

玉米淀粉合成相关基因 *ZmAGP2* 的序列变异分析

张 萌,张晗菡,徐 敏,何 欢,吴雪怡,吕莹莹,张恩盈

(青岛农业大学 农学院,青岛市主要农作物物质创新与应用重点实验室,山东 青岛 266109)

摘要:玉米 *ZmAGP2* 基因编码 ADP 葡萄糖焦磷酸化酶大亚基,参与玉米淀粉的合成途径,从而影响玉米的产量和品质。由多基因控制的玉米的产量性状,具有非常复杂的遗传机理。拟通过基因序列变异分析充分挖掘优异等位变异,这些变异与产量性状密切相关,为玉米在分子水平上的育种提供理论基础。利用 NimbleGen 平台对 103 份供试自交系试验群体 *ZmAGP2* 基因进行目标序列捕获并测序。利用 ClustalX 2.0、DnaSP 6.0 等软件分析目的基因的核苷酸多态性,进行变异位点的单倍型分析和中性进化检验。对 103 份玉米自交系的 *ZmAGP2* 基因重测序,获得了 4 111 bp 的全长序列,经过核苷酸多态性分析,共检测到全长序列中 112 个变异位点,均为 SNP,其中 5' UTR 区段的核苷酸多态性明显高于其他区段。基因的 CDS 区的多态性将 *ZmAGP2* 基因分为 8 个单倍型,编码 5 种不同的 *ZmAGP2* 蛋白。中性检验未检测到明显的人工选择的信号。玉米 *ZmAGP2* 基因表现出较高的变异,符合中性进化模型假设,未受到明显的人工选择作用,且蛋白序列发生改变,这表明该基因的序列变异引起了蛋白序列的改变,导致该基因功能可能发生了变化。

关键词:玉米;*ZmAGP2* 基因;淀粉合成;核苷酸多态性;中性检验;单倍型

中图分类号:S513.03;Q78 **文献标识码:**A **文章编号:**1000-7091(2020)增刊-0040-07

doi:10.7668/hbxb.20190786



Sequence Variation Analysis of Maize Starch Synthesis Related Gene *ZmAGP2*

ZHANG Meng, ZHANG Hanhan, XU Min, HE Huan, WU Xueyi, LÜ Yingying, ZHANG Enying

(College of Agronomy, Qingdao Agricultural University, Key Laboratory of Germplasm Innovation and Application of Major Crops in Qingdao, Qingdao 266109, China)

Abstract: The *ZmAGP2* gene encoding ADP large subunit of glucose pyrophosphory is involved in the synthesis of corn starch, thus affecting the yield and quality of corn. Maize yield traits are controlled by multiple genes, and its genetic mechanism is very complex. This study aims to fully explore the excellent allelic variation of genes closely related to yield traits through gene sequence variation analysis, which provides a theoretical basis for improving the molecular level of maize breeding. The *ZmAGP2* genes sequences of 103 inbred test populations were captured and sequenced using NimbleGen platform. ClustalX 2.0, DnaSP 6.0 and other software were used to analyze the nucleotide polymorphism of the target gene, and the neutral evolution test and haplotype analysis of the mutation site were carried out. The full-length sequence of 4 111 bp was obtained by resequencing the *ZmAGP2* genes of 103 maize inbred lines. After nucleotide polymorphism analysis, a total of 112 mutation sites in the full-length sequence were detected, all of which SNP, nucleotide polymorphisms in 5' UTR segments were significantly higher than those in other segments. Polymorphism in the CDS region of the gene driven the *ZmAGP2* gene into eight haplotypes encoding five different *ZmAGP2* proteins. Neutral test did not detect obvious manually selected signals. The *ZmAGP2* gene showed high variation, which accorded with the neutral evolution model hypothesis, was not subject to obvious artificial selection, and the protein sequence changed, which indicated that the sequence variation of the gene causes the change of the protein sequence, which lead to the change of the function of the gene.

Key words: Maize; *ZmAGP2* gene; Starch synthesis; Nucleotides polymorphism; Neutral test; Haplotype

收稿日期:2020-06-14

基金项目:山东省农业科学院农业科技创新工程项目(CXGC2017B01)

作者简介:张 萌(1995-),女,山东德州人,硕士,主要从事玉米遗传育种及统计分析研究。

通讯作者:张恩盈(1976-),男,山东陵城人,副教授,博士,主要从事玉米遗传育种及生物信息学研究。

玉米 (*Zea mays* L.) 在人类生产生活中发挥着巨大的作用, 又是我国主要的经济作物和部分饲料的原材料。玉米籽粒中包括淀粉、蛋白质、脂肪等重要组分, 而玉米胚乳约占玉米籽粒重的 80% ~ 85%, 而淀粉又是胚乳细胞中的主要成分, 是玉米籽粒的重要组成成分^[1]。籽粒淀粉是决定玉米产量和品质的关键因素, 提高玉米淀粉的含量和淀粉的品质是玉米育种工作者最重要的任务^[2]。淀粉是在一系列淀粉合成相关酶的催化作用下完成的, 在淀粉的合成过程中, 有 4 种酶起主要催化作用, 包括淀粉去分支酶 (Starch debranching enzyme, SDBE)、ADP 葡萄糖焦磷酸化酶 (ADP-glucose pyrophosphorylase, AGPase)、淀粉分支酶 (Starch branching enzyme, SBE) 以及淀粉合成酶 (Starch synthase, SS)^[3-4]。

玉米 *ZmAGP2* 基因 (ADP glucose pyrophosphorylase 2) 编码 ADP 葡萄糖焦磷酸化酶大亚基。ADP 葡萄糖焦磷酸化酶是一个异源四聚体^[5], 由 2 个相同的大亚基 (Sh2) 和 2 个相同的小亚基 (Bt2) 组成^[6]。ADP 葡萄糖焦磷酸化酶催化 ATP 和 Glc-1-p^[7], 合成 ADP-葡萄糖和二磷酸, 在 α -1, 4-葡聚糖合成的过程中, ADP-葡萄糖作为活化的葡萄糖供体^[8-9]。该反应是玉米胚乳淀粉合成的关键限速步骤。*Shrunken2* (*Sh2*) 基因编码玉米 ADP 葡萄糖焦磷酸化酶大亚基, *Brittle2* (*Bt2*) 基因编码玉米 ADP 葡萄糖焦磷酸化酶小亚基。Deng 等^[4] 分离了一个编码小麦 (*Triticum aestivum* L.) 中 AGP 酶质体小亚基 (*Ta-*

AGPS1b) 的 cDNA 序列, *TaAGPS1b* 与水稻原生质体细胞中的绿色荧光蛋白 (GFP) 融合, 共聚焦荧光显微镜观察表明, *TaAGPS1b*-GFP 定位于叶绿体中。Smidansky 等^[10] 将玉米中编码 ADP 葡萄糖焦磷酸化酶大亚基的 *Shrunken2* 基因改造后 (*Sh2r6hs*) 转入小麦中, 发现转基因小麦株系平均每株比对照多出 38% 的粒质量。Zhang 等^[3] 将小麦中的 *TaAGP* 突变, 结果发现, 导致其他 4 个淀粉合成基因的表达模式发生变化, 导致淀粉和支链淀粉含量降低。Rose 等^[11] 发现在小麦自然群体中 *TaAGP* 的 SNPs 引起千粒质量和种子数量的变化, 但总淀粉含量没有显著变化。灌浆时间和速率对最终的晶粒质量至关重要。谷物发育初期的干旱胁迫会干扰器官建立和淀粉生物合成所需的细胞分裂和分化过程。在籽粒灌浆过程中, 在水分亏缺条件下, AGPase 在蛋白质和转录水平上的活性下调是淀粉积累显著下降的原因^[12-13]。*ZmAGP2* 基因在优质玉米自交系的核苷酸多态性尚不清楚, 本研究的目的就是对 103 份玉米自交系的 *ZmAGP2* 基因的序列进行定点捕获, 分析该基因的核苷酸多态性和单倍型密度。

1 材料和方法

1.1 试验材料

本研究选用了 103 份优良玉米自交系, 其中包括来自于国外的种质材料、热带、亚热带种质以及我国温带玉米的 5 个杂种优势群的代表性种质 (表 1)。

表 1 本研究所用玉米自交系列表

Tab. 1 List of maize inbred lines of the study

编号 Number	自交系 Maize inbred lines	编号 Number	自交系 Maize inbred lines	编号 Number	自交系 Maize inbred lines
1	QH19612	36	铁 9206	70	QBL-5
2	昌 7-2	37	本育 15	71	QMXG-1
3	981F	38	QZ01	72	JND-1
4	107	39	纯 2433	73	JND-2
5	黄 518	40	478 选	74	BJ-2
6	k12	41	郑 58	75	F8111
7	NX335	42	7922	76	FN-6
8	s80	43	8605-2	77	M3564
9	H21	44	178	78	M1
10	吉 853	45	QP1712	79	WT262
11	412	46	Ex 韩	80	WT26
12	杂 107	47	xy335	81	QZN01F
13	黄早四	48	P138	82	QKN-1M
14	502	49	EX319	83	QD-1T
15	鲁原 92	50	丹 988	84	QLY-2
16	10168	51	319B	85	JKN2000F
17	丹 598	52	BQ319	86	QZBN-1

表 1 (续)

编号 Number	自交系 Maize inbred lines	编号 Number	自交系 Maize inbred lines	编号 Number	自交系 Maize inbred lines
18	综 3	53	齐 318	87	KH-1F
19	4CV	54	沈 137	88	QBN488
20	E28	55	11099	89	QB1029
21	H·丹 340	56	Suwan	90	BJ-4
22	自 330	57	11118	91	BEM
23	良 88M	58	11200	92	QHNH-4
24	340 改	59	10533-1	93	BJ-1
25	丹黄 25	60	GB28	94	扬 BK
26	JH-5F	61	CML-15	95	LY16F
27	丹 99	62	QSY-7	96	QBNL189
28	齐 232	63	DK3110-65	97	M1132
29	OH43	64	热 BS11	98	JKN2000M
30	MO17	65	FLB-1	99	BJ-3
31	8112	66	Fc521	100	QJZB-3
32	K8112	67	kw456-2	101	BJ-5
33	武 314	68	QF-11	102	Y53
34	DH1M	69	QF01	103	JB
35	纯硬 189				

1.2 DNA 提取和 *ZmAGP2* 基因测序

通过水培的方式,将所有的自交系种子种植在 28 °C/25 °C,16 h 光照、8 h 黑暗的昼夜交替的温室里。各样本的 DNA 的提取是利用改良过的 CTAB 法,利用天根®超微量分光光度计来测定 DNA 的浓度,各样本 DNA 的质量的检测是用 1% 的琼脂糖胶。以 B73 的 *ZmAGP2* 基因序列为模板,103 份自交系材料的 *ZmAGP2* 基因的定点捕获和测序是通过 NimbleGen 平台进行的,华大公司完成该序列的定点捕获和测序。

1.3 序列数据分析

本研究用到的软件有 ClustalX 2.0^[14] 和 DnaSP 6.0^[15] 等。分别对获得的 *ZmAGP2* 的基因序列进行比对和进一步人工校对,估算各区段以及全长的核苷酸多态性。基因的中性进化检测利用 Tajima^[16]、Fu 与 Li^[17] 的 D* 和 F* 的参数来衡量。

2 结果与分析

2.1 玉米 *ZmAGP2* 基因的序列变异

ZmAGP2 基因序列的全长是以玉米自交系 B73 的 *ZmAGP2* 基因为模板在 103 个玉米供试自交系中获得的。经过多序列比对后,获得了 *ZmAGP2* 基因的全长序列为 4 111 bp。包括启动子序列 246 bp,5'UTR 区段序列 1 334 bp,3'UTR 区段序列 370 bp,以及内含子 1 106 bp 和外显子 1 055 bp。以 100 bp 为滑动框,25 bp 为步长,分析得到 *ZmAGP2* 基因的核苷酸多态性图(图 1)。结果显示,5'UTR 区段与其

他区段相比,核苷酸多态性较高。

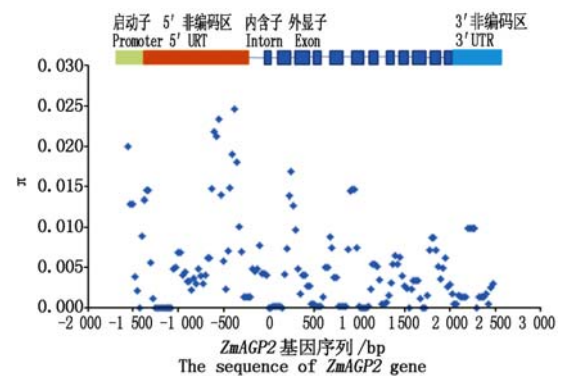


图 1 *ZmAGP2* 基因序列核苷酸多样性

Fig. 1 The nucleotide diversity estimated along the *ZmAGP2* gene sequence

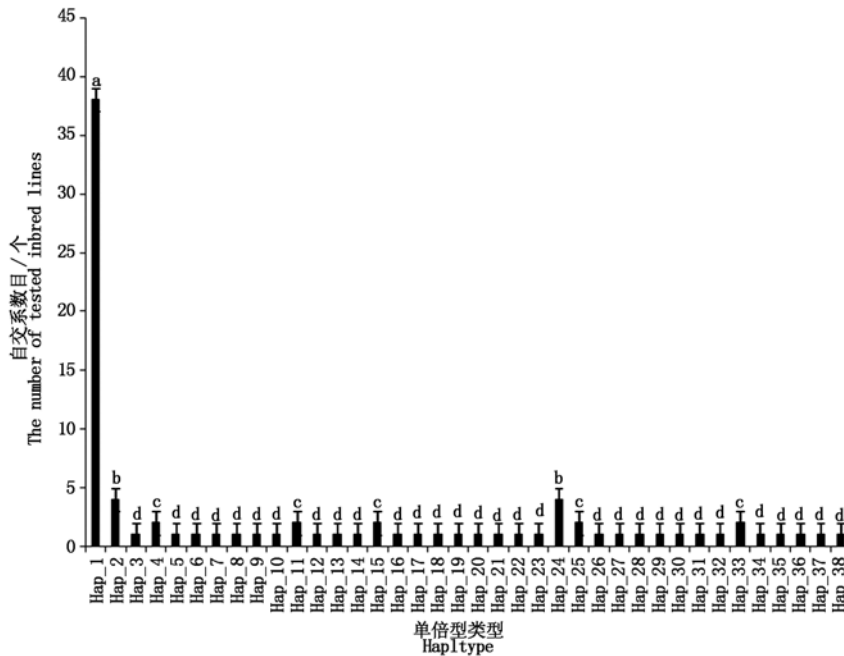
对 *ZmAGP2* 基因全长序列进一步分析发现,共检测到 112 个变异位点,没有检测到 Indel 变异,均为 SNP 位点(表 2)。非编码区含有 103 个 SNP,其中内含子区段有 42 个 SNP,5'UTR 区段有 40 个 SNP。编码区含有 9 个 SNP。说明内含子区段和 5'UTR 区段是主要的遗传变异来源。变异频率最高的是第六内含子区段,为 0.077。在 6 个区段没有发现变异位点。

常用参数 θ 和 π 分析核苷酸多态性。在本研究的 103 份玉米自交系中,*ZmAGP2* 基因全长序列的 π 值和 θ 值分别为 0.051 40 和 0.005 52(表 3)。Tajima's D 测验 *ZmAGP2* 基因序列中的多态性位点。结果显示,Tajima's D 没有达到显著水平,说明整段序列没有经过明显的选择作用。而只有启动子区段的 Fu 与 Li 的 D* 值达到了显著水平。

2.2 ZmAGP2 基因的单倍型分析

单倍型分析可以为育种工作者提供丰富的遗传信息。本研究通过对 103 份自交系的 *ZmAGP2* 基因的全长进行单倍型分析,共检测到 38 个单倍型(图 2),单倍型 Hap_1 的供试材料数目最多,共 38 个自交系,单倍

型多态性达到 0.80。同时,单倍型 Hap_2-Hap_38 均与 Hap_1 表现显著性差异。Hap_2 与 Hap_24 差异不显著,Hap_4、Hap_11、Hap_15、Hap_25 和 Hap_33 没有显著性差异。表明自交系在每个单倍型中的分布并不均匀,单倍型 Hap_1 中含有的单倍型最多。



显著水平为 $P < 0.05$ 。

Significance at $P < 0.05$ level.

图 2 供试玉米自交系中 *ZmAGP2* 基因全长单倍型分布

Fig. 2 Distribution of major haplotypes of the *ZmAGP2* gene total sequences among tested inbred lines sequence

为进一步深入研究,针对 *ZmAGP2* 基因的编码区对单倍型进行分析(图 3)。共检测到 8 个单倍型,CDS Hap_1 和 CDS Hap_2 包括 75 个自交系,占供试材料的 87.2%。其中 CDS Hap_1 包含 63 个自交系,数目最多;CDS Hap_3、CDS Hap_5、CDS Hap_6、CDS Hap_7 和 CDS Hap_8 包含自交系数目最少为 1 个;CDS Hap_2 包含 12 个自交系;CDS Hap_4 包含 6 个自交系。单倍型 CDS Hap_2-8 与 CDS Hap_1 具有显著差异,CDS Hap_2 与 CDS Hap_3-8 也具有显著差异性,单倍型 CDS Hap_3、

CDS Hap_5、CDS Hap_6、CDS Hap_7 和 CDS Hap_8 没有显著差异,说明自交系在编码区的单倍型分布也不均匀,CDS Hap_1 含有的玉米自交系数最多。

ZmAGP2 基因共编码了 5 种蛋白质序列,氨基酸全长为 334 aa(图 4)。*ZmAGP2* 基因编码区有 9 个 SNP,包括 3 个同义突变,6 个为非同义突变,后者导致氨基酸序列的突变,产生 5 种不同类型的 *ZmAGP2* 蛋白。

3 讨论

丰富的遗传变异使植物育种家能够创造出新的植物基因组合,并可以选择更适合不同农业系统需要的作物品种。作物功能基因遗传多样性的分析对于理解表型变异的遗传背景至关重要,进而为作物改良提供了极大的帮助^[18-19]。本研究在玉米 *ZmAGP2* 基因全长序列中检测到 112 个变异位点,变异类型均为 SNP。在这些 SNP 中,有 9 个位于编码区,其中 3 个为同义突变,未改变氨基酸序列,6 个为非同义突变,是改变蛋白序列的最主要的原因。编码区的 9 个 SNP 将被检测的自交系划分为 8 个单倍型,编码 5 种 *ZmAGP2* 蛋白。然而发挥主要作

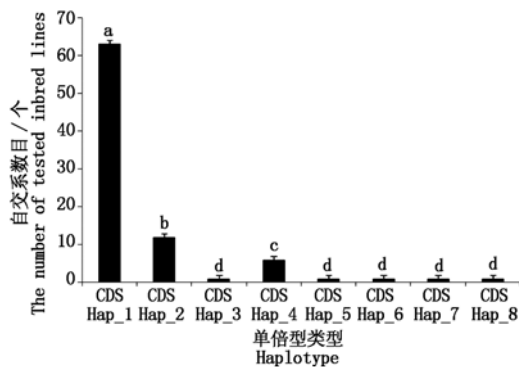
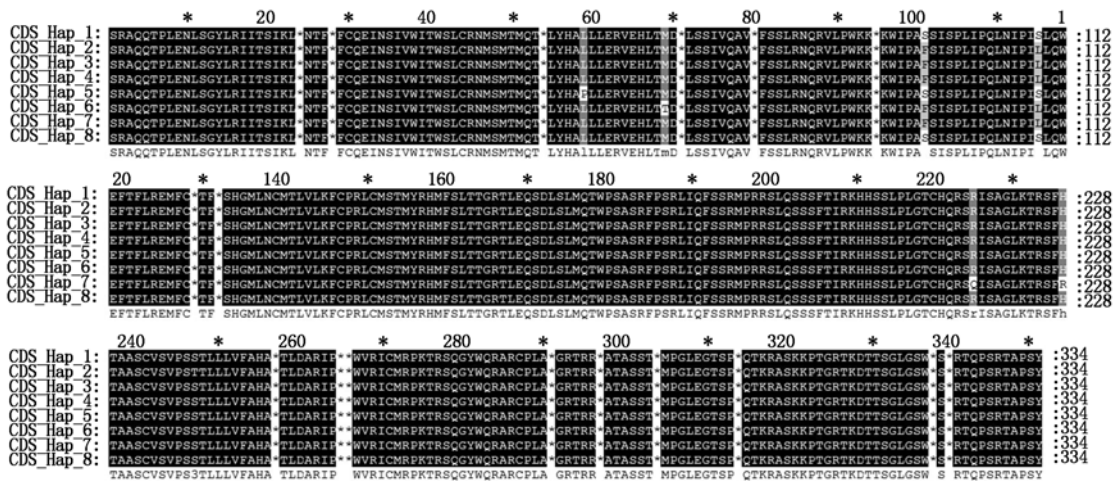


图 3 供试玉米自交系中 *ZmAGP2* 基因编码区单倍型分布

Fig. 3 Distribution of major haplotypes of the *ZmAGP2* gene coding sequences among tested inbred lines

图 4 不同 CDS 单倍型编码的 *ZmAGP2* 蛋白序列比对Fig. 4 Sequence alignment of the *ZmAGP2* proteins encoded by different CDS haplotypes

用的可能是非编码区段的位点变异,尤其是 5'UTR 和内含子区段,分别检测到 40,42 个 SNP,其中的原因可能是非编码区段的序列长度要远远大于编码区。经分析发现 B73 中该基因含有 15 个内含子和 15 个外显子,但对 103 份自交系 *ZmAGP2* 基因的全长序列进行重测序后,发现前 4 个内含子和前 3 个外显子丢失,其原因可能因为在转录过程中基因发生可变剪切或者测序过程中出现的错误导致,其碱基的大量缺失产生的性状和表型以及是否对淀粉的合成产生影响尚不清楚,需进一步的研究。Ching 等^[20]对 36 个玉米自交系进行序列变异分析,发现在非编码区每 31 bp 就发生一个变异位点,而在编码区每 124 bp 会产生一个变异位点。Zhang 等^[21]对 73 个中国优良自交系的 *RTCS* 基因的核苷酸多态性进行分析,编码区每 45.9 bp 有一个 SNP,在非编码区每 10.9 bp 就有一个 SNP。本研究在非编码区每 30 bp 就会有一个 SNP;在编码区每 117 bp 就有一个 SNP。

本研究对玉米 103 份自交系的 *ZmAGP2* 基因各区段以及全长进行了 Tajima^[16]以及 Fu 与 Li^[17]的 D* 和 F* 测验,没有检测到明显的人工选择的信号。因此,*ZmAGP2* 基因虽然参与淀粉的合成,但是可能不是影响淀粉合成过程的关键基因。同时,若要研究该基因对产量和品质的影响时,应该考虑更大的群体。

参考文献:

- [1] Qu J Z, Xu S T, Zhang Z Q, Chen G Z, Zhong Y Y, Liu L S, Zhang R H, Xue J Q, Guo D W. Evolutionary, structural and expression analysis of core genes involved in starch synthesis[J]. *Sci Rep*, 2018, 8(1):12736. doi:10.1038/s41598-018-30411-y.
- [2] Yu G W, Lü Y N, Shen L Y, Wang Y B, Qing Y, Wu N, Li Y P, Huang H H, Zhang N, Liu Y H, Hu Y F, Liu H M,

- Zhang J J, Huang Y B. The proteomic analysis of maize endosperm protein enriched by Phos-tag (tm) reveals the phosphorylation of Brittle-2 subunit of ADP-Glc pyrophosphorylase in starch biosynthesis process[J]. *Int J Mol Sci*, 2019, 20(4):986. doi:10.3390/ijms20040986.
- [3] Zhang S L, Guo H J, Irshad A, Xie Y D, Zhao L S, Xiong H C, Gu J Y, Zhao S R, Ding Y P, Liu L X. The synergistic effects of TaAGP. L-B1 and TaSSIVb-D mutations in wheat lead to alterations of gene expression patterns and starch content in grain development[J]. *PLoS One*, 2019, 14(10):e0223783. doi:10.1371/journal.pone.0223783.
- [4] Deng X, Liu Y, Xu X X, Liu D M, Zhu G R, Yan X, Wang Z M, Yan Y M. Comparative analysis of AGPase genes and encoded proteins in eight monocots and three dicots with emphasis on wheat[J]. *Front Plant Sci*, 2017, 8:19. doi:10.3389/fpls.2018.00425.
- [5] Slattery C J, Kavakli I H and Okita T W, Engineering starch for increased quantity and quality[J]. *Trends Plant Sci*, 2000, 5(7):291-298. doi:10.1016/s1360-1385(00)01657-5. c.
- [6] Hou J, Li T, Wang Y M, Hao C Y, Liu H X, Zhang X Y. ADP-glucose pyrophosphorylase genes, associated with kernel weight, underwent selection during wheat domestication and breeding[J]. *Plant Biotechnol J*, 2017, 15(12):1533-1543. doi:10.1111/pbi.12735.
- [7] Miao H X, Sun P G, Liu Q, Liu J H, Xu B Y, Jin Z Q. The AGPase family proteins in banana: Genome-wide identification, phylogeny, and expression analyses reveal their involvement in the development, ripening, and abiotic/biotic stress responses[J]. *Int J Mol Sci*, 2017, 18(8):1581. doi:10.3390/ijms18081581.
- [8] Zeeman S C, Kossmann J S, Alison M K, Smith A M. Starch: its metabolism, evolution, and biotechnological modification in plants[J]. *Annu Rev Plant Biol*, 2010, 61:209-234. doi:10.1146/annurev-arplant-042809-112301.
- [9] Tiessen A, Nerlich A, Faix B, Hummer C, Fox S, Trafford K, Weber H, Weschke W, Geigenberger P. Sub cellular analysis of starch metabolism in developing barley seeds using a non-aqueous fractionation method[J]. *J Exp Bot*, 2012, 63(5):2071-2087. doi:10.1093/jxb/err408.
- [10] Smidansky E D, Clancy M, Meyer F D, Lanning S P, Blake

- N K, Talbert L E, Giroux M J. Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield[J]. *Proc Natl Acad Sci USA*, 2002, 99(3): 1724 – 1729. doi:10.1073/pnas.022635299.
- [11] Rose M K, Huang X Q, Brule B A. Molecular characterization and sequence diversity of genes encoding the large subunit of the ADP-glucose pyrophosphorylase in wheat (*Triticum aestivum* L.) [J]. *J Appl Genet*, 2016, 57(1): 15 – 25. doi:10.1007/s13353-015-0298-1.
- [12] Deng X, Liu Y, Xu X X, Liu D M, Zhu G R, Yan X, Wang Z M, Yan Y M. Comparative proteome analysis of wheat flag leaves and developing grains under water deficit[J]. *Front Plant Sci*, 2018, 9: 425. doi:10.3389/fpls.2018.00425.
- [13] Chen G X, Zhen S M, Liu Y L, Yan X, Zhang M, Yan Y M. *In vivo* phosphoproteome characterization reveals key starch granule-binding phosphoproteins involved in wheat water-deficit response [J]. *BMC Plant Biol*, 2017, 17(1): 168. doi:10.1186/s12870-017-1118-z.
- [14] Murray M G, Thompson W F. Rapid isolation of high molecular weight plant DNA [J]. *Nucleic Acids Res*, 1980, 8(19): 4321 – 4325. doi:10.1093/nar/8.19.4321.
- [15] Rozas J, Ferrer-Mata A, Sánchez-DelBarrio J C, Guirao-Rico S, Librado P, Ramos-Onsins S E, Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism analysis of large data sets [J]. *Mol Biol Evol*, 2017, 34(12): 3299 – 3302. doi:10.1093/molbev/msx248.
- [16] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism [J]. *Genetics*, 1989, 123(3): 585 – 595.
- [17] Fu Y X, Li W H. Statistical tests of neutrality of mutations [J]. *Genetics*, 1993, 133(3): 693 – 709.
- [18] Xiao Y J, Liu H J, Wu L J, Warburton M, Yan J B. Genome-wide association studies in maize: praise and stargaze [J]. *Mol Plant*, 2017, 10(3): 359 – 374. doi:10.1016/j.molp.2016.12.008.
- [19] Mafra G S, Amaral J, Antonio T D, Almeida F, Janeo E D, Vivas M, Santos P H A D, Santos J S, Pena G F, Lima V J D, Kamphorst S H, Oliveira FT D, Souza Y P D, Schwantes I A, Santos T D, Oliveira B, Rosimeire B, Maldonado C, Mora F. SNP-based mixed model association of growth-and yield-related traits in popcorn [J]. *PLoS One*, 2019, 14(6): e0218552. doi:10.1371/journal.pone.0218552.
- [20] Ching A, Caldwell K S, Jung M, Dolan M, Smith O S, Tingey S, Morgante M, Rafalski A J. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines [J]. *BMC Genet*, 2002, 3: 19. doi:10.1186/1471-2156-3-19.
- [21] Zhang E Y, Yang Z F, Wang Y F, Hu Y Y, Song X Y, Xu C W. Nucleotide polymorphisms and haplotype diversity of RTCS gene in China elite maize inbred lines [J]. *PLoS One*, 2013, 8(2): e56495. doi:10.1371/journal.pone.0056495.