

番茄 AT-hook 基因家族的鉴定及胁迫条件下的表达分析

丁丽雪^{1,2,3}, 李涛^{2,3}, 李植良², 徐小万^{2,3}, 李颖², 王恒明², 王永飞¹, 马三梅¹, 黎振兴²

(¹暨南大学生命科学技术学院, 广州 510532; ²广东省农业科学院蔬菜研究所, 广州 510640;

³广东省蔬菜新技术研究重点实验室, 广州 510640)

摘要: AT-hook 蛋白家族在植物生长发育、器官构建及逆境胁迫和激素信号应答中发挥重要作用。本研究在番茄基因组范围内, 利用生物信息学方法对番茄 AT-hook 基因家族的成员、分布、结构和功能进行分析。结果表明, 番茄 AT-hook 家族包含 32 个成员, 分为 3 种类型, 其中类型 I 含有 13 个成员; 遗传进化分析表明番茄 AT-hook 基因成员与拟南芥家族基因具有相似分类。利用实时荧光定量 PCR 对番茄 32 个基因开展组织表达分析, 结果表明 AT-hook 基因具有表达差异, 主要在根和花中表达较高。氧化胁迫分析结果表明, 32 个基因受 ABA、SA、盐、高温和低温诱导表达, 其中部分基因显著上调或下调表达, 很可能参与了番茄逆境胁迫条件下的防御应答反应。本研究结果将为番茄 AT-hook 家族基因的深入研究提供依据, 为进一步解析番茄 AT-hook 基因的功能奠定基础。

关键词: 番茄; AT-HOOK 基因家族; 生物信息学; 基因表达; 氧化胁迫

Genome-wide Identification and Expression Analysis in Oxidative Stress of AT-hook Gene Family in Tomato

DING Li-xue^{1,2,3}, LI Tao^{2,3}, LI Zhi-liang², XU Xiao-wan^{2,3},

LI Ying², WANG Heng-ming², WANG Yong-fei¹, MA San-mei¹, LI Zhen-xing²

(¹Department of Biotechnology, Jinan University, Guangzhou 510632;

²Vegetable Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640;

³Guangdong Key Laboratory for New Technology Research of Vegetables, Guangzhou 510640)

Abstract: AT-hook proteins played an important role in regulation of eukaryotic gene expression, and implicated in plant development, organ formation, stress response, and hormonal signaling. In this study, our analysis revealed the presence of 32 genes encoding AT-hook proteins in the tomato genome, and the distribution on the chromosome, the structure and function of proteins were also analyzed. AT-hook proteins could be classified into 3 type based on phylogenetic analysis, 13 of them belonged to Type I, and other members had the similar classification with *Arabidopsis thaliana*, which was supported with organization of predicted conserved putative motifs in AT-hook proteins. Real-time fluorescence quantitative PCR were used to gain insights into the function of AT-hook gene family members during various stages of development and in response to abiotic stress and phytohormones. Some of the AHL genes were expressed in specific organs such as roots and flowers. 32 SLAHL genes showed distinct expression

收稿日期: 2015-04-16 修回日期: 2015-06-01 网络出版日期: 2016-01-28

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20160128.1524.016.html>

基金项目: 国家自然科学基金项目(31301776); 广东省省级科技计划项目(2014A020208044, 2015A030302040, 2015B090906012); 农业部现代产业体系项目(CARS-25-G-36); 广东省对外科技合作项目(2013B050800012); 广东省科技厅条件建设项目(粤财教[2013]112号)

第一作者研究方向为番茄基因功能研究。E-mail: 15264311931@126.com; 李涛为共同第一作者, 研究方向为番茄分子植物育种。E-mail: tianxing84@163.com

通信作者: 黎振兴, 主要从事番茄遗传育种研究。E-mail: lizhngaas@163.com

马三梅, 主要从事蔬菜作物的转基因与应用。E-mail: msmwdw@163.com

patterns in response to stresses of salt, SA, ABA, high temperature and cold, the expressions of several genes were significantly up/down regulated, implying that these members might participate in regulating the defense response against abiotic stresses. The results will provide a very useful reference for cloning and functional analysis of each member of AT-hook gene family in Solanaceae crops.

Key words: tomato; AT-hook gene family; bioinformatics; gene expression; oxidative stress

AT-hook 是一种新的 DNA 结合蛋白基序,普遍存在于细菌、真菌、哺乳动物和植物的 DNA 结合蛋白中,是以精氨酸-甘氨酸-精氨酸-脯氨酸(RGRP)4 个残基为中心保守基序^[1],这个短的保守氨基酸序列对 AT-hook 蛋白结合 DNA 和核定位是必需的^[2-3]。1973 年 G. H. Goodwin 等^[4]首次在哺乳动物非组蛋白的染色体高迁移率蛋白(HMG)中发现,后续研究表明 AT-hook 蛋白在染色质结构组装和调控目标基因转录活性方面具有重要作用^[5]。

AT-hook 蛋白主要包含 2 个保守的结构域,AT-hook 基序和植物与原核生物保守结构域(PPC, plants and prokaryotes conserved, DUF296)^[6]。根据序列相似程度和 AT-hook 序列 C 端氨基酸的类型不同及其与 DNA 结合的亲和力大小,将 AT-hook 蛋白分为 3 个类型(I 型, C 端第 2 位置为甘氨酸; II 型, C 端为赖氨酸代替了 I 型的甘氨酸; III 型为赖氨酸出现在 RGRP 下游第 4 位)^[2]。AT-hook 蛋白 PPC 结构域的长度约为 120 个氨基酸,且在细菌和古细菌以单个蛋白的形式存在^[7]。几个细菌和古细菌的 PPC 蛋白的晶体结构为三聚体^[8-9]。在陆生植物中, AHL 蛋白中已经确定含有 PPC 域,它定位在相对于 AT-hook 基序的羧基末端^[7]。例如在转录因子中 PPC 结构域负责 AHL 蛋白的细胞核定位, AHL 蛋白之间相互影响并与其他转录因子存在相互作用进而发挥功能,这可能暗示 AHL 蛋白在调节植物转录激活过程中发挥重要作用^[6]。

近年来,随着许多物种基因组测序计划的完成,在模式植物拟南芥^[6-7]、水稻^[10]、小麦、大豆、豌豆、木薯、蓖麻、毛果杨、蒺藜状苜蓿、黄瓜、碧桃、苹果、芜菁、番木瓜、葡萄、高粱、玉米、二穗短柄草、江南卷柏、小立碗藓中发现 AT-hook 蛋白^[11]。AT-hook 蛋白在模式植物拟南芥中含有 29 个成员,被称为 AHL(AT-hook nuclear localized proteins)蛋白家族^[7],研究发现其主要参与赤霉素(GA)^[12]、茉莉酸甲酯(JA)^[13]和细胞分裂素(CK)^[14]等植物体内激素平衡,开花时间^[15]、下胚轴的伸长^[6]和衰老^[16]等的调控。其中 A. Matsushita 等^[12]研究发现 *AGF1*

(*AtAHL25*)调控赤霉素代谢 *AtGA3ox1* 基因的表达控制植物体内 GA 的动态平衡。*AtAHL22* 通过 FT 和 PIF4 控制开花和下胚轴伸长,过量表达 *AtAHL22* 导致开花延迟,沉默 *AtAHL22*、*AtAHL18*、*AtAHL27* 和 *AtAHL29* 能够促进开花时间提早及增强 *ahl22-1* 突变体下胚轴生长^[15,17]。*GIK/AtAHL21* (GIK, GIANT KILLER)作为 MADS 转录因子 *AGAMOUS* (AG)的下游基因通过调控转录因子如 *ARF3* 调节组织分化和生殖器官的发育,过量表达或突变 *GIK* 基因导致生殖器官模式形成和分化相关的缺陷^[18]。过量表达 *SOB3/AtAHL29* 或突变 *ESC/AtAHL27* 基因会导致下胚轴表型变化及延迟叶片衰老表型的出现^[19]。进一步研究发现 AT-hook 蛋白家族成员之间存在相互作用进而调控下胚轴的伸长^[6]。张贵慰等^[10]在水稻中鉴定出 45 个 AT-hook 蛋白,组织表达分析表明主要在幼穗中表达。Y. Jin 等^[20]通过图位克隆方法获得 1 个 AT-hook 蛋白 *dp1* (depressed palea1) 主要控制内稃形成和花器官数目。玉米 BAF1 在叶芽分生组织诱导中发挥作用,是玉米雌穗形成所必须的^[21]。S. K. Delaney 等^[22]在棉花中克隆的 AT-hook 蛋白 GhAT1 通过抑制 *FS1p4* 基因的表达进而影响棉花纤维的生长。在茄果类研究方面, S. Kim 等^[23]通过 cDNA 文库克隆到辣椒 *CaATL1* (*Capsicum annuum* L. Bukang AT-hook-like gene 1),该基因受细菌或病毒诱导表达,且在 SA 和乙烯处理后上调表达,在番茄中过量表达 *CaATL1* 基因能够提高对细菌和真菌的抗性。

番茄 (*Solanum lycopersicum* L.) 是世界上重要的经济作物和模式植物之一。2012 年番茄基因组测序计划的完成^[24],使国内外学者在基因组水平对番茄重要基因家族及重要功能基因的分析 and 分子标记的开发成为可能^[25-28],为功能基因组学提供重要的研究数据。鉴于 AT-hook 基因家族的重要生理功能,本研究利用生物信息学方法对番茄 AT-hook 基因家族进行预测、挖掘及表达分析,为番茄 AT-hook 基因家族的功能分析提供基础信息,从而为分子植物育种提供候选基因。

1 材料与方法

1.1 供试材料

将番茄测序品种 Heinz 1706 种植于广东省农业科学院蔬菜研究所温室,培养条件为温度 28 ℃,湿度 70%,每天光照时间为 16 h,黑暗时间为 8 h。

1.2 番茄 AT-hook 基因家族成员的识别、保守结构域预测及基因结构分析

Pfam 数据库中标号为 PF02178 的结构域^[29]是 AT-hook 蛋白的特征结构域。利用 Hmmer3.1b1 软件构建隐马氏模型,从 SGN^[24] (SGN, <http://solgenomics.net>, release v2.40) 下载番茄蛋白数据进行检索和去除冗余。

利用 Pfam(<http://www.sanger.ac.uk/Software/Pfam/search.shtml>)^[29]和 SMART(<http://smart.embl-heidelberg.de/>)^[30]对所有候选基因的氨基酸序列结构域进行鉴定,把显示不含 AT-hook 基序的蛋白序列剔除,最后得到编码 AT-hook 蛋白的基因序列即为番茄 AT-hook 基因家族成员。

在番茄基因组数据库 SGN 中检索获得番茄 AT-hook 基因家族的 DNA 及其 cDNA 序列,并查出所有基因的开放阅读框长度、染色体位置、内含子个数等基本信息。利用在线工具 Expasy (http://web.expasy.org/compute_pi/) 和 PSORT (<http://psort.ims.u-tokyo.ac.jp/form.html>)^[31]对番茄 AT-hook 基因成员等电点和亚细胞定位进行分析。

1.2 番茄 AT-hook 的染色体定位分析

根据检索到的番茄 AT-hook 基因组信息,确定各 AT-hook 基序基因在染色体组中的位置,利用番茄基因组数据 (SGN, <http://solgenomics.net>, release v2.40) 对番茄 AT-hook 基因进行染色体定位分析。

1.3 蛋白质系统发生分析与系统进化树的构建

用 ClustalX2.1^[32] 软件将候选番茄 AT-hook 家族蛋白序列与拟南芥蛋白序列进行多重序列比对分析,基于分析结果,利用 MEGA5^[33] 构建系统进化树。采用相邻连接法 (NJ, neighbor-joining), 校验参数 Bootstrap 重复 1000 次。保守性分析采用 DNAMAN 生物学软件进行保守域序列比对。

1.4 保守基序的鉴定

根据已经检索到的番茄 AT-hook 基因组氨基酸序列,利用 Meme4.6.1 (<http://meme.nbcr.net/meme/intro.html>) 在线平台进行番茄 AT-hook 蛋白中保守基序 (motif) 分析,最大 motif 检索数值定为 25。

1.5 AT-hook 基因的 Q-PCR 表达分析

分别选取在营养生长阶段的根、茎、叶,在生殖生长阶段的花和成熟果实;选取生长一致的 14 d 的幼苗进行胁迫处理 (200 mmol/L NaCl, 5 mmol/L SA 处理; 10 μmol/L ABA; 42 ℃ 高温; 4 ℃ 低温), 将以上材料用液氮速冻后放入 -80 ℃ 冰箱保存备用。

分子生物学试验在广东省蔬菜新技术研究重点实验室完成。采用 Trizol (Invitrogen) 法提取总 RNA,经 DNaseI 处理去除基因组 DNA,使用 Nano-Drop 2000c 分光光度计 (Thermo) 检测 RNA 样品的浓度和纯度。吸取 3 μg RNA 经 M-MLV 反转录合成第 1 链 cDNA,稀释后做 qRT-PCR 模板。根据番茄数据库基因序列,利用 Oligo 6.0 设计引物 (表 1),由英潍捷基 (上海) 生物公司合成。荧光实时定量 PCR 反应试剂盒为 SYBR Green Realtime PCR Master mix (Takara, 大连),反应体系为 SYBR Premix Ex TaqTM II (TliRNaseH Plus) (2 ×) 10 μL,引物 (10 μmol/L) 各 0.8 μL, DNA 模板 2.0 μL, 无菌双蒸馏水 6.4 μL,反应总体积 20.0 μL。反应条件为: 95 ℃ 预变性 30 s; 95 ℃ 变性 5 s, 60 ℃ 退火 20 s, 循环 40 次; 溶解曲线分析 95 ℃ 15 s, 65 ℃ 30 s, 95 ℃ 15 s。在 Roche Light Cycler 480 仪器上进行。采用 2^{-ΔΔCt} 方法计算基因相对表达量,每个基因的表达反应重复 3 次,以 Actin 作为内参基因。Roche Light Cycler 480 软件根据 PCR 扩增曲线计算实时荧光定量 PCR 的扩增效率。

2 结果与分析

2.1 番茄 AT-hook 基因家族成员的识别

以 Pfam 数据库中的 AT-hook 特征基序 (PF02178)^[29] 检索 SGN 数据库,将得到的候选序列用 SMART 在线分析预测其保守结构域,剔除非 AT-hook 基序的蛋白序列。结果共获得 32 条蛋白质序列,基因命名为 *SIAHL1* ~ *SIAHL32* (表 2)。蛋白质生化属性分析发现,蛋白序列长度为 161 ~ 418 aa,分子量为 17.29 ~ 42.85 kD,等电点为 5.31 ~ 9.68, CDS 长度为 483 ~ 1254 bp,含有 0 ~ 5 个内含子,其中 *SIAHL1*、*SIAHL12*、*SIAHL14*、*SIAHL15*、*SIAHL21*、*SIAHL24*、*SIAHL27* 和 *SIAHL31* 基因无内含子,而其他番茄 AT-hook 基因则含有 1 ~ 5 个不等的外显子。番茄中大多数的 AT-hook 基因在细胞中大多位于细胞核,也有少部分位于叶绿体中。所有基因不均匀的分布在 1 ~ 12 号染色体上,其中 10 号染色体上不含有 AT-hook 基因。

表 1 本研究所用引物

Table 1 Q-PCR primers used in this study

引物名称 Primer name	序列(5'-3') Squence(5'-3')	引物名称 Primer name	序列(5'-3') Squence(5'-3')
SIAHL1 – QF	TTAGCGGCTGGTCAAGGTCAAG	SIAHL17 – QF	ACGGGAGGAATGAGAAACAGGTC
SIAHL1 – QR	AACCGGTCCTGATGCAATTAGCG	SIAHL17 – QR	AGGAAGCTCCCAACAACAATCTGG
SIAHL2 – QF	TAGCTGCTAGTCCTGTCCAGGTTG	SIAHL18 – QF	TTTAGCTGGCGGACAAGGTCAG
SIAHL2 – QR	TGTTTCTTGGGCTTGACCTCTTGG	SIAHL18 – QR	ACCGGACCAGATGCCATAAGAG
SIAHL3 – QF	ACAATTAGACAACCCGCGTCTCC	SIAHL19 – QF	CATCAGCGTAGGAGAGGATGTTGC
SIAHL3 – QR	TCCGGCCAAGTATATCGTCAGC	SIAHL19 – QR	TAAAGCCCTAGGCCTCTGTTCTGC
SIAHL4 – QF	CTCAGAAGAGAGGTAGAGGAAGGC	SIAHL20 – QF	AGGCTCATTTGTTGCTACCAGACG
SIAHL4 – QR	TTGGACAACCATCCACCAAAGG	SIAHL20 – QR	CACGTTGAGACCACCCGTTATTCC
SIAHL5 – QF	TAGGATCTGCAAGTGCTGGGACTG	SIAHL21 – QF	CAAGGTTCTGATGGTGGTGATGGG
SIAHL5 – QR	AGCTGACATGATGCAGACAGCTTG	SIAHL21 – QR	AACGGAAGCCCAGGATTAGACG
SIAHL6 – QF	GCAAAGCGTTGATTTGGGATCAG	SIAHL22 – QF	AATGGTGGTGGAGTTTCAGGAG
SIAHL6 – QR	CCATTGGCTGACAGAACACACAC	SIAHL22 – QR	GAAGCTACCCACAATTACCTGCAC
SIAHL7 – QF	AGGACCTGATGGTCGGGTATG	SIAHL23 – QF	ATCCGGCCAGTTGGGTTAGTTAG
SIAHL7 – QR	CAACCACAATCTGAATGGGACCAG	SIAHL23 – QR	CATGAAACCCATTGCGCTGGAG
SIAHL8 – QF	TGCTAAGGCGCAGGTAATCCTTG	SIAHL24 – QF	TAGGTGCACTCATGGCATCTGG
SIAHL8 – QR	TTGCAGACTCTTTGGAGGGTGAC	SIAHL24 – QR	GCAGCATTACTGAAAGACGCTGAC
SIAHL9 – QF	TAGTTGGTAAGGCCGAGGTAATCC	SIAHL25 – QF	GTCCTGATGGCCATGTAATTGGAG
SIAHL9 – QR	TTGGAGGGCGACGTTACGTATG	SIAHL25 – QR	AAAGCTGCATACGACTACCTGAAC
SIAHL10 – QF	CTGTCATTTGCCGTCGGATGTG	SIAHL26 – QF	ACAGCAGGAGGTGACTTTATGCC
SIAHL10 – QR	AGTAAGGCGGTGCTTGTGCTTG	SIAHL26 – QR	ATAAACCTTTCTGCCACATCCTC
SIAHL11 – QF	ACGACTTCCCTTGGAGGAAGATG	SIAHL27 – QF	ACCACCAGGATCCAAGAACAAGC
SIAHL11 – QR	AGCTGGCTGTACCTGTGTAGTAGG	SIAHL27 – QR	ATGTGGCTACGGAGGGCATTAG
SIAHL12 – QF	TGAGCGGCAGCGGTATAGTAAC	SIAHL28 – QF	CCCTCCTGGATCAACAGGTTTG
SIAHL12 – QR	ATCTCGAACCTACCGTGAGCG	SIAHL28 – QR	CCCTACTACACTTCTCCCAACAAC
SIAHL13 – QF	AGCTGGTCCTGTTCAGGTTGTTG	SIAHL29 – QF	TGGTTTGAGCCATGTTGCTGGAG
SIAHL13 – QR	ACCTCTGCTTCTTGGGATTCTGC	SIAHL29 – QR	GCATTGCCCAACCGAGCTAATTG
SIAHL14 – QF	GACACCGCCACCAAATCTTCTC	SIAHL30 – QF	AACCAACAACAGATACCAGGATCG
SIAHL14 – QR	AGGTCGTCCCGTTGACCAATATC	SIAHL30 – QR	ACTCACCAGTCGCTGACACAATAC
SIAHL15 – QF	TATGCGTGCTGAGTGCAACAGG	SIAHL31 – QF	TCACAAATTTCCGGCGGAGGAC
SIAHL15 – QR	TGGACCCTGATGTAGCTGGTTG	SIAHL31 – QR	TGCCCATATCACATCCGATGGC
SIAHL16 – QF	TGCAAGGACCACCGGTAGTTTG	SIAHL32 – QF	GCTGTCTGACATTGGTGGTCAAC
SIAHL16 – QR	ACCTTGTAAGCCTGCTTGAGG	SIAHL32 – QR	AGCTACCTACGATAACCTGGACTG

表 2 AHL 基因家族分布及理化性质分析

Table 2 The distribution of AHL gene family and physic-chemical analysis in tomato

基因 Gene	SGN 编号 SGN code	位置 Chromosome location	长度 (bp) CDS length	内含子 数量 Introns number	氨基酸 (aa) Protein length	等电点 PI	分子量 (kD) Mw	亚细胞 定位 Subcellular localization	染色体 Chromo some
<i>SIAHL1</i>	Solyc01g080960.2.1	72800762..72800497	798	0	266	6.37	28.125	Nuc	1
<i>SIAHL2</i>	Solyc01g091370.2.1	76767771..76767453	957	4	319	9.54	33.480	Nuc	1
<i>SIAHL3</i>	Solyc01g091470.2.1	76860856..76861158	909	1	303	7.83	31.881	Nuc	1
<i>SIAHL4</i>	Solyc01g094460.2.1	7769440..77694068	1020	4	340	10.40	34.866	Nuc	1
<i>SIAHL5</i>	Solyc02g038760.2.1	22246946..22247272	981	4	327	6.97	33.088	Nuc	2
<i>SIAHL6</i>	Solyc02g082300.2.1	40588459..40588720	786	4	262	4.91	26.955	Nuc	2
<i>SIAHL7</i>	Solyc03g007150.2.1	1735261..1735586	978	4	326	9.23	33.701	Nuc	3
<i>SIAHL8</i>	Solyc04g053090.1.1	50662361..50662530	510	4	170	6.81	17.506	Chlo	4
<i>SIAHL9</i>	Solyc04g053100.1.1	50767217..50767383	501	4	167	8.38	17.446	Chlo	4
<i>SIAHL10</i>	Solyc04g076220.2.1	58757542..58757878	1011	1	337	7.32	36.245	Nuc	4
<i>SIAHL11</i>	Solyc04g082810.2.1	63940809..63941054	738	1	246	6.08	25.037	Nuc	4
<i>SIAHL12</i>	Solyc05g053000.1.1	62313115..62313428	942	0	314	5.31	33.095	Nuc	5
<i>SIAHL13</i>	Solyc06g035430.2.1	21003305..21002959	1041	4	347	6.54	36.409	Nuc	6
<i>SIAHL14</i>	Solyc06g060500.1.1	34916326..34916106	663	0	221	6.51	23127.200	Nuc	6
<i>SIAHL15</i>	Solyc07g056300.1.1	61475095..61474844	756	0	252	9.16	26.577	Nuc	7
<i>SIAHL16</i>	Solyc08g006950.2.1	1516472..1516081	1176	4	392	6.50	40.209	Chlo	8
<i>SIAHL17</i>	Solyc08g007530.2.1	2078161..2077833	987	4	329	9.40	33.564	Chlo	8
<i>SIAHL18</i>	Solyc08g007560.1.1	2106197..2106516	960	1	320	6.09	34.336	Nuc	8
<i>SIAHL19</i>	Solyc08g008030.2.1	2528486..2528140	1041	4	347	9.70	35.386	Nuc	8
<i>SIAHL20</i>	Solyc08g059690.1.1	41863886..41863726	483	2	161	6.10	17.294	Nuc	8
<i>SIAHL21</i>	Solyc08g077010.1.1	58127804..58128092	867	0	289	5.77	30.228	Nuc	8
<i>SIAHL22</i>	Solyc08g077030.2.1	58143474..58143065	1230	4	410	7.45	42.851	Nuc	8
<i>SIAHL23</i>	Solyc08g079630.2.1	60280510..60280189	966	4	322	9.05	33.385	Chlo	8
<i>SIAHL24</i>	Solyc08g079720.2.1	60370198..60369906	879	0	293	6.05	31.085	Nuc	8
<i>SIAHL25</i>	Solyc08g080960.2.1	61259864..61259530	1005	4	335	9.68	34.167	Nuc	8
<i>SIAHL26</i>	Solyc09g008940.2.1	2333443..2333860	1254	4	418	6.51	41.506	Nuc	9
<i>SIAHL27</i>	Solyc09g089620.1.1	64729527..64729241	861	0	287	5.53	30.132	Nuc	9
<i>SIAHL28</i>	Solyc11g006090.1.1	876985..877239	765	3	255	5.85	25.885	Nuc	11
<i>SIAHL29</i>	Solyc11g006190.1.1	941588..941251	1014	4	338	9.25	34.855	Nuc	11
<i>SIAHL30</i>	Solyc12g019430.1.1	10266362..10266170	579	5	193	9.42	19.889	Chlo	12
<i>SIAHL31</i>	Solyc12g087950.1.1	61893580..61893314	801	0	267	6.65	28.754	Nuc	12
<i>SIAHL32</i>	Solyc12g094710.1.1	63194536..63194179	1074	4	358	9.59	36.077	Nuc	12

Chlo; Chloroplast, Nuc; Nucleus

2.2 番茄 AT-hook 基因在染色体上的分布

根据基因位置信息得到 32 个 AT-hook 基因在番茄染色体上的定位图,分布在番茄 11 条染色体上。

由图 1 可见,番茄 AT-hook 基因在不同染色体上呈不均匀分布,其中第 8 号染色体基因数目最多,含有 10 个基因,其次是第 1,4 号染色体,分别含有 4 个成员,

第 12 号染色体上有 3 个基因,第 2、6、9、11 号染色体上分别有 2 个基因,第 3、5、7 号染色体上仅有 1 个基因,第 10 号染色体没有 AT-hook 基因的分布(图 1)。

AT-hook 在番茄基因组中存在 8 个基因簇,第 1、9、11 号染色体包含 1 个基因簇,第 4 号染色体上包含 2 个基因簇,第 8 号染色体上基因簇多达 3 个。

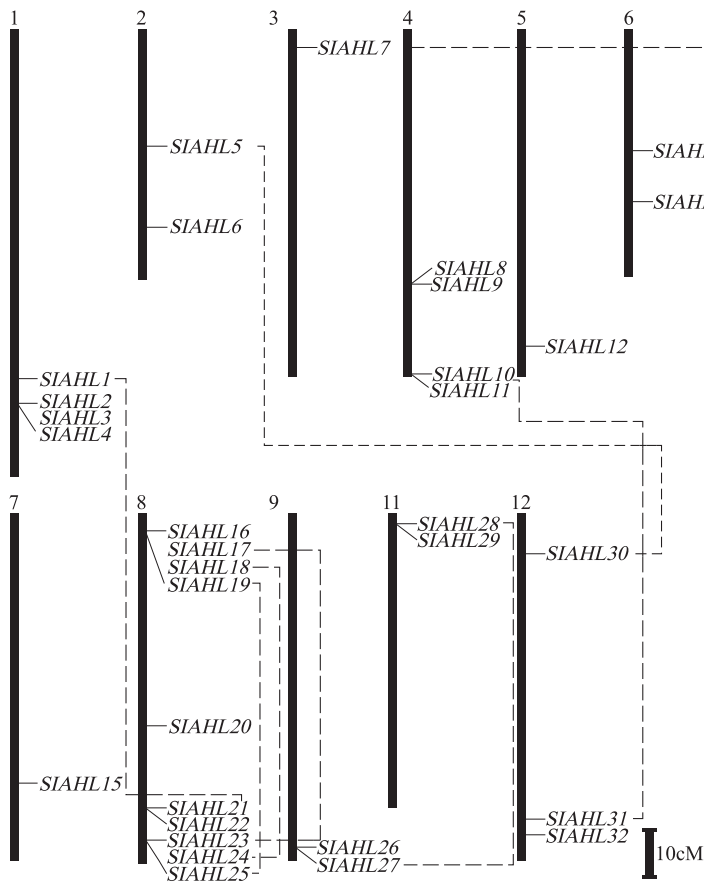


图 1 番茄 AHL 基因在染色体上的分布
Fig. 1 Chromosomal location of tomato AHL gene family

2.3 AT-hook 家族基因系统进化分析

利用 MEGA 6 软件采用邻接法将 32 个番茄与 34 条拟南芥 AT-hook 氨基酸序列生成系统进化树(图 2)。根据基因定位分析及系统发育树发现,AT-hook 在进化过程中存在片段和串联复制现象,在番茄片段复制区域找到 9 对旁系同源基因,第 1 号染色体上的 *SIAHL1* 与第 8 号染色体上的 *SIAHL21*,第 2 号染色体上的 *SIAHL5* 与第 12 号染色体上的 *SIAHL30*,第 3 号染色体上的 *SIAHL7* 与第 6 号染色体上的 *SIAHL13*,第 4 号染色体上的 *SIAHL8* 与 *SIAHL9*,第 4 号染色体上的 *SIAHL10* 与第 12 号染色体上的 *SIAHL31*,第 8 号染色体上的 *SIAHL17* 与 *SIAHL23*,*SIAHL18* 与 *SIAHL24*,*SIAHL19* 与 *SIAHL25*,第 9 号染色体上的 *SIAHL27* 与第 11 号染色体上的 *SIAHL28*,这些同源对在图 1 中用虚线连接,它们序列同源性高,基因组染色体复制信息与系统进化树

结果相一致。

2.4 番茄 AT-hook 家族基因基序分析

为了进一步分析番茄 AT-hook 家族基因氨基酸的保守性,利用 Meme 4. 6. 1 在线工具开展保守基序(motif)分析(图 3A),结果表明在 AT-hook 基因家族中含有 15 个保守基序(表 3),通常聚类分析在一个亚族的蛋白质有着相似的 motif 组成,其中基序 1、2 和 3 是 AT-hook 家族保守基序,均出现在 31 个基因中,且基序 3 中甘氨酸-精氨酸-苯丙氨酸-谷氨酸-异亮氨酸-亮氨酸(GRFEIL)6 个残基为 PPC/DUF296 结构域核心保守氨基酸(图 3B、C 和 D);前人根据 AT-hook 结构域及 PPC/DUF296 结构域将 AT-hook 基因家族分为 3 类^[2],其中 AT-hook 结构域根据精氨酸-甘氨酸-精氨酸-脯氨酸(RGRP)4 个残基前后氨基酸不同分为 AT-hook 结构域 1(基序 4,核心保守氨基酸为:脯氨酸-精氨酸-甘氨酸-精氨

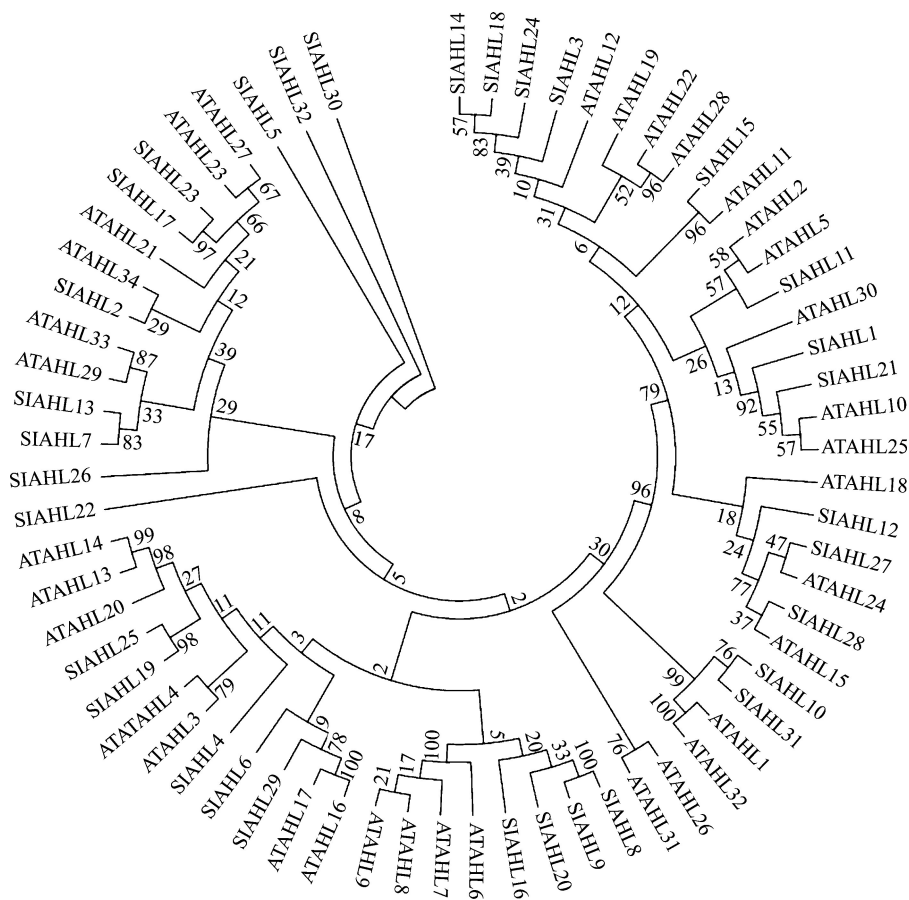


图 2 番茄与拟南芥中 AHL 基因家族蛋白质系统进化分析

Fig.2 Phylogenetic tree of AHL gene family in tomato and *Arabidopsis thaliana*

酸-脯氨酸 PRGRP) 和结构域 2(基序 5, 核心保守氨基酸为: 赖氨酸-赖氨酸-赖氨酸-精氨酸-甘氨酸-精氨酸-脯氨酸 KKKRGRP) (表 3, 图 3B、C 和 D); PPC/DUF296 结构域根据其起始氨基酸不同分为 A (亮氨酸-精氨酸-丝氨酸-组氨酸 LRSH) 和 B (基序 6, 苯丙氨酸-苏氨酸-苯丙氨酸-组氨酸 FTPH) 两类 (表 3, 图 3B、C 和 D); 本研究发现番茄中有 13 个成员为第 I 类型 (Type I) 含有 AT-hook 结构域 1 和 PPC/DUF296 结构域 A (图 3B); 番茄中有 7 个成员为第 II 类型 (Type II) 含有 AT-hook 结构域 1、2 和 PPC/DUF296 结构域 B (图 3C); 有 7 个成员为第 III 类型 (Type III) 含有 AT-hook 结构域 2 和 PPC/DUF296 结构域 B (图 3D); 剩余 5 个基因含有部分 AT-hook 或 PPC/DUF296 结构域, 但因氨基酸保守性不同无法归为 3 个类型当中。

2.5 SIAHL 基因组织表达分析

选取番茄 AT-hook 基因家族中的 32 个基因开展根、茎、叶、花以及成熟果实中的表达分析, 结果显示该家族绝大多数成员在番茄各组织或发育阶段中基

本都有表达 (图 4)。SIAHL1、SIAHL3、SIAHL6、SIAHL8、SIAHL11、SIAHL15、SIAHL16、SIAHL20 和 SIAHL28 在根的表达量相对于其他家族成员相对较低, SIAHL4、SIAHL5 的表达水平在茎中受到限制, 水平较低, 而 SIAHL1、SIAHL3、SIAHL8、SIAHL18、SIAHL28 在茎中的表达水平较高。SIAHL15 和 SIAHL20 在根茎中表达模式没有明显差异。在叶片中, 所有基因的表达水平均较低, 而且 SIAHL1、SIAHL3、SIAHL15、SIAHL16、SIAHL18、SIAHL19、SIAHL20、SIAHL21、SIAHL28、SIAHL30 在叶中几乎不表达。SIAHL10、SIAHL12、SIAHL13、SIAHL29、SIAHL31 在叶片中的表达量相对较高。SIAHL2、SIAHL6、SIAHL7、SIAHL8、SIAHL14、SIAHL17、SIAHL23、SIAHL24、SIAHL25 在营养器官中特异性表达, 表达量相对较高。以上结果表明, 有 8 个番茄 SIAHLs (SIAHL2、SIAHL6、SIAHL7、SIAHL19、SIAHL24、SIAHL25、SIAHL29、SIAHL32) 基因在花中的表达较高和 4 个 (SIAHL15、SIAHL16、SIAHL20、SIAHL30) 基因在果实中表达较高, 这些基因可能参与番茄花和果实的发育过程。

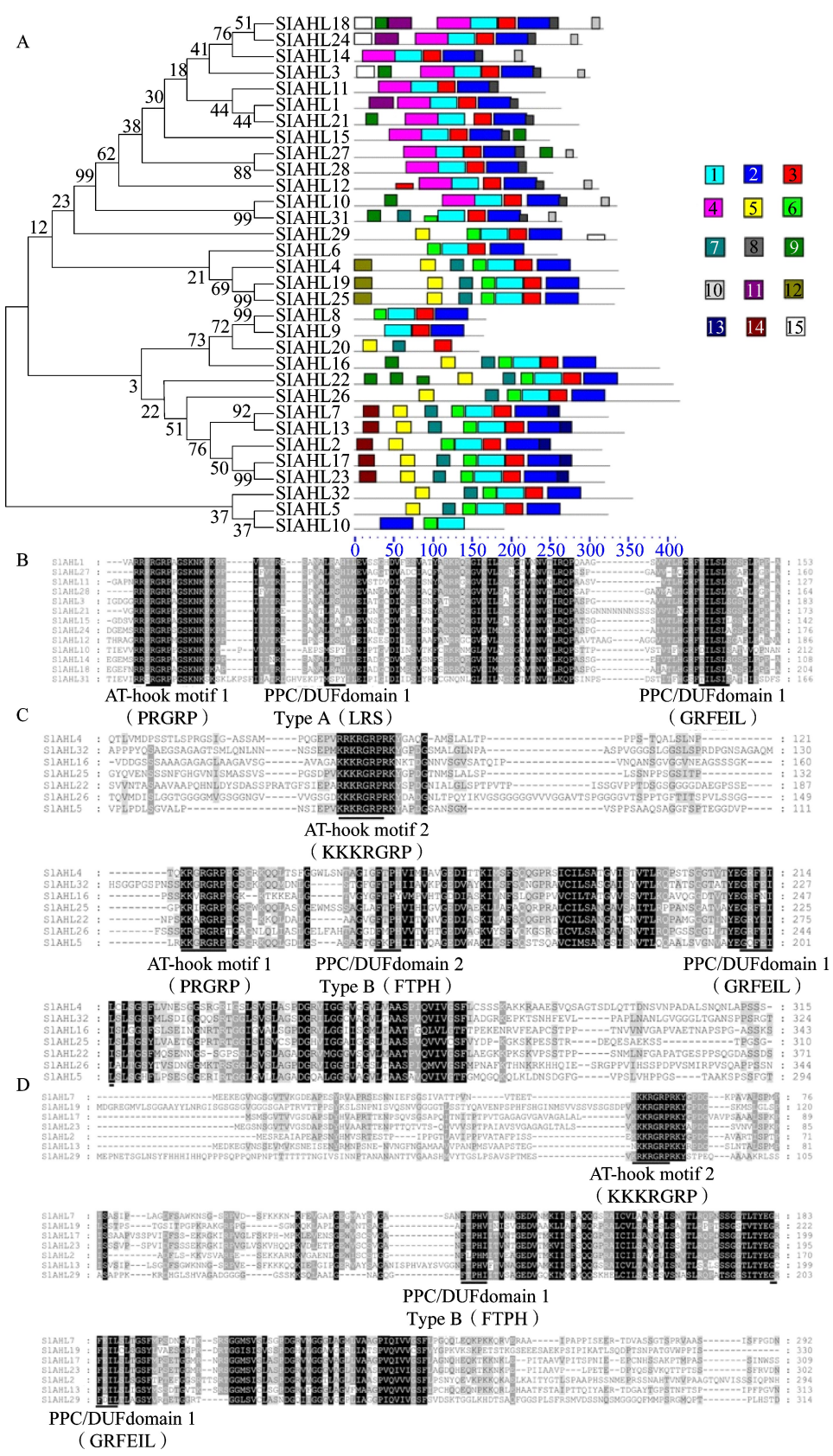


图3 番茄 AHL 基因保守基序组成分析及家族分类

Fig. 3 Classification and conserved motif of tomato AHL gene family

表 3 利用 MEME 对番茄 AHL 蛋白的基序分析

Table 3 List of putative conserved motifs in tomato AHL proteins predicted by MEME

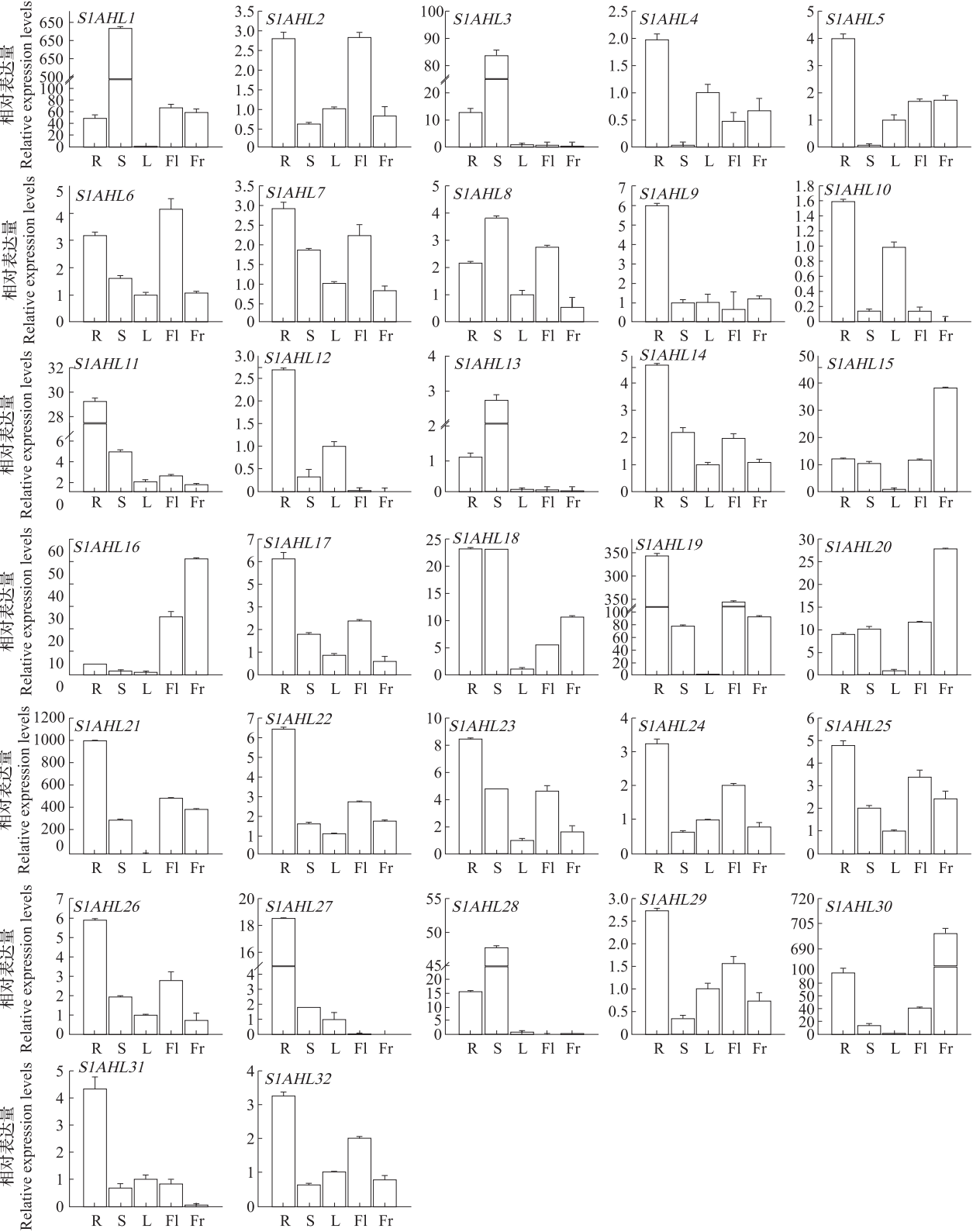
基序编号	基序序列	番茄基因数目	E 值	基序长度
Motif	Motif sequence	Number of site in tomato	E-value	(bp)
code				Motif length
1	IMNFCQQRQRAICILSATGCVSNVTLRQPASSG	31	3.6e-431	33
2	GLTIYLAGPDGQVWGCVVGMLIAAGPVQVIVCSFLNAKYE	31	3.3e-506	41
3	TYHGRFEILSLSGSFLPPEAP	31	7.2e-311	21
4	PRGRPAGSKNKPKPPIIITRDSANALRTHIMEIADGCDIME	12	1.8e-252	41
5	EPVKKKRCRPRKYGPDG	15	9.5e-115	17
6	NFTPHVITVKTGEDV	17	1.8e-095	15
7	KRGRGRPPGSGWKQQ	14	2.5e-040	15
8	RLPLEEDDE	9	1.7e-026	9
9	QHHHYQQQQHQQHHH	8	2.6e-016	15
10	QLPHEAYW	3	4.9e-008	8
11	EDEQDGTSLNMSGQKRDRDDNNNNNNNN	3	5.2e-008	29
12	MDRREGMMLSGSAAYYLNIRGI	5	1.8e-007	21
13	EQPKKKQKLEPIAI	5	3.2e-006	14
14	VTVKGEAPEDYHMAPRTE	2	8.4e-004	19
15	MDPHNHSLPPPFHTRDFNLQH	3	1.3e-003	21

2.6 SIAHL 基因氧化胁迫条件下的响应分析

为了检测 *SIAHL* 基因在番茄中对胁迫的应答反应,本文利用 RT-qPCR 技术,分析了 32 个 *SIAHL* 基因在 SA、ABA、盐胁迫、高温(42 ℃)和低温(4 ℃)处理下的表达模式。选取生长良好的番茄幼苗,分别进行 1 h、2 h、8 h 的胁迫处理,并且以 0 h 未处理条件下 *SIAHLs* 基因相对 *Slactin* 的表达量设为 1。结果表明,基因在激素处理条件下,由 ABA 处理过的植株,几乎所有的基因表达量逐渐上升,而且在 8 h 表达量达到峰值。*SIAHL1*、*SIAHL2*、*SIAHL5*、*SIAHL6*、*SIAHL7*、*SIAHL11*、*SIAHL15*、*SIAHL17*、*SIAHL19*、*SIAHL22*、*SIAHL25*、*SIAHL32* 表现比较明显,*SIAHL13* 随着胁迫时间延长表达水平一直降低(图 5A、B)。*SIAHL16*、*SIAHL18*、*SIAHL28*、*SIAHL3*、*SIAHL9*、*SIAHL26*、*SIAHL29* 表达量呈现出先降后升的趋势。*SIAHL2*、*SIAHL4*、*SIAHL10*、*SIAHL13*、*SIAHL17*、*SIAHL19*、*SIAHL21*、*SIAHL22*、*SIAHL25*、*SIAHL26*、*SIAHL31*、*SIAHL32* 则相反,表达量的变化趋势呈现出先升后降(图 5A、B)。由 SA 处理过的植株,如图 5C 所示,基因相对表达量变化趋势呈现出先升后降的趋势,在 2 h 时达到最高值。10 个基因受 SA 诱导相对表达量逐渐提高(图 5D),5 个基因相对表达量最高,到达 8 h 时的表达量达到最高值,且相对表达量随

着处理时间的增长持续上升(图 5E)。

在盐胁迫条件下,几乎所有基因的表达水平都是逐渐降低的(图 6A、B)。*SIAHL1*、*SIAHL2*、*SIAHL4*、*SIAHL8*、*SIAHL13*、*SIAHL14*、*SIAHL18*、*SIAHL20*、*SIAHL24*、*SIAHL28*、*SIAHL30*、*SIAHL31* 的表达趋势是先上升后下降的,*SIAHL5*、*SIAHL6*、*SIAHL17*、*SIAHL19*、*SIAHL26*、*SIAHL29* 的表达量从 0 h 开始一直上升,而 *SIAHL3*、*SIAHL7*、*SIAHL11*、*SIAHL16*、*SIAHL21*、*SIAHL22*、*SIAHL23*、*SIAHL25*、*SIAHL32* 却呈现出先下降后上升的趋势。只有 *SIAHL9*、*SIAHL10* 的表达量一直在下降。在高温条件下(图 6C、D),32 个基因整体表达量受到抑制,大多数基因表达趋势是上升的,例如 *SIAHL1*、*SIAHL2*、*SIAHL5*、*SIAHL8*、*SIAHL10*、*SIAHL11*、*SIAHL12*、*SIAHL26*、*SIAHL29*、*SIAHL32*,只有 *SIAHL28*、*SIAHL31* 的表达水平是下降的,其他基因的表达量没有明显变化。在低温条件下(图 6E、F),所有基因的相对表达量下调,但是 *SIAHL7*、*SIAHL10*、*SIAHL11*、*SIAHL12*、*SIAHL15*、*SIAHL16*、*SIAHL17*、*SIAHL19*、*SIAHL20*、*SIAHL22*、*SIAHL23*、*SIAHL25*、*SIAHL26*、*SIAHL27*、*SIAHL28*、*SIAHL30*、*SIAHL32* 表达趋势是先降后升的,而 *SIAHL2*、*SIAHL4*、*SIAHL5*、*SIAHL6*、*SIAHL14* 的表达水平则稳步提高。



R:根;S:茎;L:叶;Fl:花;Fr:果实(花后 20 d)

R:Root,S:Stem,L:Leaf,Fl:Flower,Fr:Fruit(20 days after flower)

图 4 AHL 基因在番茄器官中的表达分析

Fig.4 Expression profiles of tomato AHL genes in various tissues

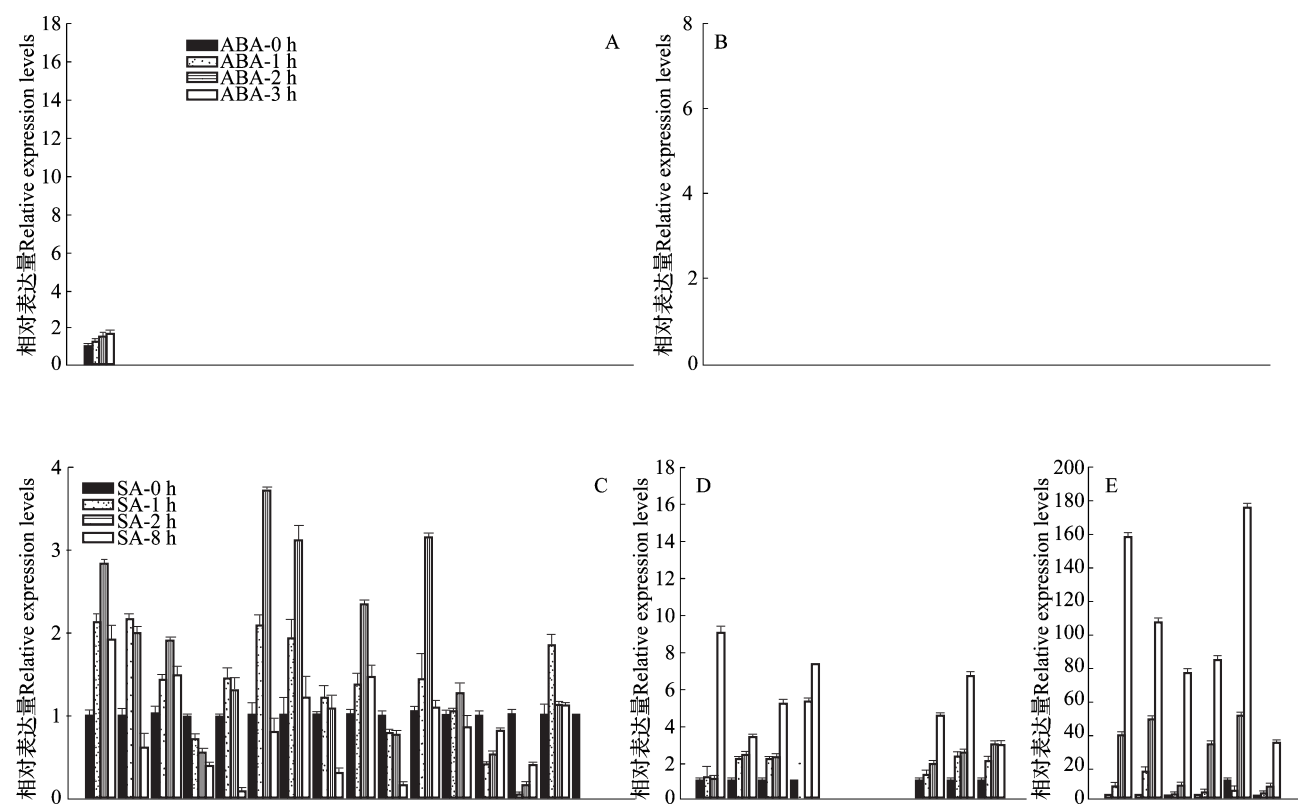


图5 AHL 基因在激素 ABA 和 SA 处理条件下的表达分析
Fig.5 Expression profiles of AHL genes under ABA and SA treatment

3 讨论

随着番茄基因组计划的完成^[24],从全基因组层面鉴定和研究基因家族的分类、进化特征和功能预测是番茄功能基因研究的热点^[25-28]。AT-hook 蛋白在染色质结构组装和对目标基因转录活性调控中起着重要作用^[5]。国内外学者在模式植物拟南芥和水稻 AT-hook 基因家族及功能研究方面开展了较多的研究,发现 AT-hook 蛋白主要参与植物激素调控、光诱导成花、下胚轴的伸长等生长发育过程^[6,11]。J. Zhao 等^[11] 2014 年对其他作物如小麦、大豆、豌豆、木薯、蓖麻、毛果杨、蒺藜状苜蓿、黄瓜、碧桃、苹果、芜菁、番木瓜、葡萄、高粱、玉米、二穗短柄草、江南卷柏、小立碗藓的 AT-hook 家族基因及成员进化做了较为深入的研究。水稻和拟南芥分别含有 24 个和 29 个 AT-hook 成员,其中第 I 类型 (Type I) 分别还有 13 个和 15 个,第 II 类型 (Type II) 分别含有 9 个和 7 个,第 III 类型 (Type III) 分别含有 2 个和 7 个,前人通过对 AT-hook 家族成员进化分析发现,在长期的进化过程中,第 I 类型成员较多,而第 III 类型

在双子叶植物中成员较多,而单子叶植物成员较少如水稻含有 2 个,玉米有 1 个,而高粱则没有^[11]。本研究发现番茄含有 32 个 AT-hook 家族基因成员,3 种类型分别为 13 个、7 个和 7 个,这和拟南芥 3 种类型成员的个数相近,另外 5 个成员则仅含有 AT-hook 或 PPC/DUF296 结构域,张贵慰等^[10] 2014 年研究发现水稻 AT-hook 家族含有 45 个成员,其中包含 J. Zhao 等^[11] 发现水稻有 24 个成员,及其他一些仅含部分结构域的成员。通过比较番茄和拟南芥 AT-hook 成员各类型的进化关系发现,二者进化特征相似;番茄相比拟南芥多 3 个成员,其中在第 8 条染色体上存在 *SlAHL16* ~ *SlAHL25* 共 10 个成员及 3 处串联复制现象。

在番茄 32 个 AT-hook 基因中含有 15 个基序,其中基序 1、2、3 分别出现在 31 个基因中,且基序 3 为 PPC/DUF296 结构域核心保守氨基酸,基序 4 和 5 分别为 AT-hook 结构域分类的依据基序。根据 MEME 基序分析及序列比对,结果表明番茄 AT-hook 家族基因与拟南芥具有相似的保守基序及特征分类。

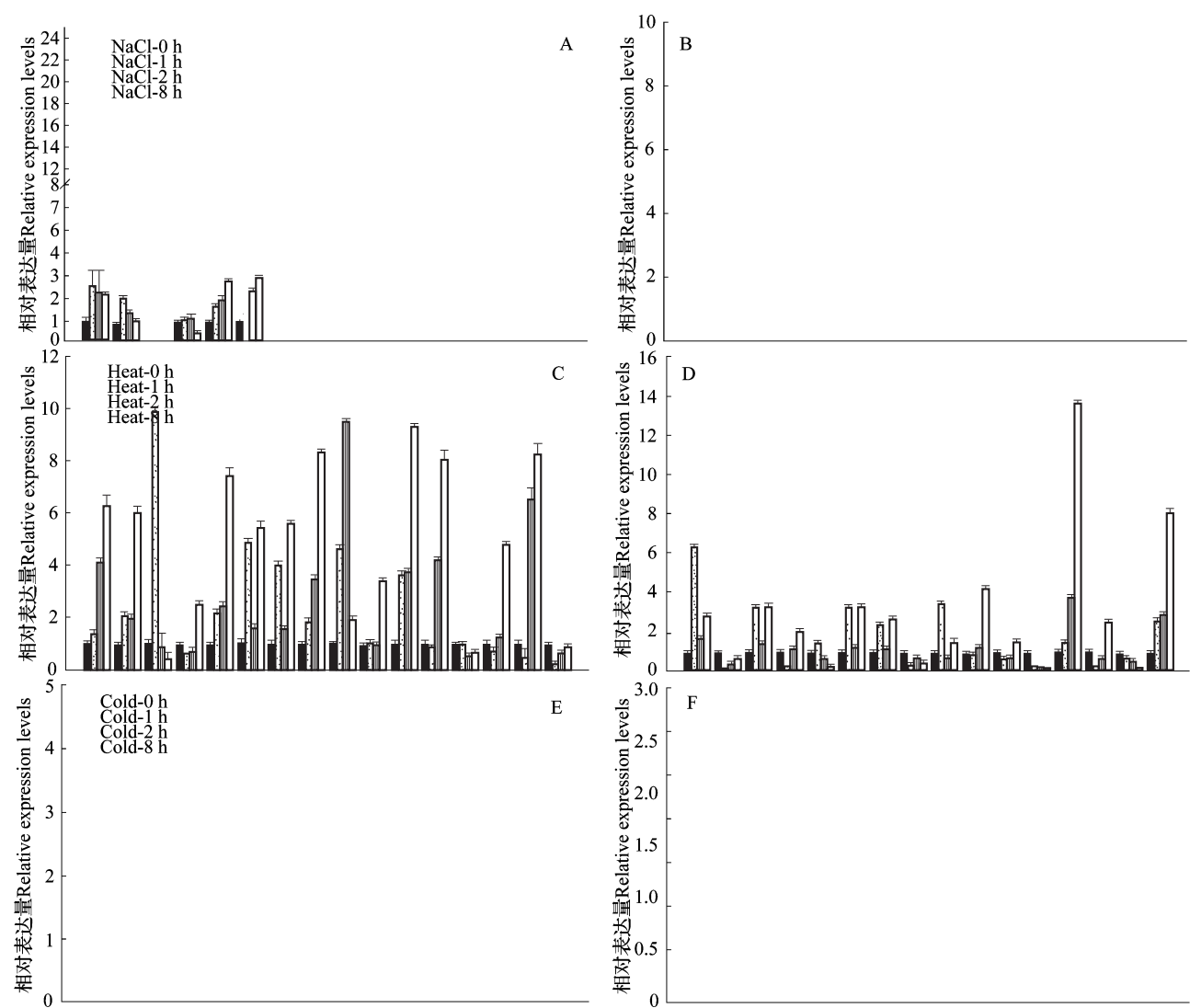


图6 AHL 基因在不同胁迫处理条件下的表达分析

Fig. 6 Expression profiles of AHL genes under different abiotic stress treatments

经亚细胞定位分析,在番茄 32 个成员中,26 个成员定位于细胞核中,这与前人研究 AT-hook 蛋白大多定位于细胞核的结果基本一致^[6]。组织表达分析结果表明,AT-hook 家族基因在番茄各组织中基本都有表达,且在不同成员组织的表达存在差异,大多数基因在根和花器官中表达较高,这与前人报道的 AT-hook 蛋白参与成花调控相关^[17,20],可能预示番茄 AT-hook 基因也参与了花器官的发育过程。对 32 个基因开展氧化胁迫响应分析发现,大多数基因受胁迫诱导表达,在 ABA 处理条件下 18 个成员表达持续上调,4 个基因表达下调;在 SA 处理条件下, *SlAHL8*、*SlAHL9*、*SlAHL14*、*SlAHL15*、*SlAHL20* 和 *SlAHL30* 受 SA 诱导表达显著上调, *SlAHL12* 则显著

下调;32 个基因在盐、高温和低温处理条件下均有响应上调或下调。这为进一步研究番茄 AT-hook 家族基因的功能及机制提供了重要的依据。

参考文献

[1] Eckner R, Birnstiel M L. Cloning of cDNAs coding for human HMG I and HMG Y proteins; both are capable of binding to the octamer sequence motif [J]. Nucleic Acids Res, 1989, 17 (15) : 5947-5959

[2] Aravind L, Landsman D. AT-hook motifs identified in a wide variety of DNA-binding proteins [J]. Nucleic Acids Res, 1998, 26 (19) : 4413-4421

[3] Do H J, Song H, Yang H M, et al. Identification of multiple nuclear localization signals in murine E1B, an ETS transcription factor [J]. FEBS Lett, 2006, 580 (7) : 1865-1871

[4] Goodwin G H, Sanders C, Johns E W. A new group of chromatin associated proteins with a high content of acidic and basic amino

- acids[J]. Eur J Biochem, 1973, 38(1): 14-19
- [5] 肖朝文, 傅永福. AT-hook 蛋白的研究进展[J]. 中国农业科技导报, 2009(5): 12-16
- [6] Zhao J, Favero D S, Peng H, et al. *Arabidopsis thaliana* AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain[J]. Proc Natl Acad Sci USA, 2013, 110(48): 4688-4697
- [7] Fujimoto S, Matsunaga S, Yonemura M, et al. Identification of a novel plant MAR DNA binding protein localized on chromosomal surfaces[J]. Plant Mol Biol, 2004, 56(2): 225-239
- [8] Lin L, Nakano H, Nakamura S, et al. Crystal structure of *Pyrococcus horikoshii* PPC protein at 1.60 Å resolution[J]. Proteins, 2007, 67(2): 505-507
- [9] Lin L, Nakano H, Uchiyama S, et al. Crystallization and preliminary X-ray crystallographic analysis of a conserved domain in plants and prokaryotes from *Pyrococcus horikoshii* OT3[J]. Acta Crystallogr F-Struct Biol, 2005, 61(4): 414-416
- [10] 张贵慰, 曾珏, 郭维, 等. 水稻 AT-hook 基因家族生物信息学分析[J]. 植物学报, 2014, 49(1): 49-62
- [11] Zhao J, Favero D S, Qiu J, et al. Insights into the evolution and diversification of the AT-hook Motif Nuclear Localized gene family in land plants[J]. BMC Plant Biol, 2014, 14(1): 266
- [12] Matsushita A, Furumoto T, Ishida S, et al. AGF1, an AT-hook protein, is necessary for the negative feedback of AtGA3ox1 encoding GA 3-oxidase [J]. Plant Physiol, 2007, 143(3): 1152-1162
- [13] Vom Endt D, e Silva M S, Kijne J W, et al. Identification of a bipartite jasmonate-responsive promoter element in the *Catharanthus roseus* ORCA3 transcription factor gene that interacts specifically with AT-Hook DNA-binding proteins[J]. Plant Physiol, 2007, 144(3): 1680-1689
- [14] Rashotte A M, Carson S D B, To J P C, et al. Expression profiling of cytokinin action in *Arabidopsis* [J]. Plant Physiol, 2003, 132(4): 1998-2011
- [15] Yun J, Kim Y S, Jung J H, et al. The AT-hook motif-containing protein AHL22 regulates flowering initiation by modifying FLOWERING LOCUS T chromatin in *Arabidopsis* [J]. J Biol Chem, 2012, 287(19): 15307-15316
- [16] Lim P O, Kim Y, Breeze E, et al. Overexpression of a chromatin architecture - controlling AT - hook protein extends leaf longevity and increases the post - harvest storage life of plants[J]. Plant J, 2007, 52(6): 1140-1153
- [17] Xiao C, Chen F, Yu X, et al. Over-expression of an AT-hook gene, *AHL22*, delays flowering and inhibits the elongation of the hypocotyl in *Arabidopsis thaliana* [J]. Plant Mol Biol, 2009, 71(1-2): 39-50
- [18] Ng K H, Yu H, Ito T. AGAMOUS controls GIGANT KILLER, a multifunctional chromatin modifier in reproductive organ patterning and differentiation[J]. PLoS Biol, 2009, 7(11): e1000251
- [19] Street I H, Shah P K, Smith A M, et al. The AT - hook - containing proteins SOB3/AHL29 and ESC/AHL27 are negative modulators of hypocotyl growth in *Arabidopsis* [J]. Plant J, 2008, 54(1): 1-14
- [20] Jin Y, Luo Q, Tong H, et al. An AT-hook gene is required for palea formation and floral organ number control in rice[J]. Develop Biol, 2011, 359(2): 277-288
- [21] Morisawa G, Han-yama A, Moda I, et al. AHM1, a novel type of nuclear matrix - localized, MAR binding protein with a single AT hook and a J domain - homologous region[J]. Plant Cell, 2000, 12(10): 1903-1916
- [22] Delaney S K, Orford S J, Martin-Harris M, et al. The fiber specificity of the cotton *FS1tp4* gene promoter is regulated by an AT-rich promoter region and the AT-hook transcription factor GhAT1 [J]. Plant Cell Physiol, 2007, 48(10): 1426-1437
- [23] Kim S, Kim Y C, Seong E S, et al. The chili pepper CaATL1: an AT - hook motif - containing transcription factor implicated in defence responses against pathogens[J]. Mol Plant Pathol, 2007, 8(6): 761-771
- [24] The tomato genome consortium. The tomato genome sequence provides insights into fleshy fruit evolution[J]. Nature, 2012, 485(7400): 635-641
- [25] 孙保娟, 李植良, 金庆敏, 等. 60 份优质番茄自交系遗传多样性 AFLP 分析[J]. 植物遗传资源学报, 2013, 14(5): 788-793
- [26] 钟晓武, 付强, 邹颀, 等. 普通烟草八氢番茄红素脱氢酶基因的原核表达及表达谱分析[J]. 植物遗传资源学报, 2014, 15(4): 838-843
- [27] 王柏柯, 杨生保, 余庆辉, 等. 加工番茄种质资源的 SSR 分析[J]. 植物遗传资源学报, 2014, 15(1): 196-200
- [28] 刘磊, 宋燕, 郑峰, 等. 利用 *Solanum pennellii* LA0716 渐渗系群体初步定位番茄果实硬度 QTL[J]. 植物遗传资源学报, 2015, 16(2): 323-329
- [29] Punta M, Coggill P C, Eberhardt R Y, et al. The Pfam protein families database[J]. Nucleic Acids Res, 2012, 40: 290-301
- [30] Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource[J]. Nucleic Acids Res, 2012, 40(1): 302-305
- [31] Horton P, Park K J, Obayashi T, et al. WoLF PSORT: protein localization predictor [J]. Nucleic Acids Res, 2007, 35(2): 585-587
- [32] Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. Nucleic Acids Res, 1997, 25(24): 4876-4882
- [33] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. Mol Biol Evol, 2011, 28(10): 2731-2739