



分子植物育种
Molecular Plant Breeding
ISSN 1672-416X, CN 46-1068/S

《分子植物育种》网络首发论文

题目： 大麦基因组 GRAS 基因家族的全基因组鉴定与表达分析
作者： 孔豆豆, 毛成志, 王蕾, 王寒冬, 张怀刚, 徐金青, 沈裕虎
网络首发日期： 2019-11-28
引用格式： 孔豆豆, 毛成志, 王蕾, 王寒冬, 张怀刚, 徐金青, 沈裕虎. 大麦基因组 GRAS 基因家族的全基因组鉴定与表达分析. 分子植物育种.
<http://kns.cnki.net/kcms/detail/46.1068.S.20191128.0957.002.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

研究报告

Research Report

大麦基因组 GRAS 基因家族的全基因组鉴定与表达分析

孔豆豆^{1,2} 毛成志^{1,2} 王蕾^{1,3} 王寒冬^{1,3} 张怀刚^{1,3} 徐金青^{1,3*} 沈裕虎^{1,3*}

1 中国科学院西北高原生物研究所, 中国科学院高原适应与进化重点实验室, 青海省作物分子育种重点实验室, 西宁, 810001;

2 中国科学院大学, 生命科学学院, 北京, 100049; 3 中国科学院种子创新研究院(筹), 西宁, 810001

*共同通信作者, shenyuhu@nwipb.cas.cn; xjq1088@126.com

摘要 GRAS 基因家族是一类仅存在于植物并广泛参与其生长发育调控的转录因子, 根据其序列结构和系统发育树分化特征, GRAS 转录因子包含 PAT1, DELLA, HAM, SCR, SHR 等多个亚家族成员。本研究利用大麦最新的基因组数据库, 采用生物信息学的方法筛选鉴定出 41 条 GRAS 基因序列, 其中有 34 条序列具有完整的 GRAS 家族蛋白特有的 GRAS 结构域, 可定位到大麦 7 条染色体上且呈不均匀分布。与拟南芥和水稻 GRAS 蛋白进行的系统发育分析, 可将大麦 GRAS 家族蛋白进一步划分为 10 个亚家族。对大麦 GRAS 基因的表达丰度分析, 发现部分基因在发育阶段高表达, 这可能暗示着这些基因在相应发育阶段起到重要作用。本研究可为后续挖掘和验证大麦 GRAS 基因提供参考。

关键词 大麦, GRAS 基因家族, 系统发育分析, 表达谱分析

Genome-Wide Identification and Phylogenetic Analysis of the GRAS Gene Family in Barley (*Hordeum vulgare* L.)

Kong Doudou^{1,2} Mao Chengzhi^{1,2} Wang Lei^{1,3} Wang Handong^{1,3} Zhang Huaigang^{1,3} Xu Jinqing^{1,3*} Shen Yuhu^{1,3*}

1 Key Laboratory of Adaptation and Evolution of Plateau Biota (AEPB), Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, 810001; 2 University of Chinese Academy of Sciences, College of Life Sciences, Beijing, 100049; 3 Innovation Academy for Seed Design, Chinese Academy of Sciences, Xining, 810001

* Co-Corresponding author: shenyuhu@nwipb.cas.cn; xjq1088@126.com

Abstract The GRAS gene family is a kind of transcription factor that exists only in plants and is widely involved in the regulation of its growth and development. According to its sequence structure and functional differentiation characteristics, GRAS transcription factors include PAT1, DELLA, HAM, SCR, SHR and other subfamily members. In this study, we used the latest genomic database of barley to screen and identify 41 GRAS gene sequences by bioinformatics methods, among which 34 sequences have complete GRAS domain specific to GRAS family proteins. Among them, 40 genes can be mapped to 7 chromosomes of barley and are unevenly distributed. Phylogenetic analysis with Arabidopsis and rice GRAS proteins further divides the barley GRAS family proteins into 10 subfamilies. Analysis of the abundance of barley GRAS gene revealed that some genes are highly expressed during development, which may imply that these genes play an important role in the corresponding developmental stages. This study can provide a reference for subsequent mining and verification of barley GRAS genes.

Keywords Barley, GRAS gene family, Phylogenetic analysis, Expression profiling analysis

大麦(*Hordeum vulgare* L.)是非常古老的作物之一，早在新石器时代就已开始栽培。由于大麦具有优良的耐干旱、耐紫外辐射、耐低温和耐瘠薄能力，大麦(青稞)在我国青藏高原和云贵高原广泛分布，几乎是唯一能够在海拔 4000 米以上保持优良性状的作物。得益于已发表的大麦基因组数据(Mayer et al., 2012; Mascher et al., 2017)，人们能够在基因组水平上对大麦生长发育和逆境适应调控的分子机制进行研究。然而由于大麦基因组较大(5.1 Gb)，约为水稻基因组的 11 倍，很多重要基因尚未定位或克隆。

转录因子(Transcription factor)是一类可与目的基因序列结合、保证其进行特异性时空表达的蛋白质因子。GRAS 转录因子家族命名来源于三个最早被研究的家族成员：GAI (Gibberellic acid insensitive) (Peng et al., 1997)、RGA (repressor of GA1-3 mutant) (Silverstone et al., 1998)和 SCR (Scarecrow) (Di et al., 1996)。GRAS 蛋白家族成员基因在序列长度和结构有较大差异。典型的 GRAS 蛋白一般由 400~770 个氨基酸组成，包括一个可变的 N 端序列和一个较为保守的 C 端序列。C 末端含有五个保守的基序(motif):LR I (Leucine

heptad region I)、VHIID、LR II (Leucine heptad region II)、PFYRE 和 SAW, 其中 VHIID、PFYRE 和 SAW 基序相对更为保守(Pysh et al., 2010)。前人研究表明 GRAS 家族蛋白除了 N 端序列自由多变外, 绝大多数蛋白 C 端都仅有一个结构域, 即 GRAS 结构域(李亚飞等, 2019)。根据 GRAS 转录因子的不同功能, 早期将拟南芥 GRAS 转录因子划分为 DELLA、SCL3、LISCL、SCR、HAM、SHR、PATI 和 LS 8 个亚家族(Cordelia, 2004)。后来在研究分子信号识别时, 又进一步将拟南芥 GRAS 转录因子划分为 10 个亚家族: DELLA、HAM、LISCL、AtSCR、AtPAT1、DLT、AtSCL3、AtSCL4/7、AtLAS、AtSHR (Sun et al., 2011)。但是在这 10 个亚家族中并未包含水稻 *OsGRAS4* 和 *OsGRAS19* 的同源基因。Liu 和 Widmer (2014)对水稻、拟南芥、杨树、葡萄和番茄等多物种的 GRAS 蛋白序列进行了比对分析和鉴定, 将 GRAS 蛋白划分为 DELLA、HAM、LISCL、AtSCR、AtPAT1、DLT、AtSCL3、AtSCL4/7、AtLAS、AtSHR、Pt20、Os19 和 Os4 等 13 个亚家族。其中, Pt20 为杨树特有亚家族, Os19 和 Os4 为新鉴定的亚家族。不同亚家族转录因子拥有不同的生理功能。水稻中 DLT 亚家族参与 BR 信号传导, GSK2 可直接作用并磷酸化 DLT (Tong et al., 2012), *GS6* 基因属于 GRAS 家族 DLT 亚家族, 它可通过调控颖片中细胞的数目影响内外稃的发育, 从而负调控水稻籽粒大小(Sun et al., 2013)。Torres-Galea 等(2013)、牛义岭等(2016)发现 PATI 亚家族的 *PATI* 和 *SCL21* 参与 phy A 介导的光信号转导过程, *SCL13* 作为正调节因子参与 phy B 介导的光信号转导过程。Dill 等(2001)发现 DELLA 亚家族可以负调控 GA 信号通路, 调节植物的生长。拟南芥 DELLA 亚家族参与 GA 信号调控的主要有: GAI、RGA、RGL1、RGL2 和 RGL3。

目前已经在多种植物中对 GRAS 基因家族进行了鉴定和分析, 拟南芥中存在 34 个 GRAS 转录因子 (Sun et al., 2013), 水稻中存在 60 个 GRAS 转录因子(Sun et al., 2013), 小麦中鉴定得到 153 个 GRAS 转录因子(李亚飞等, 2019)。大麦是一种集饲用、啤用和粮食于一体的重要农作物, 其 GRAS 转录因子家族并未鉴定, 此基因家族在植物生长发育过程中起着重要作用。此外在其它植物如百脉根(Heckmann et al., 2006)、松树(Solé et al., 2009)、佛肚竹(Zhou et al., 2011)、佛手(石瑞等, 2011)和葡萄(孙欣等, 2011)等物种中也鉴定得到了大量 GRAS 家族转录因子。本研究利用最新的大麦基因组数据库, 对大麦 GRAS 家族转录因子进行

了筛选和鉴定，构建了 34 条具有完整 GRAS 结构域的 GRAS 转录因子的系统发育树，并对其基因结构和表达丰度进行了分析，可为后续大麦 GRAS 基因家族的功能研究提供参考。

1 结果与分析

1.1 大麦 GRAS 基因的鉴定

在大麦基因组注释信息中以“GRAS family”为关键词共搜索到 41 条蛋白质序列，使用 NCBI-CDD 工具筛选保留 34 条具有完整 GRAS 结构域的序列。以 *HvGRAS* 命名大麦 GRAS 基因，将对应的 41 条 DNA 序列作为大麦 GRAS 转录因子家族的候选基因。染色体定位结果显示，除 *HvGRAS0H_1* 基因未定位到具体染色体外，其余基因呈不均匀状态分布在大麦 1~7 号染色体上。

从染色体定位结果来看，4 号染色体上分布的 *HvGRAS* 基因最多，共有 13 个，6 号染色体上基因最少，只有两个。*HvGRAS* 基因在染色体上分布位置不均，大多数位于染色体两端(图 1)。

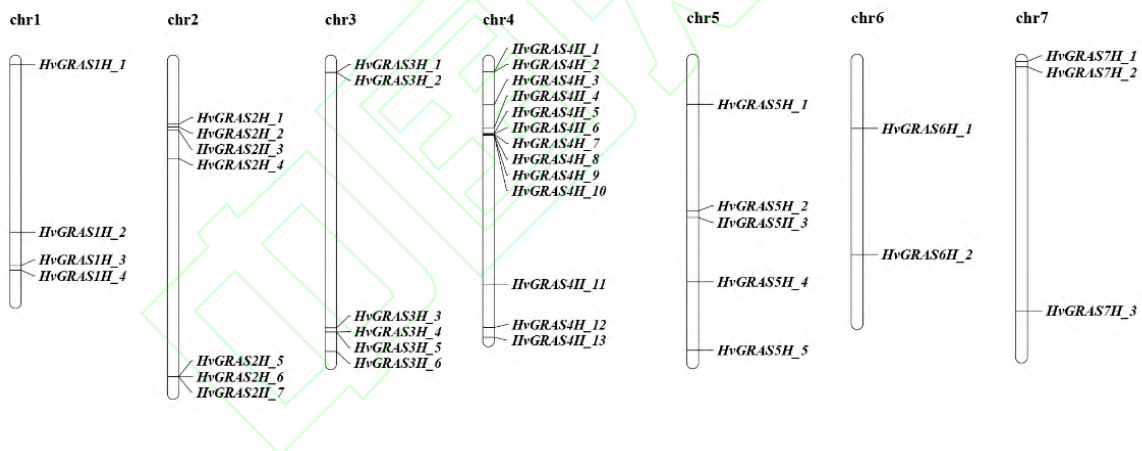


图 1 *HvGRAS* 基因家族的染色体定位

Figure 1 chromosomal locations of the barley GRAS gene family

HvGRAS 基因长度(length)范围为 947bp(*HvGRAS2H_7*) ~ 36544bp(*HvGRAS2H_5*), 多肽大小(size)范围为 179 aa (*HvGRAS3H_6*) ~ 793 aa (*HvGRAS4H_10*), 理论等电点(pI)范围为 4.51 (*HvGRAS5H_4*) ~ 9.56 (*HvGRAS3H_2*) (表 1)。

表 1 *HvGRAS* 基因编码蛋白的基本特征

Table 1 Basic characteristics of the proteins encoded by *HvGRAS* genes

基因名称	基因号	染色体	染色体位置	基因长度	等电点	多肽大小
gene name	gene ID	chr	Location (bp)	Length (bp)	pI	size (aa)
<i>HvGRAS0H_1</i>	HORVU0Hr1G004640.1	-	26283817~26285757	1 940	5.72	479
<i>HvGRAS1H_1</i>	HORVU1Hr1G004440.1	1H	9163281~9166923	3 642	7.23	719
<i>HvGRAS1H_2</i>	HORVU1Hr1G053510.1	1H	395873135~395875994	2 859	5.91	620
<i>HvGRAS1H_3</i>	HORVU1Hr1G066390.1	1H	472662646~472664296	1 650	6.17	415
<i>HvGRAS1H_4</i>	HORVU1Hr1G069140.2	1H	483084352~483085318	966	5.18	216
<i>HvGRAS2H_1</i>	HORVU2Hr1G035120.1	2H	146366257~146368654	2 397	5.45	458
<i>HvGRAS2H_2</i>	HORVU2Hr1G035730.2	2H	152815540~152818096	2 556	5.51	402
<i>HvGRAS2H_3</i>	HORVU2Hr1G036280.1	2H	160352129~160355538	3 409	6.21	557
<i>HvGRAS2H_4</i>	HORVU2Hr1G043780.1	2H	225382581~225387077	4 496	5.79	570
<i>HvGRAS2H_5</i>	HORVU2Hr1G113360.1	2H	728433055~728469599	36 544	6.41	475
<i>HvGRAS2H_6</i>	HORVU2Hr1G113480.1	2H	728896619~728897566	947	6.76	315
<i>HvGRAS2H_7</i>	HORVU2Hr1G113490.1	2H	728906883~728907830	947	6.76	315
<i>HvGRAS3H_1</i>	HORVU3Hr1G012520.1	3H	27025471~27026840	1 369	9.56	456
<i>HvGRAS3H_2</i>	HORVU3Hr1G012550.6	3H	27046202~27054168	7 966	8.82	596
<i>HvGRAS3H_3</i>	HORVU3Hr1G086210.8	3H	616216740~616219702	2 962	5.69	791
<i>HvGRAS3H_4</i>	HORVU3Hr1G088780.2	3H	625825826~625827709	1 883	6.39	531
<i>HvGRAS3H_5</i>	HORVU3Hr1G088820.1	3H	625945682~625946867	1 185	-	260
<i>HvGRAS3H_6</i>	HORVU3Hr1G106260.1	3H	670706464~670707477	1 013	6.76	179
<i>HvGRAS4H_1</i>	HORVU4Hr1G008930.3	4H	25525326—25526610	1 284	-	353

<i>HvGRAS4H_10</i>	HORVU4Hr1G027920.1	4H	171923610~171927659	4 049	6.02	793
<i>HvGRAS4H_11</i>	HORVU4Hr1G061310.1	4H	515361930~515364171	2 241	5.41	606
<i>HvGRAS4H_12</i>	HORVU4Hr1G080460.5	4H	615736049~615737871	1 822	5.53	582
<i>HvGRAS4H_13</i>	HORVU4Hr1G087700.8	4H	638044256~638050012	5 756	6.52	330
<i>HvGRAS4H_2</i>	HORVU4Hr1G008980.3	4H	25658574~25660462	1 888	-	553
<i>HvGRAS4H_3</i>	HORVU4Hr1G020490.28	4H	101454651~101465909	11 258	4.97	474
<i>HvGRAS4H_4</i>	HORVU4Hr1G026010.3	4H	154386614~154391276	4 662	7.25	446
<i>HvGRAS4H_5</i>	HORVU4Hr1G027340.3	4H	167899833~167901758	1 925	6.2	374
<i>HvGRAS4H_6</i>	HORVU4Hr1G027480.1	4H	169007327~169009840	2 513	6.16	707
<i>HvGRAS4H_7</i>	HORVU4Hr1G027520.2	4H	169282885~169284846	1 961	5.48	653
<i>HvGRAS4H_8</i>	HORVU4Hr1G027610.7	4H	169472615~169476708	4 093	5.91	677
<i>HvGRAS4H_9</i>	HORVU4Hr1G027900.2	4H	171834377~171837455	3 078	6.04	662
<i>HvGRAS5H_1</i>	HORVU5Hr1G020900.4	5H	97832937~97835642	2 705	5.38	745
<i>HvGRAS5H_2</i>	HORVU5Hr1G043310.1	5H	333546455~333548332	1 877	6.63	550
<i>HvGRAS5H_3</i>	HORVU5Hr1G045020.17	5H	348282527~348288674	6 147	7.31	776
<i>HvGRAS5H_4</i>	HORVU5Hr1G063270.3	5H	490660211~490662014	1 803	4.51	396
<i>HvGRAS5H_5</i>	HORVU5Hr1G113620.1	5H	642237666~642240639	2 973	6.14	644
<i>HvGRAS6H_1</i>	HORVU6Hr1G033280.1	6H	151045446~151046788	1 342	5.26	446
<i>HvGRAS6H_2</i>	HORVU6Hr1G063650.1	6H	430342265~430345365	3 100	5.9	703
<i>HvGRAS7H_1</i>	HORVU7Hr1G001300.3	7H	2726997~2729398	2 401	7.23	225
<i>HvGRAS7H_2</i>	HORVU7Hr1G010620.1	7H	14944547~14946958	2 411	5.78	563
<i>HvGRAS7H_3</i>	HORVU7Hr1G091000.1	7H	555109073~555110524	1 451	6.43	426

1.2 HvGRAS 基因的系统进化分析

典型的 GRAS 蛋白至少含有 350 个氨基酸残基(Liu and Widmer, 2014, Huang et al., 2015), 依据 GRAS 结构域序列的相似性, 使用 NCBI-CDD 工具筛选保留 34 条具有完整 GRAS 结构域的蛋白序列。使用 MEGA 7.0 构建拟南芥、水稻和大麦 GRAS 蛋白保守域序列的系统进化树(图 2)。根据前人已发表的拟南芥 GRAS 和水稻 GRAS 亚家族分类情况和建树结果, 拟南芥、水稻和大麦三个物种的 GRAS 蛋白可进一步划分为 12 个亚家族, OsGRAS43 未被聚类到任何一个亚家族内(Liu and Widmer, 2014) (图 2)。HvGRAS 在 12 个亚家族中的分布不均匀, 除 AtSCR 和 DELLA 亚家族外, 其余 10 个亚家族中均有 HvGRAS 蛋白的分布。其中, LISCL 亚家族中 HvGRAS 蛋白分布最多有 14 个, 而 Os4、Os19 和 DLT 亚家族中仅有 1 个 HvGRAS 蛋白。AtPAT 亚族中的 *OsCIGR1* 和 *HvGRAS2H_4*, Os4 亚族中的 *OsGRAS4* 与 *HvGRAS3H_4*, SCL4/7 亚族中的 *OsGRAS18* 和 *HvGRAS4H_1*、*HvGRAS4H_2*, Os19 亚族中的 *OsGRAS19* 与 *HvGRAS0H_1*, DLT 亚族中的 *OsGRAS29* 与 *HvGRAS7H_2*, AtLAS 亚族中的 *OsMOCI* 与 *HvGRAS7H_3* 的自展值均为 100, 推测其可能为直系同源蛋白, 其功能也较为保守。从不同亚家族间拟南芥和水稻、大麦的 GRAS 蛋白分布情况来看, SCL3、Os4 和 Os19 三个亚家族中仅有水稻和大麦 GRAS, 没有拟南芥 GRAS 的分布。其余 9 个亚家族中均有 3 种植物 GRAS 的分布。说明 GRAS 蛋白亚家族在单双子叶植物间存在一定程度的分化。在 LISCL 亚族和 SCL4/7 亚族内, *OsGRAS* 和 *HvGRAS* 蛋白混杂分布, 与 AtGRAS 独立成枝, 推测可能是植物的单双子叶分化导致的(李亚飞等, 2019)。因此, 植物 GRAS 基因的单双子叶分化既存在于亚家族间也存在于亚家族内部。

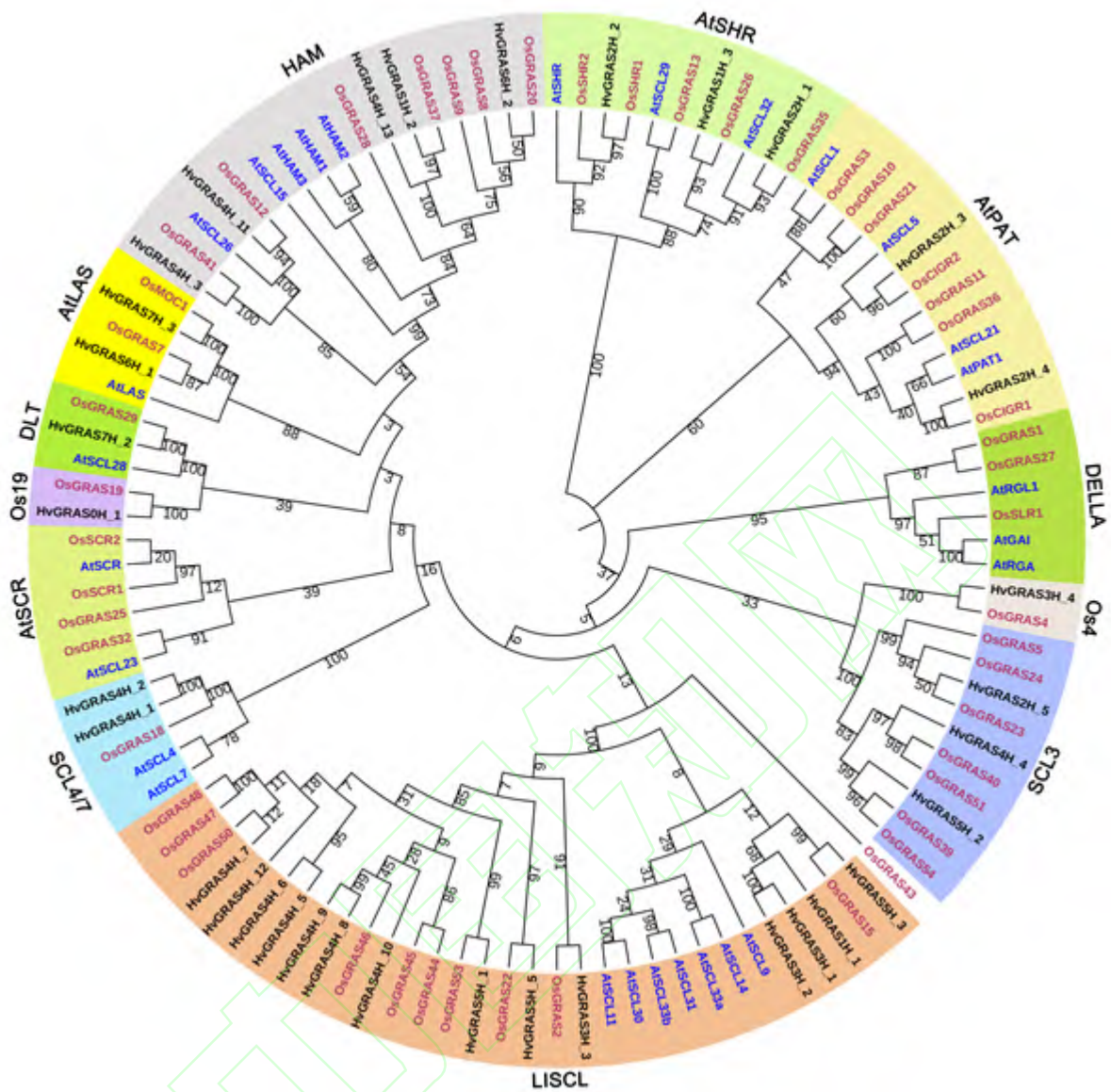


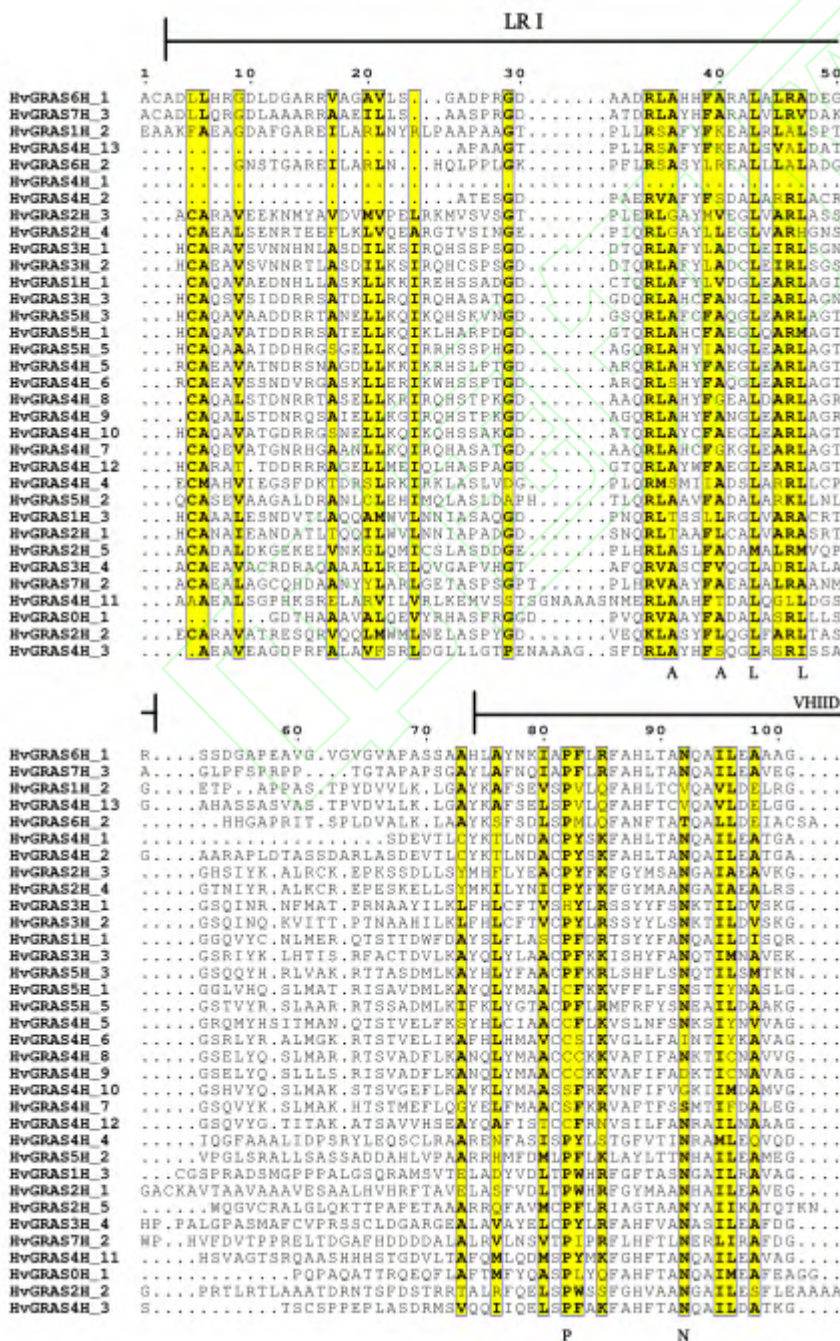
图2 拟南芥, 水稻和大麦 GRAS 蛋白保守域系统进化树

Figure 2 Phylogenetic analysis of GRAS conserved domains in *HvGRAS* proteins

1.3 *HvGRAS* 蛋白家族保守元件分析

GRAS 基因家族的编码产物在 C 端具有高度的相似性, 由五个基序按照 LR I – VHIIID – LR II – PFYRE – SAW 顺序组成。LR I (Leucine Heptad I) 和 LR II (Leucine Heptad II) 富含亮氨酸, 为分别含 3~5 个和 2~3 个亮氨酸的重复七肽结构。VHIIID 基序在较大区域内含较为保守的 P – N – H – D – Q – L 氨基酸残基, PFYRE 基序的特征是含有保守的 P – F – Y – R – E 氨基酸残基, SAW 基序含三对保守残基 R – E, W – G, W – W (Pysh et al., 2010)。

对 34 个 HvGRAS 蛋白保守结构域的多序列比对结果显示, 不同序列中的 LR I 和 LR II 基序都有较大的差异, 但所有的 LR I 区域内均呈现亮氨酸富集状态; 其次, 在所有蛋白序列中 VHIID 基序中的保守位点都易于识别, 其中谷氨酰胺位点(Q)绝对保守; 同时, 在 PFYRE 基序内均检测到了 P-F-Y-R-E 五个保守氨基酸残基。虽然 PFYRE 基序保守性不如 VHIID 基序, 但也表现出高度的相似性; 在 SAW 基序内也同样检测出了三对保守残基 R-E, W-G, W-W(图 3)。以上结果表明, 大麦 HvGRAS 家族蛋白在 C-端同样存在高度的保守性。



VHIIID

110 120 130 140

```

HvGRAS6H_1 ...ARRVHIVDQAAHG...VQPPLLQATADR...AD...AVGPFPEVRIITGAGPDI
HvGRAS7H_3 ...SRRVHIVDQAAHG...VQPPLLQATAER...ADP...ALGPFPEVRIITGAGADR
HvGRAS1H_2 ...AGCIHVLDQDQIGMG...EHWASLMQELAOR...RPA...AALKVITLVFSPSSHHP
HvGRAS4H_13 ...AGCIHVLDQDQIGVG...EHWASLMQELAOR...RPA...AALKVITLVFSPSSHHP
HvGRAS6H_2 ...ASCIVHVDQDLVGG...GHWASFLQELAAR...RGTGGVAL...LLKLTAFVSDASHHP
HvGRAS4H_1 ...ATKIVHVDQIVHG...IHWALLQALATR...PEGKPSRI...ISGVSPSYLGP...QPA
HvGRAS4H_2 ...ATKIVHVDQIVHG...IHWALLQALATR...PEGKPSRI...ISGVSPSYLGP...QPA
HvGRAS2H_3 ...EDRIVHVDQIACG...AWISLLQALAA...PGGPTVRI...ITGIDDSVSAAYARGGG
HvGRAS2H_4 ...EDKIVHVDQIACG...TMITLQALAA...PGGPHVRI...ITGIDDPVSEYARGEG
HvGRAS3H_1 ...KSVHIVDQICFG...FQWPSLLKQFADR...EGGPKLRI...ITGIELPLPGFQPHGR
HvGRAS3H_2 ...KPRVHIVDQIGFYG...FQWPSLLKQFADR...EGGPKLRI...ITGIELPMPGFRPDEM
HvGRAS1H_1 ...KPRVHIVDQIGFYG...FQWPLMIOKFALO...EEGAPKLR...ITGIDVQPQGFPRPCEM
HvGRAS3H_3 ...AKKIVHVDQGVYFG...FQWPCLLQRLGKR...PGGPELRI...ITADITPQPGFRPAER
HvGRAS5H_3 ...ASKVHIVDQGIYFG...LQWPCLLRRLSKR...EGGPKLRI...ITGIDVPEPQGFPRTER
HvGRAS5H_1 ...KKKIVHVDQGIYFG...FQWPCFLRRISOR...PGGPNVRI...ITGIDLQPGGFRPTER
HvGRAS5H_5 ...VTSVHIVDQIDMG...FQWPIFLQRLSKR...DGGPRIRI...ITGIDLQPGGFRPAER
HvGRAS4H_5 ...RKKLHIVHGVNDG...FQWPCLLRRLAAR...EGGPEVRI...ITGIIISQPGCLAKQ
HvGRAS4H_6 ...RRTLHIVHGVNDG...FQWPCLLRRLAAR...EGGPEVRI...ITGINTPRPGLRPAQL
HvGRAS4H_8 ...RSRLHIVDQVLSOG...LQWPCLLRRLAAR...EGGPEVRI...ITGIDLQPGGFRPAER
HvGRAS4H_9 ...KTRLHIVDQVLSOG...LQWPCLLRRLAAR...EGGPEVRI...ITGIDLQPGGFRPAER
HvGRAS4H_10 ...KSRHIVDQVYVYG...FQWPCLLQALAA...EGGPDVRI...ITGIDLQPGGFRPAER
HvGRAS4H_7 ...KSLHIVDQVYVYG...FQWPCLLQALAA...EGGPDVRI...ITGIDLQPGGFRPAER
HvGRAS4H_12 ...RSRLHIVDQVYVYG...FQWSELLRSLAAR...DGGSEVRI...ITHIDLQPGGLHPRGR
HvGRAS4H_4 ...QKVIVHVDQSSCTSH...PWQWKLKLRHDFHGR...PGGPELRI...ITVHEDI...EF
HvGRAS5H_2 ...ERFVHIVDQSGPAAN...PAQWIALPHAFKGR...RGTGNLRI...ITAVHESK...EF
HvGRAS1H_3 ...RNAVHIVDQGVYFG...MCPPTLIDALSKR...PGGAPALRI...ITVPSVRPAG...PPLL
HvGRAS2H_1 ...FAVHIVDQSGYFG...MCPPTLIDALSKR...VEGPKLRI...ITVADVCSGPPPALD
HvGRAS2H_5 ...NAVHIVDQSGYFG...MCPPTLIDALSKR...VEGPKLRI...ITVADVCSGPPPALD
HvGRAS3H_4 ...ESNVHIVDQMTMGNR...HWRALLDGLATR...VQGGPARV...ITCVSARVDTMR...
HvGRAS7H_2 ...HRRVHIVDQIKCG...LQWPSLLQALAA...RPPPAHVRI...ITG...VGSRR
HvGRAS4H_11 ...DRRLHIVDQVLSOG...LQWASLMQAMTSR...PDGSPPHRI...ITAITRSGGGGARA
HvGRAS0H_1 ...TRRLHIVDQVLSOG...FQWPSLLQALAA...AATSSLSL...ITGSDGSDCEPVSILR
HvGRAS2H_2 ...SSEFQRVHILDVSNFC...TQWPTLLEALATRS...PHLSIT...ITVVSAPASSETSASVQ
HvGRAS4H_3 ...DAHVHIVDQVLSOG...LQWPSLLQALAA...GGISFRL...ITAVTADADYSDAVHH

```

VHIIID Q L LRITG

150 160 170 180 190

```

HvGRAS6H_1 ...VLLRTGDRLEAFASLNLPPFHTHPLLLPCTAQLAADP...AACLELHP...ETDA
HvGRAS7H_3 ...TLLRTGNRLRAFARSIQLPFHTHPLLLPCTAQLAADP...AACLELHP...ETDA
HvGRAS1H_2 ...LELQLIHENLSSFAAELGVVFFQFTVFNIDTLDPAELVAIAG...GPA
HvGRAS4H_13 ...LELQLIHENLSSFAAELGVVFFQFTVFNIDTLDPAELVAIAG...GPA
HvGRAS6H_2 ...LELHLTRDNLTQFAADLGGVFFQFTVFNIDTLDPAELVAIAG...GPA
HvGRAS4H_1 ...ASLAATSARLRDPAQLLGVVFFQFTVFNIDTLDPAELVAIAG...GPA
HvGRAS4H_2 ...ASLAATSARLRDPAQLLGVVFFQFTVFNIDTLDPAELVAIAG...GPA
HvGRAS2H_3 ...LD...LVGRRLSHIAGLCKVPPFRSVA...MAGEEVEEGH...LVVPGALAVNFTL
HvGRAS2H_4 ...LD...LVGNMLKNSSEFNIPLEPTPLS...VYATQVITREML...IRPGEAVAVNFTL
HvGRAS3H_1 ...KN...YTQLRLVEYASMPKVPLEYR...QIS...SKWESISIEEL...NINKGEVLIVNCEL
HvGRAS3H_2 ...KN...NTRLRLTEYANIFQVPLEYR...QIS...SKWESISIEEL...NINKGEVLIVNCEL
HvGRAS1H_1 ...IE...ETGKRLADYANMPKVPLEYR...QIS...SKWESISIEEL...NINKGEVLIVNCEL
HvGRAS3H_3 ...ID...EIGRYLSDYATFKVPPKYGIA...SQPEAVRVED...LIEKDELIVNSMF
HvGRAS5H_3 ...IE...ETGQRLLAEYAKKFGVPPKYGIA...SQPEAVRVED...LIEKDELIVNSMF
HvGRAS5H_1 ...IE...ETGRRLLKYYAHEFNVPFQYRVIAR...AKLESLRKED...LIDPDEVLIIVNSLL
HvGRAS5H_5 ...LE...ATGRRRLSEYAKKFGVPPKYGIA...AKMDAIRVED...LIDKDDLIVNCLF
HvGRAS4H_5 ...AK...ETKRRRLSYCARQFGVPPKYGIA...AKLEAVHAED...LIDPDEVLVVNSMF
HvGRAS4H_6 ...MD...EAGYRLSNYARQFGVPPKYGIA...SKLEDVVRVED...LIDPDEVLVVNSMF
HvGRAS4H_8 ...IE...ETGRRRLSNFARVFGVPPKYGIA...AKRETVKPED...LIDPDEVLVVNSMF
HvGRAS4H_9 ...IE...ETGRRRLSNFARVFGVPPKYGIA...AKRETVKPED...LIDPDEVLVVNSMF
HvGRAS4H_10 ...IE...ETGRRRLSKAREFGVPPKYGIA...AKFETVHAED...LIDPDEVLVVNSMF
HvGRAS4H_7 ...LE...ETGRRLSNCALQFGVPPKYGIA...AKWETIRAKD...LIDPDEVLVVNSMF
HvGRAS4H_12 ...ME...EVGGRLLTDIARELGVPPKYGIA...AQMQTVCAED...LIDPDEVLVVNSMF
HvGRAS4H_4 ...LD...NMQALLCKCAANLKLFFHFDORVIGK...LETLDPSNLEIL...INFGA...VVISCAL
HvGRAS5H_2 ...LA...GMSAVLVREAEAFDIPFQNAVEAR...LEDMPDALRHL...LWRS...GALAVSVAL
HvGRAS1H_3 ...VDEELGLRLANFAKSKGVQLETRNVVNRSTATSTPSPAKLQTVCC...LGS...VLSDF...PAL
HvGRAS2H_1 ...VSEELGARLVNFAKSKGVQLETRNVVNRSTATSTPSPAKLQTVCC...LGS...VLSDF...PAL
HvGRAS2H_5 ...GT...AALLASEARLHVGFQTHPVKLN...INQLLSIEFL...QVTRTEALVIVS...L
HvGRAS3H_4 ......TVGLEIAYAEEDLGMSETRAVDR...TLESVHVD...LIDDAHEAVINS...VL
HvGRAS7H_2 ...QELQDTGARLHVAAGLGLAFSTHVVDR...LEDVRLMML...LVKROER...VAVNCEL
HvGRAS4H_11 ...VO...EAGRRLAFAAGSIGQPPSTGHORLD...SDERFRPAT...VMMKGET...LVAVNCEL
HvGRAS0H_1 ...ITG...FGASADELRDTEAR...ADGCPNLRPEFEGVLSNGPEDD...RRHRI...ITDPAATV
HvGRAS2H_2 ...RVMREIQRMEKFAFLMGVPPKYGIA...GDLAELDLS...LDRREGANT...GIAVN
HvGRAS4H_3 ...AS...ARNLSEFAKSLNLPFYSSLCLH...SDELDHEFS...KNSKGS...VIFSC...DTHS

```

LAVN

PFYRE

```

200      210
HvGRAS6H_1 VNCVFLHRLSGDG.....
HvGRAS7H_3 VNCVFLHKLGGQD.....
HvGRAS1H_2 LAVHLPVG.AAHAA.....
HvGRAS4H_13 IAVKLPVG.SVHAA.....
HvGRAS6H_2 VAVSLPVGCSARAP.....
HvGRAS4H_1 VNFMLXXXXXXXXXX.....
HvGRAS4H_2 VNFMLQLYHLLGDSDE.....
HvGRAS2H_3 ELHHPDETIVSTAN.....
HvGRAS2H_4 QLHHPDESVDVNN.....
HvGRAS3H_1 RMKNLPDEIISINS.....
HvGRAS3H_2 RMKNVGDEIISINS.....
HvGRAS1H_1 RMKNLGHETEAINS.....
HvGRAS3H_3 RFKTLMDESVAES.....
HvGRAS5H_3 RFRNLIDETVAVDS.....
HvGRAS5H_1 QFNKLMDESVVLES.....
HvGRAS5H_5 RMRHLMDETVTES.....
HvGRAS4H_5 HFRTLMDESLTFDME.....
HvGRAS4H_6 HFRTLMDESLTFDMV.....
HvGRAS4H_8 HFRLMDENLGFDT.....
HvGRAS4H_9 HFRLMDESLGFDD.....
HvGRAS4H_10 GFSNLMDESIVMDRQD.....
HvGRAS4H_7 NFNTLMDESIVMDR.....
HvGRAS4H_12 NFRTLMDDEVVVGSP.....
HvGRAS4H_4 QMHRLLAVDDSD.....VTRDGIQLQOMANMARLQKMACPARSPASTLNYFQ
HvGRAS5H_2 QLHRLLAADDTGGRQGGTLPQIITARSSPSSFGELLERELNTRLQLSPDASAFSSMSFQ
HvGRAS1H_3 RDSEALVVNCQSWLRHVAP.....
HvGRAS2H_1 LLHTVPDETAGSVSLTQSV.....
HvGRAS2H_5 QLHRLLADEFAEVAANPHDRKGGK.....
HvGRAS3H_4 ELHCYVKESEKRGALN.....
HvGRAS7H_2 AAHRLRDETEGGALS.....
HvGRAS4H_11 HQAAAATTVRRPTG.....
HvGRAS0H_1 VNLVFPAAQARAGS.....
HvGRAS2H_2 CMNSLRGVVPGGAR.....
HvGRAS4H_3 MPYKLLSKLPALLP.....

```

PFYRE

```

220      230      240
HvGRAS6H_1 .....ELAAFLRWVRSNPAVYIAEREGSASRG.....DDDDDE
HvGRAS7H_3 .....ELAAF LKWKAMAPAVYVAERASGGGI.....DFIOE
HvGRAS1H_2 .....AMPAVLRIVKSLGAKVWVSVDRGC.....DRTE
HvGRAS4H_13 .....AVPSVLHIVRRIGAKLIVSVDRSG.....DHGE
HvGRAS6H_2 .....FLALLQIVKQIGPKIVVANDYGA.....DRAD
HvGRAS4H_1 .....FVRRVLR LAKSDGPAVVT LGEYEVS.....LNRAG
HvGRAS4H_2 .....FVRRVLR LAKSDGPAVVT LGEYEVS.....LNRAG
HvGRAS2H_3 .....HRDRILRIVKQDRPKVLT LVEQESNTN.....TAP
HvGRAS2H_4 .....PRDGLRMVKGSLSPKVT LVEQESHTN.....TTP
HvGRAS3H_1 .....ARNRVLNTIRMMKPEVFIHGVTNGSYN.....TPF
HvGRAS3H_2 .....ARNRVLNTIRMMKPEVFIHGVTNGSYN.....TPF
HvGRAS1H_1 .....TRDKVLKTRMRMNPKIFISGTVNGLRS.....SPF
HvGRAS3H_3 .....PRNMVLNTIRKMNPEVFIHGVTNGSYN.....APF
HvGRAS5H_3 .....PRNRVLNTIRQVNPATFIHGIVNGSYS.....VPF
HvGRAS5H_1 .....PRDVVLRNIRKMRPHFIHAIVNGSFS.....APF
HvGRAS5H_5 .....PRMTVLNTIRRMNPEFIHAIVNGTFN.....APF
HvGRAS4H_5 .....NPRDLVNTIRKMKPSMLIHAAVNGAYS.....SAF
HvGRAS4H_6 .....SPRDMVLNNIRKMPVFIHQSLVNGPYS.....AAF
HvGRAS4H_8 .....PSPRDVLNNIRKMRPFIHQIMNGSYG.....ATY
HvGRAS4H_9 .....PSPRDVLNNIRKMRPFIHQIMNGSYG.....ATS
HvGRAS4H_10 .....IPSPRDMVLSNIRKMRPFIHQCVVNGTYG.....APF
HvGRAS4H_7 .....PSPRDVLNIRKMRPFIHQGVVNGSS.....GPF
HvGRAS4H_12 .....NPRDVVLSNVCRMKPDVFIQSTVNGSY.....GTF
HvGRAS4H_4 .....TPS.....PQRQIPRLLVSELSAIRALEPKIIVMMEQDADHN.....TPL
HvGRAS5H_2 .....SPIGHHFPAGAGGQRPKIGSELSAVKALSPKIMVVTQEANHN.....GAA
HvGRAS1H_3 .....GSRDGFMDAVRALDPCLVITVDEADLD.....SPS
HvGRAS2H_1 .....SLRTMLLKSIRTHDPTIVVVAEEDAFT.....AGD
HvGRAS2H_5 .....KQVHATMTRADALLRDLAQLSPKIMVVTQEADHN.....GAE
HvGRAS3H_4 .....SVLQTIKLSPKAFV LVEQDAGHN.....GPF
HvGRAS7H_2 .....DFLGLVRSFGASVLLLGEHEAAGLN.....GGR
HvGRAS4H_11 .....SVASFLTGMAISGAKVVT VVEEAGEABKNEEASD...AAAGG
HvGRAS0H_1 .....WDFRSALAHIVSLSPLVFLIDKVDDEGAGRCG.....GSAS
HvGRAS2H_2 .....RRGVFAASLRRLDPRIVTVVEEADLVATDPDASDEGCDEAA
HvGRAS4H_3 .....ACVELLQALMVTISELIVRIGKEYSLC.....KAS

```

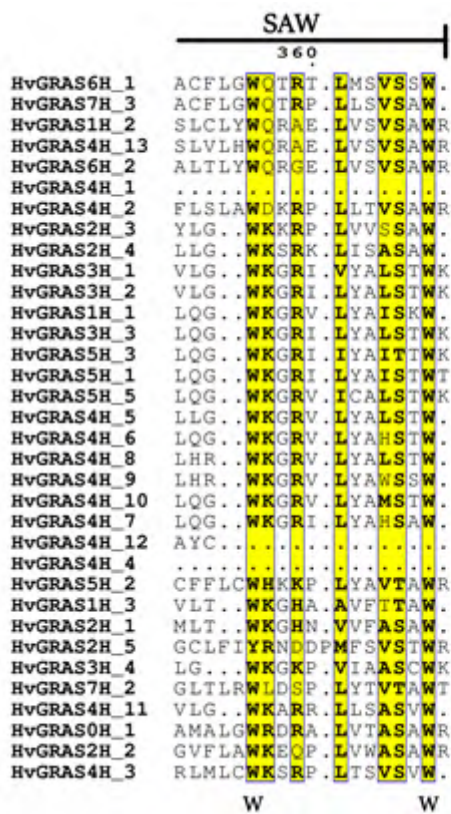



图 3 大麦 GRAS 结构域多序列比对

Figure 3 Multiple sequence alignment of GRAS domain in barley

使用 MEME 软件对大麦 GRAS 蛋白保守域进行分析,共发现 10 个保守基序(motif) (表 2)。在 NCBI-CDD 工具中检索所有 motif, 其中 motif 1~ motif 9 属于 GRAS 结构域的一部分, motif 10 功能未知。motif 2 中包含 GRAS 结构域中的 VHIID 片段, motif 3 包含 PFYRE 片段, motif 4 包含 SAW 片段。

表 2 HvGRAS 家族蛋白序列中的保守基序

Table 2 Conserved amino acid motifs in barley GRASs' protein sequences

保守基序	E 值	长度	氨基酸残基序列
Conserved motifs	E-value	Length	Sequence of amino acid
motif 1	2.0E-450	41	FGREILNVIACEGADRVERPETYKQWQVRNQRAGFRQLPLS
motif 2	9.5E-416	29	VIVDFGIAYGFQWPSQLQALAKRPGGPP
motif 3	3.4E-329	29	YNAPFFLTRFREALFYYSALFDALDATVP

motif 4	1E-192	14	LGWKGRPLYAV[S]A[W]
motif 5	6.2E-234	21	CPYLRF AHFTANQAILEAVAG
motif 6	2.9E-219	24	LEETGRRLSNFARELGVPF EFRAI
motif 7	4.5E-189	21	GDPLQRLAHYFAEGLEARLAG
motif 8	1.9E-225	29	VRVEDLNIDPDEVLA VNSLLQLHRLMDES
motif 9	1.6E-183	21	DSVLR TIRKLSPKVFVVGEVE
motif 10	4.4E-136	29	PDVVKLARDKVKKIYHKDFVIDEDQRWLL

在系统进化树(图 2)中, 根据已知的 AtGRAS 和 OsGRAS 蛋白可将 HvGRAS 蛋白划分为 10 个亚族。对每个亚族中 HvGRAS 蛋白进行保守域结构分析(图 4)。结果表明, 所有的 HvGRAS 蛋白都含有 motif 2、motif 3、motif 5、motif 6 和 motif 9。除 HvGRAS4H_1 外, 其余所有的 HvGRAS 蛋白序列 N 端都为 motif 7。此外, motif 10 为 LISCL 亚族蛋白特有保守域。总体来看, 不同亚族 HvGRAS 间的保守域结构组成差异较大, 同一亚族内各成员拥有基本一致的保守基序, 这与系统进化树分析结果一致。但是, 个别亚族内不同 HvGRAS 的蛋白序列结构也不完全一致, 如 HAM 亚族中的 *HvGRAS1H_2* 缺少 motif 3, 这种结构上的差异提示同一亚族中的 HvGRAS 蛋白也可能存在不同的功能分化。

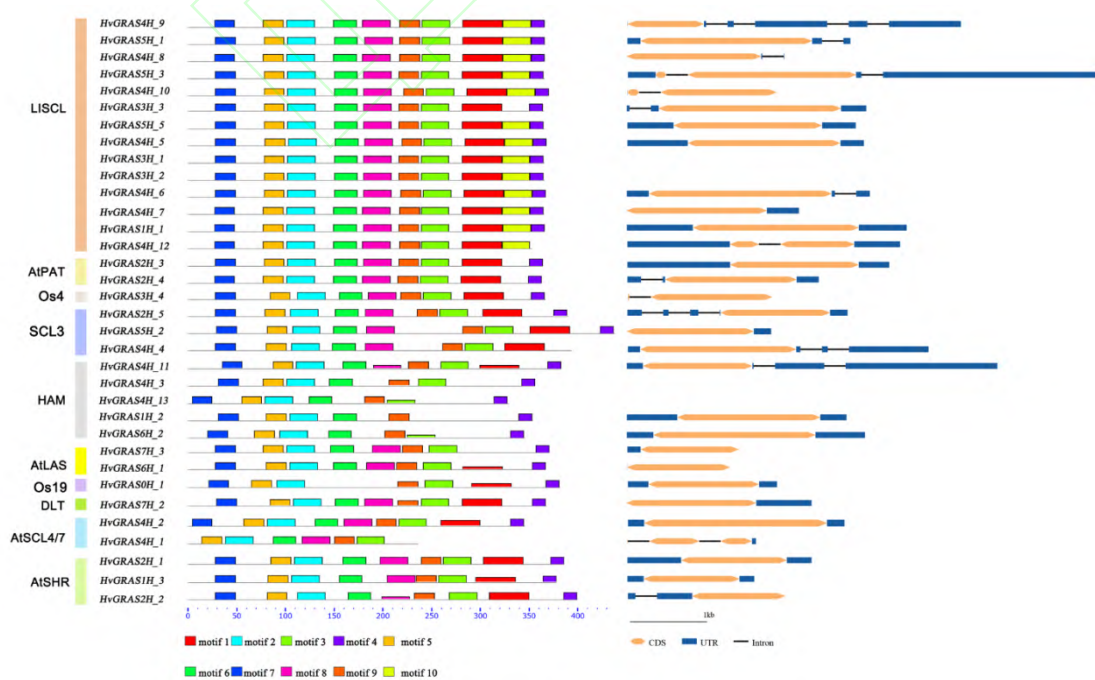


图 4 *HvGRAS* 基因家族保守基序分布和基因结构

Figure 4 Distribution of the conserved motifs and structure of genes in the *HvGRAS* gene family

1.4 *HvGRAS* 基因的时空表达

从大麦 BARLEX 数据库下载了 34 个 *HvGRAS* 基因的表达丰度数据(时丽洁等, 2019), 该数据包含了胚、幼根、幼苗、发育早期花序、发育期花序、节间、发育前期籽粒、灌浆期籽粒、外稃、浆片、内稃、穗轴、灌浆期根部和衰老叶片等不同发育时期共 14 个器官或组织中的 *HvGRAS* 基因表达情况(表 3)。

表 3 取样组织及时期

Table 3 sampling organization and period

取样组织及时期	简写
Sampling organization and period	Abbreviation
胚 4-day embryos	EMB
幼根(苗高 10cm 时期) Roots from seedlings (10 cm shoot stage)	ROO1
叶(苗高 10cm 时期) Shoots from seedlings (10 cm shoot stage)	LEA
发育早期花序(5 mm) Young developing inflorescences (5mm)	INF1
发育期花序(1-1.5 cm) Developing inflorescences (1-1.5 cm)	INF2
第三节间(开花后 42 天) Developing tillers, 3rd internode (42 DAP)	NOD
发育前期籽粒(开花后 5 天) Developing grain (5 DAP)	CAR5
灌浆期籽粒(开花后 15 天) Developing grain (15 DAP)	CAR15
外稃(开花后 42 天) Inflorescences, lemma (42 DAP)	LEM

浆片(开花后 42 天) Inflorescences, lodicule (42 DAP)	LOD
内稃(开花后 42 天) Dissected inflorescences, palea (42 DAP)	PAL
穗轴(开花后 35 天) Inflorescences, rachis (35 DAP)	RAC
根(开花后 28 天) Roots (28 DAP)	ROO2
衰老叶片(开花后 56 天) Senescing leaves (56 DAP)	SEN

利用 pheatmap R 包绘制的基因表达热图表明, 在不同时期或不同组织中, 不同 *HvGRAS* 基因的表达存在很大差异(图 5)。其中, *HvGRAS2H_3*、*HvGRAS2H_4* 和 *HvGRAS4H_12* 除了在早期发育花序未见有明显表达外, 在其余时期尤其是成熟期穗部器官中和根中高表达, 说明上述三个基因同籽粒成熟和植株衰老密切相关。所有 34 个 *HvGRAS* 基因在早期发育花序中的表达量很低甚至接近于 0, 说明 *HvGRAS* 家族基因并未参与到花序的早期发育中。同时, *HvGRAS2H_5*、*HvGRAS3H_1* 和 *HvGRAS3H_4* 三个基因在不同时期不同器官中的表达量均很低, 各类生理过程的参与度较低。

对同一组织不同时期表达丰度进行比较, 发育初期的根(ROO1)和发育后期的根(ROO2) *HvGRAS* 基因存在差异表达, 差异表达基因有 *HvGRAS2H_3*、*HvGRAS2H_4*、*HvGRAS5H_1*。在 ROO1 和 ROO2 时期 *HvGRAS2H_3* 的表达量分别为 13.03 和 87.91; *HvGRAS2H_4* 的表达量分别为 14.96 和 136.1; *HvGRAS5H_1* 的表达量分别为 31.11 和 9.74。其中 *HvGRAS2H_3*、*HvGRAS2H_4* 的差异表达最为明显, 二者均属于 AtPAT 亚家族且分别与 *OsCIGR2*、*OsCIGR1* 相似度极高(图 2)。Day 等人发现 *OsCIGR2*、*OsCIGR1* 通过磷酸化/去磷酸化参与了 α -淀粉酶活性的调控且受到外源 GA 的诱导, 是 GA 信号的应答因子(Day et al., 2003), 因此推测 *HvGRAS2H_3*、*HvGRAS2H_4* 可能具有相似功能, 也是 GA 信号的应答因子。

对发育早期的花序(INF1)和发育期的花序(INF2)的基因表达量进行比较, INF1 时期除 *HvGRAS1H_1* 微量表达(0.71)外其他基因均未表达。INF2 时期表达量较高的基因有 *HvGRAS2H_3*(13.95)、*HvGRAS2H_4*(17.36)、*HvGRAS5H_1*(27.36)和 *HvGRAS6H_2*(26.68), 但这些基因的表达量相较于 ROO1、

ROO2 中高表达基因的表达量还是较低，并不能说明在相应的时期内它们起到了调控作用。

对开花后 5 天的籽粒(CAR5)和开花后 15 天的籽粒(CAR15)的基因表达量进行比较， *HvGRAS2H_3* 和 *HvGRAS2H_4* 同样地在两个时期表现了明显的差异表达。结合二者在其它组织(除 INF1)中也存在高表达的情形，猜测二者在大麦的整个发育进程中都参与了 GA 信号传导。

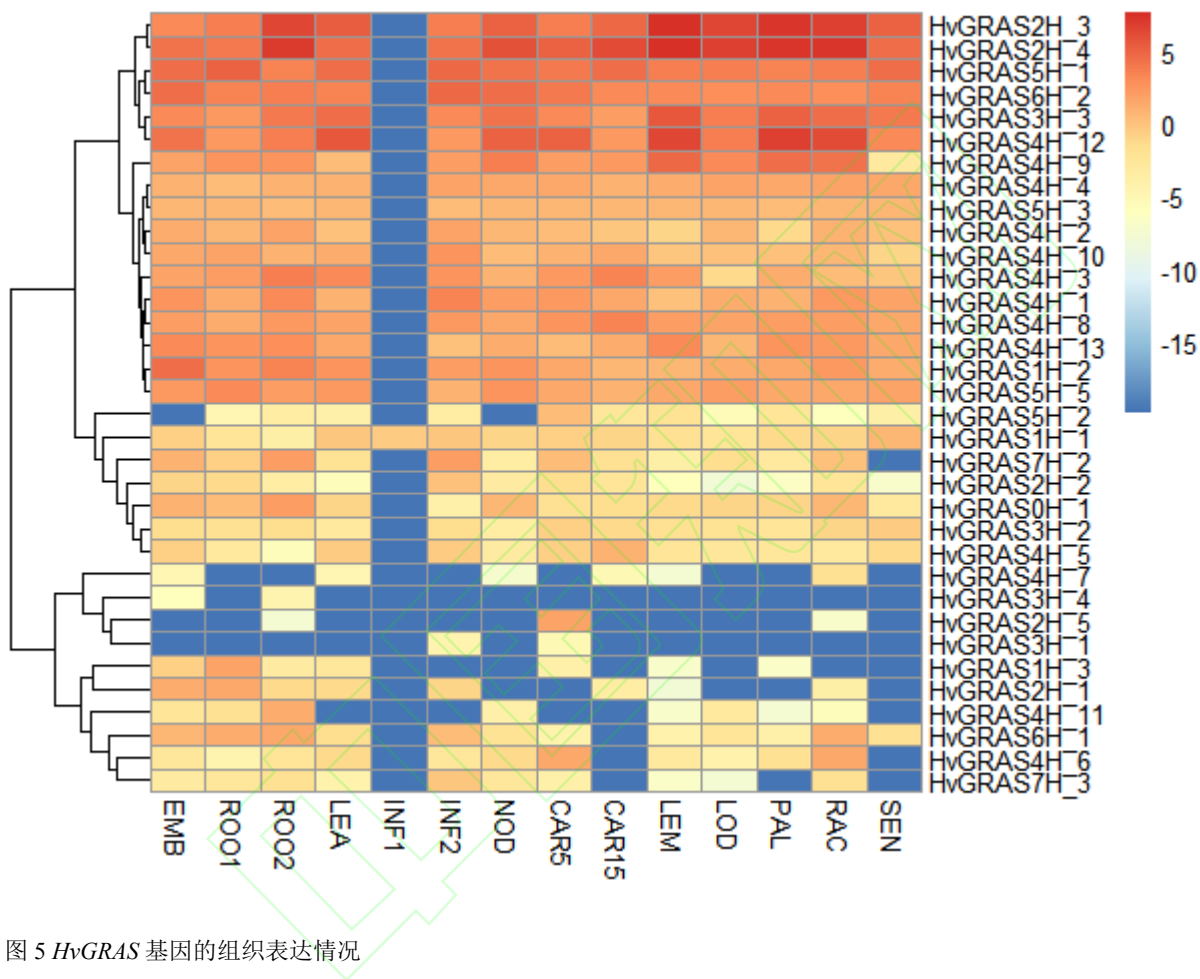


图 5 *HvGRAS* 基因的组织表达情况

Figure 5 Transcriptional abundance of *HvGRAS* genes in different tissues

2 讨论

本文使用最新发布的大麦基因组数据对大麦 GRAS 蛋白家族进行鉴定和保守元件分析。共挖掘出 34 个含 GRAS 完整结构域的 *HvGRAS* 蛋白(图 4)。以往研究表明，GRAS 蛋白家族在单双子叶植物中广泛存在，且每种植物中所含基因数目与基因组大小无关(张焕欣等, 2017)。与拟南芥和水稻中的 GRAS 蛋白进行系统发育树构建，可将大麦 GRAS 蛋白家族进一步划分为 10 个亚族，系统进化树聚类结果与前人在拟南

芥和水稻中的聚类结果基本一致(图 2)。一些 HvGRAS 蛋白和水稻、拟南芥 GRAS 蛋白具有很高的同源性，这种进化的稳定性对维持蛋白功能的稳定具有重要意义。同时，大麦 GRAS 蛋白在 AtSCR 和 DELLA 亚族中没有分布，其功能的互补性有待进一步研究。另外，多重序列比对结果显示，大麦 GRAS 蛋白家族中 VHIID、PFYRE 和 SAW 基序相较于 LR I 和 LR II 基序更为保守(图 3)。

一般而言，处于系统发育树同一亚族中的不同成员往往具有相似的功能(牛义岭等，2016)。HvGRAS2H_1、HvGRAS1H_3 同属 AtSHR 亚族(图 2)，热图中二者具有类似的表达模式且都在根中高表达(图 5)。AtSHR 蛋白是一个调控根发育的 GRAS 蛋白，相较于 AtSCR 仅负责根组织细胞的分裂而言，AtSHR 还负责调控内皮层细胞的分化。HvGRAS2H_1 和 HvGRAS1H_3 的组织表达模式分析也与前人研究结果一致。

同属 AtSHR 亚族的 HvGRAS2H_2 同 HvGRAS2H_1 和 HvGRAS1H_3 存在表达模式的差异，主要在发育花序中表达(图 5)。推测即使是同一亚族的 HvGRAS 基因也会存在不同的功能分化。HvGRAS2H_3、HvGRAS2H_4 表达模式聚类结果和系统聚类一致均属于 AtPAT 亚家族，在除发育早期花序(INF1)外所有组织中均有表达，外稃 (LEM)、浆片 (LOD)、内稃 (PAL)、穗轴 (RAC)等组织中高表达(图 5)。前人研究表明 PATI 亚家族中的成员在光信号转导中发挥重要作用，且 PATI 在大部分组织中均有表达 (牛义岭等，2016)，此结果与本研究热图所绘结果一致(图 5)。

HvGRAS5H_1、HvGRAS3H_3 和 HvGRAS4H_12 属 LISCL 亚族(图 2)，热图中表达模式也较为相近(图 5)。LISCL 亚族的成员研究较少，在百合花粉母细胞中，减数分裂相关启动子的活性通过 LISCL 蛋白共表达而特异性地增强 (Morohashi et al., 2003)。我们注意到 LISCL 亚族中 HvGRAS 分布最多，不同 HvGRAS 蛋白也存在功能分化，如 HvGRAS3H_1 和 HvGRAS3H_2 蛋白在系统发育树上的分支节点自展值达到 100，但其表达模式却具有较大差异(图 2)。HvGRAS3H_1 在除了发育早期花序(INF1)样本外的各样本中均有表达，而 HvGRAS3H_2 在除了发育期花序(INF2)和发育前期籽粒(CAR5)样本外的其他样本中表达量均为 0(图 5)。同样地，HAM 亚族中的 HvGRAS1H_2 和 HvGRAS4H_13，AtLAS 亚族中的 HvGRAS7H_3 和 HvGRAS6H_1 均存在明显的表达模式分化(图 5)；Os4、Os19 和 DLT 亚族中均只有一个 HvGRAS 基因——HvGRAS0H_1、

HvGRAS3H_4 和 *HvGRAS7H_2*(图 2), 三者热图中也未与其它基因聚类在一起, 说明其序列与表达模式的一致性。HAM 亚族已被证明参与了调控分生组织生长过程 (牛义岭等, 2016), 而 *HvGRAS7H_2* 蛋白是 *OsGRAS29* 的直系同源蛋白, *OsGRAS29(DLT/GS6)* 调控 BR 信号转导, 突变体存在碱基缺失的结构变异, 导致粒重粒宽显著增加, 内外稃和表皮的厚壁细胞增多(Sun et al., 2013)。

本研究利用生物信息学方法在全基因组水平上分析了大麦 GRAS 基因家族, 并对其基因结构、染色体分布、蛋白保守结构域、系统进化关系及表达谱进行了分析, 至于大麦 GRAS 基因对各个生理过程的调控机制还有待进一步研究。

3 材料与方法

3.1 大麦全基因组 GRAS 基因的鉴定与染色体定位

为了鉴定大麦 GRAS 蛋白质, 从大麦基因组数据库(http://webblast.ipk-gatersleben.de/barley_ibsc/)中下载了高可信度(High-Confidence)的基因序列及含有注释信息的蛋白序列, 以“GRAS family”作为关键词搜索大麦基因组中被注释为 GRAS 转录因子的蛋白序列, 然后用 NCBI-CDD 工具(<https://www.ncbi.nlm.nih.gov/cdd/?term=>)和 SMART 软件(<http://smart.embl-heidelberg.de/>)在线检测所筛选的蛋白序列, 去除 GRAS 结构域不完整的序列。同时在大麦基因组数据库中搜索筛选后的 GRAS 基因的基因组定位信息, 以 *HvGRAS* 命名大麦 GRAS 基因, 使用 MapChart 软件绘制 *HvGRAS* 基因的染色体图谱。使用 ExPaSy 软件的 pI/Mw(http://web.expasy.org/compute_pi/)工具在线预测 *HvGRAS* 蛋白质的分子量和等电点信息。

3.2 构建 *HvGRAS* 基因的系统发育树

使用 NCBI-CDD 工具获得的 34 个大麦 GRAS 蛋白(*HvGRAS*)、60 个水稻 GRAS 蛋白(*OsGRAS*)和 34 个拟南芥 GRAS 蛋白(*AtGRAS*)蛋白中完整的 GRAS 结构域序列, 用 ClustalW 对拟南芥、水稻和大麦 GRAS 结构域序列进行多重序列比对, 将比对结果在 MEGA7.0 中采用邻接法(Neighbor-Joining)构建系统发育树, 自展值(Bootstrap)设为 1000。参考 *OsGRAS* 和 *AtGRAS* 蛋白的分类信息 (Liu and Widmer, 2014)对系统发

育树进行聚类分析。

3.3 *HvGRAS* 基因结构和蛋白保守结构域分析

从 Ensembl Plants 数据库下载最新的大麦基因组的 GFF3 描述文件，从中找到每个 *HvGRAS* 基因的注释信息。采用 GSDS 2.0 软件(<http://gsds.cbi.pku.edu.cn/>)在线绘制基因结构图。利用保守结构域在线预测软件 MEME (<http://eme.nbcrc.net/meme/tools/meme>) (Bailey et al., 2009)预测蛋白的结构域，结合 NCBI-CDD 工具进行基因结构分析。

3.4 *HvGRAS* 基因表达热图绘制

从 BARLEX 数据库(<https://apex.ipk-gatersleben.de/apex/f?p=284>)下载的不同组织或器官中 *HvGRAS* 基因的时空表达数据，采用 pheatmap R 软件包绘制基因表达热图。

作者贡献

孔豆豆和毛成志是本研究的实验设计和实验研究的执行人；孔豆豆完成数据分析和论文初稿的写作；王蕾、王寒冬、张怀刚参与实验设计；沈裕虎、徐金青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

本研究由青海省科技成果转化专项(2019-NK-17)、青海省创新平台建设专项项目(2017-ZJ-Y14)和中科院科技服务网络(STS)计划区域重点项目“青海高寒农牧交错区饲草品种培育与推广”共同资助。

参考文献

- Bailey T.L., Boden M., Buske F.A., Frith M., Grant C.E., Clementi L., Ren J., Li W.W., and Noble W.S., 2009, MEME SUITE: tools for motif discovery and searching, *Nucleic Acids Res.*, 37: 202-208
- Cordelia B., 2004, The role of GRAS proteins in plant signal transduction and development, *Planta*, 218(5): 683-692
- Day R.B., Shibuya N., and Minami E., 2003, Identification and characterization of two new members of the GRAS gene family in rice responsive to N-acetylchitooligosaccharide elicitor, *Biochimica Et Biophysica Acta-Gene Structure and Expression*,

1625(3): 261-268

- Di L.L., Wysockadiller J., Malamy J.E., Pysh L., Helariutta Y., Freshour G., Hahn M.G., Feldmann K.A., and Benfey P.N., 1996, The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root, *Cell*, 86(3): 423
- Dill A., Jung H.S., and Sun T.P., 2001, The DELLA motif is essential for gibberellin-induced degradation of RGA, *Proc. Natl. Acad. Sci. USA*, 98(24): 14162-14167
- Heckmann A.B., Fabien L., Hiroki M., Perry J.A., Sue B., Martin P., Wang T.L., and Allan D., 2006, Lotus japonicus nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume, *Plant Physiol.*, 142(4): 1739-1750
- Huang W., Xian Z.Q., Kang X., Tang N., and Li Z.G., 2015, Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato, *BMC Plant Biol.*, 15(1): 209
- Li Y.F., Yang W.L., Gu J.J., Zhang A.M., and Zhan K.H., 2019, Genome-wide identification and characterization of the GRAS gene family in bread wheat (*Triticum aestivum* L.), *Mailei Zuowu Xuebao (Journal of Triticeae Crops)*, (5): 1-11 (李亚飞, 阳文龙, 顾晶晶, 张爱民, 詹克慧, 2019, 小麦 GRAS 基因家族的全基因组鉴定与分析, *麦类作物学报*, (5): 1-11)
- Liu X., and Widmer A., 2014, Genome-wide comparative analysis of the GRAS gene family in populus, Arabidopsis and rice, *Plant Molecular Biology Reporter*, 32(6): 1129-1145
- Mascher M., Gundlach H., Himmelbach A., Beier S., Twardziok S.O., Wicker T., Radchuk V., Dockter C., Hedley P.E., Russell J., Bayer M., Ramsay L., Liu H., Haberer G., Zhang X.Q., Zhang Q., Barrero R.A., Li L., Taudien S., Groth M., Felder M., Hastie A., Šimková H., Staňková H., Vrána J., Chan S., Muñoz-Amatriaín M., Ounit R., Wanamaker S., Bolser D., Colmsee C., Schmutzer T., Aliyeva-Schnorr L., Grasso S., Tanskanen J., Chailyan A., Sampath D., Heavens D., Clissold L., Cao S., Chapman B., Dai F., Han Y., Li H., Li X., Lin C., Mccooke J.K., Tan C., Wang P., Wang S., Yin S., Zhou G., Poland J.A., Bellgard M.I., Borisjuk L., Houben A., Doležel J., Ayling S., Lonardi S., Kersey P., Langridge P., Muehlbauer G.J., Clark M.D., Caccamo M.,

- Schulman A.H., Mayer K.F.X., Platzer M., Close T.J., Scholz U., Hansson M., Zhang G., Braumann I., Spannagl M., Li C., Waugh R., and Stein N., 2017, A chromosome conformation capture ordered sequence of the barley genome, *Nature*, 544: 427
- Mayer K.F.X., Robbie W., Brown J.W.S., Alan S., Peter L., Matthias P., Fincher G.B., Muehlbauer G.J., Kazuhiro S., and Close T.J., 2012, A physical, genetic and functional sequence assembly of the barley genome, *Nature*, 491(7426): 711-6
- Morohashi K., Minami M., Takase H., Hotta Y., and Hiratsuka K., 2003, Isolation and characterization of a novel GRAS gene that regulates meiosis-associated gene expression, *Journal of Biological Chemistry*, 278(23): 20865-20873
- Niu Y.L., Jiang X.M., Xu X.Y., and Li J.F., 2016, Research progress of transcription factors GRAS proteins in plant, *Jiyinzuxue Yu Yingyong Shengwuxue (Genomics and Applied Biology)*, 35(9): 2519-2524 (牛义岭, 姜秀明, 许向阳, 李景富. 2016, 植物转录因子 GRAS 蛋白的研究进展, *基因组学与应用生物学*, 35(9): 2519-2524)
- Peng J., Carol P., Richards D.E., King K.E., Cowling R.J., Murphy G.P., and Harberd N.P., 1997, The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses, *Genes Development*, 11(23): 3194
- Pysh L.D., Wysocka-Diller J.W., Camilleri C., Bouchez D., and Benfey P.N., 2010, The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes, *Plant Journal for Cell & Molecular Biology*, 18(1): 111-119
- Shi L.J., Jiang C.C., Wang F.M., Yang P., and Feng Z.Y., 2019, Genome-wide characterization and transcriptional analysis of the protein disulfide isomerase (-like) genes in barley (*Hordeum vulgare*), *Zuowu Xuebao (Acta Agronomica Sinica)*, 45(9):1365-1374 (时丽洁, 蒋枫隽, 王方梅, 杨平, 冯宗云. 2019, 大麦蛋白质二硫键异构酶基因家族的鉴定与表达分析, *作物学报*, 45(9):1365-1374)
- Shi R., Cao Y.B., Chen W.R., and Guo W.D., 2011, On cDNA cloning and expression analysis of GRAS gene in fingered citron, *Zhejiang Shifan Daxue Xuebao (Journal of Zhejiang Normal University (Natural Sciences))*, 34(4): 446-451 (石瑞, 曹诣斌, 陈文荣, 郭卫东, 2011, 佛手 GRAS 基因的克隆及表达分析, *浙江师范大学学报(自然科学版)*, 34(4): 446-451)
- Silverstone A.L., Ciampaglio C.N., and Sun T., 1998, The Arabidopsis RGA gene encodes a transcriptional regulator repressing the

gibberellin signal transduction pathway, *Plant Cell*, 10(2): 155-169

Solé A., Sánchez C., Vielba J.M., Valladares S., Abarca D., and Díaz-Sala C., 2009, Characterization and expression of a *Pinus radiata* putative ortholog to the *Arabidopsis* SHORT-ROOT gene, *Plant Signaling & Behavior*, 28(8): 1629-39

Sun L.J., Li X.J., Fu Y.C., Zhu Z.F., Tan L.B., Liu F.X., Sun X.Y., and Sun C.Q., 2013, GS6, A member of the GRAS gene family, negatively regulates grain size in rice, *J. Int. Plant Biol.*, 55(10): 938-949

Sun X., Wang C., Fang J.G., Mu Q., and Wang X.C., 2011a, Bioinformatics analysis of GRAS gene family in grapevine, *Jiangxi Nongye Xuebao (Acta Agriculturae Jiangxi)*, 23(7): 1-8 (孙欣, 王晨, 房经贵, 慕茜, 王西成. 2011, 葡萄 GRAS 基因家族生物信息学分析, *江西农业学报*, 23(7): 1-8)

Sun X.L., Xue B., Jones W.T., Rikkerink E., Dunker A.K., and Uversky V.N., 2011b, A functionally required unfoldome from the plant kingdom: intrinsically disordered N-terminal domains of GRAS proteins are involved in molecular recognition during plant development, *Plant Molecular Biology*, 77(3): 205-223

Tong H.N., Liu L.C., Jin Y., Du L., Hai Y.Y., Qian Q., Zhu L.H., and Chu C.C., 2012, DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice, *Plant Cell*, 24(6): 2562-2577

Torres-Galea P., Hirtreiter B., and Bolle C., 2013, Two GRAS proteins, SCARECROW-LIKE21 and PHYTOCHROME A SIGNAL TRANSDUCTION1, function cooperatively in phytochrome A signal transduction, *Plant Physiol.*, 161(1): 291-304

Zhang H.X., Dong C.J., and Shang Q.M., 2017, Genome-wide identification and expression analysis of GRAS gene Family in Pepper, *Yuanyi Xuebao (Acta Horticulturae Sinica)*, (12): 2305-2317 (张焕欣, 董春娟, 尚庆茂, 2017, 辣椒 GRAS 家族全基因组鉴定与表达分析, *园艺学报*, (12): 2305-2317)

Zhou M.B., Yan Z., and Tang D.Q., 2011, Characterization and primary functional analysis of BvCIGR, a Member of the GRAS gene family in *Bambusa ventricosa*, *Botanical Review*, 77(3): 233-241