

郭 琪, 徐珍珍, 黄 芳, 等. 棉花 *HKT* 基因家族的全基因组分析[J]. 江苏农业学报, 2017, 33(5): 975-981.
doi: 10.3969/j.issn.1000-4440.2017.05.003

棉花 *HKT* 基因家族的全基因组分析

郭 琪¹, 徐珍珍¹, 黄 芳², 徐 鹏¹, 张香桂¹, 沈新莲¹

(1. 江苏省农业科学院经济作物研究所/农业部长江下游棉花和油菜重点实验室, 江苏 南京 210014; 2. 南通大学生命科学院, 江苏 南通 226019)

摘要: *HKT* (high-affinity K^+ transporter) 转运蛋白在调节植物体内 K^+ / Na^+ 转运, 参与逆境胁迫调控中发挥重要作用。本研究利用生物信息学方法, 系统分析了棉花 *HKT* 基因家族的分类、系统进化树、组织表达以及 *HKT* 蛋白的亚细胞定位、跨膜结构、基序情况等。结果显示, 棉花中共有 8 个 *HKT* 基因家族成员, 编码的 8 个 *HKT* 蛋白都定位到了质膜上, 具有 7~10 个不等的跨膜数; 系统进化树分析结果表明, 棉花 8 个 *HKT* 基因家族成员大致分为 2 个小组, 在进化树上分支比较近的 *HKT* 基因家族成员编码的 *HKT* 蛋白具有完全相同或者比较类似的 motif 类型和跨膜数; 组织表达分析发现, 克劳茨基棉仅有 *GrHKT1;1* 和 *GrHKT1;2* 2 个成员在 200 mmol/L 的 NaCl 溶液处理 0 h、3 h、12 h、72 h 和 144 h 后具有不同程度的响应。

关键词: *HKT* 基因家族; 生物信息; 棉花

中图分类号: S562 **文献标识码:** A **文章编号:** 1000-4440(2017)05-0975-07

Genome-wide analysis of high-affinity potassium transporter gene family in cotton

GUO Qi¹, XU Zhen-zhen¹, HUANG Fang², XU Peng¹, ZHANG Xiang-gui¹, SHEN Xin-lian¹

(1. Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences/Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture, P.R. China, Nanjing 210014, China; 2. School of Life Science, Nantong University, Nantong 226019, China)

Abstract: High-affinity potassium transporter (*HKT*) is involved in K^+ / Na^+ transportation, and plays an important role in adaptation to abiotic stress. The classification of *HKT* gene family, *HKT* protein subcellular localization, transmembrane domain, and tissue expression pattern in cotton were systematically analyzed by bioinformatics in this study. Eight *HKT* genes were identified, and all proteins encoded were localized on the plasma membrane, with 7–10 transmembrane domains. Phylogenetic tree revealed that 8 *HKT* genes could be divided into 2 groups, and identical or similar motif type and transmembrane domain were found in closely-related branches. Only *GrHKT1;1* and *GrHKT1;2* were differentially expressed in *Gossypium Klotzschianum* subjected to 200 mmol/L NaCl treatment for 0 h, 3 h, 12 h, 72 h and 144 h.

Key words: *HKT* gene family; bioinformatics; cotton

收稿日期: 2017-04-01

基金项目: 国家转基因生物新品种培育重大专项 (2016ZX08005-004-002); 江苏省协同创新中心项目; 江苏省农业科技自主创新基金项目 [CX (14) 2065]; 江苏省科技项目 (BK20150540)

作者简介: 郭 琪 (1987-), 女, 山西太原人, 硕士, 助理研究员, 主要从事植物基因分子生物学研究。 (Tel) 025-84390291; (E-mail) sevenguo86@163.com

通讯作者: 沈新莲, (E-mail) xlshen68@126.com

钾 (K^+) 是高等植物的主要营养元素之一, 在植物光合作用、体内矿质元素代谢、向性生长、渗透调控以及产量形成中发挥重要作用^[1-5]。*HKT* (high-affinity K^+ transporter) 类蛋白是高亲和 K^+ 转运体, 也是一种 Na^+ 转运载体, 具有 Na^+ 和 K^+ 转运的双重功能, 即低 Na^+ 时充当 Na^+ 和 K^+ 的协同转运体, 高 Na^+ 时仅运输 Na^+ ^[6]。值得注意的是, 这类蛋白对 Na^+ 和 K^+ 的选择性不同, 因此, 在调节植物体内 K^+ 和

Na^+ 平衡中发挥着重要作用^[7-9]。

HKT 类蛋白首先在小麦 (*Triticum aestivum* L.) 中被克隆得到^[9], 命名为 TaHKT1。随后在拟南芥^[10] (*Arabidopsis thaliana* L.)、水稻^[11-12] (*Oryza sativa* L.)、大麦^[13] (*Hordeum vulgare* L.)、桉树^[14] (*Eucalyptus camaldulensis*) 及冰日水中花^[15] (*Mesembryanthemum crystallinum* L.) 中都陆续发现了 HKT 类蛋白。研究结果表明, HKT 蛋白的分子结构中包括多个跨膜 (Transmembrane domain) 区和孔环 (Pore-loop, P-loop)^[16-17]。通过拟南芥 *AtHKT1* 和水稻 *OsHKT1* 基因编码的氨基酸序列相似性分析发现, 第 1 个 P-loop 中存在氨基酸的突变。根据这一特征, HKT 蛋白分为 2 种类型。类型 I: 第 1 个 P-Loop 中含有丝氨酸, 丝氨酸是 Na^+ 特异性载体。该亚家族成员主要来自双子叶植物^[10], 如拟南芥、桉树 (*Eucalyptus camaldulensis*)、盐地碱蓬 (*Suaeda salsa* L.) 和杨树 (*Acer buergerianum* Miq L.)。也包括单子叶植物水稻中的部分 HKT 蛋白家族成员; 类型 II: 第 1 个 P-loop 中含有甘氨酸^[18], 是 K^+ 选择性载体。该亚家族成员主要来自单子叶植物, 如芦苇 (*Phragmites australis* L.)、大麦、小麦 (*Triticum aestivum* L.) 以及水稻中的部分成员^[11, 18-19]。

研究结果表明, 植物 HKT 基因家族成员参与抗逆境胁迫^[20-22]。拟南芥和水稻的 HKT 蛋白家族成员通过维持细胞内 Na^+ 含量以及 Na^+/K^+ 动态平衡调控其耐盐性能^[19, 23]。例如 *OsHKT1;5* 基因通过控制根部木质部 Na^+ 的移动来平衡水稻中 Na^+/K^+ , 从而在盐胁迫条件下维持根中高浓度的 K^+ 和低浓度的 Na^+ ^[12]; 芦苇耐盐品种的 HKT1 基因在盐胁迫下识别并专一地吸收 K^+ , 对保持 Na^+/K^+ 平衡发挥重要作用^[24]; 大豆 (*Glycine max* L.) 的 *GmHKT1;2* 基因对低钾、干旱、低温、高盐和 ABA 几种不同逆境具有较强的响应^[25]。棉花中, 鲜有 HKT 基因家族响应逆境胁迫的报道。

棉花作为一种重要的经济作物, 盐渍、高温、低温和干旱等逆境胁迫严重影响棉花的产量^[25]。本研究利用公开发表的雷蒙德氏棉 (*Gossypium raimondii*)^[26]、亚洲棉石系亚 1 号^[27] (*G. arboreum*) 和陆地棉 (*G. hirsutum* L.) 遗传标准系 TM-1 的基因组信息^[28-29], 对 HKT 基因家族进行了全基因组水平的鉴定和生物信息学分析, 系统分析了棉花 HKT 基因家族的分类、系统进化树、组织表达情况以及 HKT 蛋白的亚细胞定位、跨膜结构、基序, 从而为深

入探索 HKT 基因家族的功能奠定一定的基础。

1 材料和方法

1.1 材料

棉属二倍体 D 组野生种克劳茨基棉 (*Gossypium Klotzschianum*) 种子由中国农业科学院棉花研究所海南岛野生棉种植园提供, 拟南芥和水稻的 HKT 基因家族成员 (*ATKHT1;1*、*OSHKT1;1*、*OSHKT1;2*、*OSHKT1;3*、*OSHKT2;1* 和 *OSHKT2;2*) 来自 <http://www.phytozome.com> 网站, 雷蒙德氏棉、亚洲棉和陆地棉遗传标准系 TM-1 的 CDS 和氨基酸序列来自 <http://www.phytozome.com> 和 <http://cgp.genomics.org.cn/page/species/index.jsp> 网站; 雷蒙德氏棉、亚洲棉和 TM-1 的 EST 序列来自 NCBI。

1.2 棉花 HKT 基因家族的鉴定

首先在 SMART (<http://smart.embl-heidelberg.de/>) 网站对拟南芥和水稻基因组中的 6 个 HKT 成员进行结构域预测, 然后在 Pfam (<http://pfam.janelia.org/>) 网站下载预测的 HKT 蛋白的保守结构域 Trk 的种子文件, 利用 HMMER 程序在雷蒙德氏棉、亚洲棉和 TM-1 氨基酸数据库中搜索 HKT 蛋白的同源序列。同时, 统计雷蒙德氏棉、亚洲棉和 TM-1 基因组中 HKT 基因的特征, 包括对应的基因组、基因 ID 以及编码的氨基酸长度等。

1.3 棉花 HKT 基因家族的系统进化树分析

利用 MEGA5.1 软件对雷蒙德氏棉、亚洲棉和 TM-1 基因组中 HKT 基因家族的氨基酸酸序列进行多序列比对。然后采用邻接 (Neighbor-Joining, NJ) 算法进行系统进化树分析, 并进行 500 次 Boot strap 抽样。

1.4 棉花 HKT 蛋白家族的特征分析

亚细胞定位: CELLO (<http://cello.life.nctu.edu.tw/>) 在线软件; 多序列比对: Vector NTI Advance 11.5 AlignX 软件; 跨膜结构域预测网站: <http://www.cbs.dtu.dk/services/TMHMM/> 和 <http://smart.embl-heidelberg.de/>; Motif 分析: MEME 在线软件 (<http://meme-suite.org/tools/meme>)。参数设置如下: 基序最多显示 10 个, 其他参数为默认值。

1.5 棉花 HKT 基因家族的表达分析

在 NCBI 下载雷蒙德氏棉、亚洲棉和陆地棉的 EST 序列, 分别得到 64 903、63 642 和 330 520 条序列。然后以雷蒙德氏棉、亚洲棉和陆地棉的 HKT 基因家族的 CDS 序列为 Query, 以下载好的 EST 为数

数据库,通过 Blastn 搜索,获得棉花 *HKT* 基因在不同组织的表达情况。

克劳茨基棉的种子用小刀破壳后置于清水浸泡 3 h 左右,然后将种子置于铺有 4 层卷纸的玻璃培养皿中,浇适量的蒸馏水后,用 2 层卷纸覆盖,置于 28 ℃ 的培养箱中 24 h。将发芽的种子种植于培养室中,昼夜温度分别为 28 ℃ 和 25 ℃,光照周期为 14 h 光照/10 h 黑暗。待棉花幼苗长到两叶一心时,浇 80~100 ml 浓度为 200 mmol/L 的 NaCl 溶液,分别于 NaCl 溶液中处理 0 h、3 h、12 h、72 h 和 144 h,取根部提取 RNA,进行数字基因表达谱(Digital gene expression profiling, DGE)分析。然后将 3 h、12 h、72 h 和 144 h 处理的数字基因表达谱数据分别与 0 h 处理的数据进行差异表达分析,参考基因组:雷蒙德氏棉的基因组数据和混合样品 RNA-Seq 测序得到的 Unigenes。利用 CROSS_MATCH 软件搜索 *HKT* 基因家族成员的 CDS 序列与转录组数据 Unigenes 的匹配,参数设定为匹配长度>100 bp,相似度>98%。

2 结果与分析

2.1 棉花 *HKT* 基因家族的鉴定和分类

利用同源搜索的方法,分别在雷蒙德氏棉鉴定到 2 个,亚洲棉石系亚 1 号鉴定到 2 个和陆地棉遗传标准系 TM-1 基因组中鉴定到 4 个 *HKT* 基因家族成员。根据 Platten 等^[30]提出的命名方法,我们首先对 8 个 *HKT* 蛋白家族进行了跨膜结构的预测,跨膜数为 7 个、9 个和 10 个不等(表 1)。然后,我们将棉花 8 个、拟南芥 1 个和水稻的 5 个 *HKT* 蛋白家族成员的第 1 个跨膜氨基酸序列提取出来,进行多序列比对,结果显示,除了水稻中 2 个成员 OSHKT2;1 和 OSHKT2;3 第 1 个 P-Loop 中含有甘氨酸外,剩余的其他成员第 1 个 P-Loop 中均含有丝氨酸(图 1)。因此,棉花的 *HKT* 基因家族成员属于类型 I,分别命名为:*GaHKT1;1*、*GaHKT1;2*、*GrHKT1;1*、*GrHKT1;2*、*GhHKT1;1*、*GhHKT1;2*、*GhHKT1;3* 和 *GhHKT1;4*。我们进一步统计了 8 个 *HKT* 基因家族成员对应的基因组、基因 ID 号以及编码的氨基酸长度。结果表明,棉花 8 个 *HKT* 蛋白家族成员的氨基酸长度从 478 到 650 个氨基酸不等(表 1)。通过亚细胞定位在线软件 CELLO (<http://cello.life.nctu.edu.tw/>) 分析,8 个 *HKT* 蛋白家族成员均定位在细胞质膜上。

表 1 棉花 *HKT* 基因家族

Table 1 Summary of the identified *HKT* gene family in cotton

基因名称	基因号	基因组	编码的蛋白质氨基酸长度 (aa)	编码的蛋白质跨膜数
<i>GaHKT1;1</i>	Cotton_A_40914	AA	650	10
<i>GaHKT1;2</i>	Cotton_A_11692	AA	478	7
<i>GrHKT1;2</i>	Gorai.004G096400.1	DD	556	9
<i>GrHKT1;1</i>	Gorai.012G180600.1	DD	499	9
<i>GhHKT1;3</i>	Gh_A04G1216	AADD	499	9
<i>GhHKT1;4</i>	Gh_A08G2473	AADD	552	9
<i>GhHKT1;1</i>	Gh_D04G1853	AADD	499	9
<i>GhHKT1;2</i>	Gh_D08G0868	AADD	556	9

	(1)	1	10	*	28
ATHKT1;1	(1)	DLFFTSVSAT	TVSSMSTV	DMEVF	SNTQL
GaHKT1;1	(1)	DLFFTSVSAT	TVSSMSTV	EMEIL	SNPQL
GhHKT1;3	(1)	DLFFTSVSAT	TVSSMSTV	EMEIL	SNPQL
GhHKT1;1	(1)	DLFFTSVSAT	TVSSMSTV	EMEIL	SNPQL
GrHKT1;1	(1)	DLFFTSVSAT	TVSSMSTV	EMEIL	SNPQL
OsHKT1;2	(1)	DLFFTSVSAT	TVSSM	VAVEMES	SFNSQL
OsHKT1;3	(1)	DVFFTSVSAT	TVSSMAT	VEMED	FSSAQL
OsHKT1;1	(1)	DLFFTSVSAT	TVSSMAT	VEMED	LSDRQL
GaHKT1;2	(1)	DVFFTSVSAT	TVSSMSTV	EMEIV	SNTQL
GhHKT1;2	(1)	DVFFTSVSAT	TVSSMSTV	EMEIV	SNTQL
GrHKT1;2	(1)	DVFFTSVSAT	TVSSMSTV	EMEIV	SNTQL
GhHKT1;4	(1)	DVFFTSVSAT	TVSSMSTV	EMEIV	SNTQL
OsHKT2;1	(1)	DMFFLST	SAVTVT	GLATIQ	MEDLSSSQI
OsHKT2;2	(1)	DMFFLST	SAVTVT	GLATIQ	MEDLSSSQI

* 表示 *HKT* 蛋白家族成员的第 1 个跨膜区氨基酸序列。

图 1 棉花、拟南芥和水稻 *HKT* 蛋白第 1 个跨膜区的多序列比对

Fig.1 The first transmembrane domain sequence alignment of *HKT* proteins in cotton, *Arabidopsis* and rice

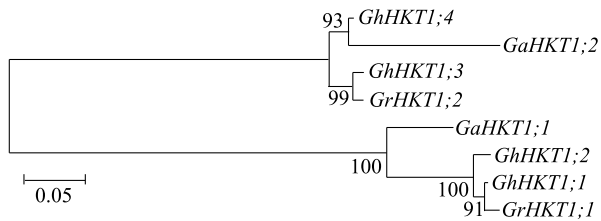
2.2 棉花 *HKT* 基因家族的多序列比对和系统进化树分析

将棉花中 8 个 *HKT* 蛋白家族成员(2 个雷蒙德氏棉的 *HKT* 蛋白家族成员,2 个亚洲棉石系亚 1 号的 *HKT* 蛋白家族成员和 4 个陆地棉遗传标准系 TM-1 的 *HKT* 蛋白家族成员)的氨基酸全长进行多序列比对,结果显示,*GhHKT1;4*、*GaHKT1;2*、*GhHKT1;3* 和 *GrHKT1;2* 4 个成员的序列相似度较高,*GaHKT1;1*、*GhHKT1;2*、*GhHKT1;1* 和 *GrHKT1;1* 序列相似度较高(图 2)。为了进一步研究棉花中 8 个 *HKT* 基因家族成员的亲缘关系,我们利用 MEGA 构建了系统进化树,结果显示,棉花 8 个 *HKT* 基因大致分为 2 个组,*GhHKT1;4*、*GaHKT1;2*、*GhHKT1;3* 和 *GrHKT1;2* 4 个成员属于一个小组,*GaHKT1;1*、*GhHKT1;2*、*GhHKT1;1* 和 *GrHKT1;1* 4 个成员属于一个小组(图 3)。

GhHKT1_4	: MSNTIIICFGRKLEHFSNSRSKLSCFNQSYRGLISTCFRFLVFRVNPFTVQLAYFIVSLVGFAGAKMSNPKSGSVSPNN :	80
GrHKT1_1	: MSNTIIICFGRKLEHFSNSRSKLSCFNQSYRGLISTCFRFLVFRVNPFTVQLAYFIVSLVGFAGAKMSNPKSGSVSPNN :	80
GhHKT1_2	: MSNTIIICFGRKLEHFSNSRSKLSCFNQSYRGLISTCFRFLVFRVNPFTVQLAYFIVSLVGFAGAKMSNPKSGSVSPNN :	80
GaHKT1_1	: MSNTIIICFGRKLEHFSNSRSKLSCFNQSYRGLISTCFRFLVFRVNPFTVQLAYFIVSLVGFAGAKMSNPKSGSVSPNN :	80
GhHKT1_3	: MPKFVLSLS----YSSSKLRLCH-----VLSVNPFTTRILYFIFLSLLGFWGGLNVLEPRTDSFKPKK :	59
GrHKT1_2	: MPKFVLSLS----YSSSKLRLCH-----VLSVNPFTTRILYFIFLSLLGFWGGLNVLEPRTDSFKPKK :	59
GhHKT1_1	: MPKFVLSLS----YSSSKLRLCH-----VLSVNPFTTRILYFIFLSLLGFWGGLNVLEPRTDSFKPKK :	59
GaHKT1_2	: MPKFVLSLS----YSSSKLRLCH-----VLSVNPFTTRILYFIFLSLLGFWGGLNVLEPRTDSFKPKK :	39
GhHKT1_4	: IDVFFTSVSATTVSSMSTVMEVFSNTQLIIITILMLVGGEVFTSMGLLGRFRFETPPSPFGDIVNSISNRGLQSDCS :	160
GrHKT1_1	: IDVFFTSVSATTVSSMSTVMEVFSNTQLIIITILMLVGGEVFTSMGLLGRFRFETPPSPFGDIVNSISNRGLQSDCS :	160
GhHKT1_2	: IDVFFTSVSATTVSSMSTVMEVFSNTQLIIITILMLVGGEVFTSMGLLGRFRFETPPSPFGDIVNSISNRGLQSDCS :	160
GaHKT1_1	: IDVFFTSVSATTVSSMSTVMEVFSNTQLIIITILMLVGGEVFTSMGLLGRFRFETPPSPFGDIVNSISNRGLQSDCS :	160
GhHKT1_3	: FDLFFTSVSATTVSSMSTVMEILSNPQLIIMTILMFIIGGEVFTSMAGLFFSNFISEKNNVID----- :	122
GrHKT1_2	: FDLFFTSVSATTVSSMSTVMEILSNPQLIIMTILMFIIGGEVFTSMAGLFFSNFISEKNNVID----- :	122
GhHKT1_1	: FDLFFTSVSATTVSSMSTVMEILSNPQLIIMTILMFIIGGEVFTSMAGLFFSNFISEKNNVID----- :	122
GaHKT1_2	: FDLFFTSVSATTVSSMSTVMEILSNPQLIIMTILMFIIGGEVFTSMAGLFFSNFISEKNNVID----- :	102
GhHKT1_4	: LTSKNISQEVELGRITCSALANEKPITDLESNNLTSDS-KSLKYKATGILGYLVGYLLVHVAVGSSVSMYVSLVPSA :	239
GrHKT1_1	: LTSKNISQEVELGRITCSALANEKPITDLESNNLTSDS-KSLKYKATGILGYLVGYLLVHVAVGSSVSMYVSLVPSA :	239
GhHKT1_2	: LTSKNISQEVELGRITCSALANEKPITDLESNNLTSDS-KSLKYKATGILGYLVGYLLVHVAVGSSVSMYVSLVPSA :	239
GaHKT1_1	: LTSKNISQEVELGRITCSALANEKPITDLESNNLTSDS-KSLKYKATGILGYLVGYLLVHVAVGSSVSMYVSLVPSA :	239
GhHKT1_3	: -----DGEIELGNPKPENLPEQ-----LAQNGGEFNGEHLHYHSIMLGCVIIGYLLIVNIGSATVFLYISFVSSA :	192
GrHKT1_2	: -----DGEIELGNPKPENLPEQ-----LAQNGGEFNGEHLHYHSIMLGCVIIGYLLIVNIGSATVFLYISFVSSA :	192
GaHKT1_1	: -----DGEIELGNPKPENLPEQ-----LAQNGGEFNGEHLHYHSIMLGCVIIGYLLIVNIGSATVFLYISFVSSA :	192
GhHKT1_2	: -----DGEIELGNPKPENLPEQ-----LAQNGGEFNGEHLHYHSIMLGCVIIGYLLIVNIGSATVFLYISFVSSA :	172
GhHKT1_4	: RKILKTKGIDIRTFSEFTVSTLANCGFVPTNENMIVFKKNSGLLLLPQILIGNLTPACLRALIIWVFDKMTKKV--E :	317
GrHKT1_1	: RKILKTKGIDIRTFSEFTVSTLANCGFVPTNENMIVFKKNSGLLLLPQILIGNLTPACLRALIIWVFDKMTKKV--E :	317
GhHKT1_2	: RKILKTKGIDIRTFSEFTVSTLANCGFVPTNENMIVFKKNSGLLLLPQILIGNLTPACLRALIIWVFDKMTKKV--E :	317
GaHKT1_1	: RKILKTKGIDIRTFSEFTVSTLANCGFVPTNENMIVFKKNSGLLLLPQILIGNLTPACLRALIIWVFDKMTKKV--E :	317
GhHKT1_3	: RSTLKSGLNLFTFSEFTTISSTNCGFIPTNENMVFKNKNSGLLVIIALLGNPLFPMCLRFSIWVMGKCVSKVGDY :	272
GrHKT1_2	: RSTLKSGLNLFTFSEFTTISSTNCGFIPTNENMVFKNKNSGLLVIIALLGNPLFPMCLRFSIWVMGKCVSKVGDY :	272
GhHKT1_1	: RSTLKSGLNLFTFSEFTTISSTNCGFIPTNENMVFKNKNSGLLVIIALLGNPLFPMCLRFSIWVMGKCVSKVGDY :	272
GaHKT1_2	: RSTLKSGLNLFTFSEFTTISSTNCGFIPTNENMVFKNKNSGLLVIIALLGNPLFPMCLRFSIWVMGKCVSKVGDY :	252
GhHKT1_4	: FSVILKNWKEMGYSHLLSGLHSFIAATVFGTLVQFILFCSEMEWNETMDGFSSYQKVVGSFQGVNSRYAGESVVDLS :	397
GrHKT1_1	: FSVILKNWKEMGYSHLLSGLHSFIAATVFGTLVQFILFCSEMEWNETMDGFSSYQKVVGSFQGVNSRYAGESVVDLS :	397
GhHKT1_2	: FSVILKNWKEMGYSHLLSGLHSFIAATVFGTLVQFILFCSEMEWNETMDGFSSYQKVVGSFQGVNSRYAGESVVDLS :	397
GaHKT1_1	: FSVILKNWKEMGYSHLLSGLHSFIAATVFGTLVQFILFCSEMEWNETMDGFSSYQKVVGSFQGVNSRYAGESVVDLS :	397
GhHKT1_3	: CDYLLKNTREIGFHHLFTRRSQCVVGTALVFVVQVALLFAMEWNSASLKEKNPFERTIGVLFQSVNTSQAQETIVNLP :	352
GrHKT1_2	: CDYLLKNTREIGFHHLFTRRSQCVVGTALVFVVQVALLFAMEWNSASLKEKNPFERTIGVLFQSVNTSQAQETIVNLP :	352
GhHKT1_1	: CDYLLKNTREIGFHHLFTRRSQCVVGTALVFVVQVALLFAMEWNSASLKEKNPFERTIGVLFQSVNTSQAQETIVNLP :	352
GaHKT1_2	: CDYLLKNTREIGFHHLFTRRSQCVVGTALVFVVQVALLFAMEWNSASLKEKNPFERTIGVLFQSVNTSQAQETIVNLP :	332
GhHKT1_4	: TISSAILVLFVVMYLPYTSIPIKYQKKEIEKEKVKGSQNESAGKSIIEFVLSQLSYLAIFILLICISERQKLKEDP :	477
GrHKT1_1	: TISSAILVLFVVMYLPYTSIPIKYQKKEIEKEKVKGSQNESAGKSIIEFVLSQLSYLAIFILLICISERQKLKEDP :	477
GhHKT1_2	: TISSAILVLFVVMYLPYTSIPIKYQKKEIE-----KGSQNESAGKSILDCVLSQLSYLAIFILLICISERQKLKEDP :	473
GaHKT1_1	: TISSAILVLFVVMYLPYTSIPIKYQKKEIE-----KGSQNESAGKSILDCVLSQLSYLAIFILLICISERQKLKEDP :	473
GhHKT1_3	: AMSTVILVITIMYFPYTSIPFVKDEKKEQA-----NQEKREGK--TAKBLLSGLASICVFVLCITERKNNKEDP :	424
GrHKT1_2	: AMSTVILVITIMYFPYTSIPFVKDEKKEQA-----NQEKREGK--TAKBLLSGLASICVFVLCITERKNNKEDP :	424
GhHKT1_1	: AMSTVILVITIMYFPYTSIPFVKDEKKEQA-----NQEKREGK--TAKBLLSGLASICVFVLCITERKNNKEDP :	424
GaHKT1_2	: AMSTVILVITIMYFPYTSIPFNDDEEQH-----QQQKRKK--TAKBLLSGLASICVFVLCITERKNNKEDP :	403
GhHKT1_4	: LNFNVLNITIEVVS-----AYGNVGSTGY-----CKRRLNGKQWCTDSCAGFAGRWSTG--- :	529
GrHKT1_1	: LNFNVLNITIEVVS-----AYGNVGSTGY-----CKRRLNGKQWCTDSCAGFAGRWSTG--- :	529
GhHKT1_2	: LNFNVLNITIEVVS-----AYGNVGSTGY-----CKRRLNGKQWCTDSCAGFAGRWSTG--- :	525
GaHKT1_1	: LNFNVLNITIEVVSNMSILRKSQHSIPLSPYPLLFHTLLFNPGPIDPTTPHRPTTNGFIKPAIVGLVNLKSARDGSF :	553
GhHKT1_3	: FNFTPFNLFVVS-----AYGNVGYSLGY-----CKRLRLDEANCVDKSYGFAGRWSDAG--- :	476
GrHKT1_2	: FNFTPFNLFVVS-----AYGNVGYSLGY-----CKRLRLDEANCVDKSYGFAGRWSDAG--- :	476
GhHKT1_1	: FNFTPFNLFVVS-----AYGNVGYSMGY-----CKRLRLDEANCVDKSYGFAGRWSDAG--- :	476
GaHKT1_2	: FNFTPFNLFVVIS-----AYGNVGYTMGY-----CKRLRLDGANCVDKWYGFAGRWSDAG--- :	455
GhHKT1_4	: -----KFLILVVMFFGRFN-----FSFKYGA*KLK----- :	556
GrHKT1_1	: -----KFLILVVMFFGRFN-----FSFKYGA*KLK----- :	556
GhHKT1_2	: -----KFLILVVMFFGRFN-----FSFKYGA*KLK----- :	552
GaHKT1_1	: NTKIKTFNFIINTEYSLSIFPLSRDLKGLMRNHFNLQIACLYRIERIEESSCVLDDENGGLKATTFHHIVIMLTCKTK :	633
GhHKT1_3	: -----KTVLIVVMMLGRLKR-----YNMV*WKS----- :	499
GrHKT1_2	: -----KTVLIVVMMLGRLKR-----YNMV*WKS----- :	499
GhHKT1_1	: -----KTVLIVVMMLGRLKR-----YNMV*WKS----- :	499
GaHKT1_2	: -----KTVLIVVMMLGRLKR-----YNMG*WKS----- :	478
GhHKT1_4	: ----- : -	
GrHKT1_1	: ----- : -	
GhHKT1_2	: ----- : -	
GaHKT1_1	: MDDGSDSRGVSLDVIT : 650	
GhHKT1_3	: ----- : -	
GrHKT1_2	: ----- : -	
GhHKT1_1	: ----- : -	

图 2 棉花 HKT 蛋白家族的多序列比对

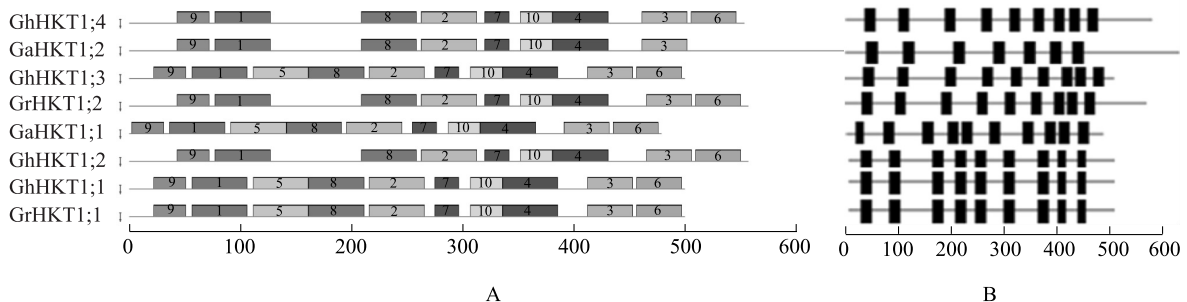
Fig.2 The sequence alignment of HKT protein family in cotton

图3 棉花 *HKT* 基因家族的系统进化树Fig.3 Phylogenetic tree of *HKT* gene family in cotton

2.3 棉花 *HKT* 蛋白家族的特征

亚细胞定位显示:棉花 8 个 *HKT* 蛋白家族成员都定位到了质膜上。棉花 8 个 *HKT* 蛋白家族成员

的基序 (motif) 分析结果表明,在进化树上分支比较近的 *HKT* 蛋白家族成员具有完全相同或者比较类似的 motif 类型和排列顺序。进化树分析结果显示,棉花 8 个 *HKT* 蛋白家族成员大致分为 2 个小组, GhHKT1;4、GaHKT1;2、GhHKT1;3 和 GrHKT1;2 为 1 个小组,4 个成员都具有 8~10 个相似的 motif,且 motif 具有相似的排列数序;GaHKT1;1、GhHKT1;2、GhHKT1;1 和 GrHKT1;1 4 个成员属于 1 个小组,4 个成员都具有 9~10 个 motif,具有相似的 motif 类型和排列数序。进一步分析发现,除了成员 GhHKT1;2 具有 9 个 motif 外,其他 3 个成员具有完全相同的 10 个 motif 类型和排列顺序 (图 4A)。

图4 棉花中 *HKT* 蛋白家族的基序类型 (A) 和棉花中 *HKT* 蛋白家族的跨膜结构 (B)Fig.4 The motif compositions of the *HKT* protein family (A) and the transmembrane domain (B) in cotton

跨膜结构预测分析发现,在系统进化树上分支比较近的 *HKT* 蛋白家族成员具有完全相同或者比较类似的跨膜数目。GhHKT1;4、GaHKT1;2、GhHKT1;3 和 GrHKT1;2 为 1 个小组,4 个成员都具有 7~10 个不等的跨膜结构。进一步分析发现,除了成员 GaHKT1;2 具有 7 个跨膜结构外,其他 3 个成员具有 9 个跨膜结构;GaHKT1;1、GhHKT1;2、GhHKT1;1 和 GrHKT1;1 4 个成员属于 1 个小组,4 个成员都具有 9~10 个不等的跨膜结构。进一步分析发现,除了成员 GaHKT1;1 具有 9 个跨膜外结构,其他 3 个成员具有完全相同的 10 个跨膜结构 (图 4B)。

2.4 棉花 *HKT* 基因家族的组织表达分析

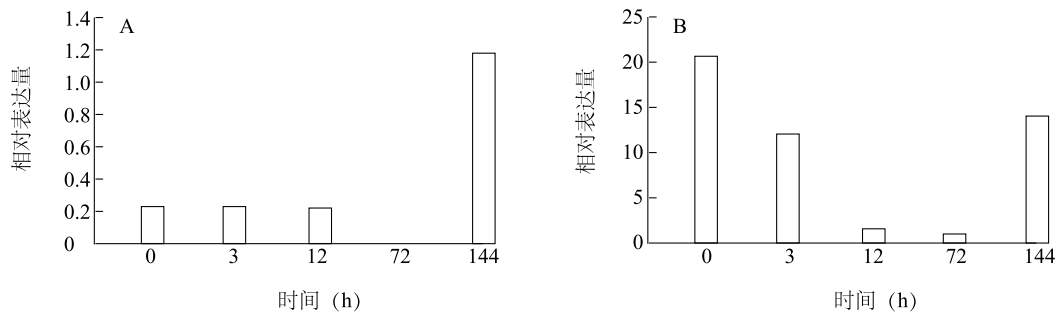
棉花 8 个 *HKT* 基因家族成员的组织表达分析结果表明:在 NCBI 下载的 EST 数据库中没有匹配到对应的 EST 序列,这在一定程度上意味着在现有的棉花 EST 数据库中检测不到棉花 8 个 *HKT* 基因家族成员的表达。

为了进一步研究棉花 *HKT* 基因家族成员是否

响应盐胁迫,我们用 200 mmol/L 的 NaCl 溶液处理了棉属二倍体 D 组野生种克劳茨基棉 0 h、3 h、12 h、72 h 和 144 h,取根部提取 RNA,进行了转录组测序,并与棉花的 8 个 *HKT* 基因家族成员进行了比对,结果显示:在 200 mmol/L 的 NaCl 溶液处理 0 h、3 h、12 h、72 h 和 144 h 后 *GrHKT1;1* 和 *GrHKT1;2* 基因具有不同程度的响应。*GrHKT1;1* 基因在 NaCl 溶液处理 0 h、3 h 和 12 h 时表达量较低,处理 72 h 时 *GrHKT1;1* 基因下调表达,144 h 上调表达 (图 5A);*GrHKT1;2* 基因在处理 0 h、3 h 和 12 h 的时,随着 NaCl 的处理时间延长,下调表达,当 NaCl 处理 72 h 和 144 h 的时候,随着 NaCl 的处理时间增加,又上调表达 (图 5B)。

3 讨论

本研究中,我们利用同源搜索的方法在 AA 组亚洲棉石系亚 1 号、DD 组雷蒙德氏棉和 AADD 组陆地棉遗传标准系 TM-1 基因组共鉴定了 8 个 *HKT*



A: *GrHKT1;1* 基因; B: *GrHKT1;2* 基因。

图 5 盐胁迫下克劳茨基棉中 *HKT* 基因的表达分析

Fig.5 The expression patterns of *HKT* genes in salt-stressed cotton *Gossypium klotzschianum*

基因家族成员。典型的 *HKT* 蛋白由 8 个跨膜和 4 个 P-loop 构成^[19,23,25], 本研究鉴定到的 8 个 *HKT* 蛋白家族成员亚细胞定位都定位到质膜上, 跨膜数为 7 个、9 个和 10 个, 没有任何一个成员符合典型特征, 推测可能 8 个 *HKT* 成员对 K^+ 的转运没有更高的亲和性。在大豆研究中, 同样也发现克隆到的 4 个 *HKT* 蛋白家族成员, 仅有 *GmHKT1;2* 符合 8 个跨膜的典型结构, 推测 *GmHKT1;2* 基因可能对大豆 K^+ 的转运具有更高的亲和性^[31-34]。对 8 个成员第一个跨膜区的氨基酸进行多序列比对发现, 第一个 P-Loop 中均含有丝氨酸, 8 个成员都属于类型 I, 这跟以往报道也是相吻合的, 类型 I 亚组主要来自双子叶植物^[10,14,16-17]。系统进化树和 motif 分析结果表明, 棉花 8 个 *HKT* 蛋白家族成员大致分为 2 个组, 在进化树上分支比较近的 *HKT* 蛋白家族成员具有完全相同或者比较类似的 motif 类型和跨膜数, 因此, 我们可以推测每个小组尤其是分支较近的成员之间具有相似的功能。

棉花中 8 个 *HKT* 基因的表达分析结果表明, 仅有 *GrHKT1;1* 和 *GrHKT1;2* 2 个成员在 200 mmol/L 的 NaCl 溶液处理 0 h、3 h、12 h、72 h 和 144 h 后具有不同程度的响应。尽管在 NCBI 公布的 EST (包括 64 903、63 642 和 330 520 条的雷蒙德氏棉、亚洲棉和陆地棉 EST 序列) 序列中, 有盐胁迫相关的 EST 序列, 但是, 在我们的研究中, 没有检测到棉花中 8 个 *HKT* 基因的表达。在一定程度上意味着, 棉属二倍体 D 组野生种克劳茨基棉中 *GrHKT1;1* 和 *GrHKT1;2* 基因响应了盐胁迫。我们在前期的研究中, 也发现克劳茨基棉具有较好的耐盐性。因此, 在以后的研究中, 可以进一步研究 *GrHKT1;1* 和 *GrHKT1;2*

基因的功能, 对深入研究耐盐机理以及拓展棉花种质资源有重要意义。

参考文献:

- [1] 张辉, 朱绿丹, 安霞, 等. 水分和钾肥耦合对甘薯光合特性和水分利用效率的影响[J]. 江苏农业学报, 2016, 32(6): 1294-1301.
- [2] 周峰, 李平华, 王宝. K^+ 稳态与植物耐盐性的关系[J]. 植物生理学通讯, 2003, 39(1): 67-70.
- [3] 张明乾, 范建立, 郭传滨. 不同钾肥对烤烟主要化学成分的影响[J]. 山东农业科学, 2015, 47(5): 68-71.
- [4] 杜海岩, 柳新伟, 徐双, 等. 钾素营养对盐碱地棉花中后期生长及生理特性的影响[J]. 江苏农业科学, 2016, 44(3): 114-117.
- [5] 林子龙. 种植密度与钾肥对甘薯新品种龙薯 14 号产量的影响[J]. 南方农业学报, 2015, 46(6): 1002-1006.
- [6] RUBIO F, GASSMANN W, SCHROEDER J I. Sodium-driven potassium uptake by the plant potassium transporter HKT1 and mutations conferring salt tolerance[J]. Science, 1995, 270(5242): 1660-1663.
- [7] FAIRBAIRN D J, LIU W H, SCHACHTMAN D P, et al. Characterisation of two distinct HKT1-like potassium transporters from *Eucalyptus camaldulensis*[J]. Plant Molecular Biology, 2000, 43(4): 515-525.
- [8] MUNNS R, TESTER M. Mechanisms of salinity tolerance[J]. Annual Review of Plant Physiology, 2008, 59: 651-681.
- [9] SCHACHTMAN D P, SCHROEDER J I. Structure and transport mechanism of a high-affinity potassium uptake transporter from higher plants[J]. Nature, 1994, 370(6491): 655-658.
- [10] UOZUMI N, KIM E J, RUBIO F, et al. The Arabidopsis *HKT1* gene homolog mediates inward Na^+ currents in *Xenopus laevis* oocytes and Na^+ uptake in *Saccharomyces cerevisiae*[J]. Plant Physiology, 2000, 122(4): 1249-1260.
- [11] HORIE T, YOSHIDA K, NAKAYAMA H, et al. Two types of HKT transporters with different properties of Na^+ and K^+ transport

- in *Oryza sativa* [J]. Plant Journal, 2001, 27(2): 129-138.
- [12] REN Z H, GAO J P, LI L G, et al. A rice quantitative trait locus for salt tolerance encodes a sodium transporter [J]. Nature Genetics, 2005, 37(10): 1141-1146.
- [13] HARO R, BANUELOS M A, SENN M E, et al. HKT1 mediates sodium uniport in roots; pitfalls in the expression of HKT1 in yeast [J]. Plant Physiology, 2005, 139(3): 1495-1506.
- [14] FAIRBAIRN D J, LIU W, SCHACHTMAN D P, et al. Characterisation of two distinct HKT1-like potassium transporters from *Eucalyptus camaldulensis* [J]. Plant Molecular Biology, 2000, 43(4): 515-525.
- [15] SU H, BALDERAS E, ROSARIO V-E, et al. Expression of the cation transporter McHKT1 in a halophyte [J]. Plant Molecular Biology, 2003, 52(5): 967-980.
- [16] DURELLS R, GUY H R. Structural models of the KtrB, TrkH and Trk1, 2 symporters based on the structure of the K⁺ channel [J]. Biophysical Journal, 1999, 77(2): 789-807.
- [17] DURELL S R, HAO Y, NAKAMURA T, et al. Evolutionary relationship between K⁺ channels and symporters [J]. Biophysical Journal, 1999, 77(2): 775-788.
- [18] GARCIASDEBLAS B, SENN M E, BANUELOS M A, et al. Sodium transport and HKT transporters: the rice model [J]. Plant Journal, 2003, 34(6): 788-801.
- [19] MASER P, ECKELMAN B, VAIDYANATHAN R, et al. Altered shoot/root Na⁺ distribution and bifurcating salt sensitivity in Arabidopsis by genetic disruption of the Na⁺ transporter AtHKT [J]. FEBS Letters, 2002, 531(2): 157-161.
- [20] WANG Y, WU W. Potassium transport and signaling in higher plants [J]. Annual Review of Plant Physiology, 2013, 64(1): 451-476.
- [21] VERY A A, SEBTENAC H. Molecular mechanisms and regulation of transport in higher plants [J]. Annual Review of Plant Physiology, 2003, 54(1): 575-603.
- [22] WATERS S, GILLIHAN M, HRMOVA M. Plant high-affinity potassium (HKT) transporters involved in salinity tolerance: structural insights to probe differences in ion selectivity [J]. International Journal of Molecular Sciences, 2013, 14(14): 7660-7680.
- [23] KATO Y, SAKAGUCHI M, MORI Y, et al. Evidence in support of a four transmembrane-pore-transmembrane topology model for the Arabidopsis thaliana Na⁺/K⁺ translocating AtHKT1 protein a member of the superfamily of K⁺ transporters [J]. Proceedings of the National Academy of Sciences of the United States of America, 2001, 98(11): 6488-6493.
- [24] RODRIGUEN A, RUBIO F. High-affinity potassium and sodium transport systems in plants [J]. Journal of Experimental Botany, 2006, 57(5): 1149-1160.
- [25] 叶武威, 刘金定. 棉花种质资源耐盐性鉴定技术与应用 [J]. 中国棉花, 1998, 25(9): 37-38.
- [26] WANG K B, WANG Z W, LI F G, et al. The draft genome of a diploid cotton *Gossypium raimondii* [J]. Nature Genetics, 2012, 44(10): 1098-1103.
- [27] LI F G, FAN G Y, WANG K B, et al. Genome sequence of the cultivated cotton *Gossypium arboreum* [J]. Nature Genetics, 2014, 46(6): 562-574.
- [28] LI F G, FAN G Y, LU C R, et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution [J]. Nature Biotechnology, 2015, 33(5): 524-530.
- [29] ZHANG T Z, HU Y, JIANG W K, et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement [J]. Nature Biotechnology, 2015, 33(5): 531-754.
- [30] PLATTEN J D, COTSIFOS O, BERTHOMIRU P, et al. Nomenclature for HKT transporters, key determinants of plant salinity tolerance [J]. Trends in Plant Science, 2006, 11(8): 372-374.
- [31] CHEN H, CHEN X, GU H P, et al. *GmHKT1;4*, a novel soybean gene regulating Na⁺/K⁺ ratio in roots enhances salt tolerance in transgenic plants [J]. Plant Growth Regulation, 2014, 73(3): 299-308.
- [32] KOLTUNOW A M, TRUETTNER J, COX K H, et al. Different temporal and spatial gene expression patterns occur during Anther Development [J]. Plant Cell, 1990, 2(12): 1201-1224.
- [33] SCHAUERR A, RANES M, SANTAMARIA R, et al. Visualizing gene expression in time and space in the filamentous bacterium *Streptomyces coelicolor* [J]. Science, 1988, 240(4853): 768-772.
- [34] MCGUIRE S E, ROMAN G, DAVIS R L. Gene expression systems in *Drosophila*: a synthesis of time and space [J]. Trends in Genetics, 2004, 20(20): 384-391.

(责任编辑:陈海霞)