

谷子抽穗期基因 *SiGI* 的表达与单倍型变异分析

姚琦^{1,2}, 智慧², 孟强², 袁迪², 张彬¹, 刁现民², 贾冠清^{1,2}

(¹山西农业大学农学院, 晋中 030801; ²中国农业科学院作物科学研究所, 北京 100081)

摘要: 谷子抽穗期是决定品种适应性的关键性状, 解析抽穗期关键基因的表达规律及其单倍型变异特点可为提升品种改良效率奠定基础。本研究通过全基因组关联分析, 发现了一个与抽穗期紧密相关的关联信号, 该信号定位于谷子5号染色体的11062649 bp处, 该位点附近存在一个拟南芥抽穗期基因 *AtGI* 的同源基因 *SiGI*。使用 qRT-PCR 技术研究短日照(10 h 光照/14 h 黑暗)条件下 *SiGI* 24 h 节律变化规律, 并对其进行了组织时空表达及亚细胞定位分析。利用 697 份谷子品种, 分析 *SiGI* 编码区和启动子区的遗传多态性和单倍型变异规律, 并对单倍型的表型效应进行鉴定。结果显示, *SiGI* 在光周期响应组织(根、茎、叶等)中高表达, 亚细胞定位于细胞核, 在傍晚表达量上调, 呈现出 24 h 节律性表达模式。*SiGI* 在不同谷子品种中存在丰富的多态性, 其中, 启动子 Hap-6 单倍型较 Hap-3 单倍型品种的相对表达量显著上调了约 1.5 倍 ($P=0.0083$)。在 8 个不同种植环境下, Hap-6 单倍型品种的抽穗期显著提前, 并在 4 个环境中株高显著降低。结果表明 *SiGI* 基因的 Hap-6 单倍型具有早熟且对产量影响较小的优势, 可作为优异单倍型用于分子育种选择。

关键词: 全基因组关联分析; 单倍型分析; 节律性表达; 抽穗期; 谷子

Characterizations of Transcriptional and Haplotypic Variations of *SiGI* in Foxtail Millet

YAO Qi^{1,2}, ZHI Hui², MENG Qiang², YUAN Di², ZHANG Bin¹, DIAO Xianmin², JIA Guanqing^{1,2}

(¹College of Agriculture, Shanxi Agricultural University, Jinzhong 030801; ²Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081)

Abstract: The heading date of foxtail millet is a key trait that determined the adaptability of varieties, and dissection of the transcriptional pattern and haplotype variations of key genes are important for promoting cultivars improvement. In this research, *SiGI* was identified by genome-wide association study (GWAS), which is a key regulator responsible for foxtail millet flowering. The expression profile of *SiGI* was analyzed, as well as subcellular localization of *SiGI* was performed by using protoplasts. The 24-hour rhythm expression pattern analysis of *SiGI* was performed by qRT-PCR under short-day (10 h light/14 h dark) conditions. Genetic and haplotypes diversity of *SiGI* coding and promoter regions were analyzed by using representative 697 millet varieties, and haplotype morphological effects of *SiGI* was also summarized. As a result, a significant correlation signal was identified at 11062649 bp on chromosome 5, which is closely related to heading date. *SiGI* was found near this locus, whose homologue is *AtGI*. *SiGI* was highly expressed in photoperiod-responsive tissues (roots, stems, leaves, etc.), and subcellular localization was located in the nucleus, and the expression level was up-regulated in the evening, showing a 24-hour rhythmic expression pattern. *SiGI* exhibits rich polymorphism in

收稿日期: 2023-11-30 网络出版日期: 2024-01-16

URL: <https://doi.org/10.13430/j.cnki.jpgr.20231130001>

第一作者研究方向为谷子基因多样性鉴定, E-mail: yaoqi5857@163.com

通信作者: 刁现民, 研究方向为谷子遗传育种, E-mail: diaoxianmin@caas.cn

贾冠清, 研究方向为谷子基因资源发掘, E-mail: jiaguanqing@caas.cn

基金项目: 国家重点研发计划(2023YFD1200704, 2023YFD1200700); 国家谷子高粱产业技术体系(CARS06-14.5-A04); 中国农业科学院创新工程(01-ICS-07)

Foundation projects: National Key Research and Development Program of China (2023YFD1200704, 2023YFD1200700); China Agricultural Research System (CARS06-14.5-A04); Technology Innovation Program of the Chinese Academy of Agricultural Sciences (01-ICS-07)

different foxtail millet varieties, and the relative expression of promoter haplotype Hap-6 was significantly upregulation by approximately 1.5-fold compared to Hap-3 ($P=0.0083$). Heading date of Hap-6 containing varieties were significantly earlier than other haplotypes under 8 environments, and the plant height of Hap-6 containing varieties were significantly reduced under 4 environments. *SiGI* haplotype Hap-6 has no obvious impact on yield, and can be used as the main haplotype for molecular breeding selection.

Key words: genome-wide association analysis; haplotype analysis; rhythmic expression; heading date; foxtail millet

抽穗期是作物重要的农艺性状之一^[1],不仅决定了作物由营养生长开始转向生殖生长的时间节点,同时间接反映了不同作物生育期的长短,是决定作物高产稳产的关键因素之一^[2]。植物抽穗期的调控机制长期以来都是生物学家和育种家关注的研究热点。近年来,国内外学者在抽穗期基因克隆方面取得了一些初步的研究进展^[1],并在水稻中定位到了大量与抽穗期相关的遗传位点,并克隆了 *Ghd7*、*Hd3a*、*Ehd1*、*Hd1* 等关键基因^[3-6]。其中, *Hd1* 和 *Hd3a* 分别为拟南芥 *CO* (*CONSTANS*) 和 *FT* (*FLOWERING LOCUS T*) 的同源基因。在拟南芥中的研究发现,由 *GI* (*GIGANTEA*)、*CO* 和 *FT* 基因组成的生物钟调控途径决定了拟南芥在长日照下的开花时间,并且发现在此通路中 *GI* 比 *CO* 和 *FT* 更早地响应光周期^[7],这表明 *GI* 应该是调节植物抽穗期的关键上游基因。

已有研究发现拟南芥 *GIGANTEA* 是一种植物特异性核蛋白^[8],也是光周期开花途径的核心组成部分,参与光周期开花、生物钟控制和光受体信号传导等,其转录水平受生物钟和光信号的调控^[9], *AtGI* 基因突变可改变光周期不敏感材料的开花期和昼夜节律^[10-11],除此之外, *AtGI* 在蔗糖信号传导、淀粉积累以及耐冷的敏感性等多种生理过程中也发挥作用^[8]。在长日照单子叶模式植物二穗短柄草 (*Brachypodium distachyon* (L.) P. Beauv.) 中, *BdGI* 在寒冷条件和光照变化时表达都会发生改变,此外,将 *BdGI* 基因转入拟南芥的一个 *GI* 突变体中,结果有效恢复了拟南芥的晚开花表型,说明 *BdGI* 参与了二穗短柄草对光照变化的适应性反应,这与拟南芥中诱导开花的作用相类似^[12]。水稻中 *OsGI* 也是一个昼夜节律响应基因, *OsGI* 的突变不仅导致水稻在短日照条件下晚抽穗^[13],还增加了突变体中的脯氨酸和蔗糖含量,加速了气孔运动,增加了对聚乙二醇 (PEG, polyethylene glycol) 产生的渗透胁迫的耐受性,表明 *OsGI* 在水稻渗透胁迫响应中具有负调节作用^[14]。此外,已有研究在豌豆和小麦中验证了

GI 蛋白功能存在保守性^[15-16]。在大豆中, *E2* 为拟南芥 *GI* 的同源基因,且和水稻 *OsGI* 有相似功能, *E2* 的功能丧失导致开花期缩短^[17]。

目前,拟南芥及水稻 *GI* 基因的生物钟特征及光周期调控模式已被初步解析^[10,13],但 *GI* 在旱生禾谷类作物中的表达规律和自然变异情况尚不清晰,如何通过生物钟特征拓宽农作物的适种性,挖掘优异单倍型,提高禾谷类作物产量和品质,是当前急需解决的重要应用问题^[18]。谷子 (*Setaria italica* (L.) Beauv.) 是起源于我国的禾谷类粮饲兼用的古老作物,在我国旱作可持续农业发展中发挥了重要作用。我国的谷子种质资源丰富,已经完成了单倍型物理图谱的构建^[19],本研究通过全基因组关联分析定位到与抽穗期紧密相关的位点,进一步筛选得到谷子基因 *SiGI*,并利用时空表达、亚细胞定位、光周期响应模式以及单倍型变异情况与抽穗期、株高、主穗重等表型相关性进行研究,将为谷子的高产育种和广适型分子育种提供理论依据。

1 材料与方法

1.1 试验材料

根据 He 等^[20] 研究数据,选取了山西长治 (2011 年)、山西太原 (2012 年、2013 年)、河南安阳 (2011 年、2016 年)、甘肃定西 (2018 年、2019 年) 和黑龙江齐齐哈尔 (2017 年) 8 个不同环境下谷子的 697 份抽穗期、701 份株高、286 份主穗重和 292 份主穗粒重数据。本研究所用的 4 项农艺性状均根据《谷子种质资源描述规范和数据标准》^[21] 测量,所用株高数据即为主茎长度,主穗重和主穗粒重为单株的主茎穗重和主茎穗粒重。

1.2 全基因组关联分析及生物信息学分析

参照张林林等^[18] 所用的全基因组关联分析方法,对 2012 年山西太原材料的抽穗期数据进行分析处理^[20],采用混合线性 MLM 模型进行数据运算,再利用 R 软件包的 qqman 完成最后的数据分析。所使用的单倍型图谱和遗传结构数据均基于已经发布

的研究成果^[19]。谷子基因组注释文件来源于Phytozome 数据库 (<https://phytozome.jgi.doe.gov/pz/portal.html/>)。

通过Photozome 数据库获取 *SiGI* (*Seita.5G129500*) 的蛋白序列,提交至 NCBI (<https://www.ncbi.nlm.nih.gov/>) 进行 BLAST 比对,获得水稻、高粱、拟南芥、柳枝稷、二穗短柄草等不同物种的同源蛋白序列,整理后导入到 MEGA 6 中进行多序列比对,并利用邻接法 (N-J, neighbor-joining) 构建系统发育进化树, Bootstrap 值设置为 1000。根据 Lescot 等^[22] 的研究方法,以 *SiGI* 上游 2000 bp 作为启动子,通过数据库 Plant Care (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 对该基因进行顺式作用元件分析。

1.3 表达分析及亚细胞定位

参照桑璐曼等^[23] 方法,对 *SiGI* 的时空组织表达特异性进行分析及可视化。谷子 Ci846 是一种光敏感品种,由于其高转化率常被选做试验材料。称取 0.2 g 左右鲜叶片,按照 RNA 提取试剂盒 (TransZol Up, 北京全式金生物) 和反转录试剂盒 (PrimeScriptTMII 1st

Strand cDNA Synthesis Kit, TaKaRa) 说明书的步骤对样品进行 RNA 提取并反转录,以得到实验所需的 cDNA。以谷子 Ci846 成熟叶片的 cDNA 为模板,设计亚细胞定位 N 端载体引物及检测引物 (表 1)。PCR 反应体系为: 12.5 μ L Primer STAR[®] HS DNA Polymerase with GC Buffer (TaKaRa), 1.0 μ L 正向引物, 1.0 μ L 反向引物, 1.0 μ L DNA 模板, 9.5 μ L H₂O。PCR 扩增程序为: 95 $^{\circ}$ C 预变性 3 min; 95 $^{\circ}$ C 变性 15 s, 72 $^{\circ}$ C 退火 15 s, 72 $^{\circ}$ C 延伸 5 s/kb, 35 次循环; 72 $^{\circ}$ C 彻底延伸 5 min。结合张林林等^[18] 和张慧等^[24] 的方法构建 35S :: SiGI: GFP 的亚细胞定位载体以及诱导质粒表达。采用限制性内切酶 *Spe* I 和 *Xba* I 双酶切 PUC-N 端载体, 利用同源重组酶 (Phanta[®] Max Super-Fidelity DNA Polymerase, 南京诺唯赞生物科技股份有限公司) 连接回收得到的基因片段及酶切载体, 构建 35S :: SiGI: GFP 的亚细胞定位载体。将 Ci846 处于避光条件下长至 2~3 叶, 利用幼嫩黄化苗的叶片组织为材料, 在室温条件下提取原生质体并利用 40% PEG 诱导质粒, 随后暗培养 14~20 h, 在激光共聚焦显微镜下观察转化后的谷子原生质体。

表 1 *SiGI* 亚细胞定位载体及 qRT-PCR 相关引物

Table 1 *SiGI* subcellular localization vector and qRT-PCR related primers

引物用途 Primer action	引物名称 Primer name	引物序列 (5'-3') Primer sequence (5'-3')	产物长度 (bp) Product length
GFP-N 端载体 GFP-N terminal vector	580-5G129500F 580-5G129500R	GCCCAGATCAACTAGTATGTCAGCTTCAAATGAGAAGTGG TCGAGACGTCTCTAGAGCAAGGGAGGGGGCAGCC	3512
检测引物 Detection primers	PAN580-Test-F PAN580-Test-R	ATGACGCACAATCCCACTATCC AAGACCGGCAACAGGATTCAAT	4400
qRT-PCR 引物 qRT-PCR primers	5G-2-F 5G-2-R	AAGTGCCGTCTATCACCCAC GGGGGTTATGTGTCCGTTGT	111
内参基因 (Photozome 基因号: <i>Seita.3G037700</i>)	Cullin-F Cullin-R	TATGGGTCATCAACAGCTTGTC GTAGTCCCTCGTGATGAGATCC	112

1.4 表达节律分析

参照张林林等^[18] 的方法, 选取谷子 Ci846 的种子, 点播在小花盆中 (长 8 cm, 宽 8 cm, 高 9.5 cm), 每盆点播 8~12 粒, 并置于短日条件 (10 h 光照/14 h 黑暗) 下的培养间中培养 2 周, 进行苗期 24 h 取样 (7:00 至次日 7:00), 每隔 2 h 取样一次 (取地上部分), 共取样 13 次。提取样本总 RNA 并反转录成

cDNA, 进行 qRT-PCR (设置 3 组生物学重复), 所用引物见表 1。qRT-PCR 反应体系为: 2.0 μ L cDNA, 10.0 μ L 2 \times SYBR qPCR Master Mix (北京聚合美生物科技有限公司), 0.2 μ L 正向引物, 0.2 μ L 反向引物, 7.6 μ L ddH₂O。qRT-PCR 扩增程序为: 95 $^{\circ}$ C 预变性 1 min; 95 $^{\circ}$ C 15 s, 58 $^{\circ}$ C 15 s, 72 $^{\circ}$ C 30 s, 40 次循环; 溶解曲线程序为 95 $^{\circ}$ C 15 s, 60 $^{\circ}$ C 1 min, 95 $^{\circ}$ C

1 s。利用 $2^{-\Delta\Delta Ct}$ 公式^[25] 计算 13 个时间点 *SiGI* 的相对表达量, 并对其进行节律性表达分析。

1.5 单倍型变异分析

本研究在 1173 份谷子和狗尾草材料的重测序数据^[20] 的基础上, 参照 Zhang 等^[26] 的方法, 利用 RStudio 软件 (<https://www.rstudio.com/products/rstudio/download/>) 及 Prism 软件进行单倍型分析和结果可视化。参照桑璐曼等^[23] 的方法, 用 Excel 对所有得到的变异位点进行聚类分型, 在 1173 份材料中去除狗尾草以及没有抽穗期数据的谷子材料, 剩余 697 份谷子资源。结合抽穗期、株高等表型数据对不同单倍型进行分析, 并通过 *t*-检验进行显著性分析, 结果以箱式图的形式呈现; 进行单倍型之间的衍生关系分析, 并生成单倍型 Network 图。根据单倍型分析结果, 挑选出 Hap-3 单倍型材料:

Ci0162、Ci0203、Ci0227、Ci0230、Ci0779 和 Hap-6 单倍型材料: Ci0085、Ci0841、Ci0851、Ci0862、Ci0892。在中国农业科学院作物科学研究所北京市北圃场试验基地的自然条件下, 播种 Hap-3 单倍型材料和 Hap-6 单倍型材料, 取拔节期成熟叶片提取 RNA, 对 *SiGI* 进行 qRT-PCR。

2 结果与分析

2.1 谷子抽穗期基因 *SiGI* 的关联分析定位

通过全基因组关联分析, 定位到谷子第 5 号染色体物理位置 11062649 bp 处的一个与抽穗期紧密相关的位点, 该位点上下游 50 kb 区间内共存在 8 个编码基因(图 1), 分析 8 个候选基因在拟南芥中同源基因的功能注释(表 2), 筛选到与抽穗期密切相关的 *GI* 基因, 在谷子中的同源基因为 *Seita.5G129500*。

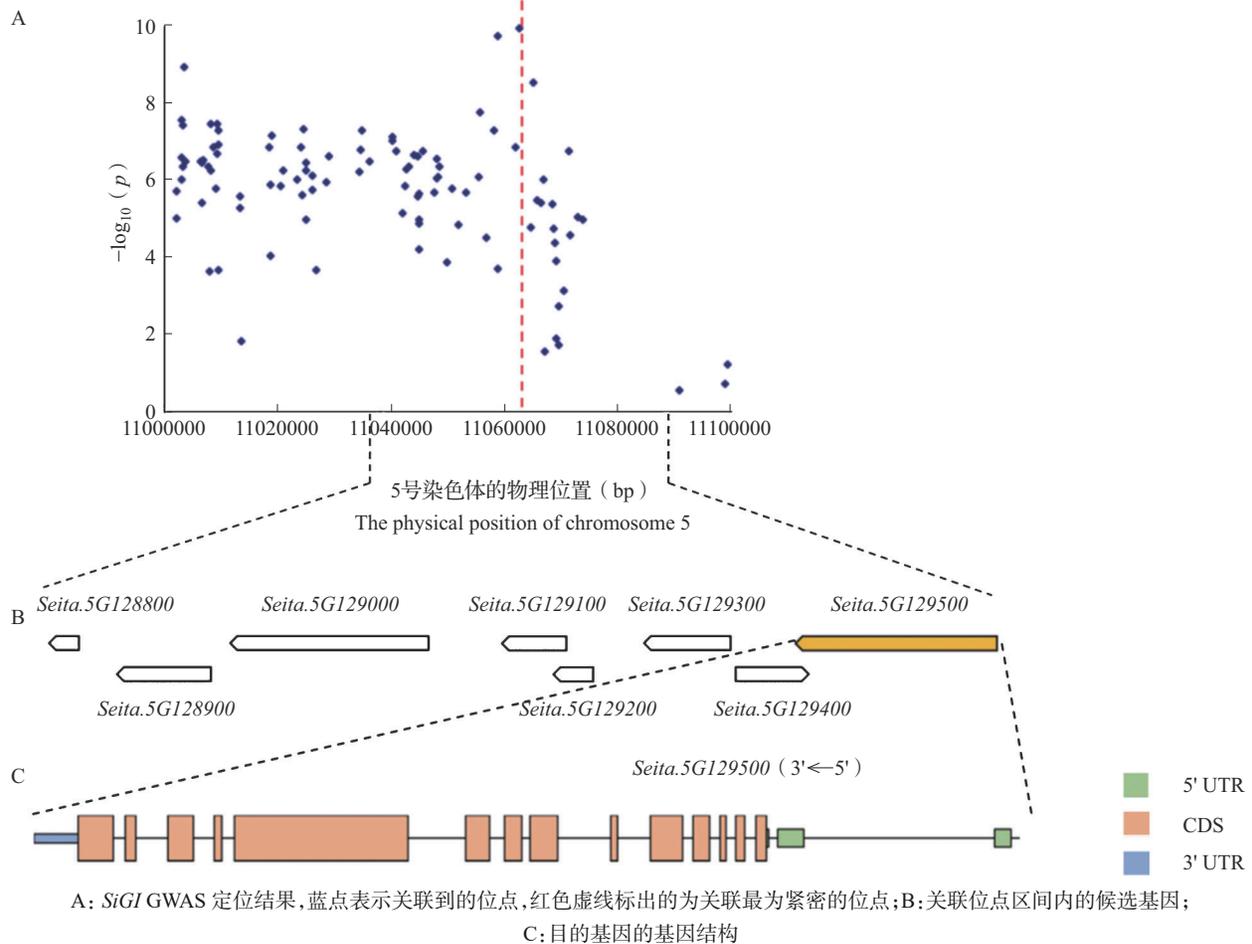


图 1 *SiGI* 的 GWAS 分析

Fig.1 GWAS analysis of *SiGI*

表2 候选基因注释信息

Table 2 Annotation of candidate gene

位点名称 Locus name	拟南芥同源基因 Arabidopsis homologous genes	拟南芥基因功能注释 Arabidopsis annotation
<i>Seita.5G128800</i>	<i>AT4G05497</i>	RNI-like superfamily protein
<i>Seita.5G128900</i>	<i>AT3G22440</i>	FRIGIDA-like protein
<i>Seita.5G129000</i>	<i>AT5G26830</i>	苏氨酸 tRNA 合成酶
<i>Seita.5G129100</i>	无	无
<i>Seita.5G129200</i>	<i>AT2G25735</i>	未知蛋白
<i>Seita.5G129500</i>	<i>AT2G36960</i>	TSL 激酶相互作用蛋白 1
<i>Seita.5G129400</i>	<i>AT5G64780</i>	未鉴定的保守蛋白 UCP009193
<u><i>Seita.5G129500</i></u>	<u><i>AT1G22770 (AtGI)</i></u>	<u>Gigantea protein (GI)</u>

下划线表示 *GI* 同源基因 *SiGI* (*Seita.5G129500*)

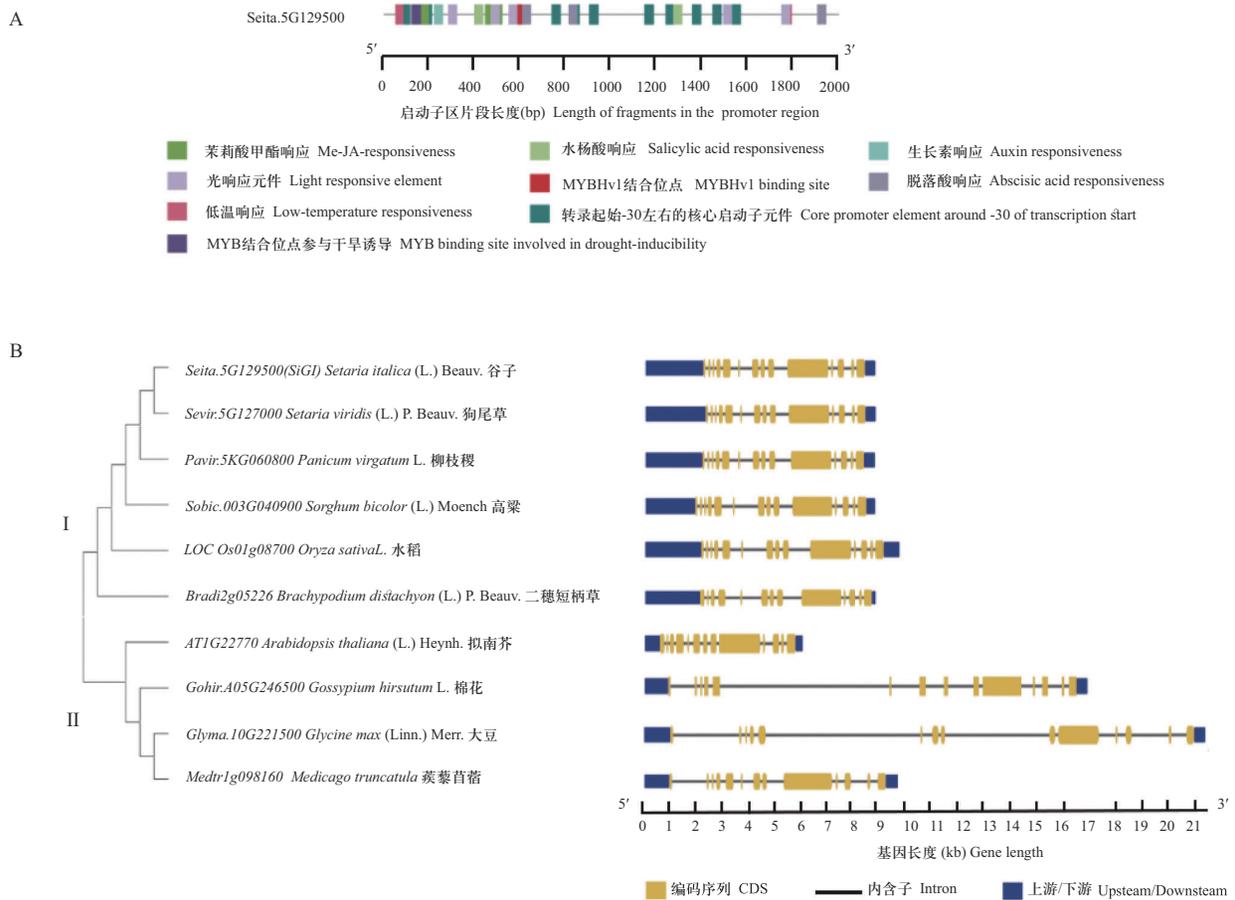
GI homologous gene *SiGI* (*Seita.5G129500*) is marked in underline

2.2 *SiGI* 的基因结构与同源行化

谷子 *Seita.5G129500* (*SiGI*) 基因的启动子顺式

作用元件分析结果显示(图2A),除存在启动子区核心元件外,还存在光响应元件以及低温响应、脱落酸响应等与逆境胁迫相关的元件,表明 *SiGI* 可以对多种非生物胁迫做出响应。

SiGI 基因包含 14 个外显子和 15 个内含子,共编码 1161 个氨基酸。*SiGI* 蛋白 C-末端第 896~1036 位氨基酸(第 14 外显子)存在一个四环素转录抑制(Tetracycline transcriptional repressor)保守结构域。系统发育进化树将水稻、高粱、拟南芥、柳枝稷、二穗短柄草等不同物种中划分为两个主要的分枝:分枝 I 均为单子叶植物(从上至下分别为谷子、狗尾草、柳枝稷、高粱、水稻、二穗短柄草);分枝 II 均为双子叶植物(从上至下分别为拟南芥、棉花、大豆、蒺藜苜蓿)。谷子 *SiGI* 与狗尾草、柳枝稷、高粱、水稻、二穗短柄草具有更近的亲缘关系(图 2B)。通过比较基因结构发现,不同物种中的 *GI* 基因结构较为保守,表明该基因在谷子中的功能保守。



A: *SiGI* 启动子区上游 2000 bp 顺式作用元件分析; B: *SiGI* 同源蛋白系统进化树和基因结构

A: Analysis of 2000 bp *cis*-acting elements upstream of the *SiGI* promoter region;

B: *SiGI* homologous protein phylogenetic tree and gene structure

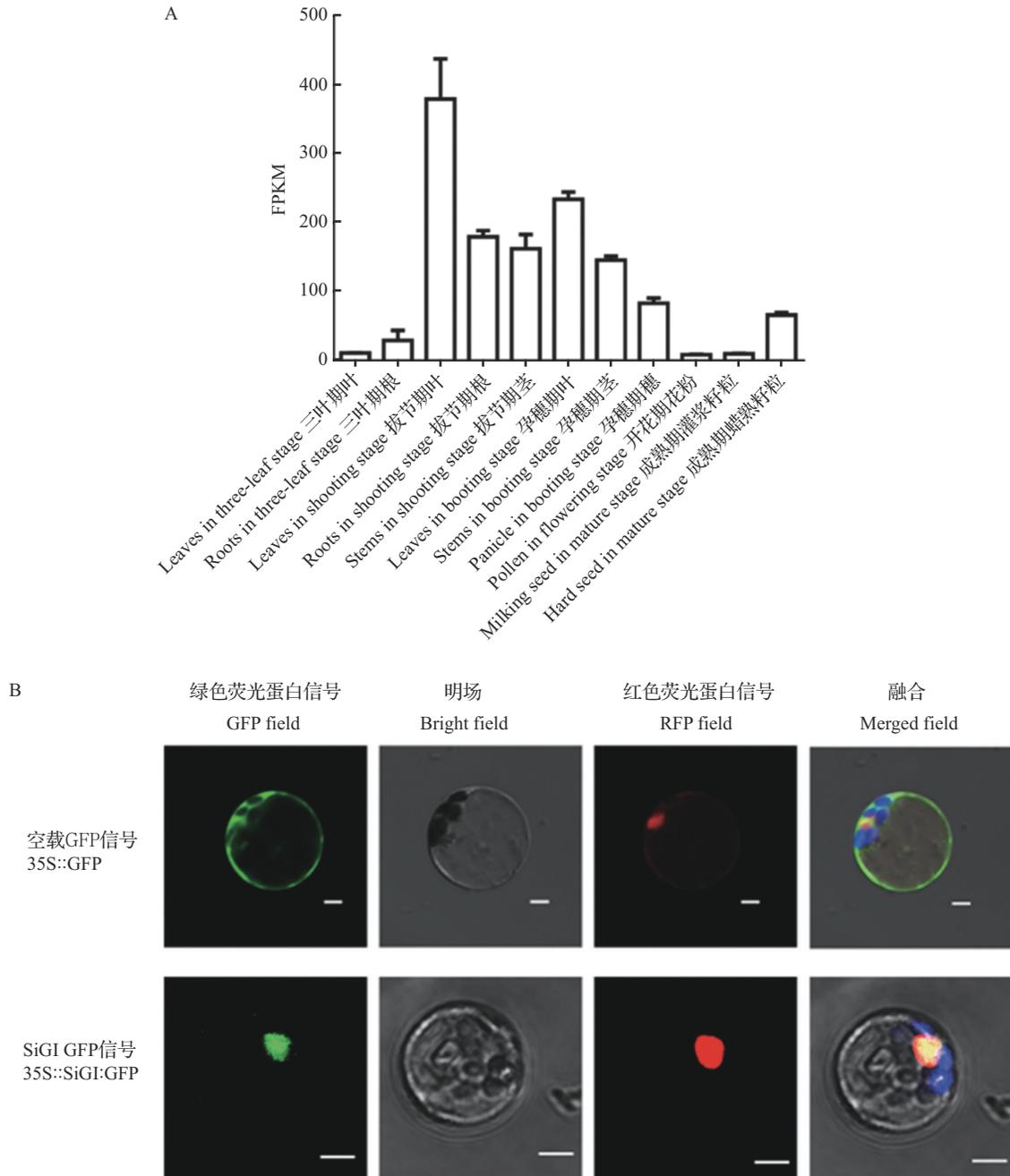
图2 *SiGI* 的启动子区顺式作用元件分析及基因编码区同源蛋白系统进化树

Fig.2 Analysis of *cis*-acting elements in the promoter region of *SiGI* and phylogenetic tree of homologous proteins in the gene coding region

2.3 *SiGI* 的表达规律及亚细胞定位

利用已发表的谷子基因表达谱^[20]对 *SiGI* 进行组织特异性表达分析,结果显示,*SiGI* 在幼苗阶段的地上部分以及成熟期的根、茎、叶、穗等多个组织中都有表达,而且随着谷子生长发育过程,其表达水平呈先上升后下降的趋势。此外,*SiGI* 在拔节期的根、茎、叶

以及孕穗期的叶、茎、幼穗中高表达,相反在三叶期的根和叶、开花期的花粉、成熟期灌浆籽粒以及成熟期蜡熟籽粒中表达量较低,说明该基因参与谷子的全生育期,并在生长发育中后期起重要作用(图 3A)。*SiGI* 的亚细胞定位显示,该基因定位在细胞核中(图 3B),这与拟南芥 *GI* 基因一致。



A: *SiGI* 的组织特异性表达分析; B: *SiGI* 的亚细胞定位, 红色荧光蛋白信号通道显示核定位信号肽的信号, 比例尺 bar=5 μm
A: Tissue-specific expression analysis of *SiGI*; B: Subcellular localization of *SiGI*, the RFP field shows RFP signals of nuclear localization signal peptide, bar=5 μm; GFP: Green fluorescent protein; RFP: Red fluorescent protein

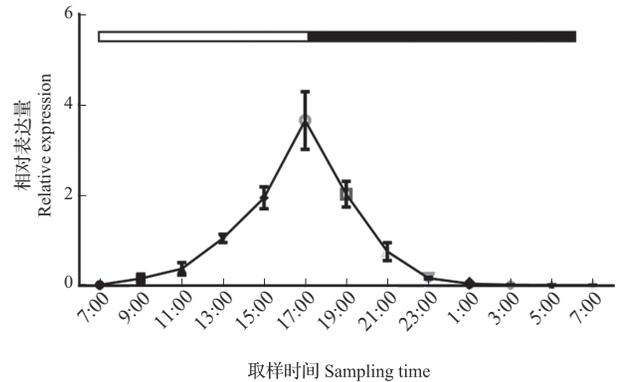
图3 *SiGI* 的组织特异性表达分析及亚细胞定位
Fig.3 Tissue-specific expression analysis and subcellular localization of *SiGI*

2.4 *SiGI* 节律表达模式分析

SiGI 节律表达模式分析该基因呈现周期性振荡特征(图4),整体呈现先升高后降低的趋势。可以看出,光照初期4 h(7:00-11:00)内 *SiGI* 的表达量较低,而从 13:00 开始表达量急剧升高,15:00 到 17:00 的增长速率最快,且在 17:00 时表达量达到峰值(光照 10 h 后);随后 6 h 的基因表达量急速下降(17:00-23:00),次日凌晨 1:00 至 7:00 该基因几乎不表达。这种变化可以通过观察 24 h 趋势来发现,以更好地理解其节律表达模式。

2.5 *SiGI* 核苷酸序列遗传多态性分析

对 697 份谷子品种的 *SiGI* 基因及其启动子区变异位点进行分析,结果显示共 70 个变异位点,包括 64 个 SNP 位点和 6 个 Indel 位点,分为 16 种单倍型,包括 5 种主要单倍型 Hap-1、Hap-2、Hap-3、Hap-4 和 Hap-6(图 5)。在所有材料的编码区变异中,只存在一个非同义 SNP 位点,即 460A/C(Ser154Arg),在 5 种主要单倍型中只有 Hap-6 与其他单倍型在该位点存在差异。相对于基因编码区,启动子区存在



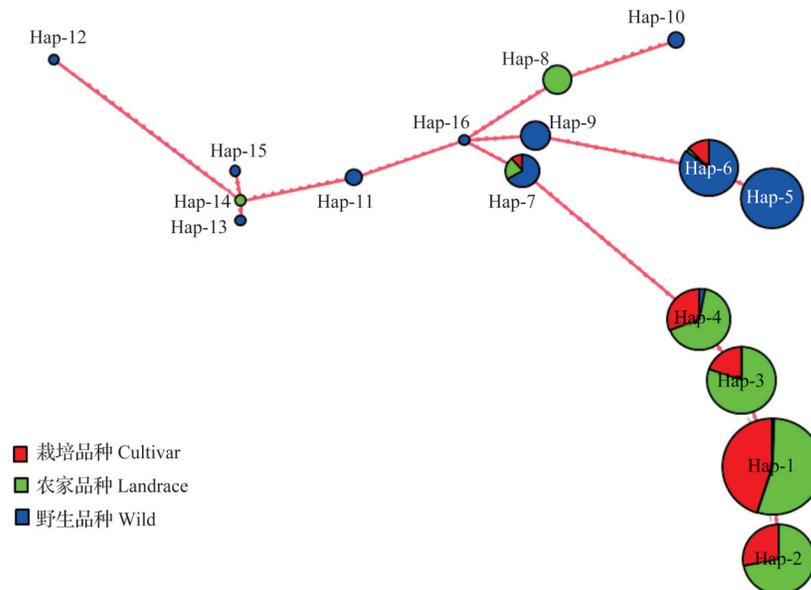
顶部条框的白色和黑色部分分别对应光照时间段(7:00-17:00)和黑暗时间段(17:00-7:00)

The white and black parts of the top bar correspond to the light(7:00-17:00)and dark (17:00-7:00) time periods, respectively

图4 *SiGI* 短日条件 24 h 节律表达模式分析

Fig.4 Rhythm expression pattern of *SiGI* in short day

丰富的变异情况,Hap-6 与其他单倍型存在多个 SNP 和 Indel 变异(图 6,详见 <https://doi.org/10.13430/j.cnki.jpgr.20231130001>,附表 1),说明 *SiGI* 的转录水平可能存在差异。其中,Hap-3 和 Hap-6 两个单倍型



SiGI 编码区及启动子区单倍型 Network 图,不同圆代表不同单倍型,圆的面积和对应单倍型所包含的品种数量成正比,红色连接线代表不同单倍型的突变步骤,连接线上的红色圆点代表一次突变。其中,有表型数据的主要分型为 Hap-1(单倍型频次为 499)、

Hap-2(单倍型频次为 85)、Hap-3(单倍型频次为 61)、Hap-4(单倍型频次为 45)和 Hap-6(单倍型频次为 7)

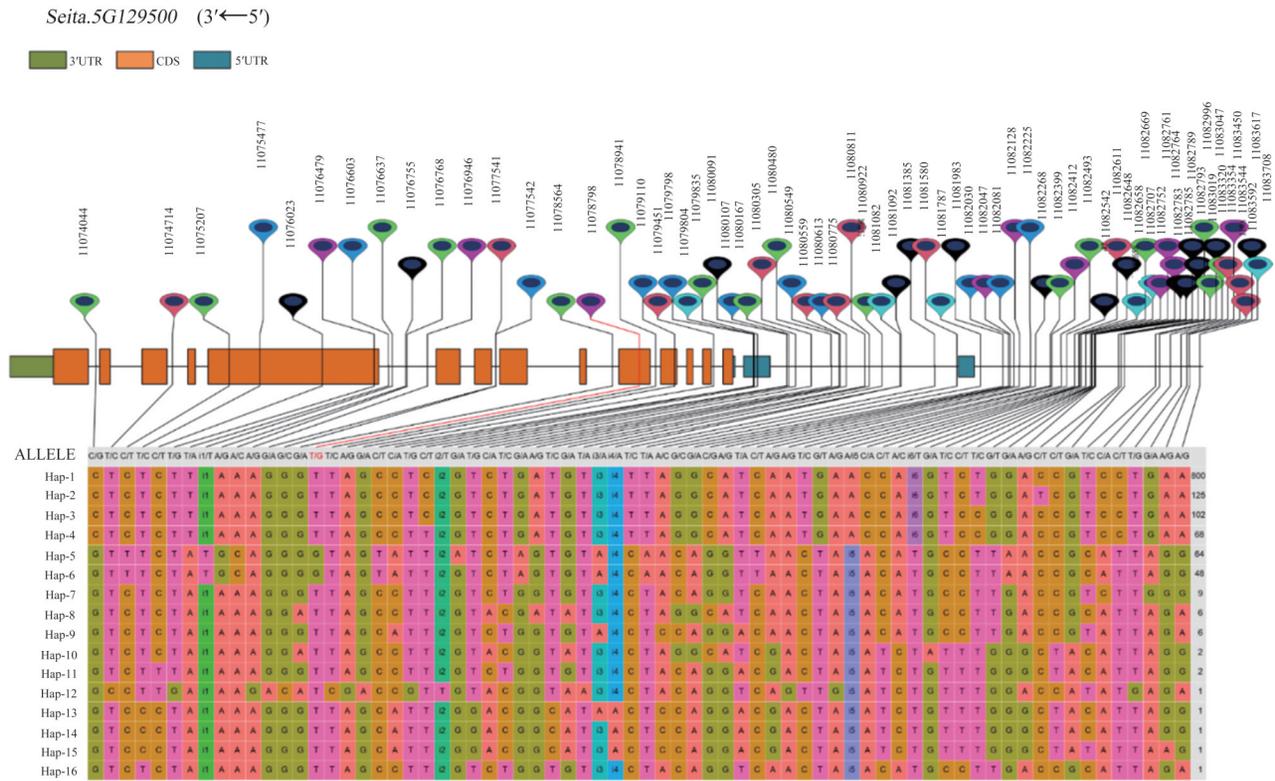
SiGI haplotype network diagram of coding region and promoter region, different circles represent different haplotypes, the area of the circle is proportional to the number of species contained in the corresponding haplotype, the red connecting lines represent the mutation steps of different haplotypes, and the red circles on the connecting lines dots represent a mutation. Among them, the main haplotypes with phenotype data are Hap-1 (haplotype frequency is 499), Hap-2 (haplotype frequency is 85), Hap-3 (haplotype frequency is 61), Hap-4 (haplotype frequency is 45), and Hap-6 (haplotype frequency is 7)

图5 *SiGI* 编码区及启动子区单倍型衍生关系

Fig.5 Relationships between *SiGI* haplotypes detected in coding region and promoter region

之间的变异情况最为丰富,共存在 31 个 SNP 和 4 个 Indel,表明这两个单倍型的表型效应可能存在差异。对 5 个主要单倍型的启动子序列进行了比较,发现两个关键光响应顺式作用元件(AGAAACAA, GGGCGG)存在两个突变位点(启动子上游 1511 bp、

上游 1768 bp 处),且该基因的 Hap-6 分型与其他单倍型在这两处位点都存在差异,说明该基因可能受光响应元件的差异诱导从而影响谷子抽穗期。



上方为 *SiGI* 编码区基因结构及启动子区 2000 bp,其中,框格代表外显子,框格之间的连接线代表内含子。下方对应不同单倍型组合信息表,连接上方 *SiGI* 基因结构图的实心线指示变异位点,其中,红色连接线指示为错义突变,表格中不同颜色单元格代表不同碱基

The upper part shows the gene structure and promoter region of the *SiGI* coding region at 2000 bp. The frame represents the exon, the connecting lines between the frames represent the intron. The lower part corresponds to the information table of different haplotype combinations. The solid line connecting the *SiGI* gene structure map above indicates the mutation site. The red connecting line indicates the missense mutation, and the cells with different colors in the table represent different bases. i1:TATAACAA; i2:TA; i3:AT; i4:AG; i5:TG

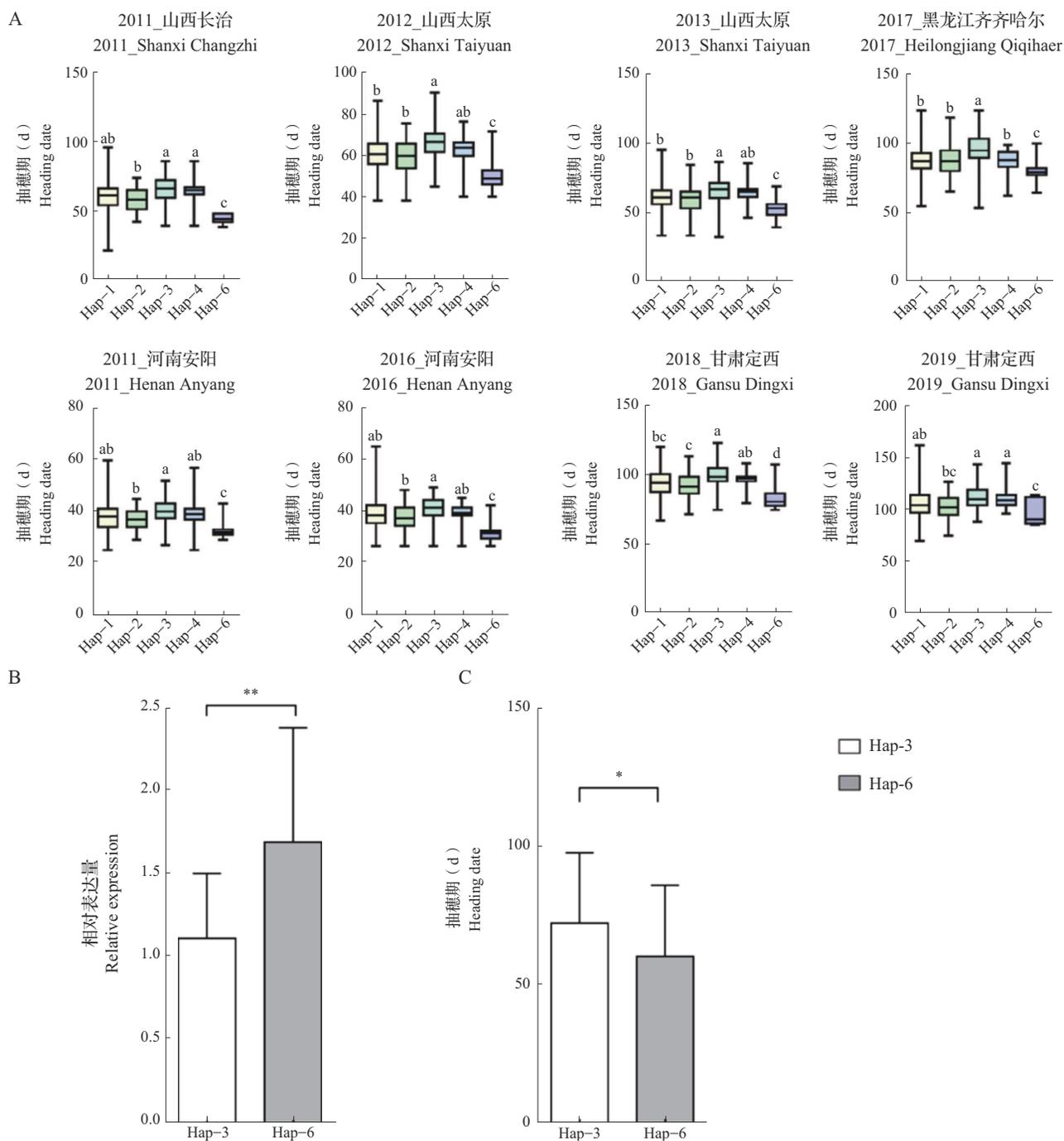
图 6 *SiGI* 编码区及启动子区结构变异

Fig.6 Structural variation of *SiGI* coding region and promoter region

2.6 单倍型的表型效应

SiGI 单倍型的抽穗期、株高、主穗重以及主穗粒重效应分析(图 7、图 8)结果显示,除 2019 年甘肃定西 Hap-2 和 Hap-6 的抽穗期之间无显著差异外,其他 7 个环境下的 Hap-1、Hap-2、Hap-3 和 Hap-4 抽穗期都显著晚于 Hap-6(图 7A)。由于环境变异影响植物生长发育,因此将 *SiGI* 的 Hap-3 和 Hap-6 两个主要单倍型在 8 个环境下对抽穗期的影响进行了比较。其中在 8 个环境下,Hap-3 的抽穗期较 Hap-6 均显著延迟,山西 3 个环境下(2011 年长治、2012 年太原、2013 年太原)分别平均延迟 21 d、15 d 和 13 d,

甘肃 2 个环境下(2018 年定西、2019 年定西)分别延迟了约 10 d,河南 2 个环境下(2011 年安阳、2016 年安阳)分别平均延迟 7 d 和 8 d,以及在黑龙江 1 个环境下(2017 年齐齐哈尔)平均延迟了 16 d。鉴于 Hap-3 单倍型和 Hap-6 单倍型在编码区和启动子区存在多个变异位点共分离,推测 *SiGI* 在 8 个环境下抽穗期的显著差异由编码区和启动子区的变异共同决定。对挑选出的 Hap-3 单倍型材料和 Hap-6 单倍型材料进行 *SiGI* 表达量和抽穗期分析,结果显示,Hap-3 表达量低于 Hap-6,下调近 1.5 倍($P=0.0083$)(图 7B),而抽穗期则平均延迟了约 12 d(图 7C)。



A: *SiGI* 与不同环境抽穗时间关联的箱式图(697份抽穗期数据),不同小写字母表示显著性差异;B: *SiGI* 启动子单倍型相对表达量分析, $P=0.0083$;C: *SiGI* 启动子单倍型关联抽穗期分析;*,** :分别在 $P < 0.05$ 和 0.01 水平达到显著差异;Hap-3包括Ci0162、Ci0203、Ci0227、Ci0230、Ci0779; Hap-6包括Ci0085、Ci0841、Ci0851、Ci0862、Ci0892

A: Box plot of *SiGI* associated with heading time in different environments(697 pieces of heading date data), different lowercase letters represents significant difference; B: Relative expression analysis of *SiGI* promoter haplotypes, $P=0.0083$; C: Analysis of haploid association heading date in *SiGI* promoter; *,** : Reach a significant difference at $P < 0.05$ and 0.01 , respectively; Hap-3: Ci0162、Ci0203、Ci0227、Ci0230、Ci0779, Hap-6: Ci0085、Ci0841、Ci0851、Ci0862、Ci0892

图7 *SiGI* 单倍型分型及转录分析

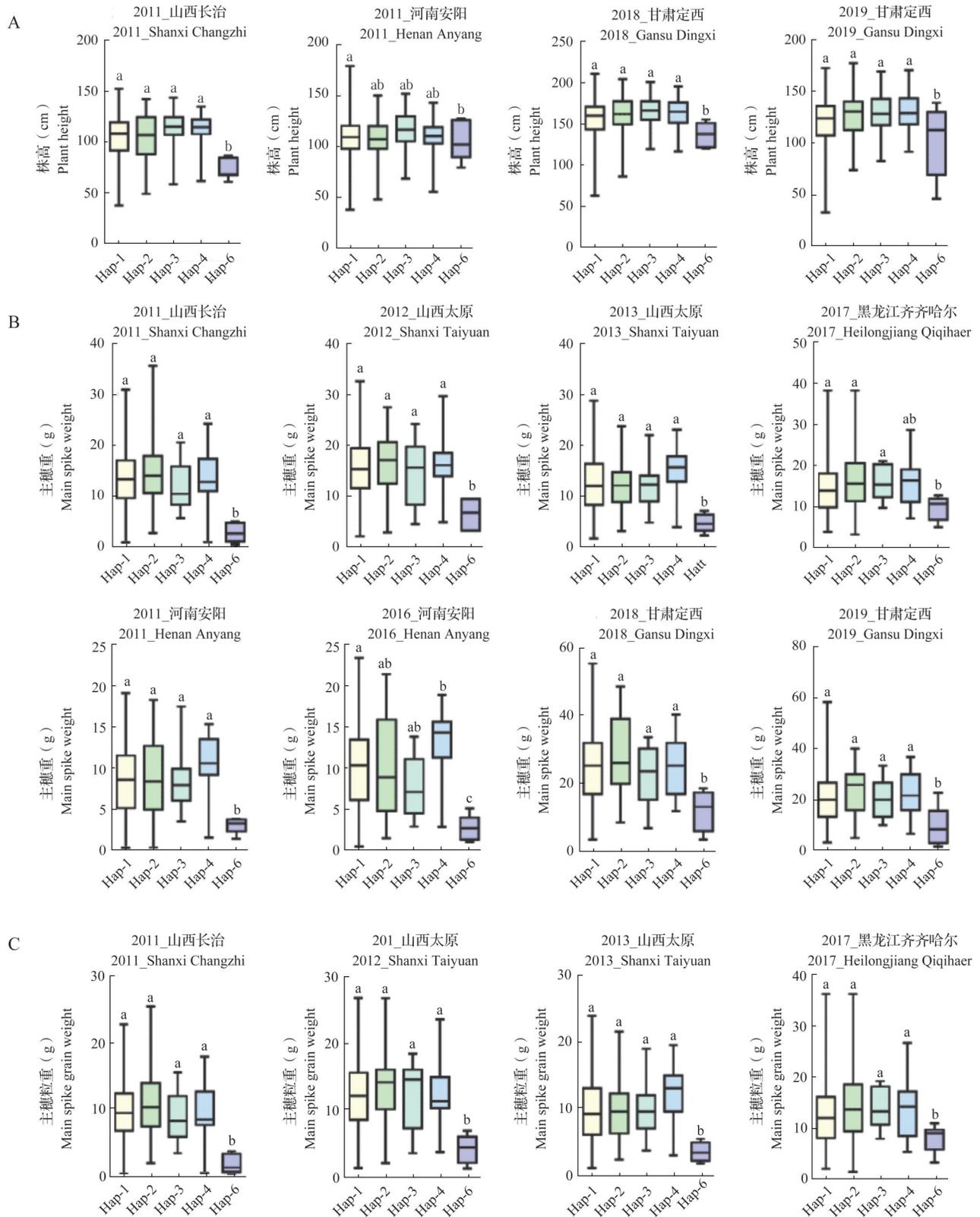
Fig.7 Haplotype typing and transcriptional analysis of *SiGI* promoter region

*SiGI*单倍型在4个环境下的株高差异显著性分析结果显示(图8A),2011年河南安阳的5个单倍型间主穗重没有明显差异,而2011年山西长治、2018

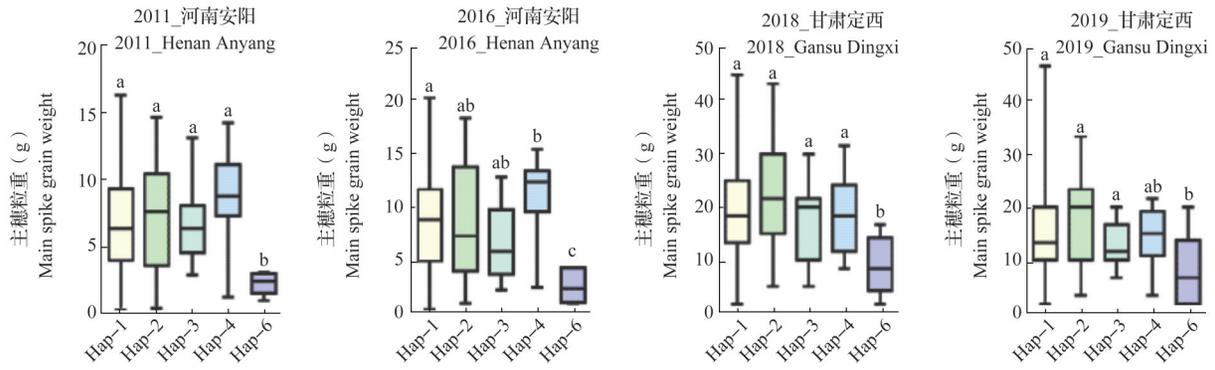
年和2019年的甘肃定西的Hap-1、Hap-2、Hap-3和Hap-4都显著高于Hap-6。*SiGI*单倍型的主穗重和主穗粒重效应分析结果显示(图8B、C),Hap-4和

Hap-6 两个单倍型在 2017 年黑龙江齐齐哈尔的主穗重和 2019 年甘肃定西的主穗粒重两个环境下无显著性差异, Hap-6 与其他单倍型在其他 6 个环境下主

穗重和主穗粒重均存在显著性差异。值得注意的是, Hap-3 和 Hap-6 无论是株高、主穗重还是主穗粒重均存在显著性差异。



(图 8)



A: *SiGI* 与不同环境株高关联的箱式图 (701 份株高数据); B: *SiGI* 与不同环境主穗重关联的箱式图 (286 份主穗重数据);

C: *SiGI* 与不同环境主穗粒重关联的箱式图 (292 份主穗粒重数据)

A: Box plot of *SiGI* associated with plant height in different environments (701 plant height data); B: Box plot of *SiGI* associated with main spike weight in different environments (286 main spike weight data); C: Box plot of *SiGI* associated with main spike grain weight in different environments (292 main spike grain weight)

图8 *SiGI* 单倍型分型关联表型

Fig.8 *SiGI* promoter region haplotype associated phenotype

3 讨论

3.1 *GI* 在禾谷类作物中具有功能保守性

GI 同源基因进化分析结果显示, 不同物种中的 *GI* 具有类似的基因结构, 表明不同物种的 *GI* 基因可能具有相似的功能, 尤其是在单子叶作物中, *GI* 可能具有更高的结构保守性和更相似的功能。在拟南芥中, 由 *GI-CO-FT* 等 3 个基因组成的生物钟调控途径促进了长日照下的开花, 并且在此通路中 *GI* 比 *CO* 和 *FT* 更早地响应光周期变化, 而之前的研究成果表明, 在拟南芥中 *GI* 作为 *CO* 的正调控因子, 通过 *CO* 整合了来自生物钟中的光信号从而调节了 *FT* 的表达, 促进植物开花^[2]; 在水稻中, *OsGI* 同样可以激活 *Hdl* (*CO*), 在长日条件下, *Hdl* 可以抑制 *Hd3a* (*FT*) 的表达。在水稻中过量表达 *OsGI*, 在长、短日照条件下均导致水稻延迟开花^[3], 而 *OsGI* 突变体在实验室短日条件下表现为晚开花, 长日条件下则没有显著差别^[13]。

谷子 *SiGI* 单倍型变异结合抽穗期和基因相对表达量分析显示, 该基因表达量的升高可导致田间环境下抽穗期提前, 这与拟南芥中 *GI* 基因的表达量升高促进提早开花的结果相一致^[2]。虽然在不同作物中该基因的调控机制不尽相同, 但是水稻和拟南芥 *GI* 基因 (*OsGI* 和 *AtGI*) 的表达在长短日照下都以类似的模式受到生物钟的调节^[3]。本研究的结果表明在大田环境中, 谷子中 *SiGI* 表达量升高可引起显著的早熟效应, 具有潜在的育种价值。

此外, *GI* 表达受生物钟调控, 拟南芥 *GI* 在光照后 8 h 左右转录水平达到峰值, 这个高峰的时间、表

达丰度和持续时间受日照时长的影响^[7,27]。谷子 *SiGI* 基因的时空表达规律也显示出了基因振荡表达模式的特点, 在黑暗来临前表达量达到峰值, 而在黑暗中迅速下降。因此, *SiGI* 基因也受到生物钟的控制。谷子 *SiGI* 基因亚细胞定位结果显示定位于细胞核, 表明 *SiGI* 蛋白是一种核蛋白, 可能作为核定位的转录因子行使功能。本研究发现 *SiGI* 基因在谷子全生育期的不同组织均有表达, 其启动子区的顺式作用元件预测和功能分析发现存在与激素调控、光响应和逆境胁迫等相关的作用元件, 这表明 *SiGI* 基因对谷子生长发育、光响应和抗逆性均有一定的调控作用。

3.2 *SiGI* 具有主要单倍型并在多环境表现早熟效应

本研究中, 通过对 *SiGI* 编码区和启动子区的单倍型分析, 发现其存在丰富的遗传多样性, 共检测到 70 个变异位点, 划分为 16 种单倍型。其中, Hap-6 与其他主要单倍型在编码区和启动子区存在多个差异位点, 包括编码区唯一的错义突变位点, 共同导致谷子抽穗期在不同环境下的差异。由于 *SiGI* 启动子区的变异位点可能是影响基因表达量的重要位点, 本研究选取了 Hap-3 和 Hap-6 两个分型作为实验材料, 比较了两种分型材料中 *SiGI* 基因相对表达量以及抽穗期。结果表明, Hap-3 和 Hap-6 分型之间的表达量和抽穗期均存在显著差异。因此, 可以将这些变异位点作为设计分子标记选育早抽穗品种的候选位点。综上, Hap-6 分型具有更高的表达量和较短的抽穗期, 可作为早抽穗单倍型用于育种选择。

3.3 谷子广适性育种中平衡抽穗期和农艺表现的技术路径

通过分析多个环境下的株高、主穗重和主穗粒重发现,4个环境下Hap-6分型的材料株高明显降低,但是在2017年黑龙江齐齐哈尔的主穗重和2019年甘肃定西的主穗粒重并无显著性差异,这表明在抽穗期提前以及株高降低的前提下,可能选育出对主穗重和主穗粒重影响较小的材料。

开花和发育阶段变化的协调调节对作物产量至关重要,因此提高这些发育性状正成为提高产量的主要研究目标。从这个角度来看,谷子的*SiGI*基因可以被编辑以延迟开花或生殖生长,这将有助于农作物在不同光周期环境下提高适应性,而单倍型分析可为分子育种的研发提供关键变异信息。谷子是一种对不同光周期适应性较差的光敏作物^[3,28],而本研究发现谷子*SiGI*存在主效单倍型,这将有助于未来谷子新品种的改良。

4 结论

该研究发现谷子*SiGI*基因具有四环素转录抑制保守结构域,此外,*SiGI*基因在响应光周期的组织(如根、茎、叶)中高表达,具有典型的昼夜节律性表达模式,并定位于细胞核。*SiGI*基因单倍型分析结果表明,谷子的主要单倍型Hap-3和Hap-6具有不同的表达水平,并且与抽穗期表现出显著相关性,即单倍型Hap-6的表达量较Hap-3显著上调,导致谷子品种在8个环境下抽穗期显著提前。此外,Hap-6分型在4个环境下株高显著降低,但在2个环境下主穗重和主穗粒重无显著性差异。本研究为*SiGI*基因功能提供了重要的理论基础,也为谷子分子标记辅助选择提供了新的遗传信息,在谷子分子育种中具有广阔的应用前景。综上所述,单倍型Hap-6可作为早抽穗的重要单倍型进行育种选择。

参考文献

- [1] 胡时开,苏岩,叶卫军,郭龙彪.水稻抽穗期遗传与分子调控机理研究进展.中国水稻科学,2012,26(3):373-382
Hu S K, Su Y, Ye W J, Guo L B. Advances in genetic analysis and molecular regulation mechanism of heading date in rice (*Oryza sativa* L.). Chinese Journal of Rice Science, 2012, 26(3): 373-382
- [2] Seo E, Lee H, Jeon J, Park H, Kim J, Noh Y, Lee I. Crosstalk between cold response and flowering in Arabidopsis is mediated through the flowering-time gene *SOCI* and its upstream negative regulator *FLC*. The Plant Cell, 2009, 10(21): 3185-3197
- [3] Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature, 2003, 422(6933): 719-722
- [4] Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. The Plant Cell, 2000, 12(12): 2473-2483
- [5] Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M. *Hd3a*, a Rice ortholog of the Arabidopsis *FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. Plant and Cell Physiology, 2002, 43(10): 1096-1105
- [6] Xue W Y, Xing Y Z, Weng X Y, Zhao Y, Tang W J, Wang L, Zhou H J, Yu S B, Xu G G, Li X H, Zhang Q F. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nature Publishing Group, 2008, 40(6): 761-767
- [7] Mizoguchi T, Wright L, Fujiwara S, Cremer F, Lee K, Onouchi H, Mouradov A, Fowler S, Kamada H, Putterill J, Coupland G. Distinct roles of *GIGANTEA* in promoting flowering and regulating circadian rhythms in Arabidopsis. The Plant Cell, 2005, 17(8): 2255-2270
- [8] Mishra P, Panigrahi K C S. *GIGANTEA* - An emerging story. Frontiers in Plant Science, 2015, 6(8): 8
- [9] Berns M C, Nordstroem K, Cremer F, Tóth R, Hartke M, Simon S, Klasen J R, Bürstel I, Coupland G. Evening expression of Arabidopsis *GIGANTEA* is controlled by combinatorial interactions among evolutionarily conserved regulatory motifs. The Plant Cell, 2014, 26(10): 3999-4018
- [10] Park D H, Somers D E, Kim Y S, Choy Y H, Lim H K, Soh M S, Kim H J, Kay S A, Nam H G. Control of circadian rhythms and photoperiodic flowering by the Arabidopsis *GIGANTEA* gene. Science, 1999, 285(5433): 1579-1582
- [11] Araki T, Komeda Y. Analysis of the role of the late-flowering locus, *GI*, in the flowering of Arabidopsis thaliana. Plant Journal, 1993, 3(2): 231-239
- [12] Hong S Y, Lee S, Seo P J, Yang M S, Park C M. Identification and molecular characterization of a *Brachypodium distachyon* *GIGANTEA* gene: Functional conservation in monocot and dicot plants. Plant molecular biology, 2010, 3: 485-497
- [13] Izawa T, Mihara M, Suzuki Y, Gupta M, Itoh H, Nagano A J, Motoyama R, Sawada Y, Yano M, Hirai M Y, Makino A, Nagamura Y. Os-*GIGANTEA* confers robust diurnal rhythms on the global transcriptome of rice in the field. Plant Cell, 2011, 23(5): 1741-1755
- [14] Li S, Yue W H, Wang M, Qiu W M, Zhou L, Shou H X. Mutation of *OsGIGANTEA* leads to enhanced tolerance to polyethylene glycol-generated osmotic stress in rice. Frontiers in Plant Science, 2016, 7: 465

- [15] Zhao X Y, Liu M S, Li J R, Guan C M, Zhang X S. The wheat *TaG11*, involved in photoperiodic flowering, encodes an Arabidopsis *GI* ortholog. *Plant Molecular Biology*, 2005, 58 (1): 53-64
- [16] Hecht V, Knowles C L, Vander Schoor J K, Liew L C, Jones S E, Lambert M J M, Weller J L. Pea *LATE BLOOMER1* is a *GIGANTEA* ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. *Plant Physiology*, 2007, 144(2): 648-661
- [17] Li D D, Hou Z H, Li H Y, Li Z B, Fang C, Kong L P, Li Y L, Du H, Li T, Wang L S, He M L, Zhao X H, Cheng Q, Kong F J, Liu B H. Agronomical selection on loss-of-function of *GIGANTEA* simultaneously facilitates soybean salt tolerance and early maturity. *Journal of Integrative Plant Biology*, 2022, 64 (10): 1866-1882
- [18] 张林林, 智慧, 汤沙, 张仁梁, 张伟, 贾冠清, 刁现民. 谷子抽穗时间基因 *SiTOC1* 的表达与单倍型变异分析. *中国农业科学*, 2021, 54(11): 2273-2286
Zhang L L, Zhi H, Tang S, Zhang R L, Zhang W, Jia G Q, Diao X M. Characterizations of transcriptional and haplotypic variations of *SiTOC1* in foxtail millet. *Scientia Agricultura Sinica*, 2021, 54 (11): 2273-2286
- [19] Jia G Q, Huang X H, Zhi H, Zhao Y, Zhao Q, Li W J, Chai Y, Yang L F, Liu K Y, Lu H Y, Zhu C R, Lu Y Q, Zhou C C, Fan D L, Weng Q J, Guo Y L, Huang T, Zhang L, Lu T T, Feng Q, Hao H F, Liu H K, Lu P, Zhang N, Li Y H, Guo E H, Wang S J, Wang S Y, Liu J R, Zhang W F, Chen G Q, Zhang B J, Li W, Wang Y F, Li H Q, Zhao B H, Li J Y, Diao X M, Han B. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nature Genetics*, 2013, 45: 957-961
- [20] He Q, Tang S, Zhi H, Chen J F, Zhang J, Liang H K, Alam O, Li H B, Zhang H, Xing L H, Li X K, Zhang W, Wang H L, Shi J P, Du H L, Wu H P, Wang L W, Yang P, Xing L, Yan H S, Song Z Q, Liu J R, Wang H G, Tian X, Qiao Z J, Feng G J, Guo R F, Zhu W J, Ren Y M, Hao H B, Li M Z, Zhang A Y, Guo E H, Yan F, Li Q Q, Liu Y L, Tian B H, Zhao X Q, Jia R L, Feng J W, Wei J H, Lai J S, Jia G Q, Purugganan M, Diao X M. A graph-based genome and pan-genome variation of the model plant *Setaria*. *Nature Genetics*, 2023, 55: 1232-1242
- [21] 陆平. 谷子种质资源描述规范和数据标准. 北京: 中国农业出版社, 2006
Lu P. Description specification and data standard of millet germplasm resources. Beijing: China Agricultural Press, 2006
- [22] Lescot M, Dehais P, Thijs G, Marchal Y, Peer Y V, Rouze P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, 2002, 30 (1): 325-327
- [23] 桑璐曼, 汤沙, 张仁梁, 贾小平, 刁现民. 谷子热激蛋白 HSP90 基因家族鉴定及分析. *植物遗传资源学报*, 2022, 23(4): 1085-1097
Sang L M, Tang S, Zhang R L, Jia X P, Diao X M. Identification and analysis of heat shock protein HSP90 family genes in foxtail millet. *Journal of Plant Genetic Resources*, 2022, 23 (4): 1085-1097
- [24] 张慧, 梁红凯, 智慧, 张林林, 刁现民, 贾冠清. 谷子 β -胡萝卜素异构酶家族基因的表达与变异分析. *植物学报*, 2023, 58 (1): 34-50
Zhang H, Liang H K, Zhi H, Zhang L L, Diao X M, Jia G Q. Analyses on the transcription and structure variation of β -carotene isomerase gene family in foxtail millet. *Chinese Bulletin of Botany*, 2023, 58 (1): 34-50
- [25] Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods*, 2001, 25(4): 402-408
- [26] Zhang R L, Jia G Q, Diao X M. GeneHapR: An R package for gene haplotypic statistics and visualization. *BMC Bioinformatics*, 2023, 24(1): 199
- [27] Fowler S, Lee K, Onouchi H, Samach A, Richardson K, Moris B, Coupland G, Putterill J. *GIGANTEA*: A circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *Embo Journal*, 2014, 18 (17): 4679-4688
- [28] 刁现民, 程汝宏. 十五年区试数据分析展示谷子糜子育种现状. *中国农业科学*, 2017, 50(23): 4469-4474
Diao X M, Cheng R H. Current breeding situation of foxtail millet and common millet in China as revealed by exploitation of 15 years regional adaptation test data. *Scientia Agricultura Sinica*, 2017, 50 (23): 4469-4474

附表 1 *SiGI* 启动子区单倍型变异情况Table 1:Haploid variation in the *SiGI* promoter region

5 号染色体物理位置 (bp) Chr.5 physical position	Hp-1	Hp-2	Hp-3	Hp-4	Hp-6
11081983	G	G	G	G	C
11082030	G	G	G	G	A
11082047	C	C	C	C	G
11082081	A	A	A	A	G
11082225	C	C	C	C	T
11082412	T	T	T	T	C
11082493	G	G	G	G	T
11082611	A	A	A	A	AG
11082648	C	C	C	C	A
11082707	TG	TG	TG	TG	T
11082761	T	T	T	T	C
11082783	T	T	C	C	T
11082785	G	G	G	G	T
11082789	G	G	G	G	A
11082996	C	T	C	C	C
11083320	T	T	T	T	C
11083354	C	C	C	C	A
11083450	C	C	C	C	T
11083592	G	G	G	G	A
11083617	A	A	A	A	G
11083708	A	A	A	A	G

红色字体标注 Hp-6 与其他单倍型的多个 SNP 和 Indel 变异位点

Red font annotation of multiple SNP and Indel variation sites between Hp-6 and other haplotypes