

张 斌, 陈丽娟, 李其华, 等. 栽培大豆 GRAS 转录因子家族基因鉴定及其盐胁迫下表达模式分析[J]. 江苏农业学报, 2021, 37( 2 ): 296-309.

doi: 10.3969/j.issn.1000-4440.2021.02.004

## 栽培大豆 GRAS 转录因子家族基因鉴定及其盐胁迫下表达模式分析

张 斌, 陈丽娟, 李其华, 唐满生

(湖南科技学院化学与生物工程学院, 湖南省银杏工程技术研究中心, 湖南 永州 425199)

**摘要:** 利用生物信息学手段及转录组测序方法对大豆 78 个 GRAS 家族基因进行系统分析。染色体定位结果表明 78 个 GRAS 基因不均匀地分布在 20 条染色体上。通过系统进化分析将大豆 GRAS 家族分为 11 个亚族。基因结构和保守基序分布分析结果表明 GRAS 家族成员在进化上具有保守性, 尤其是进化关系较近的成员多具有类似的基因结构和蛋白质结构。转录组数据及 qRT-PCR 结果显示, 5 个基因受盐胁迫诱导上调, 5 个基因受盐胁迫诱导下调, 其中 *GmGRAS14*、*GmGRAS33* 和 *GmGRAS69* 在盐处理 12 h 时上调倍数最高, 而 *GmGRAS17*、*GmGRAS54* 和 *GmGRAS57* 在盐处理 12 h 时表达量下调倍数最高, 说明这些基因可能在大豆响应盐胁迫方面发挥重要功能。

**关键词:** 大豆; GRAS 家族; 盐胁迫; 表达模式

中图分类号: S565.101

文献标识码: A

文章编号: 1000-4440(2021)02-0296-14

## Identification of gene of GRAS transcription factor family in cultivated soybean( *Glycine max* L.) and expression pattern analysis under salt stress

ZHANG Bin, CHEN Li-juan, LI Qi-hua, TANG Man-sheng

(Hunan Provincial Engineering Research Center for Ginkgo, College of Chemistry and Bioengineering, Hunan University of Science and Engineering, Yongzhou 425199, China)

**Abstract:** In this study, 78 genes of GRAS family in soybean were systematically investigated using bioinformatics and RNA-seq. The chromosomal distribution map showed that 78 GRAS genes were randomly located in 20 chromosomes. Phylogenetic analysis showed that the soybean GRAS family could be divided into 11 subfamilies. The gene structure and conserved motif distribution analysis suggested that the GRAS family members were conservative in evolution, and most of the GRAS members with close evolutionary relationship had similar gene and protein structures. Transcriptomal and qRT-PCR results showed that five genes were up-regulated and five genes were down-regulated under salt stress. Among them, *GmGRAS14*, *GmGRAS33* and *GmGRAS69* had the highest fold of up-regulation at 12 h of salt treatment. *GmGRAS17*, *GmGRAS54* and *GmGRAS57* showed the highest down-regulation multiple at 12 h of salt treatment, suggesting that these genes may play an important role in the response to salt stress of soybean.

**Key words:** soybean; GRAS family; salt stress; expression pattern

收稿日期: 2020-05-11

基金项目: 湖南科技学院重点项目(17XKY012); 湖南省大学生研究性学习和创新性实验计划项目[湘教通(2018)255号]

作者简介: 张 斌(1981-), 男, 湖南永州人, 博士, 讲师, 主要从事植物发育生物学研究。(E-mail) zhangbin27104@163.com

通讯作者: 唐满生, (E-mail) manshengtang@sina.com

转录因子通过结合到启动子区发挥激活或抑制作用调控特定基因的表达, 在植物生长发育以及生物或非生物胁迫响应过程中发挥着重要作用。目前很多转录因子家族基因的功能已被报道, 如 WRKY、MYB、MADS、bZIP、AP2/ERF、GRAS 等<sup>[1]</sup>, 其中 GRAS 是植物特有的转录因子家族。GRAS 家

族基因的命名是根据最早鉴定出的该家族 3 个成员的基因名而来,包括 *GAI*、*RGA* 和 *SCR*<sup>[2-4]</sup>。通常 GRAS 家族蛋白长度为 400~700 aa,并且具有序列非常保守的 C 端和序列可变的 N 端<sup>[5]</sup>。根据系统进化树可将 GRAS 家族分成多个亚家族,但是不同物种中亚家族的数量也存在差异。如 GRAS 家族最早在拟南芥中被分成 8 个亚族:*LISCL*、*PAT1*、*SCL3*、*DELLA*、*SCR*、*SHR*、*LAS* 和 *HAM*<sup>[6]</sup>;在蓖麻中被分成 13 个亚族:*SCR*、*DLT*、*OS19*、*LAS*、*SCL4/7*、*OS4*、*OS43*、*DELLA*、*PAT1*、*SHR*、*HAM*、*SCL3* 和 *LISCL*<sup>[7]</sup>;在棉花和木薯中均可分为 14 个亚族<sup>[8-9]</sup>。目前,有多种植物 GRAS 转录因子家族基因的全基因组已被鉴定,如拟南芥与水稻<sup>[10]</sup>、烟草<sup>[11]</sup>、玉米<sup>[12]</sup>、白杨<sup>[13]</sup>、甘薯<sup>[14]</sup>、棉花<sup>[8]</sup>等。

GRAS 转录因子在植物生长发育过程中发挥着多种功能,参与信号转导、调控组织发育、生物或非生物胁迫应答等过程<sup>[15]</sup>。拟南芥 *SCL3* 蛋白属于 *SCL3* 亚族,其在根的内胚层可通过减弱 *DELLA* 抑制效应以维持赤霉素(GA)途径正常功能,并在拟南芥生长过程中参与调控根细胞的伸长<sup>[16]</sup>。水稻 *SCR* 亚族基因 *OsGRAS32* 基因在赤霉素代谢过程发挥重要调控作用<sup>[17]</sup>。过表达丹参 *SmGRAS3* 基因导致水杨酸和赤霉素的合成以及丹参毛状根的生长均受到抑制<sup>[18]</sup>。此外,大量研究结果表明 GRAS 转录因子参与植物对生物和非生物胁迫的响应过程。通过 RNAi 技术将水稻 GRAS 基因 *OsCIGR2* 沉默后接种稻瘟病菌,发现细胞死亡数目较野生型材料显著增加<sup>[19]</sup>。在拟南芥中过表达甘蓝型油菜 *LAS* 亚族基因 *BnLAS*,则可通过降低过表达材料叶片失水率、减小气孔开度、合成较多的叶表皮蜡质增强转基因材料耐旱性<sup>[20]</sup>。胡杨 *SCL4/7* 亚族基因 *PeSCL7* 的表达受干旱和盐胁迫诱导上调,过表达 *PeSCL7* 基因可提高转基因拟南芥耐旱和耐盐能力<sup>[21]</sup>。在盐和甘露醇处理条件下,番茄 GRAS 家族基因 *SIGARS7* 的表达水平显著上调,*SIGARS7* 基因过表达材料相较于野生型材料更加耐旱、耐盐<sup>[22]</sup>。通过对甘薯、棉花 GRAS 家族的全基因组学鉴定,也发现大量基因在盐、干旱、热、冷胁迫条件下被诱导表达上调<sup>[8,14]</sup>。

在很多植物中已对 GRAS 家族进行了全基因组学鉴定,并且对一些基因功能进行了深入研究。大豆中也有少数 GRAS 家族基因的功能被报道,例如 *GmSCL6* 和 *GmNSP2* (即本研究中的 *GmGRAS1* 和

*GmGRAS12*) 的表达被 miRNA (*gma-miR171o* 和 *gma-miR171q*) 调控,进而影响大豆结瘤相关基因 *NIN*、*ENOD40* 和结瘤所需乙烯响应因子基因 *ERN* 的表达水平<sup>[23]</sup>。但是关于大豆 GRAS 家族的全基因组学分析及在非生物胁迫下的功能研究尚未见报道。本研究中,我们从大豆基因组数据库中鉴定 GRAS 家族成员,并对其理化性质、染色体定位、系统进化关系、基因结构、蛋白质氨基酸序列保守基序以及在不同组织中的表达模式和盐胁迫下的表达模式进行分析,为进一步研究 GRAS 家族基因在环境适应方面的功能提供理论基础。

## 1 材料与方法

### 1.1 大豆 GRAS 家族基因的获取和 GRAS 蛋白理化性质预测

在栽培大豆数据库 (<http://soykb.org/>) 检索 GRAS 转录因子家族成员,并下载检索到的所有基因编码的蛋白质氨基酸序列。将蛋白质氨基酸序列在数据库 SMART (<http://smart.embl-heidelberg.de/>) 和 NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 中检测是否含有 GRAS 结构域(PF03514),去除不含有此保守结构域的成员后即最终鉴定出的大豆 GRAS 转录因子家族成员。根据候选基因在染色体上的位置依次命名。利用在线数据库 ExPASy (<https://web.expasy.org/protparam/>) 对 GRAS 家族成员蛋白质的氨基酸序列长度、相对分子质量、等电点、不稳定系数等理化性质进行预测。

### 1.2 染色体定位、基因结构和蛋白质保守基序分析

栽培大豆所有染色体的总长度信息下载自 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/genome/?term=Glycine+max>),大豆 GRAS 家族成员基因在染色体上的位置信息从 Phytozome 数据库 (<https://phytozome.jgi.doe.gov/>) 下载整理,利用 Map Chart 2.2 软件绘制大豆 GRAS 家族成员基因在染色体上的定位图<sup>[24]</sup>。根据初步鉴定出的大豆 GRAS 家族成员的基因 ID,在大豆数据库中逐一下载对应的基因碱基序列和 CDS 序列,利用在线工具 GSDS (<http://gsds.cbi.pku.edu.cn/index.php>) 绘制基因结构图<sup>[25]</sup>。利用在线数据库 MEME (<http://meme-suite.org/>) 分析 GRAS 家族成员蛋白质氨基酸序列上的保守基序<sup>[26]</sup>,参数为:保守基序最小长度为 6,最大长度为 50,数量设置为 10。

### 1.3 系统进化树的构建

在拟南芥 TAIR 数据库 (<http://www.arabidopsis.org/index.jsp>) 下载拟南芥 GRAS 家族成员的氨基酸序列(33 个),在水稻数据库 (<http://rice.plantbiology.msu.edu/>) 下载水稻 GRAS 家族成员的氨基酸序列(50 个)。将大豆、拟南芥和水稻 GRAS 家族成员的氨基酸序列在软件 Clustal X 中进行多重比对,在 MAGA 7.0 软件中利用邻接法(Neighbor-Joining, NJ)绘制系统进化树,参数设置:Bootstrap method 1 000, Poisson model, Pairwise deletion<sup>[8]</sup>。

### 1.4 GRAS 家族基因在不同组织中的表达模式分析

栽培大豆 GRAS 家族基因在不同组织的表达数据从 Phytozome 数据库下载整理,即转录组测序数据中的 FPKM 值。包括基因在根、根毛、茎、顶端分生组织、叶、花、豆荚、种子、根瘤共 9 个部位的表达数据,所有的 FPKM 值经公式  $\ln(X+1)$  计算后得到的数值用于表示基因的相对表达水平,并利用软件 Hem1.0 绘制基因表达热图<sup>[8]</sup>。

### 1.5 大豆幼苗的培养和盐处理后的转录组测序

试验所用的栽培大豆(*Glycine max*)品种为天隆 1 号。将充分吸水的大豆种子播种于湿润的营养土中(丹麦品氏泥炭土),置于生长室生长,生长室条件:恒温 25 ℃,湿度为 50%~60%,光照为中日照(12 h 昼/12 h 夜)。大豆生长约 10 d 后,第 1 张三出复叶完全长出,将大豆分为 2 组,一组浇灌普通自来水(对照组),另一组浇等量的 200 mmol/L NaCl 水溶液,6 h 后分别取对照组和处理组大豆幼苗的地上部分(含茎和叶)于 10 ml 离心管中(均有 3 组生物重复),并迅速置于液氮中。样品交予上海欧易生物公司进行转录组学测序(RNA-seq)。从转录组数据中先初步筛选出 GO 注释为 GRAS 家族的基因,并进一步找出前期通过生物信息手段鉴定的大豆 GRAS 家族基因在正常生长和盐处理 6 h 后的基因表达 FPKM 值,绘制基因表达热图以分析盐胁迫条件下 GRAS 家族所有基因的表达模式。

### 1.6 RNA 提取和荧光定量 PCR 检测

为了验证转录组数据中 GRAS 家族基因表达模式的准确性,重新培养一批大豆幼苗并进行 NaCl 处理(幼苗生长条件、生长时间、处理方式和方法同方法 1.5)。使用 TRIzol 法提取大豆幼苗总 RNA,使用反转录试剂盒 SuperScript III First-Strand Synthesis SuperMix(Thermo)合成 cDNA。以合成的 cDNA 为

模板进行 qRT-PCR 分析,试验操作参照文献[27]。PCR 反应总体积为 20.0  $\mu$ l:10.0  $\mu$ l SYBR Green Mix(Thermo),8.0  $\mu$ l RNAase-free 水,正向和反向引物各 0.5  $\mu$ l,cDNA 模板 1.0  $\mu$ l。利用 ABI 7500 定量 PCR 仪进行 PCR 扩增反应,反应程序:94 ℃,预变性 32 min;94 ℃变性 10 s,60 ℃退火 10 s,72 ℃延伸 10 s,共 40 个循环。利用  $2^{-\Delta\Delta Ct}$  法计算基因相对表达水平。大豆 *UBI3* 基因作为内参,本研究所用引物如表 1 所示。

表 1 本研究中用到的引物

Table 1 Primers used in this study

基因	引物序列 (5'→3')
<i>GmGRAS14</i>	F:TTGAGCGGTTGAGCCTAAG R:CAGCAGCATGTCATCGAGC
<i>GmGRAS17</i>	F:TGGGTGATAGCAGTAGGCCA R:GGAATCCCTGTAGCAACGCT
<i>GmGRAS33</i>	F:CTGGGACAAGCTCGAATGGA R:TGCAACATGCAGGGTTTTCG
<i>GmGRAS54</i>	F:CACGAGCCGTCTCTCTGACT R:GGCGAAGCAAGCTCGTTAAG
<i>GmGRAS57</i>	F:ACGAGCTTGCTTCGCCTTAT R:CAGCCACGGAGGAAAGAGTT
<i>GmGRAS69</i>	F:TGCCAGGTCTTGTGGAGTTC R:TCATCTGGCATGTGTGCAA
<i>GmUBI3</i>	F:GGGGGTGTAAGTGGAAAGCA R:TTGCACGTCCTTCCATCCTC

## 2 结果与分析

### 2.1 大豆 GRAS 家族成员的理化性质及 GRAS 家族基因染色体定位

本研究最终在栽培大豆中共鉴定出 78 个 GRAS 家族成员,并根据基因在染色体上的位置将这些基因依次命名为 *GmGRAS1*~*GmGRAS78*。通过数据库 ExPASy 对大豆 GRAS 家族成员的理化性质进行分析。结果如表 2 所示,该家族成员的理化性质差异较大,蛋白质的长度在 168 aa(*GmGRAS35*)至 744 aa(*GmGRAS1*)之间,相对分子质量则介于  $1.897\ 584\times 10^4$  (*GmGRAS35*)至  $8.238\ 555\times 10^4$  (*GmGRAS9*)之间,等电点在 4.82(*GmGRAS63*)至 9.21(*GmGRAS35*)之间,不稳定系数在 39.54(*GmGRAS26*)至 61.75(*GmGRAS1*)之间,其中仅有 *GmGRAS26* 和 *GmGRAS47* 的不稳定系数小于 40,说

明这 2 个蛋白质被预测为稳定蛋白质。染色体定位分析结果表明,大豆 *GRAS* 家族 78 个基因不均匀地分布在 20 条染色体上,其中第 11、12、13、15 号染色体上分布较多,分别有 12、6、9、7 个;其次为第 16 号染色体,共有 5 个;第 1、10、17 号染色体上均包含 4

个 *GRAS* 基因,第 2、3、5、9、18 号染色体上均为 3 个 *GRAS* 基因,第 4、6、7、14、20 号染色体上则均有 2 个 *GRAS* 基因;而第 8 和第 19 号染色体上含有的 *GRAS* 基因数量最少,均为 1 个(图 1)。

表 2 鉴定出的大豆 *GRAS* 家族基因

Table 2 Soybean *GRAS* family genes identified

基因	基因 ID	蛋白质长度 (aa)	相对分子质量	等电点 (pI)	不稳定系数
<i>GmGRAS1</i>	Glyma01g18040.1	744	81 588.60	5.48	61.75
<i>GmGRAS2</i>	Glyma01g33270.1	737	81 389.76	5.76	58.33
<i>GmGRAS3</i>	Glyma01g40180.1	476	54 005.58	5.45	45.29
<i>GmGRAS4</i>	Glyma01g43620.1	465	52 340.30	6.04	52.86
<i>GmGRAS5</i>	Glyma02g06530.1	523	58 196.97	5.91	47.84
<i>GmGRAS6</i>	Glyma02g08241.1	476	54 810.36	6.27	46.66
<i>GmGRAS7</i>	Glyma02g46730.1	545	61 468.80	5.42	56.75
<i>GmGRAS8</i>	Glyma03g03760.1	732	81 292.91	5.68	58.45
<i>GmGRAS9</i>	Glyma03g10320.1	730	82 385.55	5.79	47.92
<i>GmGRAS10</i>	Glyma03g37851.1	540	61 835.07	6.03	46.42
<i>GmGRAS11</i>	Glyma04g28490.1	443	50 345.11	6.53	40.89
<i>GmGRAS12</i>	Glyma04g43090.1	502	55 183.73	5.75	40.61
<i>GmGRAS13</i>	Glyma05g03020.1	511	57 670.13	6.05	53.64
<i>GmGRAS14</i>	Glyma05g22140.1	480	53 365.74	5.65	43.10
<i>GmGRAS15</i>	Glyma05g27190.1	523	57 870.21	5.28	41.14
<i>GmGRAS16</i>	Glyma06g12701.1	600	66 778.70	4.83	48.97
<i>GmGRAS17</i>	Glyma06g41500.1	568	63 052.97	5.67	50.93
<i>GmGRAS18</i>	Glyma07g11865.1	303	34 525.10	5.15	47.75
<i>GmGRAS19</i>	Glyma07g18934.1	558	63 292.52	4.90	49.68
<i>GmGRAS20</i>	Glyma08g10140.1	517	57 260.74	5.21	40.06
<i>GmGRAS21</i>	Glyma09g01440.1	548	61 274.91	5.83	52.30
<i>GmGRAS22</i>	Glyma09g24740.1	552	60 500.40	8.84	41.32
<i>GmGRAS23</i>	Glyma09g35876.1	421	48 049.94	8.97	45.00
<i>GmGRAS24</i>	Glyma10g01571.1	451	51 852.40	8.65	49.62
<i>GmGRAS25</i>	Glyma10g04421.1	537	60 407.33	6.01	46.33
<i>GmGRAS26</i>	Glyma10g33380.1	488	54 187.50	5.28	39.54
<i>GmGRAS27</i>	Glyma10g37640.1	590	65 220.89	8.26	47.80
<i>GmGRAS28</i>	Glyma11g01850.1	473	52 975.81	5.71	47.16
<i>GmGRAS29</i>	Glyma11g05110.1	482	54 540.12	5.45	45.13
<i>GmGRAS30</i>	Glyma11g06980.1	516	57 496.36	5.56	44.46
<i>GmGRAS31</i>	Glyma11g09761.1	463	51 932.13	6.00	44.65
<i>GmGRAS32</i>	Glyma11g10220.1	442	48 212.71	5.52	51.31
<i>GmGRAS33</i>	Glyma11g14670.1	680	76 497.82	5.84	49.63
<i>GmGRAS34</i>	Glyma11g14700.1	614	70 234.35	6.00	49.15
<i>GmGRAS35</i>	Glyma11g14690.1	168	18 975.84	9.21	42.64
<i>GmGRAS36</i>	Glyma11g14710.1	696	79 482.94	6.01	57.74



续表 2 Continued 2

基因	基因 ID	蛋白质长度 (aa)	相对分子质量	等电点 (pI)	不稳定系数
<i>GmGRAS37</i>	Glyma11g20980.1	442	50 301.21	6.13	43.85
<i>GmGRAS38</i>	Glyma11g23690.1	240	27 599.38	5.03	57.02
<i>GmGRAS39</i>	Glyma11g33720.1	595	64 978.02	5.00	50.78
<i>GmGRAS40</i>	Glyma12g02060.1	481	53 805.84	5.11	53.99
<i>GmGRAS41</i>	Glyma12g02530.1	445	48 961.64	5.92	55.46
<i>GmGRAS42</i>	Glyma12g06640.1	680	77 290.87	5.42	50.85
<i>GmGRAS43</i>	Glyma12g06670.1	660	74 180.12	6.18	46.35
<i>GmGRAS44</i>	Glyma12g32350.1	460	51 464.94	6.13	42.51
<i>GmGRAS45</i>	Glyma12g34420.1	571	63 782.91	5.50	52.68
<i>GmGRAS46</i>	Glyma13g09220.1	591	66 019.24	4.97	43.83
<i>GmGRAS47</i>	Glyma13g02840.1	472	51 427.45	5.05	39.76
<i>GmGRAS48</i>	Glyma13g18680.1	526	59 132.04	5.98	40.93
<i>GmGRAS49</i>	Glyma13g36120.1	577	64 476.92	5.45	55.40
<i>GmGRAS50</i>	Glyma13g41220.1	644	73 440.00	5.86	51.48
<i>GmGRAS51</i>	Glyma13g41231.1	486	54 605.78	8.17	41.41
<i>GmGRAS52</i>	Glyma13g41251.1	610	69 347.04	6.25	49.30
<i>GmGRAS53</i>	Glyma13g41261.1	657	73 661.72	5.46	46.80
<i>GmGRAS54</i>	Glyma13g42100.1	443	50 053.57	5.58	46.98
<i>GmGRAS55</i>	Glyma14g01020.1	545	61 133.93	5.93	55.01
<i>GmGRAS56</i>	Glyma14g01960.1	545	61 526.04	5.54	54.94
<i>GmGRAS57</i>	Glyma15g03290.1	437	49 372.77	5.27	49.92
<i>GmGRAS58</i>	Glyma15g04160.1	684	76 788.11	5.50	48.25
<i>GmGRAS59</i>	Glyma15g04166.1	606	68 783.08	6.21	44.98
<i>GmGRAS60</i>	Glyma15g04173.1	727	81 902.55	5.72	52.28
<i>GmGRAS61</i>	Glyma15g12320.1	552	61 690.32	5.57	50.27
<i>GmGRAS62</i>	Glyma15g15105.1	417	47 767.22	5.20	47.75
<i>GmGRAS63</i>	Glyma15g15110.1	593	67 365.77	4.82	54.34
<i>GmGRAS64</i>	Glyma16g01020.1	543	60 153.80	5.79	43.18
<i>GmGRAS65</i>	Glyma16g05751.1	687	76 277.68	5.73	56.90
<i>GmGRAS66</i>	Glyma16g25570.1	540	60 263.06	5.87	50.93
<i>GmGRAS67</i>	Glyma16g27310.1	471	53 895.34	6.79	45.08
<i>GmGRAS68</i>	Glyma16g29900.1	657	72 353.49	6.64	44.01
<i>GmGRAS69</i>	Glyma17g01150.1	545	60 855.45	6.05	53.04
<i>GmGRAS70</i>	Glyma17g13680.1	514	57 804.08	6.05	55.89
<i>GmGRAS71</i>	Glyma17g14030.1	669	74 655.72	6.31	54.94
<i>GmGRAS72</i>	Glyma17g17400.1	503	57 086.40	5.41	52.82
<i>GmGRAS73</i>	Glyma18g04500.1	584	63 889.00	5.16	46.56
<i>GmGRAS74</i>	Glyma18g09030.1	545	61 188.90	5.20	50.87
<i>GmGRAS75</i>	Glyma18g43580.1	543	61 256.07	4.84	48.47
<i>GmGRAS76</i>	Glyma19g26735.1	681	75 382.67	5.79	59.53
<i>GmGRAS77</i>	Glyma20g30150.1	594	65 434.85	8.31	48.04
<i>GmGRAS78</i>	Glyma20g34260.1	434	48 334.17	5.71	41.85

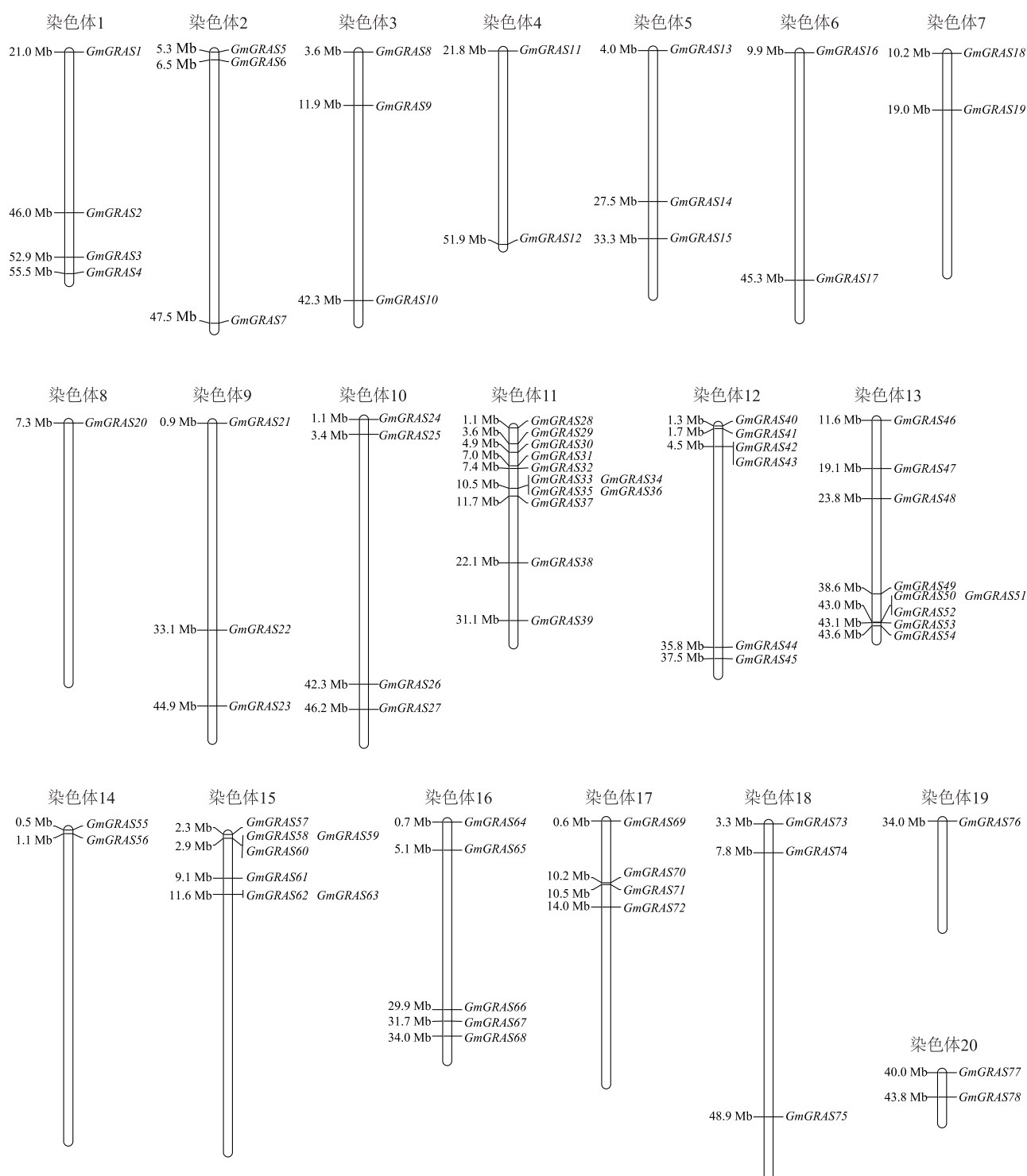


图1 大豆 GRAS 家族基因的染色体定位

Fig.1 Chromosomal location of soybean GRAS family genes

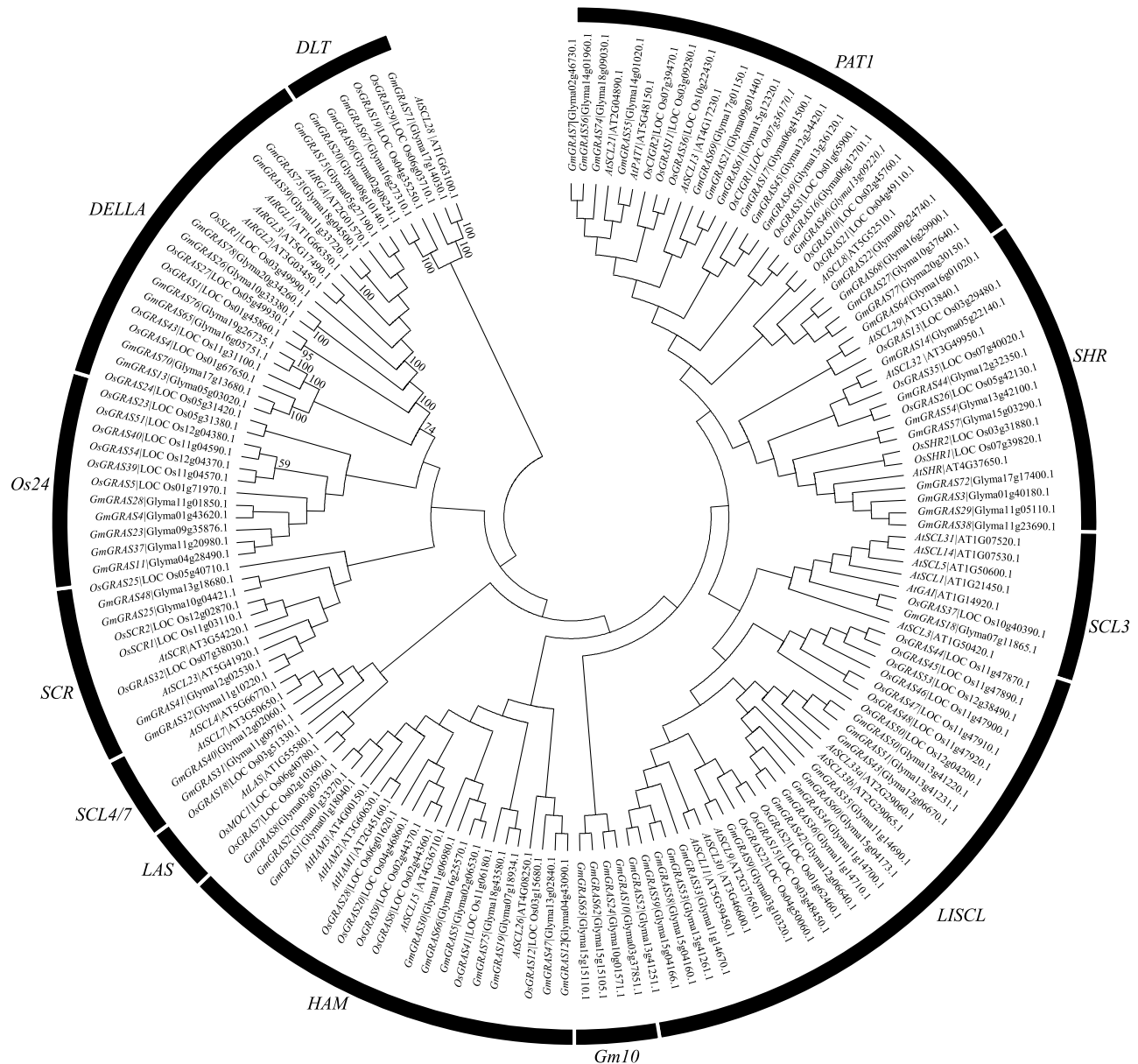
## 2.2 大豆 GRAS 家族基因的系统进化分析

为了探究 GRAS 家族基因在大豆和其他物种之间的进化关系,我们将拟南芥、水稻和大豆 GRAS 家族蛋白质氨基酸序列进行多重比对后,绘制了系统

进化树。结果如图 2 所示,参照 GRAS 家族基因在拟南芥和水稻中的分组方法,并参考蓖麻中 GRAS 亚家族的命名方式(其分为 *PAT1*、*SHR*、*SCL3*、*LISCL*、*HAM*、*SCL4/7*、*SCR*、*DELLA*、*DLT*、*Os43*、*Os4*、

Os19、Rc\_GRAS 共 13 个亚族), 将大豆 GRAS 家族分为 11 个亚家族, 分别是 *PAT1*、*SHR*、*SCL3*、*LISCL*、*HAM*、*SCL4/7*、*SCR*、*Os24*、*DELLA*、*DLT*、*Gm10*, 而进化树中 *LAS* 亚族不包含大豆 GRAS 家族成员。分析不同亚家族 GRAS 基因的数量, 发现 GRAS 基因家族成员不均匀地分布在不同的亚家族。LISCL 是最大的亚族, 家族成员最多 (共 29 个), 包括了 14 个大豆 *GmGRAS* 基因、10 个水稻 *OsGRAS* 基因和 5 个拟南芥 *AtGRAS* 基因。其次为 PAT 亚族, 共包含 27 个

GRAS 基因, 其中包括 16 个大豆 *GmGRAS* 基因、7 个水稻 *OsGRAS* 基因和 4 个拟南芥 *AtGRAS* 基因。而 *DLT*、*SCL4/7*、*Gm10* 和 *SCL3* 则为较小的亚家族, 分别包含了 6、5、4、8 个 GRAS 基因, 其中 *Gm10* 亚族包含的 4 个成员均为大豆 GRAS 基因, *SCL3* 亚族 8 个成员中仅有 1 个是大豆 GRAS 基因 (*GmGRAS18*)。而在 *HAM*、*DELLA*、*SHR*、*Os24*、*SCR* 亚族中则分别包含 10、10、9、5、4 个大豆 GRAS 基因。



进化树中包含 78 个大豆 GRAS 家族基因, 33 个拟南芥 *SERK* 家族基因, 50 个水稻 *SERK* 家族基因。

图 2 大豆和其他物种 GRAS 家族基因的系统进化树

Fig.2 Phylogenetic tree of GRAS family genes in soybean and other plants

### 2.3 大豆 GRAS 家族基因结构及 GRAS 蛋白氨基酸序列上保守基序分布

利用在线软件 GSDS 构建了 GRAS 家族基因的外显子-内含子结构图(图 3)。结果显示,大部分基因(66 个)不含有内含子,仅 12 个基因有内含子,其中 *Gm10* 亚族的 *GmGRAS62* 的内含子数量最多(6 个),*GmGRAS18* 属于 *SCL3* 亚族,其包含 4 个内含子;*GmGRAS24* 和 *GmGRAS35* 均有 3 个内含子,*GmGRAS52* 含有 2 个内含子,而 *GmGRAS10*、*GmGRAS22*、*GmGRAS23*、*GmGRAS31*、*GmGRAS38*、*GmGRAS65* 和 *GmGRAS76* 均有 1 个内含子。大多数大豆 GRAS 基因(57 个)在 5'末端和 3'末端都含有非翻译区。此外,在 78 个 GRAS 蛋白氨基酸序列上共鉴定出 10 个保守基序(Motif 1-Motif 10)(图 4)。所有 GRAS 蛋白的保守基序主要集中在 C 端,并且同一亚族蛋白质含有的保守基序类型及位置均非常相似。如 DELLA 亚族蛋白质均含有除 Motif 9 以外的其他 9 个保守基序,SCR 亚族蛋白质均含有除 Motif 2 和 Motif 9 以外的其他 8 个保守基序,这也进一步证明了相同亚族的 GRAS 家族成员进化关系较近。此外,除 *GmGRAS18* 仅含有 1 个保守基序 Motif 4、*GmGRAS35* 仅含有 1 个基序 Motif 2、*GmGRAS38* 仅含有 2 个基序 Motif 3 和 Motif 7 以外,其余大豆 GRAS 蛋白上的保守基序数量在 6 至 10 个之间。根据对不同亚族 GRAS 基因结构和 GRAS 蛋白保守基序的分析结果可知,同一亚族内的成员相对保守,推测同一亚族的基因功能可能相似。

### 2.4 大豆 GRAS 家族基因在不同组织中的表达

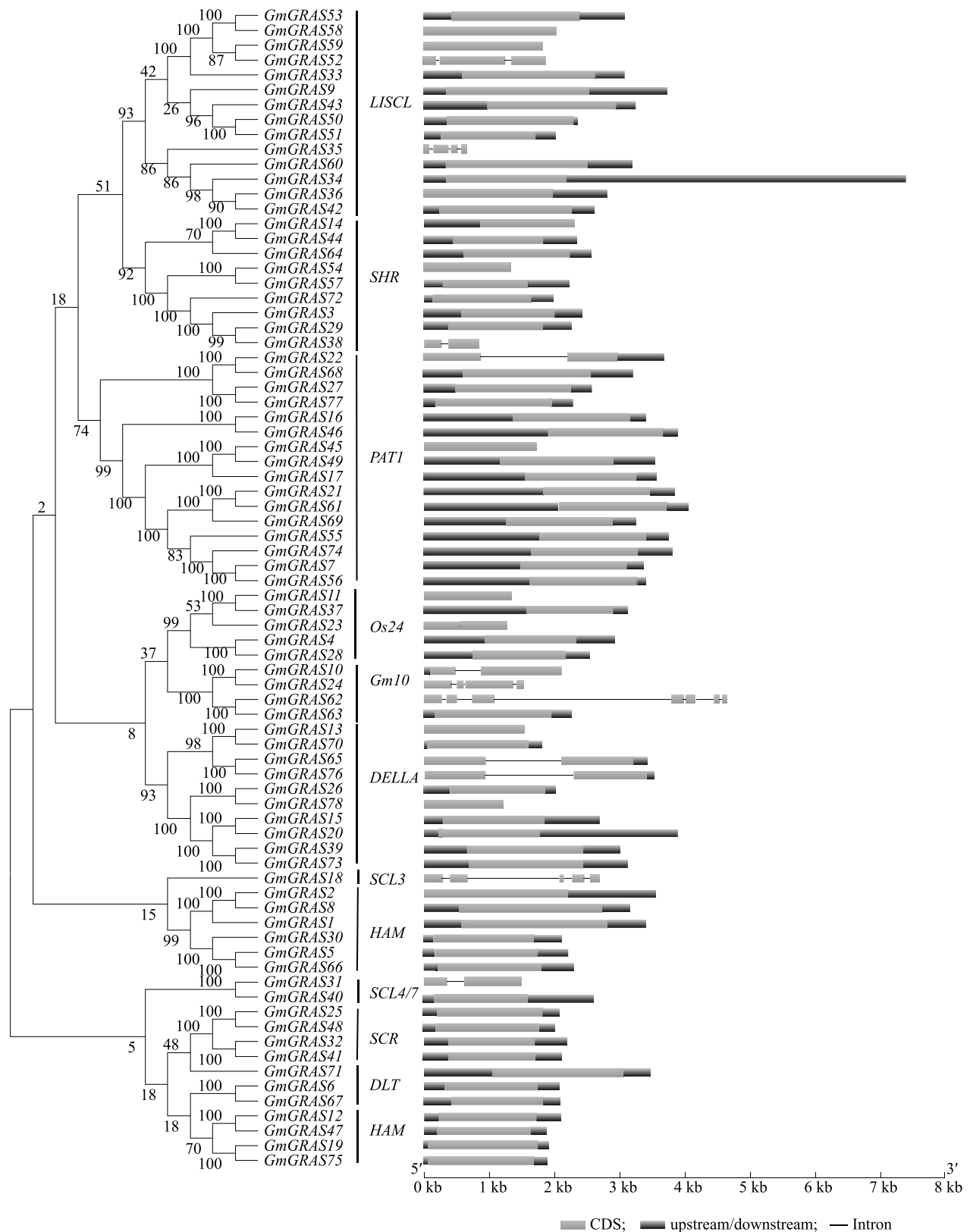
利用公共数据库 Phytozome 中大豆 GRAS 家族基因在不同组织表达的 FPKM 值绘制基因表达热图。结果如图 5A 所示,78 个大豆 GRAS 基因的表达模式呈多样化,其中 *GmGRAS18*、*GmGRAS23*、*GmGRAS25*、*GmGRAS35*、*GmGRAS38* 和 *GmGRAS62* 共 6 个基因在所有组织均未检测到表达。有 48 个基因在所有组织均检测到不同程度的表达,其中 *GmGRAS8*、*GmGRAS15*、*GmGRAS16*、*GmGRAS20*、*GmGRAS22*、*GmGRAS39* 和 *GmGRAS73* 共 7 个基因的表达水平在所有组织中均较高,而 *GmGRAS51* 在所有组织中表达均较低。此外,有些大豆 GRAS 基因表现出组织特异性高表达,如 *GmGRAS16*、*GmGRAS29*、*GmGRAS46*、*GmGRAS55* 和 *GmGRAS56*

在根中的表达水平最高,*GmGRAS39*、*GmGRAS40* 和 *GmGRAS57* 在茎中的表达水平最高,而 *GmGRAS1*、*GmGRAS26*、*GmGRAS73* 和 *GmGRAS78* 在种子中的表达量最高,可以推测这些组织特异性高表达的基因可能在特定的组织发挥特定功能。总而言之,大豆 GRAS 家族基因在不同组织中的表达模式存在较大的差异,说明该家族基因在不同组织可能发挥其独特的生物学功能。然而,同一亚族的某些基因具有类似的表达模式,如 *LISCL* 亚族的 *GmGRAS53* 和 *GmGRAS58*,*SHR* 亚族的 *GmGRAS54* 和 *GmGRAS57*,*DELLA* 亚族的 *GmGRAS15* 和 *GmGRAS20*,且它们在进化树上的进化关系也非常近(图 2),说明具有类似表达模式的基因可能也具有类似的功能。

### 2.5 GRAS 家族基因响应盐胁迫的表达模式

通过对正常生长和盐处理(200 mmol/L NaCl, 6 h)后的大豆植株进行转录组测序,获得了大豆 GRAS 家族 78 个基因在正常和盐胁迫下的表达数据 FPKM 值,并进一步绘制成表达热图(图 5B)。结果表明,大多数基因的表达水平在正常生长和盐处理的大豆中表达水平无显著差异,有 5 个基因(*GmGRAS6*、*GmGRAS14*、*GmGRAS15*、*GmGRAS33*、*GmGRAS69*)的表达水平在盐胁迫下显著上调,5 个基因(*GmGRAS17*、*GmGRAS45*、*GmGRAS54*、*GmGRAS57*、*GmGRAS71*)在盐胁迫下显著下调。此外,我们还挑选了受盐胁迫表达水平显著变化的一些基因(*GmGRAS14*、*GmGRAS33*、*GmGRAS69* 和 *GmGRAS17*、*GmGRAS54*、*GmGRAS57*)通过 qRT-PCR 进行进一步验证。结果见图 6,盐处理后这些基因的表达水平变化与转录组测序结果一致,其中 *GmGRAS14* 和 *GmGRAS69* 在盐处理 3 h、6 h 和 12 h 时表达水平均显著上调,且在 12 h 时上调倍数最高,分别为 3.67 倍和 5.82 倍。*GmGRAS33* 在盐处理 6 h 和 12 h 时表达水平显著上调,且在 12 h 时达到极显著水平,上调 5.92 倍。而 *GmGRAS17* 和 *GmGRAS57* 在盐处理 3 h、6 h 和 12 h 时表达水平均显著下调;*GmGRAS54* 在盐处理 3 h 时与对照无显著差异,在盐处理 6 h 和 12 h 时其表达水平显著下调。以上结果说明表达水平受盐胁迫上调或下调的基因在大豆盐胁迫响应过程中可能发挥调控作用。





左侧为大豆 GRAS 家族基因进化树, 右侧为大豆 GRAS 家族基因对应的内含子和外显子的分布结构。

图 3 大豆 GRAS 家族基因结构及进化树分析

Fig.3 Gene structures of soybean GRAS family genes and phylogenetic analysis

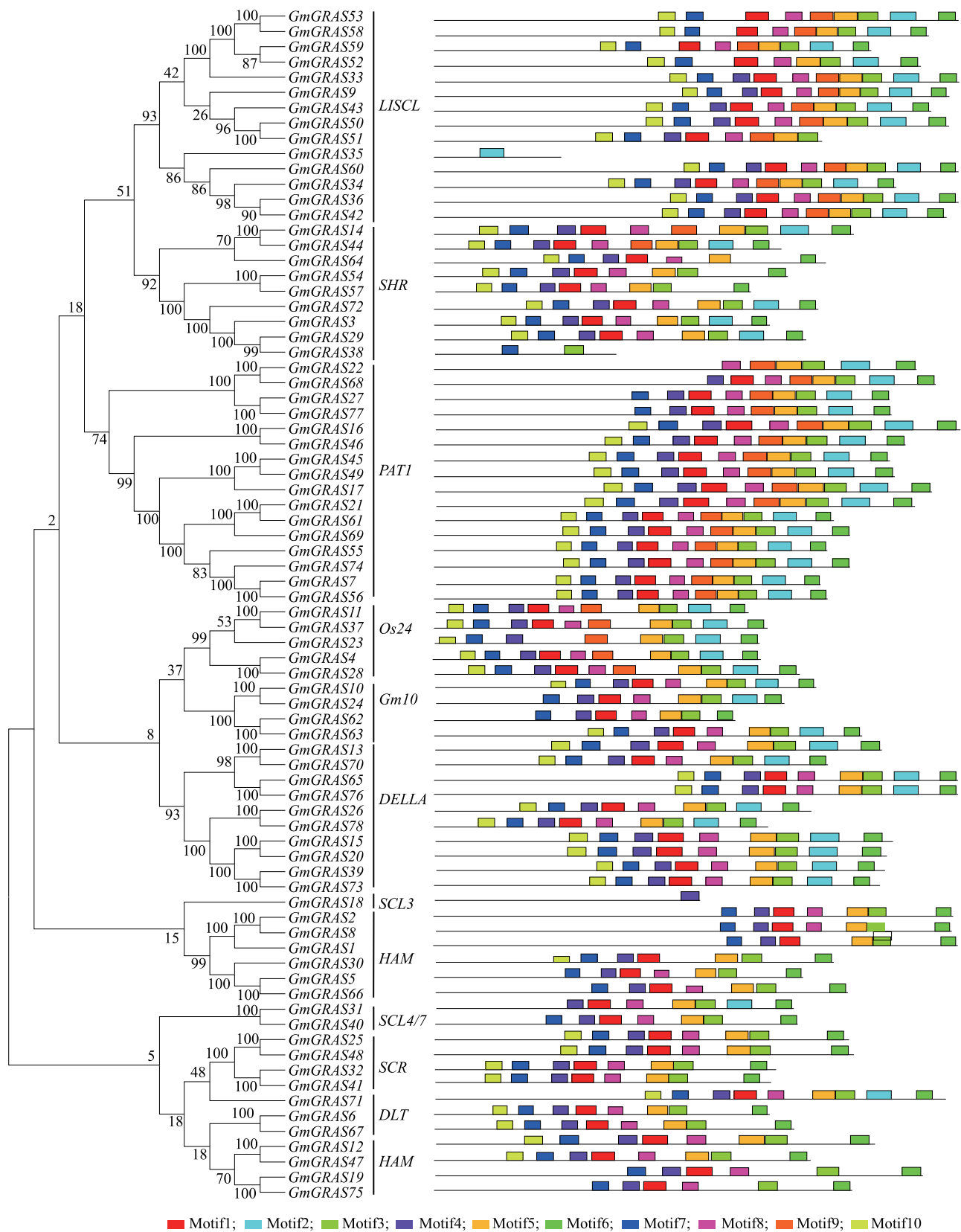
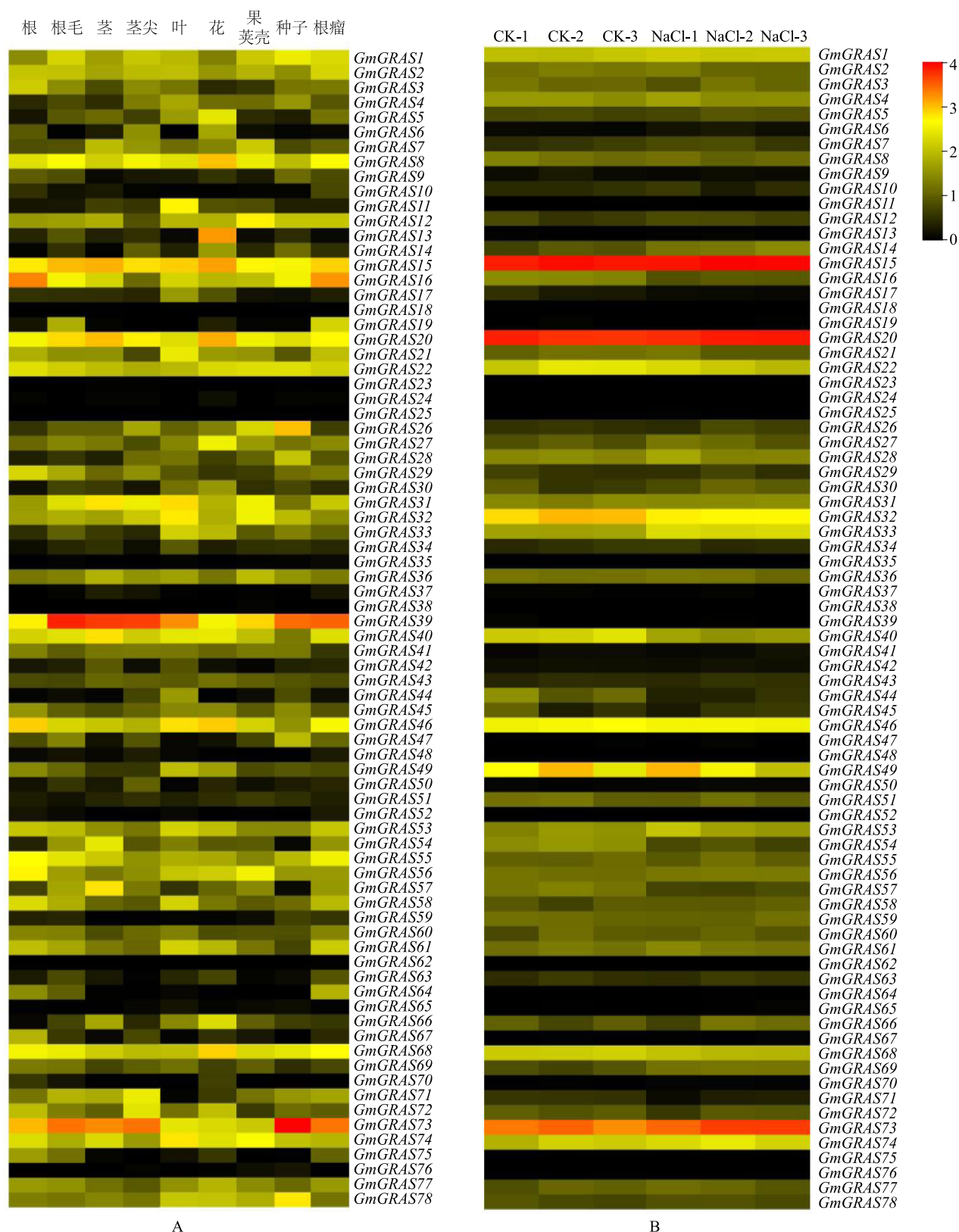


图4 大豆 GRAS 家族蛋白的保守基序分析

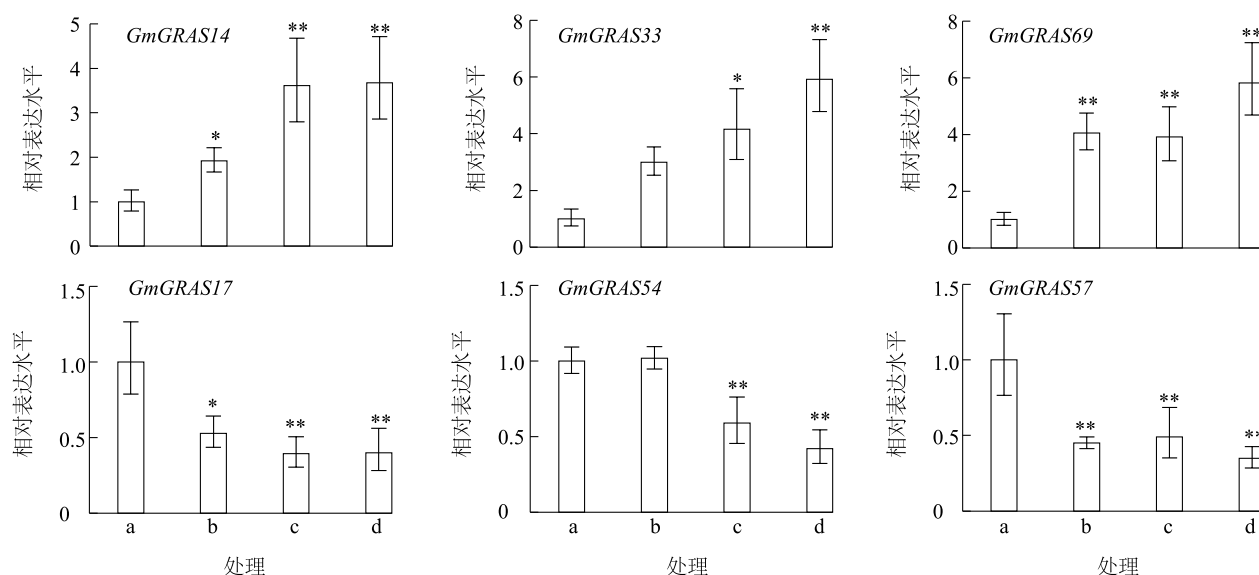
Fig.4 Analysis on the conserved motifs of soybean GRAS family proteins



图例中不同的颜色代表基因表达水平高低,红色表示表达水平最高,黑色表示表达水平最小。

图 5 大豆 GRAS 家族基因在不同组织 (A) 和盐处理后 (B) 的表达模式

Fig.5 Expression patterns of soybean GRAS family genes in different tissues (A) and after salt treatment (B)



a 为未处理对照, b、c 和 d 分别为 150 mmol/L NaCl 溶液处理 3 h、6 h 和 12 h。\*、\*\* 分别表示与对照相比差异显著和极显著。

图 6 大豆 GRAS 基因在盐胁迫下的表达水平

Fig.6 Expression levels of GmGRAS genes under salt stresses

### 3 讨论

GRAS 家族基因在植物生长发育和多种生理过程中发挥重要作用,参与赤霉素信号转导、赤霉素等激素的代谢、根的发育、响应生物和非生物胁迫的转录调控等过程。GRAS 家族的全基因组学鉴定以及它们的进化关系也已在很多植物中被研究,如模式植物拟南芥<sup>[6]</sup>和烟草<sup>[11]</sup>以及常见的农作物水稻<sup>[10]</sup>、玉米<sup>[12]</sup>、甘薯<sup>[14]</sup>、棉花<sup>[8]</sup>等。并且已有部分 GRAS 家族基因在植物中的功能及分子机制被报道。但是关于大豆 GRAS 家族的全基因组学鉴定及基因功能的研究却很少。因此,本研究从大豆基因组数据库鉴定出 GRAS 家族成员并进行全面的生物信息学分析。这将有助于我们研究大豆 GRAS 家族的进化并推测一些未知功能的大豆 GRAS 基因的生物功能。

本研究在大豆基因组中共鉴定出 78 个大豆 GRAS 家族基因,比拟南芥 (33 个)<sup>[6]</sup>、水稻 (50 个)<sup>[10]</sup>、白菜 (46 个)<sup>[28]</sup>、甘薯 (70 个)<sup>[14]</sup> 的 GRAS 家族成员数量多,但是比棉花 GRAS 家族成员少 (150 个)。造成这一现象的原因可能是不同物种间基因组大小的差异,如大豆基因组大小为 952.48 Mb,而拟南芥、水稻、白菜和甘薯的基因组大小分别为 125.0 Mb、374.0 Mb、283.3 Mb 和 492.38 Mb,棉

花的基因组大小则为 2 126.03 Mb。此外,需要特别注意的是, Wu 等<sup>[29]</sup>在大豆中鉴定出 106 个 GRAS 家族基因,而本研究鉴定出的 GRAS 基因数量与之有所差异,可能是筛选方式不同所致。本研究鉴定 GRAS 基因时主要利用公共数据库的数据,但同时参考我们自己测定的转录组结果中的 GO 注释,或许我们的转录组测序结果中注释的 GRAS 基因较实际要少一些,但本文研究的大豆 GRAS 基因均包含于已报道的 106 个基因中。Hirsch 等<sup>[6]</sup>将拟南芥 GRAS 家族基因分为 8 个亚家族,包括 LISCL、PAT1、SCL3、DELLA、SCR、SHR、LAS 和 HAM。而我们构建的拟南芥、水稻和大豆 GRAS 家族成员的进化树分为 12 个亚族,但是 LAS 分支中不含有大豆 GRAS 基因,因此将大豆 GRAS 基因家族分为 11 个亚家族,包括 LISCL、PAT1、SCL3、DELLA、SCR、SHR、HAM、SCL4/7、Os24、DLT 和 Gm10,其中 Gm10 亚族中的 4 个成员均来自大豆。说明大多数 GRAS 家族成员在进化上非常保守,而少数成员则在进化过程中具有物种特异性。此外,我们发现蛋白质上保守基序数量及分布位置相同的 GRAS 成员,其理化性质也类似,推测其基因也具有类似的功能,如 GmGRAS27 和 GmGRAS77。

通常认为在系统进化树上聚集在同一亚族的基因可能具有类似的功能<sup>[14,30-32]</sup>,因此我们分析了拟



南芥、水稻和大豆 *GRAS* 基因家族的系统进化关系, 为大豆 *GRAS* 家族基因功能的预测提供可靠的依据。进一步分析了大豆 *GRAS* 家族基因在不同组织中的表达模式, 发现大豆 *GRAS* 家族基因在不同组织的表达模式存在差异, 与之前报道的其他物种 (如甘薯、白菜) 中的表达情况类似<sup>[14]</sup>。有 7 个大豆 *GRAS* 基因 (*GmGRAS8*、*GmGRAS15*、*GmGRAS16*、*GmGRAS20*、*GmGRAS22*、*GmGRAS39* 和 *GmGRAS73*) 的表达水平在所有组织中均较高, 说明这些基因可能在大豆生长发育的多个环节都具有重要功能。尽管 Wu 等<sup>[29]</sup> 的研究已包含大豆 *GRAS* 家族全基因组学鉴定方面的工作, 但是其未关注 *GRAS* 家族基因在盐胁迫下表达模式分析。本研究通过对正常生长和 NaCl 处理的大豆进行转录组测序, 分析了大豆 *GRAS* 家族基因在盐胁迫后的表达情况, 结果表明, 有 5 个基因 (*GmGRAS6*、*GmGRAS14*、*GmGRAS15*、*GmGRAS33*、*GmGRAS69*) 的表达水平在盐胁迫下显著上调, 而 *GmGRAS17*、*GmGRAS45*、*GmGRAS54*、*GmGRAS57* 和 *GmGRAS71* 的表达水平显著下调。qRT-PCR 分析结果显示 *GmGRAS14*、*GmGRAS33* 和 *GmGRAS69* 在盐处理 6 h 时表达水平均显著上调, 与转录组测序结果一致, 且 3 个基因均在盐处理 12 h 时上调倍数最高, 分别为 3.67 倍、5.92 倍和 5.82 倍; 而 *GmGRAS17*、*GmGRAS54* 和 *GmGRAS57* 的表达水平在盐处理 6 h、12 h 时均显著下调, 与转录组结果一致。通过系统进化分析、盐胁迫下的表达模式分析以及 qRT-PCR 检测, 可初步推测大豆 *GRAS* 家族基因在生长发育过程和盐胁迫响应方面具有重要功能。本研究也为大豆 *GRAS* 基因功能的进一步研究及利用该家族基因培育大豆耐盐新品种提供了理论依据。

## 参考文献:

- [1] HOANG X L T, NHI D N H, THU N B A, et al. Transcription factors and their roles in signal transduction in plants under abiotic stresses[J]. *Current Genomics*, 2017, 18(6): 483-497.
- [2] DI LL, WY SOCKADILLER J, MALAMY JE, et al. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root[J]. *Cell*, 1996, 86(3): 423-433.
- [3] PENG J, CAROL P, RICHARDS D E, et al. The *Arabidopsis* GAI gene defines a signaling pathway that negatively regulates gibberellin responses[J]. *Genes & Development*, 1997, 11(23): 3194.
- [4] SILVERSTONE A L, CIAMPAGLIO C N, SUN T. The *Arabidopsis* RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway[J]. *Plant Cell*, 1998, 10(2): 155-169.
- [5] SUN X, XUE B, JONES W T, et al. A functionally required unfolded domain from the plant kingdom: intrinsically disordered N-terminal domains of GRAS proteins are involved in molecular recognition during plant development[J]. *Plant Molecular Biology*, 2011, 77(3): 205-223.
- [6] HISCH S, OLDROYD G E D. GRAS-domain transcription factors that regulate plant development[J]. *Plant Signaling & Behavior*, 2009, 4(8): 698-700.
- [7] XY W, CHEN Z, AHMED N, et al. Genome-wide identification, evolutionary analysis, and stress responses of the GRAS gene family in Castor beans[J]. *International Journal of Molecular Sciences*, 2016, 17(7): 1004.
- [8] ZHANG B, LIU J, YANG Z E, et al. Genome-wide analysis of GRAS transcription factor gene family in *Gossypium hirsutum* L. [J]. *BMC Genomics*, 2018, 19(1): 348.
- [9] SHAN Z Y, LUO X L, WU M Y. Genome-wide identification and expression of GRAS gene family members in cassava [J]. *BMC Plant Biology*, 2020, 20(1): 46.
- [10] TIAN C, WAN P, SUN S, et al. Genome-wide analysis of the GRAS gene family in rice and *Arabidopsis* [J]. *Plant Molecular Biology*, 2004, 54(4): 519-532.
- [11] CHEN Y Q, TAI S S, WANG D W, et al. Homology-based analysis of the GRAS gene family in tobacco [J]. *Genetics and Molecular Research*, 2015, 14(4): 15188-15200.
- [12] GUO Y, WU H, LI X, et al. Identification and expression of GRAS family genes in maize (*Zea mays* L.) [J]. *PLoS One*, 2017, 12(9): e0185418.
- [13] LIU X, WIDMER A. Genome-wide comparative analysis of the GRAS gene family in populus, *Arabidopsis* and rice [J]. *Plant Molecular Biology Reporter*, 2014, 32(6): 1129-1145.
- [14] CHEN Y, ZHU P P, WU S Y. Identification and expression analysis of GRAS transcription factors in the wild relative of sweet potato *Ipomoea trifida* [J]. *BMC Genomics*, 2019, 20(1): 911.
- [15] BOLLE C. The role of GRAS proteins in plant signal transduction and development [J]. *Planta*, 2004, 218(5): 683-692.
- [16] HEO J, CHANG K, KIM I, et al. Funneling of gibberellin signaling by the GRAS transcription regulator SCARECROW-LIKE 3 in the *Arabidopsis* root [J]. *PNAS*, 2011, 108(5): 2166-2171.
- [17] LI W, WU J, WENG S, et al. Identification and characterization of dwarf62, a loss-of-function mutation in *DLT/OsGRAS-32* affecting gibberellin metabolism in rice [J]. *Planta*, 2010, 232(6): 1383-1396.
- [18] BAI Z, XIA P, WANG R, et al. Molecular cloning and characterization of five *SmGRAS* genes associated with tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots [J]. *PLoS One*, 2017, 12(9): e0185322.
- [19] TANABE S, ONODERA H, HARA N, et al. The elicitor-respon-

- sive gene for a GRAS family protein, CIGR2, suppresses cell death in rice inoculated with rice blast fungus via activation of a heat shock transcription factor, OsHsf23[J]. Bioscience, Biotechnology and Biochemistry, 2016, 80(2): 1145-1151.
- [20] YANG M, YANG Q, FU T, et al. Overexpression of the *Brassica napus* BnLAS gene in *Arabidopsis* affects plant development and increases drought tolerance[J]. Plant Cell Reports, 2011, 30(3): 373-388.
- [21] MA HS, LIANG D, SHUAI P, et al. The salt- and drought-inducible poplar GRAS protein SCL7 confers salt and drought tolerance in *Arabidopsis thaliana* [J]. Journal of Experimental Botany, 2010, 61(14): 4011-4019.
- [22] HABIB S, WASEEM M, LI N, et al. Overexpression of *SlGRAS7* affects multiple behaviors leading to confer abiotic stresses tolerance and impacts gibberellin and auxin signaling in tomato[J]. International Journal of Genomics, 2019, 2019: 4051981.
- [23] HOSSAIN M S, HOANG N T, YAN Z. Characterization of the spatial and temporal expression of two soybean miRNAs identifies SCL6 as a novel regulator of soybean nodulation[J]. Frontiers in Plant Science, 2019, 10: 475.
- [24] VOORRIPS R E. MapChart: Software for the graphical presentation of linkage maps and QTLs[J]. Journal of Heredity, 2002, 93(1): 77-78.
- [25] HU B, JIN J, GUO A, et al. GSDS 2.0: an upgraded gene feature visualization server[J]. Bioinformatics, 2015, 31(8): 1296.
- [26] BAILEY T L, MIKAEL B, BUSKE F A, et al. MEME Suite: tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37(S): 202-208.
- [27] PI B Y, HE X H, RUAN Y, et al. Genome-wide analysis and stressresponsive expression of CCCH zinc finger family genes in *Brassica rapa*[J]. BMC Plant Biology, 2018, 18(1): 373.
- [28] SONG X, LIU T, DUAN W, et al. Genome-wide analysis of the GRAS gene family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) [J]. Genomics, 2014, 103(1): 135-146.
- [29] WU N N, ZHU Y, SONG W L, et al. Unusual tandem expansion and positive selection in subgroups of the plant GRAS transcription factor superfamily[J]. BMC Plant Biology, 2014, 14: 373.
- [30] 李 倩, 盖江涛, 白蓓蓓, 等. 茄科植物中 *HCT* 基因家族的鉴定及进化和表达分析[J]. 江苏农业科学, 2019, 47(19): 65-68.
- [31] 付瑜华, 蒙秋伊, 李秀诗, 等. 薏苡油脂合成关键基因克隆及其生物信息学分析[J]. 南方农业学报, 2020, 51(3): 485-495.
- [32] 陈亨德, 褚武英, 李玉珑, 等. 罗非鱼 *SIRT1* 基因的克隆及其表达规律分析[J]. 江苏农业科学, 2019, 47(21): 103-106.

(责任编辑:张震林)