824	7330988	185	6.25	20.93	Cytoplasmic
MaUSP31	Ma06_p20010.1	chr06	14080857	14084225	788	5.92	86.23	Nuclear
MaUSP32	Ma06_p20450.1	chr06	14544054	14546008	224	9.95	24.35	Nuclear
MaUSP33	Ma06_p21660.1	chr06	15991119	15993697	172	6.34	18.38	Cytoplasmic
MaUSP34	Ma06_p24260.1	chr06	22730901	22745993	160	8.58	17.45	Mitochondrial
MaUSP35	Ma06_p27650.1	chr06	29578350	29581407	706	9	77.03	Chloroplast
MaUSP36	Ma07_p03090.1	chr07	2407887	2408731	200	9.58	22.17	Nuclear
MaUSP37	Ma07_p24750.1	chr07	32008326	32011649	753	5.69	81.5	Chloroplast
MaUSP38	Ma08_p04250.1	chr08	2964186	2965421	227	10.22	25.02	Mitochondrial
MaUSP39	Ma08_p05530.1	chr08	3755772	3758814	779	7.61	85.69	Cytoplasmic
MaUSP40	Ma08_p06480.1	chr08	4376752	4378952	160	6.7	17.1	Cytoplasmic
MaUSP41	Ma08_p16600.1	chr08	19622604	19623973	196	9.98	21.8	PlasmaMembrane
MaUSP42	Ma08_p32810.1	chr08	43393027	43393992	224	9.12	24.27	Nuclear
MaUSP43	Ma09_p03760.1	chr09	2488807	2491334	165	8.43	18.4	Mitochondrial
MaUSP44	Ma09_p19680.1	chr09	25944173	25946801	605	6.98	66.75	Chloroplast

续表 1

Continuing table 1

基因 ID	蛋白 ID	染色体编号	起始位置	终止位置	大小(aa)	等电点	分子量(kD)	亚细胞定位
Gene ID	Protein ID	ChromosomeNumber	Gene start	Gene end	Size (aa)	pI	MW (kD)	Subcellular localization
MaUSP45	Ma09_p25070.1	chr09	36652532	36654893	264	5.02	28	Chloroplast
MaUSP46	Ma09_p30600.1	chr09	40661509	40664376	248	5.06	26.84	Chloroplast
MaUSP47	Ma10_p08820.1	chr10	23098457	23102881	268	5.08	28.28	Chloroplast
MaUSP48	Ma10_p29490.1	chr10	36056732	36057468	162	5.76	17.52	Cytoplasmic
MaUSP49	Ma11_p12790.1	chr11	16685814	16688743	162	5.8	17.73	Cytoplasmic
MaUSP50	Ma11_p15730.1	chr11	21410309	21411101	203	8.55	21.71	Extracellular
MaUSP51	Ma11_p19350.1	chr11	24243223	24250989	786	5.71	87.96	Nuclear
MaUSP52	Ma11_p24310.1	chr11	27316596	27319064	248	4.88	27.05	Cytoplasmic
MaUSP53	Ma00_t00840.1	-	-	-	736	8.49	82	Nuclear

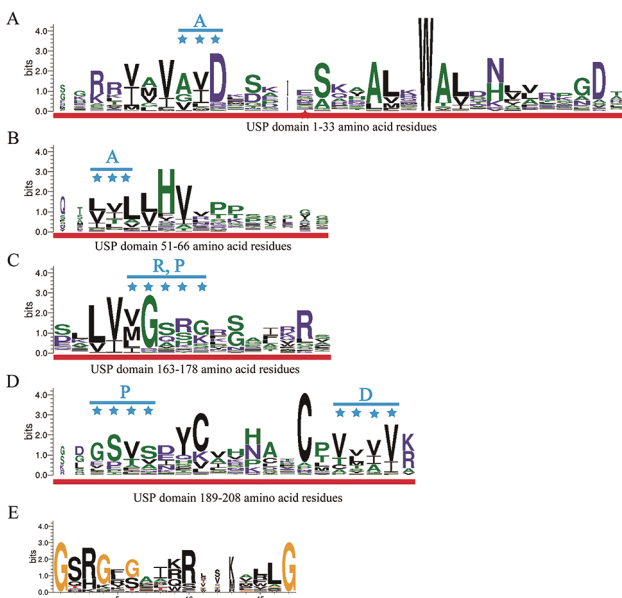


图 2 MaUSPs 中 USP 结构域氨基酸残基

注: A~D: 每个堆栈的总高度表示序列在该位置的守恒; 序列上方蓝色条纹表示来源于 MJ-0577 晶体结构的 ATP 腺嘌呤(A), ATP 磷酸(P), ATP 核糖(R)和二聚结构域(D)残基; E: 23 个 MaUSPs 的 ATP 结合基序(G-2X-G-9X-G[S/T])

Figure 2 Amino acid residues of the USP domain in MaUSPs

Note: A~D: The overall height of each stack represents the conservation of the sequence at that position; The blue stars above the sequences indicate residues that adenine of ATP (A), phosphate of ATP (P), ribose of ATP (R) and dimerization domain (D) as derived from the crystal structure of MJ-0577; E: ATP-binding motif (G-2X-G-9X-G[S/T]) in 23 MaUSPs

MaUSP28 在很早以前就已经进行了基因重复, 这些基因的序列和功能在进化过程中发生了显著的分化; 本研究未发现属于串联复制和反转录的成员。上述结果表明在香蕉 A 基因组进化过程中, MaUSPs 扩增主要是通过片段重复。

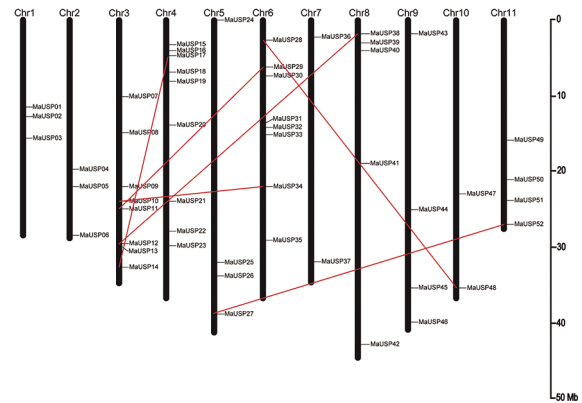


图 3 MaUSPs 基因组染色体的分布及共线分析

注: MaUSPs 的位置用垂直的黑线表示; 存在复制关系的 MaUSPs 用红线连接

Figure 3 Distribution and synteny analysis of MaUSPs chromosomes

Note: The locations of the MaUSPs are indicated by vertical black lines; The duplicate MaUSPs are connected with red lines

1.5 MaUSPs 在果实发育和采后成熟中的表达谱

对巴西蕉果实发育和采后成熟的 5 个时间点 0 DAF (Day After Flower)、20 DAF、80 DAF、0 DPH (Day After Harvest)、8 DPH 和 14 DPH 的 RNA-seq 数据进行分析, 过滤掉上述 5 个时间点 RPKM 均小于 5 的 MaUSPs 成员。我们获得 28 个 RPKM>5 的 MaUSPs 成员(图 4), 占 52.8%, 表明 MaUSPs 基因家族成员在香蕉果实发育和采后成熟过程中表达活跃。其中, 在果实发育阶段 (0 DAF、20 DAF 和 80 DAF), MaUSP49、MaUSP33、MaUSP46、MaUSP17、MaUSP24、MaUSP14、MaUSP28、MaUSP03、MaUSP47、MaUSP48、MaUSP45、MaUSP29 的 RPKM 值均大于 20, 最高的 RPKM 值为 339; 这些 MaUSPs 成员在果实采后成熟阶段(0 DPH, 8 DPH 和 14 DPH) 显著表达, RPKM 值最高达

855, 说明这些 *MaUSPs* 成员在果实发育和采后成熟阶段组成型表达;

进一步分析发现 *MaUSP49*、*MaUSP24* 的表达量在香蕉果实发育和采后成熟阶段均显著表达(RP-KM>55), 说明这些成员可能在香蕉果实和采后成熟过程中起着重要的作用。值得关注的是, 我们还发现 *MaUSP17*、*MaUSP29* 基因在果实采后成熟过程中特异高表达, 尤其在 8 DPH 和 14 DPH 两个时间点(RP-KM>800) (图 4), 说明 *MaUSP17*、*MaUSP29* 可能特异地参与香蕉果实采后成熟过程。

1.6 *MaUSPs* 在香蕉幼苗经低温、干旱、盐处理后的表达谱

在低温处理(Cold, 4°C)条件下, 8 个 *MaUSPs* 差异表达, 其中 *MaUSP05*、*MaUSP08*、*MaUSP47*、*MaUSP49* 显著上调($\log_2(\text{Cold}/\text{Control}) > 1$), *MaUSP14*、*MaUSP31*、*MaUSP03*、*MaUSP35* 显著下调。在干旱处理(Osmoti, 200 mmol/L mannitol) 条件下, 10 个 *MaUSPs* 差异表达($\log_2(\text{Cold}/\text{Control}) > 1$), 其中 *MaUSP17*、*MaUSP29*、*MaUSP28*、*MaUSP08*、*MaUSP47*、*MaUSP49* 显著上调, *MaUSP48*、*MaUSP31*、*MaUSP03*、*MaUSP35* 显著下调(图 5)。在盐 (Salt, 300 mmol/L NaCl) 处理后, 7 个 *MaUSPs* 差异表达 ($\log_2(\text{Cold}/\text{Control}) > 1$), 其中 *MaUSP17*、*MaUSP29*、*MaUSP08*、*MaUSP47*、*MaUSP49* 显著上调, *MaUSP48*、*MaUSP03* 显著下调; *MaUSP08*、*MaUSP47* 和 *MaUSP49* 在低温、干旱、盐胁迫下都显

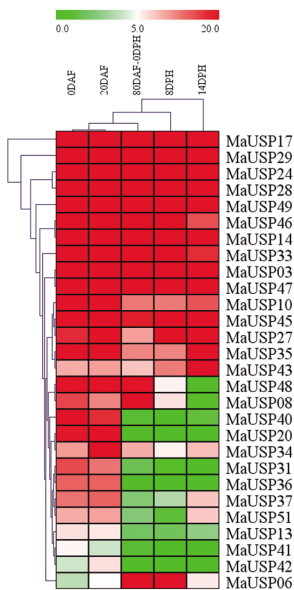


图 4 *MaUSPs* 在果实发育和成熟不同阶段的表达模式

Figure 4 Expression patterns of *MaUSPs* in different stages of fruit development and ripening

著表达; 上述结果表明 *MaUSPs* 可能参与响应香蕉抵抗逆境胁迫的过程。

1.7 *MaUSPs* 在香蕉幼苗经 Foc TR4 处理后的表达谱

12 个 *MaUSPs* 在抗病品种 GCTCV-119 (作图简称为 GCT) 和感病品种巴西蕉 (作图简称为 BX) 接种 Foc TR4 之后差异表达, 其中在巴西蕉 2DPI 时, 12 个 *MaUSPs* 基因全部显著下调; GCTCV-119 2DPI 时, 除 *MaUSP03*、*MaUSP25* 和 *MaUSP39* 下调表达, *MaUSP03*、*MaUSP25*、*MaUSP39*、*MaUSP06*、*MaUSP08*、*MaUSP24*、*MaUSP17*、*MaUSP48* 和 *MaUSP49* 基因均显著上调(图 6)。上述结果表明, *MaUSP03*、*MaUSP25*、*MaUSP39*、*MaUSP06*、*MaUSP08*、*MaUSP24*、*MaUSP17*、*MaUSP48* 和 *MaUSP49* 可能参与香蕉抗香蕉枯萎病的过程。

2 讨论

USP 是在各种生物体中保守的应激反应蛋白, 参与响应生物和非生物胁迫。UspA 超家族最早在大肠杆菌中被发现, 在多种应激和胁迫条件下, 其表达量急剧增加 (Vanbogelen et al., 1990)。在植物基因组中, USPs 的数量一般只有 20~50 个成员, 如在木豆、拟南芥、丹参和大麦中分别鉴定出 51 个、42 个、32 个和 16 个 USPs, 而在甘蓝型油菜 (*Brassica napus*) 中发现 142 个 USP 成员。在香蕉中我们鉴定出 53 个 USPs, 数量仅次于甘蓝型油菜。MaUSP 蛋白的长度包含 158~817 个氨基酸不等, 等电点分布在 4.88~10.95 之间, 分子量分布在 16.9~90.89 kD 之间(表 1), 且所有的 *MaUSPs* 都包含 USP 结构域; 这些特征与其他植物 USP 蛋白的特性一致 (Zahur et al., 2009; Tkaczuk et al., 2013)。根据其结合 ATP 的能力将 *uspA* 家族成员分为两大类 (Tkaczuk et al., 2013), 本研究在

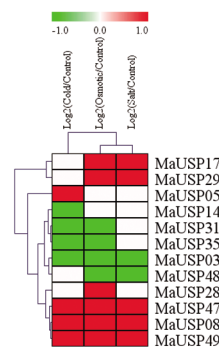


图 5 *MaUSPs* 在香蕉幼苗经低温, 干旱, 盐处理后的表达

Figure 5 The expression of *MaUSPs* in banana seedlings responding to osmotic, salt and cold treatments

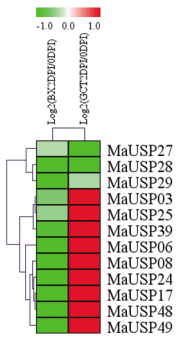


图 6 *MaUSPs* 在香蕉幼苗接种 *Foc* TR4 后的表达

Figure 6 The expression of *MaUSPs* in banana seedlings responding to *Foc* TR4 treatments

23 个 *MaUSPs* 中发现了高度保守的 ATP 特异结合位点(G-2X-G-9X-G[S/T]) (图 2), 所有的 *MaUSPs* 都属于 *USP* 基因。亚细胞定位预测结果显示, 53 个 *MaUSPs* 的位置不同, 说明 *MaUSPs* 的功能可能不同, 该结果与丹参的研究结果一致(Wang et al., 2017)。系统发育分析表明, 53 个 *MaUSPs* 分为 A、B、C、D 亚家族, 这与拟南芥和丹参的结果一致(Kerk et al., 2003; Wang et al., 2017)。香蕉 A 基因组共经过了 3 次全基因组重复事件(α -, β -, and γ -WGD), 祖先共线性区块的片段重复区域占 222 Mb, 包含 26 829 个基因(France et al., 2012)。本研究发现 12 个 *MaUSPs* 位于上述片段重复区块中(图 3), 表明片段复制是 *MaUSPs* 基因家族扩张的主要方式。

香蕉的发育和成熟过程是香蕉品质形成的一个重要的过程。呼吸跃变型果实的发育和成熟是一个氧化过程, 会产生活性氧(ROS), 如超氧自由基(O_2^-)和过氧化氢(H_2O_2) (Mondal et al., 2004; Pandey et al., 2013)。番茄 ER6 是广谱胁迫蛋白家族的一员, 在诱导果实成熟过程, 番茄完熟期的表达谱达到高峰(Zegzouti et al., 1999)。本研究发现, 在果实发育和成熟过程中, 大部分的 *MaUSPs* 基因表达量较高, 其中 *MaUSP46*、*MaUSP33* 和 *MaUSP49* 在果实发育期的表达量较高, *MaUSP24*、*MaUSP49*、*MaUSP14*、*MaUSP28* 基因在采后成熟时期差异显著(图 4), 表明这些 *MaUSPs* 在香蕉发育和采后成熟中起着重要作用。

USP 蛋白存在于古生菌、真菌、原生生物、植物和动物中(Sauter et al., 2002; Foret et al., 2011)。在大肠杆菌中, *uspA* 在营养耗尽和 DNA 损伤的生长抑制过程中会显著增加(Gustavsson and Nyström, 2002)。在植物中, *USP* 基因参与响应多种非生物胁迫(Jung et al., 2015; Udawat et al., 2016)。水稻 *OsUSP1* 在水分胁迫和乙烯诱导中上调表达(Sauter et al., 2002)。亚

洲棉 *GUSP1* 和 *GUSP2* 可以响应水分胁迫(Zahur et al., 2009)。在红杉 (*Gossypium hirsutum*) 中, *GhUSP1* 和 *GhUSP2* 参与响应植物发育和盐胁迫(Merkouropoulos and Tsiftaris, 2013)。在本研究中, 大多数 *MaUSPs* 在干旱、低温和盐等非生物胁迫下差异表达(图 5)。这些结果表明, *MaUSPs* 可能在香蕉幼苗对非生物胁迫的反应中起重要作用。

USP 基因在参与植物响应生物胁迫的过程中也有着重要作用。紫云英 *USP* 通过感染根中的根瘤菌被诱导表达 (Chou et al., 2007)。在拟南芥中, *USP* (*At3g53990*) 具有抗真菌活性, 能刺激真菌产生 ROS 和线粒体损伤。巴西蕉是香蕉主栽品种, 易感香蕉枯萎病, 而 GCTCV-119 是抗病突变体。在本研究中, 巴西蕉和 GCTCV-119 接种 *Foc* TR4 后, 12 个 *MaUSPs* 差异表达, 在巴西蕉中, *MaUSPs* 基因的表达被显著抑制, 但在 GCTCV-119 中 *MaUSP03*、*MaUSP25*、*MaUSP39*、*MaUSP06*、*MaUSP08*、*MaUSP24*、*MaUSP17*、*MaUSP48* 和 *MaUSP49* 基因均显著上调(图 6); 上述结果表明 *MaUSPs* 基因可能在响应香蕉抗枯萎病的过程中可能起着重要作用。

3 材料与方法

3.1 *MaUSPs* 鉴定和系统发育分析

本研究从香蕉 A 基因组网站(<http://musagd.biocloud.net/home>)下载香蕉 A 基因组 DH Pahang v2 编码蛋白序列的所有基因(Martin et al., 2016); 利用拟南芥的所有 *USPs* 进行 BLAST 搜索, 检索香蕉 A 基因组中心数据库中预测的 *MaUSPs*; 利用 ExPASy 数据库(<http://expasy.org/>)预测 *MaUSPs* 的分子量和等电点; 利用 Plant-mPLoc Server (<http://www.csbio.sjtu.edu.cn/bioinf/plantmulti/>) 在线预测 *MaUSPs* 蛋白亚细胞定位; 利用 NCBI 和 SMART (<https://smart.embl.de/>)对所有候选的 *MaUSPs* 蛋白序列进行 BLASTp。拟南芥的 *MaUSPs* 蛋白序列来自 TAIR (<http://www.arabidopsis.org/>)数据库; 利用 ClustalW 对香蕉和拟南芥的全长 *MaUSPs* 蛋白序列进行比对; 利用 MEGA X 软件, 采用邻接法构建进化树 (Audic and Claverie, 1997)。已上传的 *MaUSPs* 的登录号(表 1)。

3.2 染色体分布和基因重复

为确定 *MaUSPs* 的染色体分布情况, 从香蕉 A 基因组数据库中获得每条染色体上所有基因的起始和终止位置, 并按照 11 条染色体的位置按比例绘制。根据植物基因组复制数据库(plant genome duplication

database, PGDD)的方法(Lee et al., 2012)进行序列重复和片段重复。基于染色体定位确定串联重复的标准:在单个染色体上的同源的 *MaUSP* 基因,如果没有其他干预基因,且彼此之间相距 30 kbp 以内,被鉴定为串联重复(Shiu and Bleecker, 2003; Du et al., 2013)。使用 MCSCAN(参数: -a-e1e-5-s5)检测 Syntenic (Wang et al., 2019),提取 Syntenic 中所有的 *MaUSP* 基因。利用 Circos (0.63)软件绘制出 *MaUSP* 基因的位置和合成图像(<http://circos.ca/>)。

3.3 植物材料和处理

巴西蕉(*Musa acuminata* L. AAA group, cv. Cavendish)果实取自热带生物技术研究所蕉园(海南澄迈, 20N, 110E)。分别采集了 0 DAF (Day After Flower)、20 DAF、80 DAF-0 DPH (Day Poster Harvest)、8 DPH 和 14 DPH 的果实。

选取苗龄为一个月的巴西蕉幼苗,使用 Hoagland 溶液进行浇灌(0.51 g/L KNO₃, 0.82 g/L Ca(NO₃)₂, 0.49 g/L MgSO₄·7H₂O, 0.136 g/L KH₂PO₄, 0.6 mg/L FeSO₄, 2.86 mg/L H₃BO₃, 1.81 mg/L MnCl₂·4H₂O, 0.08 mg/L CuSO₄·5H₂O, 0.22 mg/L ZnSO₄·7H₂O, 0.09 mg/L H₂MoO₄·4H₂O) (pH 6.0)作为对照,在干旱(osmotic, 200 mmol/L mannitol)、盐(salt, 300 mmol/L NaCl)、低温(cold, 4°C)对香蕉苗进行胁迫处理 7 d;温室的最高温度和最低温度分别为 25°C 和 30°C,相对湿度在 55%~80%之间。

选取五叶一心的感病品种巴西蕉(作图简称为 BX)和抗病品种 GCTCV-119(作图简称为 GCT)使用蘸根法接种 Foc TR4 (Hwang and Ko, 2004),孢子浓度为 1.5×10⁶ 孢子/mL,在接种后 0 d、2 d、4 d 和 6 d (DPI)用液氮冻存香蕉根部组织,存于 -80°C 备用(Lee et al., 2012)。

3.4 转录组分析

使用多糖多酚植物总 RNA 提取试剂盒(TIAN-GEN)提取香蕉根系总 RNA, RNA 用 RevertAid First Strand cDNA Synthesis Kit 试剂盒(ThermoFisher Scientific)反转录为 cDNA,存于 -20°C 备用。使用 TruSeq RNA 文库制备试剂盒 v2 构建 cDNA 文库,随后使用 Illumina RNA-seq 协议在 Illumina HiSeq 2000 平台上进行测序。对每个样本进行两次生物重复。基因表达水平以每百万 Reads 中基因每千碱基长度的 Reads 数(RPKM)计算(Mortazavi et al., 2008)。采用每个基因读数为 2 个重复(fold change ≥ 2; FDR ≤ 0.001)。所有数据已上传至 CNGBdb 的 CNSA (<https://>

db.cngb.org/cnsa/)存储,其登录号为 CNP0000292。*MaUSPs* 基因在果实中的数据用 RPKM 值表示;在胁迫条件下的数据用 Log 值表示,利用 MeV v 4.3 作图。

作者贡献

王卓和从心黎是本研究的实验设计者;王霞是实验研究的执行人,并完成数据分析、论文初稿的写作;林秋妹、赵东方、贾彩红、王静毅、刘菊华和黄绵佳参与实验设计、试验结果分析;王卓和从心黎是项目的构思者及负责人,指导实验设计、数据分析、论文写作与修改。全体作者都阅读并同意最终的文本。

致谢

本研究由海南省自然科学基金(No.318MS090)和 CATAS (No.1630052019035)共同资助。

参考文献

- Aravind L., Anantharaman V., and Koonin E.V., 2010, Monophyly of class I aminoacyl tRNA synthetase, USPA, ETPF, photolyase, and PP-ATPase nucleotide-binding domains: Implications for protein evolution in the RNA, *Proteins*, 48 (1): 1-14
- Audic S., and Claverie J.M., 1997, The significance of digital gene expression profiles, *Genome Res.*, 7(10): 986
- Aurore G., Parfait B., and Fahrasmene L., 2009, Bananas, raw materials for making processed food products, *Trends in Food Sci. Technol.*, 20(2): 78-91
- Chou M.X., Wei X.Y., Chen D.S., and Zhou J.C., 2007, A novel nodule-enhanced gene encoding a putative universal stress protein from *Astragalus sinicus*, *J. Plant Physiol.*, 164 (6): 764-772
- D'Hont A., Denoeud F., Heslop-Harrison P., Carreel F., Franc-Christophe B., and Aury M., 2012, The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants, *Nature.*, 448: 213-217
- Du D.L., Zhang Q.X., Cheng T.R., Pan H.T., Yang W.R., and Sun L.D., 2013, Genome-wide identification and analysis of late embryogenesis abundant (LEA) genes in *Prunus mume*, *Mol. Biol. Rep.*, 40(2): 1937-1946
- Foret S., Seneca F., De Jong D., Bieller A., Hemmrich Georg., Augustin R., Hayward D., Ball E., Bosch T., Agata K., Hassel M., and Miller D.J., 2011, Phylogenomics reveals an anomalous distribution of USP genes in metazoans, *Mol. Biol. Evolution*, 28(1): 153-161
- Freeling M., 2009, Bias in plant gene content following different

- sorts of duplication: Tandem, whole-genome, segmental, or by transposition, *Ann. Rev. Plant Biol.*, 60: 433-453
- Freestone P., Trinei M., Clarke S.C., Nystrom T., and Norris V., 1998, Tyrosine phosphorylation in *Escherichia coli*, *J. Mol. Biol.*, 279(5): 1045-1051
- Gou L.M., Zhuo C.L., Lu S.Y., and Guo Z.F., 2020, A universal stress protein from *Medicago falcata* (*MfUSP1*) confers multiple stress tolerance by regulating antioxidant defense and proline accumulation, *Environmental and Experimental Botany*, 178: 104-168
- Gustavsson N., and Nyström T., 2002, The universal stress protein paralogues of *Escherichia coli* are co-ordinately regulated and co-operate in the defence against DNA damage, *Mol. Microbiol.*, 43(1): 107-117
- Gutierrez-Beltran E., Personat J.M., de la Torre F., and del Pozo O., 2017, A universal stress protein involved in oxidative stress is a phosphorylation target for protein kinase CIPK6, *Plant Physiol.*, 173: 836-852
- Hurles M., 2004, Gene duplication: The genomic trade in spare parts, *PLoS Biol.*, 2: e206
- Hwang S.C., and Ko W.H., 2004, Cavendish banana cultivars resistant to *Fusarium* wilt acquired through somaclonal variation in Taiwan, *Plant Dis.*, 88: 580-588
- Jung Y.J., Melencion S.M., and Lee E.S., 2015, Universal stress protein exhibits a redox-dependent chaperone function in *Arabidopsis* and enhances plant tolerance to heat shock and oxidative stress, *Front. Plant Science*, 6: 1141
- Kang H., Park S.J., and Kwak K.J., 2013, Plant RNA chaperones in stress response, *Trends Plant Sci.*, 18: 100-106
- Kerk D., Bulgrien J., Smith D.W., and Gribskov M., 2003, *Arabidopsis* proteins containing similarity to the universal stress protein domain of bacteria, *Plant Physiol.*, 131(3): 1209-1219
- Kvint K., Nachin L., Diez A., and Nyström T., 2003, The bacterial universal stress protein: Function and regulation, *Curr. Opin. Microbiol.*, 6(2): 140-145
- Lee T.H., Tang H.B., Wang X.Y., and Paterson A.H., 2012, PGDD: A database of gene and genome duplication in plants, *Nucleic Acids Research.*, 41: D1152-D1158
- Li W.T., Wei Y.M., Wang J.R., Liu C.J., Lan X.J., Jiang Q.T., Pu Z.E., and Zheng Y.L., 2010, Identification, localization, and characterization of putative USP genes in barley, *Theor. Appl. Genet.*, 121(5): 907-917
- Mahajan S., and Tuteja N., 2005, Cold, salinity and drought stresses: An overview, *Archives of Biochemistry and Biophysics*, 444(2): 139-158
- Martin G., Baurens F.C., Droc G., Rouard M., Cenci A., and D'Hont A., 2016, Improvement of the banana "Musa acuminata" reference sequence using NGS data and semi-automated bioinformatics methods, *BMC Genomics*, 17: 243
- Merkouropoulos G., and Tsafaris A.S., 2013, Differential expression of *Gossypium hirsutum* USP-related genes, GhUSP1 and GhUSP2, during development and upon salt stress, *Plant Mol. Biol. Rep.*, 31(6): 1539-1547
- Mondal K., Sharma N.S., Malhotra S.P., Dhawan K., and Singh R., 2004, Antioxidant systems in ripening tomato fruits, *Biol. Plant.*, 48(1): 49-53
- Mortazavi A., Williams B.A., McCue K., Schaeffer L., and Wold B., 2008, Mapping and quantifying mammalian transcriptomes by RNA-Seq, *Nature Methods.*, 5: 621
- Munns R., 2005, Genes and salt tolerance: Bringing them together, *New Phytologist*, 167(3): 645-663
- Pandey V.P., Singh S., Jaiswal N., Awasthi M., Pandey B., and Dwivedi U.N., 2013, Papaya fruit ripening: ROS metabolism, gene cloning, characterization and molecular docking of peroxidase, *J. Mol. Catalysis B: Enzymatic*, 98: 98-105
- Ploetz R.C., 2006, *Fusarium* wilt of banana is caused by several pathogens referred to as *Fusarium oxysporum* f. sp. *Cubense*, *Phytopathol.*, 96(6): 653-656
- Ploetz R.C., 2015, *Fusarium* wilt of banana, *Phytopathol.*, 105(12): 1512-1521
- Sauter M., Rzewuski G., Marwedel T., and Lorbiecke R., 2002, The novel ethylene-regulated gene *OsUsp1* from rice encodes a member of a plant protein family related to prokaryotic universal stress proteins, *J. Experimental Botany*, 53(379): 2325-2331
- Shiu S.H., and Bleecker A.B., 2003, Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in *Arabidopsis*, *Plant Physiol.*, 132: 530-543
- Sinha P., Pazhamala L.T., and Singh V.K., 2016, Identification and validation of selected universal stress protein domain containing drought-responsive genes in *Pigeonpea* (*Cajanus cajan* L.), *Front. Plant Sci.*, 6: 1065
- Tkaczuk K.L., Shumilin I.A., Chruszcz M., Evdokimova E., Savchenko A., and Minor W., 2013, Structural and functional insight into the universal stress protein family, *Evolut.*, 6: 434-449
- Udawat P., Jha R.K., Sinha D., Mishra A., and Jha B., 2016, Overexpression of a cytosolic abiotic stress responsive universal stress protein (SbUSP) mitigates salt and osmotic stress in transgenic tobacco plants, *Front. Plant Sci.*, 7: 1-21
- Vanbogelen R.A., Hutton M.E., and Neidhardt F.C., 1990, Gene-Protein database of *Escherichia coli* K-12: Edition 3, *Electrophoresis*, 11(12): 1131-1166
- Visser M., Gordon T., Fourie G., and Viljoen A., 2010, Characterisation of south african isolates of *Fusarium oxysporum* f. sp. *cubense* from cavendish bananas, *South Afr. J. Sci.*, 106: 1-6
- Vollmer A.C., and Bark S.J., 2018, Twenty-five years of investigating the universal stress protein: Function, structure, and

- applications, *Adv. Appl. Microbiol.*, 102: 1-36
- Wang X.F., Su J., Yang N., Zhang H., Cao X.Y., and Kang J.F., 2017, Functional characterization of selected universal stress protein from *Salvia miltiorrhiza* (SmUSP) in *Escherichia coli*, *Genes.*, 8: 224
- Wang Z., Miao H.X., Liu J.H., Jia C.H., Jin Z.Q., and Xu B.Y., 2019, *Musa balbisiana* genome reveals subgenome evolution and functional divergence, *Nature Plants*, 5: 810-821
- Zahur M., Maqbool A., Irfan M., Barozai Y.K., Rashid B., Rizuddin S., and Husnain T., 2009, Isolation and functional analysis of cotton universal stress protein promoter in response to phytohormones and abiotic stresses, *Mol. Biol.*, 43 (4): 578-585
- Zegzouti H., Jones B., Frasse P., Marty C., Maitre B., Latché A., Pech J.C., and Bouzayen M., 1999, Ethylene-regulated gene expression in tomato fruit: Characterization of novel ethylene-responsive and ripening-related genes isolated by differential display, *Plant J.*, 18(6): 589-600