Mapping of QTL and Identification of Candidate Genes Conferring Spontaneous Haploid Genome Doubling in Maize (*Zea mays* L.)

Jiaojiao Ren, Nicholas Boerman, Ruixiang Liu, Penghao Wu, Benjamin Trampe, Kimberly Vanous, Ursula K. Frei, Shaojiang Chen, Thomas Lübberstedt

PII: S0168-9452(19)31510-9

DOI: https://doi.org/10.1016/j.plantsci.2019.110337

Reference: PSL 110337

To appear in: *Plant Science*

Received Date: 4 August 2019

Revised Date: 7 November 2019

Accepted Date: 19 November 2019

Please cite this article as: Ren J, Boerman N, Liu R, Wu P, Trampe B, Vanous K, Frei UK, Chen S, Lübberstedt T, Mapping of QTL and Identification of Candidate Genes Conferring Spontaneous Haploid Genome Doubling in Maize (*Zea mays* L.), *Plant Science* (2019), doi: https://doi.org/10.1016/j.plantsci.2019.110337

Mapping of QTL and Identification of Candidate Genes Conferring Spontaneous Haploid Genome Doubling in Maize (*Zea mays* L.)

Jiaojiao Ren[1#], Nicholas Boerman[2#], Ruixiang Liu[3], Penghao Wu[1], Benjamin Trampe[2], Kimberly Vanous[2], Ursula K. Frei[2], Shaojiang Chen[4], Thomas Lübberstedt[2*]

[1] College of Agronomy, Xinjiang Agricultural University, Urumqi, China 830052

[2] Department of Agronomy, Iowa State University, Ames, Iowa, USA 50011

[3] Institute of Food Crops, Jiangsu Province Academy of Agricultural Sciences, Jiangsu, China 210014

[4] National Maize Improvement Center, China Agricultural University, Beijing, China 100193

*Corresponding author at: 716 Farm House Lane, Iowa State University, Ames, IA, 50011, USA

Email address: thomasL@iastate.edu (T. Lübberstedt)

#Ren J and Boerman N contributed equally to this work

Highlights:

- New QTLs related to spontaneous haploid genome doubling (SHGD) in maize were identified by selective genotyping.
- Candidate gene analysis within the *qshgd1* region was conducted by RNA-seq.
- The down-regulation of the Formin-like-5 transcript which possibly plays a role in cell division, was identified.

Abstract

In vivo doubled haploid (DH) technology is widely used in commercial maize (*Zea mays* L.) breeding. Haploid genome doubling is a critical step in DH breeding. In this study, inbred lines GF1 (0.65), GF3(0.29), and GF5 (0) with high, moderate, and poor spontaneous haploid genome doubling (SHGD), respectively, were selected to develop mapping populations for SHGD. Three QTL, *qshgd1*, *qshgd2*, and *qshgd3*, related to SHGD were identified by selective genotyping. With the exception of *qshgd3*, the source of haploid genome doubling alleles were derived from GF1. Furthermore, RNA-Seq was conducted to identify putative candidate genes between GF1 and GF5 within the *qshgd1* region. A differentially expressed formin-like protein 5 transcript was identified within the *qshgd1* region.

Keywords: maize, doubled haploid (DH), spontaneous haploid genome doubling (SHGD), RNA-seq

1. Introduction

Doubled haploid (DH) technology has been considered as one of the most innovative methods throughout maize (*Zea mays* L.) breeding history [1,2]. It is now quickly replacing other breeding methods due to the advantages seen in Europe , North America and China [3]. The routine process of DH technology based on in vivo haploid induction can be divided into three steps: haploid induction, haploid identification, and haploid genome doubling [4]. The efficiency of haploid induction and identification has been greatly improved in the past half-century [5]. This is due in part to an increase in induction rate of haploid inducers from 0.6% upon first discovery to current inducers typically averaging 8%-10% [4,6-9]. Recently, progress in understanding the genetics of haploid induction has been reported in various studies [8,10-15]. Eight QTL associated with haploid induction rate have been identified [8] and a 4-bp insertion at *ZmPLA1* has been shown to cause haploid induction [13,14]. To identify putative haploids, seeds are typically sorted manually based on embryo coloration (*R1-nj*). However, this step can be done using automated sorting procedures [4,5,16-18]. At present, haploid genome doubling routinely involves toxic reagents, such as colchicine for artificial genome doubling. Use of these chemicals can be harmful to personnel, the environment, and the plants being treated [19]. However, some lines exhibit haploid male fertility (HMF) without the use of chemical doubling

agents [20,21]. Haploid male fertility is as the most critical component of spontaneous haploid genome doubling (SHGD) [22].

Maize is a typical monoecious annual crop with separate male and female inflorescences. Genomes in haploid cells of both male and female inflorescences have to be doubled to create DH lines. The average fertility rate of female cell lines is generally high, exceeding 90% [20,21,23,24]. However, fertility rates in male cell lines typically range from 0-20% [25]. Low SHGD is a critical bottleneck for haploid genome doubling and varies substantially in different genetic backgrounds [20,21]. Wu et al. [24] reported that the range of SHGD is 9.8%-89.9% in Chinese elite inbred lines and 27.5%-85.5% in their single crosses. Maize inbred lines Yu87-1 and 4F1 showed high SHGD which were used to develop segregating populations for QTL analysis of SHGD [26]. Ma et al. [27] screened SHGD in a diversity panel of 481 maize lines crossed with "Mo17" and "Zheng58" and observed that SHGD ranged from 0-60% with a heritability of 0.65. A total of 14 significant single nucleotides polymorphisms (SNPs) associated with SHGD were identified by genome-wide association mapping (GWAS). Furthermore, De la Fuente et al. [22] evaluated SHGD in a total of 102 public and ex-PVP (expired Plant Variety Protection) inbred lines in U.S. germplasm. They observed SHGD ranging from 0-90% and significant differences in the effect of genotype, environment, and genotype by environment interactions.

Given the observations of De la Fuente et al. [22], a logical next step is to conduct a QTL mapping study. Mapping QTL can be conducted in either large populations of inbred lines that were selected to represent the overall genetic diversity of a species, or, by developing smaller bi-parental mapping populations through making specific crosses between inbred lines, resulting in related progeny where the inheritance of specific traits of interest can be measured. Within these smaller mapping populations, individuals with extreme phenotypic values are the most informative. Therefore, if using selective genotyping, a large number of individuals are phenotyped and only individuals in the lower and upper tails of the population are genotyped. A χ^2 test based on a 2 × 2 contingency table is then used to compare allele frequencies of individuals selected from the two tails [28]. In a recent QTL mapping study, four QTL controlling SHGD were identified on chromosomes 1, 3, 4, and 6 in selected haploid populations [26]. However, selective genotyping does have its limitations, such as being limited to only one

trait. To further enhance these QTL analyses, RNA-Seq can be deployed to identify differentially expressed candidate genes and/or transcripts within the region(s) identified by the QTL analysis. This approach has been used in wheat (*Triticum turgidum* ssp. *durum*) to identify candidate genes within QTL regions for grain protein content, and also in onion (*Allium cepa* L.) to identify candidate genes for cytoplasmic male sterility restoration [29,30].

In this study, QTL analysis was conducted to identify chromosomal regions affecting SHGD in two haploid populations developed by crossing a line that has a high doubling rate with two lines that have moderate and low doubling rates, respectively. Both χ^2 test and composite interval mapping (CIM) were used to detect QTL. RNA-seq analysis was used for candidate gene discovery within the major QTL, qshgd*1*. The objectives of this work were (1) to identify QTL associated with SHGD, and (2) identify candidate genes within these loci by using RNA-seq.

2. Materials and Methods

2.1. Plant materials

Three ex-PVP maize inbred lines herein referred to as: GF1 (with high SHGD ability), GF3 (with moderate SHGD ability) and GF5 with poor SHGD ability) were used to develop mapping populations. The three inbred lines were acquired from the United States Department of Agriculture Germplasm Resources Information Network (GRIN). GF1 was pollinated with GF3 and GF5 separately to form two F_1 populations (GF1/GF3 and GF1/GF5). These two F_1 populations were then induced by the 'RWS/RWK-76' inducer line. Putative haploid kernels were visually selected using the *R1-nj* color marker [31]. Haploids were grown in the field and fertile haploids were self-pollinated to produce DH lines.

2.2. Trait Scoring

In this study, SHGD was determined by anther emergence score (AES), described by Wu et al. [24] and Ren et al. [26]. AES ranges from 0 to 1, where 0 represents plants without anther emergence and 1 represents plants with more than 75% of anthers emerged across the tassel. During anthesis, haploids were evaluated on a daily basis for AES.

2.3. Field design

In the summer of 2015 in Ames, 1078 haploids of the GF1/GF3 population were grown at Iowa State University's Agricultural Engineering and Agronomy Farm (AF) in Boone, IA and SHGD was evaluated. A total of 60 haploids expressing the highest SHGD and 30 sterile haploids were selected based on their AES and were used to develop haploid populations P1 (fertile haploids) and P2 (sterile haploids), respectively. The 60 fertile haploids (P1) were self-pollinated to produce DH lines. 30 DH lines were randomly selected and pollinated with the haploid inducer 'RWS/RWK-76' during Winter 2015 in Graneros, Chile. In 2016, haploids of each of the 30 DH lines were sown at the North Central Regional Plant Introduction (PI) Station in Ames, IA, on May 13th (Environment 1) and at AF on May 26th (Environment 2) and June 2nd (Environment 3), respectively. A randomized complete block design with two replications was used. In each block, haploids were grown in 4-row plots with 25 plants/row. Supplemental water was provided as needed at PI, however, no supplemental water was supplied at AF. At the six-leaf stage, misclassified hybrids were visually removed from the plots.

2.4. Statistical analysis

Phenotypic data of the 30 DH lines were analyzed by a linear model: $Y_{ijk} = \mu + G_i + E_j + R_k + GE_{ij} + \varepsilon_{ijk}$, where μ is the overall mean, G_i is the effect of ith genotype, E_j is the effect of jth environment, R_k is the effect of kth replication, GE_{ij} is the effect of ith genotype within the jth environment, and ε_{ijk} is the effect of experimental error. Best linear unbiased prediction (BLUP) values were estimated for QTL mapping. All analyses were performed in SAS 9.2 (SAS Institute, Cary, NC, USA).

2.5. Genotyping

Young leaves of haploids and parental lines were harvested separately in the field. Genomic DNA was extracted using the method of Murray and Thompson [32] and purified. DNA concentration and quality were evaluated by NanoDrop^{TM} 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Single nucleotide polymorphism data were generated at the National Maize Improvement Center of China. Plants were genotyped using the maizeSNP6K chip and the Illumina Golden-Gate SNP genotyping platform. The maizeSNP6K chip is a subset of the Illumina MaizeSNP50 BeadChip [33,34]. Single nucleotide polymorphism (SNP) data were generated at the National Maize Improvement Center of China. Markers with more than 25% missing data, more than 10% heterozygosity, below 2.5% minor allele frequency

(MAF), and monomorphic markers between two parents were excluded by a SNP filtering process. The remaining high-quality SNPs were used for further analyses, and their genetic position was based on the IBM (intermated B73 × Mo17) genetic map [35-37].

2.6. QTL analysis

QTL analysis of SHGD was performed based on a tetrad grid χ^2 method where a 2 × 2 contingency table was used to test allele frequencies in the P1 and P2 haploid populations with Bonferroni correction [28,38]. SNPs with significantly different allele frequencies between P1 and P2 indicate putative QTL for SHGD in linkage disequilibrium with the marker. A CIM approach was also conducted to detect QTL for SHGD using WinQTL cartographer V2.5 [39] within the P1 and P2 populations. The BLUP values of AES for haploid plants representing the 30 DHs were used as phenotypic data for CIM. For the remaining 60 haploids, the phenotypic values were 0 if sterile and 1 if fertile. A LOD score of 3 obtained from 1,000 permutations at P=0.05 was used to declare QTL.

To confirm the mapping results, haploids derived from GF1/GF3 and GF1/GF5 were grown in Environments 1, 2, and 3, respectively. In each environment, 1200 plants per haploid population were planted. Haploids with AES of 1 were selected to verify the major QTL *qshgd1* on chromosome 5. A χ^2 test with Bonferroni correction was used to test the goodness of fit to the expected Mendelian segregation ratio of 1:1 of GF1 to GF3. If a marker showed segregation distortion (SD), it indicates presence of a QTL for SHGD near this marker.

2.7. RNA Isolation and RNA-seq

Haploid and diploid plants of GF1 and GF5 were used for RNA-seq. To construct RNA libraries, young tassels (~5 cm long) were collected in three biological replications, with ploidy being confirmed by flow cytometry. These tassels were then subjected to total RNA isolation using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) with no-colum DNase digestion following manufacturer's instructions. RNA concentration and RIN (RNA integrity number) were assessed using NanoDrop^{TM} 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) prior to RNA-seq. Poly(A) enrichment of total RNA was performed and the libraries were constructed using the stranded

library methods. RNA-seq was then performed using the Illumina HiSeq 3000 platform with paired-end 150-cycles.

2.8. RNA-seq analysis

Sequence data with adapters removed were subjected to trimming via sickle version 1.33 [40] with minimum window quality and read threshold set at 20 bases. These were then read into kallisto [41] and quantified using 100 bootstraps. Sequence bias correction was not used in kallisto due to the region of interest being highly repetitive. A Zea_mays.AGPv4 cDNA version 40 sequence file from Ensembl was used as a reference index for pseudoalignment. Differential expression analysis at both the gene and isoform levels were conducted using sleuth [42] and run with R software version 3.5.1 [43]. Differential expression was determined by Wald's test at P=0.05 using the Benjamini Hochberg correction for controlling false discovery rate to minimize the number of false positive candidate genes and transcripts.

3. Results

3.1. Assessment of SHGD

Field evaluation of maize SHGD confirmed that GF1 has high SHGD ability, whereas, GF3 and GF5 have poor SHGD ability (Fig. 1; Fig. 2). More than 90% of GF1 haploids showed anther emergence and most of them showed high male fertility. Whereas, about 70% of GF3 haploids showed anther emergence, most of them only exhibited few anthers emerged on a tassel. In contrast, all the haploids derived from GF5 were sterile. The AES of GF1 (0.65) was significantly higher (P = 0.05) than the values of GF3 (0.29) and GF5 (0). Haploids derived from GF1/GF3 and GF1/GF5, showed variable levels of male fertility and their AES were 0.43 and 0.23, respectively.

Anther emergence score BLUPs of haploids derived from 30 DH lines ranged from 0.02 to 0.99 (average 0.70) (Fig. 3). All DH lines showed different levels of SHGD ability, with the exception of one line. Analysis of variance (ANOVA) revealed significant genetic variance and genotype by environment interaction (P = 0.01) (Table 1).

3.2. Construction of the linkage map

The P1 and P2 haploid populations derived from the GF1/GF3 population were genotyped. A total of 5259 SNPs were detected and 2024 polymorphic markers covering all 10 chromosomes were obtained after the SNP filtering process. For these high-quality SNPs, MAF ranged from 0.23 to 0.5 (average 0.45). On average, 51.5% of the markers displayed GF1 alleles and 48.5% of the markers showed GF3 alleles, which is consistent with an expected 1:1 ratio for GF1 and GF3 alleles. For each population, the average number of recombinations were 15.6. The total length of the genetic map was 1773.9 cM with an average of 0.88 cM per marker (Supplemental Table 1).

3.3. QTL mapping for SHGD

A tetrad grid χ^2 test was used to test allele frequencies of all SNP markers using all 90 haploids. Ten regions on chromosomes 1, 3, 5, 6, 9, and 10, showed significant differences in allele frequencies between P1 and P2 (P = 0.05) (Fig. 4; Supplemental Table 2). However, only the locus on chromosome 5 (bin 5.04) remained after Bonferroni correction for multiple testing (Table 2).

Three QTL related to SHGD were detected in bin 5.03/04, 6.00/01 and 9.04/05 using CIM (Table 3). The QTL *qshgd1*, *qshgd2* and *qshgd3* explained 17.5%, 10.1% and 8.7% of the total phenotypic variation, respectively. For *qshgd1* and *qshgd2*, the SHGD increasing alleles were derived from GF1, whereas, for *qshgd3* the SHGD increasing allele came from GF3. All three QTL were detected by the tetrad grid χ^2 test without Bonferroni correction. However, only the major QTL, *qshgd1*, was confirmed by the tetrad grid χ^2 test after Bonferroni correction. To verify the *qshgd1* locus, SD of the markers (seven markers for GF1/GF3 and five markers for GF1/GF5) within the *qshgd1* region were detected using haploids with AES of 1 derived from GF1/GF3 and GF1/GF5 in three additional environments. All of these markers showed strong SD in both haploid populations across all three environments (Tables 4 and 5).

3.4. RNA-seq analysis for SHGD

The number of reads processed and mapped were all above 10,000,000 with the exception of one sample for haploid GF1, and the percentage of reads mapped were all above 80 percent (Supplemental Table 3). The sample that had an overall read count below 10,000,000 was not

removed because over 80 percent of the reads could still be mapped, and maintenance of three biological replications was desired.

At both the isoform and gene level, only when comparing GF1 with GF5 diploids, 10 genes and 8 transcripts were differentially expressed within the region of *qshgd1* (Tables 6 and 7). Of these differentially expressed genes, five were down-regulated at both the gene and transcript levels, while three transcripts were up-regulated at only the isoform level, and five genes were upregulated at only the gene level. However, at the gene level, none appeared to be directly related to the cell cycle (Table 7). Whereas, at the isoform level, the formin-like protein 5 (Zm00001d015326_T002) transcript was differentially expressed and may have a critical function during the cell cycle (Table 7).

4. Discussion

4.1. Strategies for QTL mapping of SHGD

Mixed model analysis revealed that SHGD is a complicated trait, which is affected by both environmental and genetic factors. In this study, although it is impossible to evaluate SHGD at the individual level due to incomplete penetrance, 29 out of the 30 DH lines showed variable levels of SHGD ability, indicating that the selection of haploids with AES of 1 was effective. In order to detect QTL related to SHGD efficiently, haploids with extreme phenotypic values were selected for QTL mapping. Detecting QTL by selective genotyping can decrease the number of individuals genotyped by increasing the number of individuals phenotyped. Analysis of QTL using selective genotyping has been carried out in numerous studies [26,38,44-46]. Yang et al. [38] identified two QTL, *qRfg1* and *qRfg2*, for resistance to *Gibberella* stalk rot by selecting 47 completely resistant and 47 highly susceptible plants in maize. The major QTL *qRfg1* was narrowed down to a ~500 kb interval.

In this study, 60 haploids with AES of 1 and 30 sterile haploids were genotyped. QTL analysis was performed using a tetrad grid χ^2 test to compare the frequencies of marker alleles between the fertile and sterile haploid populations and the CIM method. Three QTL, *qshgd1*, *qshgd2* and *qshgd3*, were identified on chromosome 5, 6, and 9 by the CIM method. The SHGD increasing alleles were derived from GF1 except for *qshgd3*. The major QTL *qshgd1* was detected by both methods and also confirmed in the haploid population of GF1/GF5. It explains

17.5% of the total phenotypic variation, although, this value may be overestimated [47-49]. In previous studies, Ren et al. [26] identified four QTL, *qhmf1*, *qhmf2*, *qhmf3*, and *qhmf4*, located on chromosomes 1, 3, 4, and 6, respectively, controlling SHGD in the selected haploid population of Yu87-1/Zheng58. However, none of the four QTL were detected in this study. A total of 14 significant SNPs located in bins 2.05, 2.06, 3.07, 5.05, 6.01, 7.05, 9.01, and 10.04 were detected by GWAS [27]. Only the loci in bins 5.05 and 6.01 were confirmed in our study; however, the major QTL *qshgd1* was only detected in our study. This may be due to the use of different mapping populations, different environments, and different markers. Fine mapping of *qshgd1* is required to clone the QTL and better understand the genetic basis of SHGD.

4.2. Candidate gene identification for SHGD

Although the mechanism underlying SHGD is still unknown, previous studies have shown that genes associated with meiosis may participate in SHGD in Arabidopsis (*Arabidopsis thaliana)* [50] and maize [26,51]. Cifuentes et al. [50] reported that MiMe genotypes (*spo11-1*/*osd1*/*rec8*) can restore the fertility of haploids and produce DH lines by turning meiosis into mitosis in Arabidopsis. Sugihara et al. [51] induced a first division restitution 1 (*fdr1*) mutation by sodium azide treatment in maize. Haploids derived from *fdr1* diploids exhibited high SHGD, resulting from first division restitution. In the study of Ren et al. [26], the major QTL *qhmf4* of SHGD was narrowed down to an ~800 kb region and the gene, absence of first division (*afd1*), is a candidate gene in this region.

In the present study, candidate gene analysis was conducted by RNA-seq. The down-regulation of the Formin-like-5 transcript (Table 7) could be responsible for the increasing in doubling events observed in GF1 because formins are proteins that are highly conserved among eukaryotes and are involved in numerous cellular processes requiring actin [52]. There are many formin homology (FH) families, however, FH5 is of particular interest in this study due to its down regulation in GF1 diploids (Table 7). In Arabidopsis, AtFH5 has been associated with the barbed end of actin filaments and nucleation during actin filament polymerization when studying them *in vitro* and has been isolated to the cell plate *in vivo* [53]. Actin is important for the arrangement of microtubule bundles and partially composes the preprophase band, delineating the future cell division plane, and later, the phragmoplast [54]. During the development of the cell plate there is a critical interaction between actin and formins, and it was observed that cell

plate expansion was halted between sister and non-sister chromatids in atfh5 mutant Arabidopsis lines, which suggests that AtFH5 is necessary for the completion of cytokinesis [53].

In cold stressed thermo-sensitive genic male sterile wheat, down-regulation of genes encoding formins and other critical cytoskeleton components affecting the dynamics of actin can lead to abnormal cytokinesis and restituted meiocytes [55]. It is quite possible that a reduction in ploidy is eliciting a stress response in the maize genotype GF1, similar to that observed in wheat; thus, leading to a form of overcompensation by GF1 when in a haploid state. Therefore, this in conjunction with already having the down regulated FH5 transcript at the diploid level, could promote the occurrence of an aberrant cell cycle. Furthermore, the depolymerization of actin filaments in human cancer cells by the chemical, Pectenotxin-2 has been observed to cause G2/M phase arrest and endoreduplication [56]. Thus, it is evident that actin and components influencing its behavior are critical for proper cell cycle function, and their reduction, or loss, can lead to a cell cycle stoppage and genome doubling.

To validate the functionality of candidate genes, future studies should be conducted using near-isogenic lines derived by crossing GF1 with GF5 and possibly other lines that do not possess SHGD. This should alleviate inaccuracies due to differences in RNA sequences between the non-SHGD and SHGD possessing lines. Furthermore, techniques such as gene editing could be utilized to edit the allele of our candidate genes in GF5 and possibly other non-SHGD lines, so that they possess the GF1 allele at those loci. Then, haploid progeny of the edited lines should be evaluated for SHGD and their fertility levels compared to GF1 and GF5 haploids.

In conclusion, the study of genetic basis of SHGD is important for DH breeding. In this study, three ex-pvp inbreds, GF3 and GF5 with poor SHGD and GF1 with high SHGD, were used to construct mapping populations. Three QTL, *qshgd1*, *qshgd2*, and q*shgd3,* were identified related with SHGD. The major QTL, *qshgd1* was detected by both χ^2 test and CIM. Candidate gene analysis within the *qshgd1* region was conducted by RNA-seq. The down regulation of the Formin-like-5 transcript which possibly plays a role in cell division, was identified. This study provides a better understanding of the genetic basis of SHGD.

Conflict of interest We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Author contribution statement

TL managed the project. JR, PW, SC, UF, and TL designed the experiment. JR, RX, BT, and KV performed the experiment. JR and NB performed data analysis. JR, PW, and NB wrote the paper.

Acknowledgements

We would like to thank Dr. Gerald De La Fuente for a preliminary experiment. This work was supported by USDA's National Institute of Food and Agriculture (grant numbers: IOW04314, IOW01018, IOW05510; NIFA award 2018-51181-28419), the Plant Sciences Institute, Crop Bioengineering Center, R.F. Baker Center for Plant Breeding, K.J. Frey Chair in Agronomy at Iowa State University, National Natural Science Foundation of China (grant number: 31560392), Xinjiang Natural Science Foundation (grant number: 2019D01A41), Science and Technology Assistance Xinjiang Program (grant numbers 2019E0209, 2018E02032) and Chinese Doctoral Foundation.

References

[1] L.M. Pollak, The history and success of the public-private project on germplasm enhancement of maize (GEM), Adv. Agron. 78 (2003) 45–87.

[2] G. Seitz, Interview published in Farmjournal, link provided at (2004) : http://www.greatlakeshybrids.com/performance/research-infromation/doubled-haploid-breeding-technology/.

[3] G. Seitz, The use of doubled haploids in corn breeding, Proceedings of the 41st Annual Illinois Corn Breeders' School, University of Illinois, Urbana–Champaign, IL, 2005, pp. 1-7.

[4] V. Prigge, A.E. Melchinger, Production of haploids and doubled haploids in maize, in: V.M. Loyola-Vargas, N. Ochoa-Alejo (Eds.), Plant Cell Culture Protocols, Humana Press, Totowa, NJ, 2012, pp. 161-172.

[5] A.E. Melchinger, W. Schipprack, T. Würschum, S. Chen, F. Technow, Rapid and accurate identification of in vivo-induced haploid seeds based on oil content in maize, Sci. Rep. 3 (2013) 21-29.

[6] F.K. Röber, G.A. Gordillo, H.H. Geiger, In vivo haploid induction in maize-performance of new inducers and significance of doubled haploid lines in hybrid breeding, Maydica 50 (2005) 275-283.

[7] P. Barret, M. Brinkmann, M. Beckert, A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize, Theor. Appl. Genet. 117 (2008) 581-594.

[8] V. Prigge, X.W. Xu, L. Li, R. Babu, S.J. Chen, G. Atlin, A.E. Melchinger, New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize, Genetics 190 (2011) 781-793.

[9] J. Ren, P.H. Wu, B. Trampe, X.L. Tian, T. Lübberstedt, S.J. Chen, Novel technologies in doubled haploid line development, Plant Biotechnol. J. 15 (2017) 1361-1370.

[10] X. Dong, X.W. Xu, S.J. Chen, Fine mapping of qhir1 influencing in vivo haploid induction in maize, Theor. Appl. Genet. 126 (2013) 1713-1720.

[11] X. Xu, L. Li, X. Dong, S.J. Chen, Gametophyte and zygotic selection leads to segregation distortion through in vivo induction of maternal haploid in maize, J. Exp. Bot. 64 (2013) 1083-1096.

[12] C.X. Liu, W. Li, Y. Zhong, X. Dong, H.X. Hu, X.L. Tian, L.L. Wang, B.J. Chen, C. Chen, A.E. Melchinger, S.J. Chen, Fine mapping of *qhir8* affecting in vivo haploid induction in maize, Theor. Appl. Genet. 128 (2015) 2507-2525.

[13] C.X. Liu, X. Li, D.X. Meng, Y. Zhong, C. Chen, X.W. Xu, B.J. Chen, W. li, L. Li, X.L. Tian, H.M. Zhao, W.B. Song, H.S. Luo, Q.H. Zhang, J.S. Lai, W.W. Jin, J.B. Yan, S.J. Chen, A 4 bp insertion at ZmPLA1 encoding a putative phospholipase A generates haploid induction in maize, Mol. Plant 10 (2017) 520-522.

[14] T. Kelliher, D. Starr, L. Richbourg, S. Chintamanani, B. Delzer, M.L. Nuccio, J. Green, Z.Y. Chen, J. McCuiston, W.L. Wang, T. Liebler, P. Bullock, B. Martin, MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction, Nature 542 (2017) 105-109.

[15] B.B. Wang, L. Zhu, B.B. Zhao, Y.P. Zhao, Y.R. Xie, Z.G. Zheng, Y.Y. Li, J. Sun, H.Y. Wang HY, Development of a Haploid-Inducer Mediated genome editing system for accelerating maize breeding, Mol Plant (2019) https://doi.org/10.1016/j.molp.2019.03.006.

[16] J. Liu, T. Guo, P. Yang, H. Wang, L. Liu, Z. Lu, X. Xu, J. Hu, Q. Huang, S. Chen, Development of automatic nuclear magnetic resonance screening system for haploid kernels in maize, Transactions of the Chinese Society of Agricultural Engineering 28 (2012) 233-236.

[17] B.W. Boote, D.J. Freppon, G.N. De La Fuente, T. Lübberstedt, B.J. Nikolau, E.A. Smith, Haploid differentiation in maize kernels based on fluorescence imaging, Plant Breeding 135 (2016) 439-445.

[18] G.N. De La Fuente, J.M. Carstensen, M.A. Edberg, T. Lübberstedt, J. Léon, Discrimination of haploid and diploid maize kernels via multispectral imaging, Plant Breeding 136 (2017) 50-60.

[19] K.R. Häntzschel, G. Weber, Blockage of mitosis in maize root tips using colchicine-alternatives, Protoplasma 241 (2010) 99-104.

[20] S.S. Chase, Production of homozygous diploids of maize from monoploids, Agron. J. 44 (1952) 263-267.

[21] S.T. Chalyk, Properties of maternal haploid maize plants and potential application to maize breeding, Euphytica 79 (1994) 13-18.

[22] G.N. De La Fuente, U.K. Frei, B. Trampe, J. Ren, M. Bohn, N. Yana, A. Verzegnazzi, S.C. Murray, T. Lübberstedt, A diallel analysis of a maize donor population response to in vivo maternal haploid induction II: Haploid Male Fertility, 2019.

[23] H.H. Geiger, M.D. Braun, G.A. Gordillo, Variation for female fertility among haploid maize lines, Maize Genetics Cooperation Newsletter 80 (2006) 28-29.

[24] P. Wu, J. Ren, X. Tian, T. Lübberstedt, W. Li, G. Li, X. Li, S. Chen, New insights into the genetics of haploid male fertility in maize, Crop Sci. 57 (2016) 637-647.

[25] D. Kleiber, V. Prigge, A.E. Melchinger, F. Burkard, F. San Vicente, G. Palomino, G.A. Gordillo, Haploid fertility in temperate and tropical maize germplasm, Crop Sci. 52 (2012) 623-630.

[26] J. Ren, P. Wu, X. Tian, T. Lübberstedt, S. Chen, QTL mapping for haploid male fertility by a segregation distortion method and fine mapping of a key QTL *qhmf4* in maize, Theor. Appl. Genet. 130 (2017) 1349-1359.

[27] H.L. Ma, G.L. Li, W. Tobias, Y. Zhang, D.B. Zheng, X.H. Yang, J.S. Li, W.X. Liu, J.B. Yan, S.J. Chen, Genome-wide association study of haploid male fertility in maize (*Zea mays* L.), Front. Plant Sci. 9 (2018) 974.

[28] A. Tenesa, P.M. Visscher, A.D. Carothers, S.A. Knott, Mapping quantitative trait loci using linkage disequilibrium: marker- versus trait-based methods, Behav. Genet. 35 (2005) 219-228.

[29] M. Trick, N.M. Adamski, S.G. Mugford, C.C. Jiang, M. Febrer, C. Uauy, Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat, BMC Plant Biol. 12 (2012) 14.

[30] S. Kim, C.W. Kim, M. Park, D. Choi, Identification of candidate genes associated with fertility

restoration of cytoplasmic male sterility in onion *(Allium cepa* L.) using a combination of bulked segregant analysis and RNA-seq, Theor. Appl. Genet. 128 (2015) 2289-2299.

[31] D.K. Nanda, S.S. Chase, An embryo marker for detecting monoploids of maize (*Zea mays* L.), Crop Sci. 6 (1966) 213-215.

[32] M.G. Murray, W.F. Thompson, Rapid isolation of high molecular weight plant DNA, Nucleic Acids Res. 8 (1980) 4321-4326.

[33] M.W. Ganal, G. Durstewitz, A. Polley, A. Bérard, E.S. Buckler, A. Charcosset, J.D. Clarke, E.M. Graner, M. Hansen, J. Joets, M.C. Le Paslier, M.D. McMullen, P. Montalent, M. Rose, C.C Schön, Q. Sun, H. Walter, O.C. Martin, M. Falque, A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome, Plos One, 6 (2011).

[34] Y. Meng, J. Li, J. Liu, H. Hu, W. Li, W. Liu, S. Chen, Ploidy effect and genetic architecture exploration of stalk traits using dh and its corresponding haploid populations in maize, BMC Plant Biol. 16 (2016) 50.

[35] M. Lee, N. Sharopova, W.D. Beavis, D. Grant, M. Katt, D. Blair, A. Hallauer, Expanding the genetic map of maize with the B73 x Mo17 (IBM) population, Plant Mol. Bio. 48 (2002) 453-461.

[36] F. Wei, E. Coe, W. Nelson, A.K. Bharti, F. Engler, E. Butler, H.R. Kim, J.L. Goicoechea, M. Chen, S. Lee, G. Fuks, H. Sanchez-Villeda, S. Schroeder, Z.W. Fang, M. McMullen, G. Davis, J.E. Bowers, A.H. Paterson, M. Schaeffer, J. Gardiner, K. Cone, J. Messing, C. Soderlund, R.A. Wing, Physical and genetic structure of the maize genome reflects its complex evolutionary history, PLOS Genet. 3 (2007) e123.

[37] A. Vanous, C. Gardner, M. Blanco, A. Martinschwarze, A.E. Lipka, S. Flint-Garcia, M. Bohn, J. Edwards, T. Lübberstedt, Association mapping of flowering and height traits in germplasm enhancement of maize doubled haploid (GEM-DH) lines, Plant Genome 11 (2018) 170083.

[38] Q. Yang, G.M. Yin, Y.L. Guo, D.F. Zhang, S.J. Chen, M.L. Xu, A major QTL for resistance to Gibberella stalk rot in maize, Theor. Appl. Genet. 121 (2010) 673-687.

[39] C.J. Basten, B.S. Weir, Z.B. Zeng, QTL cartographer: a reference manual and tutorial for QTL mapping. Department of Statistics, North Carolina State University, Raleigh, NC, 1997.

[40] N.A. Joshi, J.N. Fass, Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). UC Davis Bioinformatics Core, 2011.

[41] N.L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification, Nat. Biotechnol. 34 (2016) 525-528.

[42] H. Pimentel, N.L. Bray, S. Puente, P. Melsted, L. Pachter, Differential analysis of RNA-seq incorporating quantification uncertainty, Nat. Methods 14 (2017) 687-690.

[43] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2018.

[44] R. Venuprasad, M.E. Bool, C.O. Dalid, J. Bernier, A. Kumar, G.N. Atlin, Genetic loci responding to two cycles of divergent selection for grain yield under drought stress in a rice breeding population, Euphytica

167 (2009) 261-269.

[45] F. Zhang, X. Ma, Y. Gao, X. Hao, Z. Li, Genome-wide response to selection and genetic basis of cold tolerance in rice (*Oryza sativa* L.), BMC Genet. 15 (2014) 55.

[46] Y. Cui, F. Zhang, J. Xu, Z. Li, S. Xu, Mapping quantitative trait loci in selected breeding populations: a segregation distortion approach, Heredity 115 (2015) 538–546.

[47] S.D. Tanksley, Mapping polygenes, Annu. Rev. Genet. 27 (1993) 205-233.

[48] C.C. Schön, H.F. Utz, S. Groh, B. Truberg, S. Openshaw, A.E. Melchinger, Quantitative trait locus mapping based on re-sampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits, Genetics 167 (2004) 485-498.

[49] P.M. Visscher, M.E. Goddard, Prediction of the confidence interval of quantitative trait loci location, Behav. Genet. 34 (2004) 477-482.

[50] M. Cifuentes, M. Rivard, L. Pereira, L. Chelysheva, R. Mercier, Haploid meiosis in Arabidopsis: double-strand breaks are formed and repaired but without synapsis and crossovers, Plos One 8 (2013) e72431.

[51] N. Sugihara, T. Higashigawa, D. Aramoto, A. Kato, Haploid plants carrying a sodium azide-induced mutation (fdr1) produce fertile pollen grains due to first division restitution (FDR) in maize (*Zea mays* L.), Theor. Appl. Genet. 126 (2013) 2931-2941.

[52] F. Cvrckova, Formins and Membranes: anchoring cortical actin to the cell wall and beyond, Front. Plant Sci. 4 (2013) 436.

[53] M. Ingouff, J.N.F. Gerald, C. Guérin, H. Robert, M.B. Sørensen, D.V. Damme, D. Geelen, L. Blanchoin, F. Berger, Plant formin AtFH5 is an evolutionarily conserved actin nucleator involved in cytokinesis, Nat. Cell Biol. 7 (2005) 374-380.

[54] S. Li, T. Sun, H. Ren, The functions of the cytoskeleton and associated proteins during mitosis and cytokinesis in plant cells, Front. Plant Sci. 6 (2015) 282.

[55] N. De Storme, D. Geelen, Cytokinesis in plant male meiosis, Plant Signal Behav. 8 (2013) 23394.

[56] D. Moon, M. Kim, S. Kang, K. Lee, M. Heo, K. Choi, Y.H. Choi, G. Kim, Induction of G2/M arrest, endoreduplication, and apoptosis by actin depolymerization agent pextenotoxin-2 in human leukemia cells, involving activation of ERK and JNK, Biochem. Pharmacol. 76 (2008) 312-321.

Figure legends

Fig.1 Haploid male fertile tassels of parental lines GF1 (a) and GF3 (b)

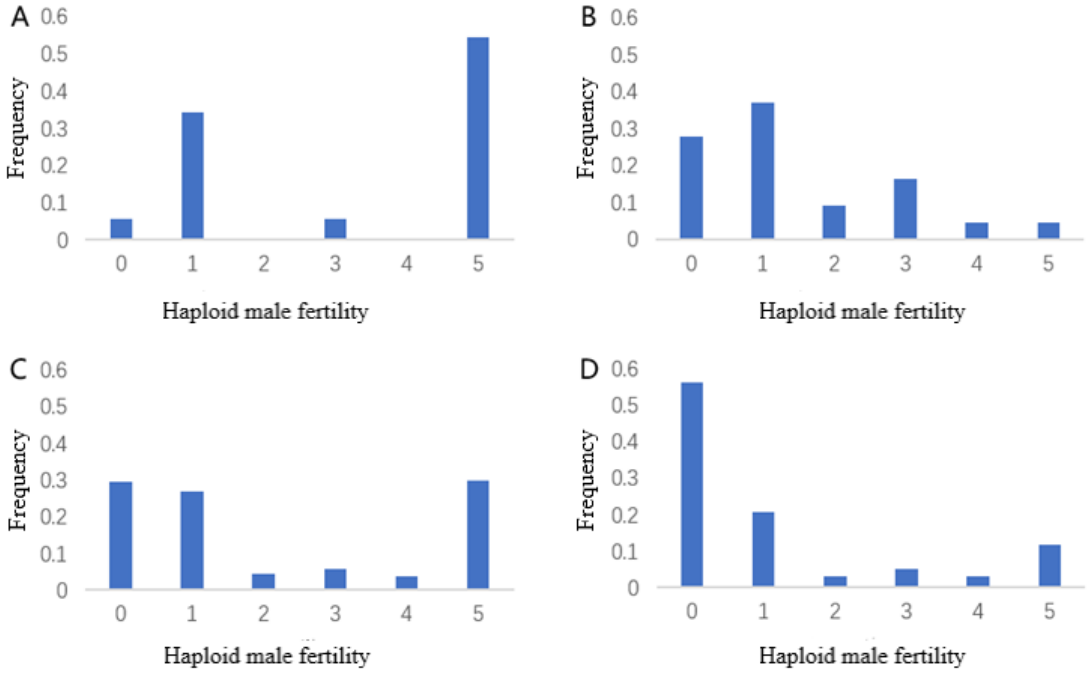

Fig. 2 Percentage of haploid male fertility for the two parental lines, GF1 and GF3, and the F_1 progeny from both the GF1/GF3 and GF1/GF5 populations. (A) High SHGD parental line GF1; (B) Low SHGD parental line GF3; (C) the GF1/GF3 F_1 hybrids; (D) the GF1/GF5 F_1 hybrids

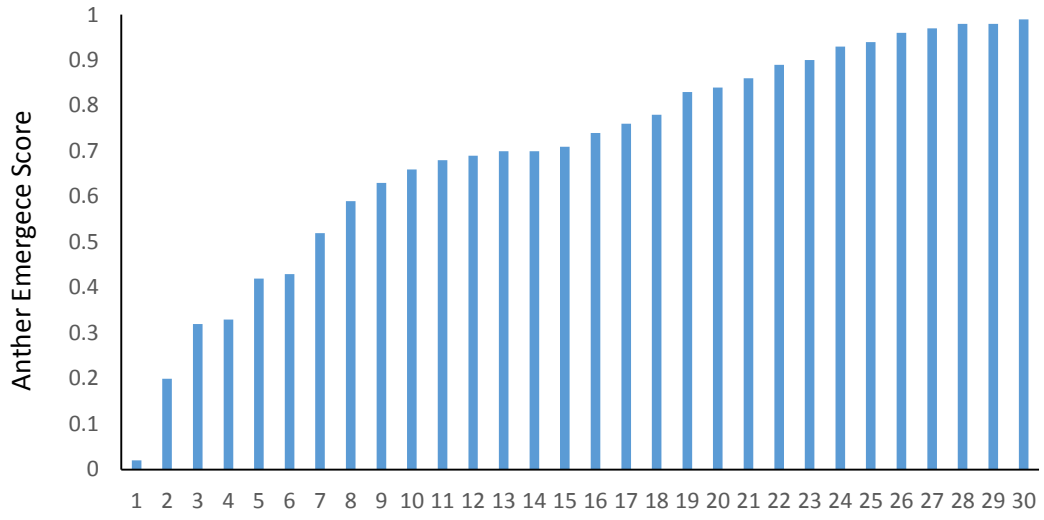

Fig. 3 BLUPs of 30 DH lines for anther emergence score (AES)

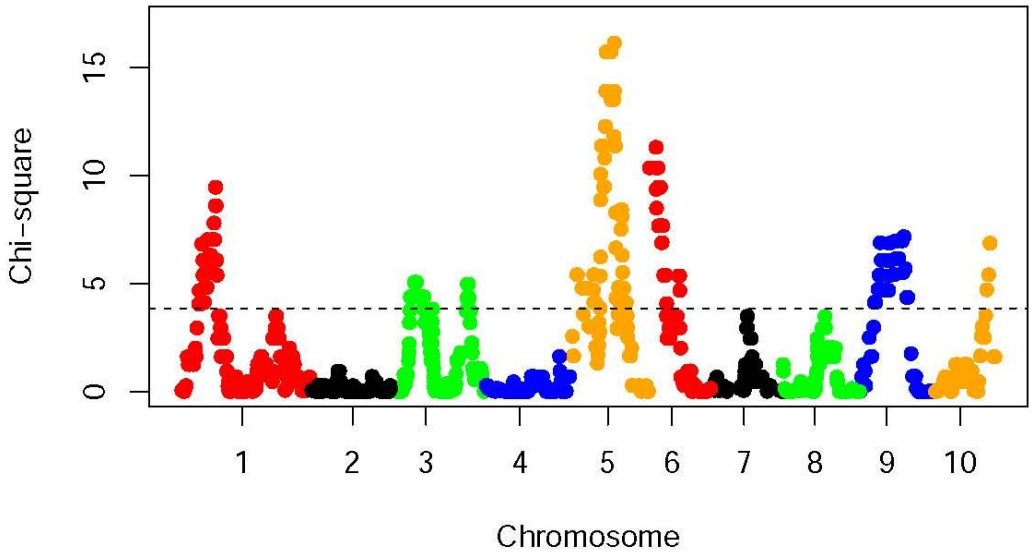

Fig. 4 X^2 test for the 90 haploids. The horizontal line represents the threshold $X^2_{0.05}(1)=3.84$.

Tables

Table 1 Analysis of variance (ANOVA) for haploid anther emergence score (AES).

Source of variation	Mean squares
Genotype	0.236**
Environment	0.013
Replication	0.004
Genotype X Environment	0.010**
Error	0.005

Note : ** p = 0.01

Table 2 SNPs showing a significant correlation of genotype and haploid male fertility after Bonferroni correction

Chromosome	Bins	SNPs	Physical position (Mb)	Allele frequencies of GF1			P value	Adjusted P value[c]
				In P1[a]	In P2[b]	χ^2		
5	5.04	PZE-105080300	92.72	0.88	0.50	13.90	0.0002	0.0923
5	5.04	PZE-105081158	94.23	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	SYN8752	96.57	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105082200	96.85	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105083471	100.02	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105083966	101.18	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105084020	101.30	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105085307	106.10	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105087106	112.00	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105087513	113.73	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105088164	117.10	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105089537	123.90	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105090165	126.08	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105090424	126.82	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105090714	127.60	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105091620	129.99	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105091854	130.50	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105092394	131.74	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105093092	134.19	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105093183	134.58	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105093360	135.02	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105093466	135.30	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105093615	135.64	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	SYN32229	135.84	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105093817	136.45	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105094031	137.15	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105094593	138.51	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105094653	138.75	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105094785	138.94	0.90	0.50	15.72	<0.0001	0.0352

5	5.04	PZE-105094796	139.08	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105094949	139.36	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105095123	139.76	0.90	0.50	15.72	<0.0001	0.0352
5	5.04	PZE-105095399	140.09	0.90	0.53	13.51	0.0002	0.1140

[a]P1: haploids with AES of 1

[b]P2: sterile haploids

[c]Adjusted p value was obtained by Bonferroni correction

Table 3 Analysis of QTL for haploid male fertility and their genetic effect by CIM method

Name of QTL	Chromosome	Bin	LOD value	Additive	R^2(%)	Source of SHGD allele
qshgd1	5	5.03/04	6.02	0.24	17.53	GF1
qshgd2	6	6.00/01	3.67	0.25	10.11	GF1
qshgd3	9	9.04/05	3.22	-0.14	8.70	GF3

Table 4 Analysis of segregation distortion (SD) in the selected haploid population of GF1/GF3 on chromosome 5

Environment[a]	Marker	Genotype[b] Number		χ^2	P value	Adjusted P value[c]
		A	B			
E1	umc2373	86	18	44.46	<0.0001	<0.0001
	umc1629	88	16	49.85	<0.0001	<0.0001
	913-4	87	17	47.12	<0.0001	<0.0001
	bnlg2323	83	21	36.96	<0.0001	<0.0001
	umc1092	83	21	36.96	<0.0001	<0.0001
	umc1348	82	22	34.62	<0.0001	<0.0001
	517-3	72	32	15.38	0.0001	0.0018
E2	umc2373	122	35	48.21	<0.0001	<0.0001
	umc1629	124	33	52.75	<0.0001	<0.0001
	913-4	124	33	52.75	<0.0001	<0.0001
	bnlg2323	117	40	37.76	<0.0001	<0.0001
	umc1092	115	42	33.94	<0.0001	<0.0001
	umc1348	115	42	33.94	<0.0001	<0.0001
	517-3	97	60	8.72	0.0031	0.0661
E3	umc2373	108	15	70.32	<0.0001	<0.0001
	umc1629	109	14	73.37	<0.0001	<0.0001
	913-4	108	15	70.32	<0.0001	<0.0001
	bnlg2323	106	17	64.40	<0.0001	<0.0001
	umc1092	101	22	50.74	<0.0001	<0.0001
	umc1348	97	26	40.98	<0.0001	<0.0001

	517-3	84	39	16.46	<0.0001	<0.0001

[a] E1:5.4 in Plant Introduction Station

E2:5.13 in Ames Agronomy Farm

E3:5.16 in Ames Agronomy Farm

[b] A: GF1 allele

B: GF3 allele

[c] Adjusted *P* value was obtained by Bonferroni corrections for multiple testing

Table 5 Analysis of segregation distortion (SD) in the selected haploid population of GF1/GF5 on chromosome 5

Environment[a]	Marker	Genotype[b] Number		χ^2	P value	Adjusted P value[c]
		A	B			
E1	umc2373	51	16	18.28	<0.0001	0.0003
	umc1629	51	16	18.28	<0.0001	0.0003
	913-7	51	16	18.28	<0.0001	0.0003
	umc1332	47	20	10.88	0.0010	0.0146
	umc1221	46	21	9.33	0.0023	0.0338
E2	umc2373	33	7	16.90	<0.0001	0.0006
	umc1629	33	7	16.90	<0.0001	0.0006
	913-7	33	7	16.90	<0.0001	0.0006
	umc1332	30	10	10.00	0.0016	0.0235
	umc1221	31	9	12.10	0.0005	0.0076
E3	umc2373	29	5	16.94	<0.0001	0.0006
	umc1629	29	5	16.94	<0.0001	0.0006
	913-7	29	5	16.94	<0.0001	0.0006
	umc1332	25	9	7.53	0.0061	0.0910
	umc1221	25	9	7.53	0.0061	0.0910

Note:

[a] E1:5.4 in Plant Introduction Station

E2:5.13 in Ames Agronomy Farm

E3:5.16 in Ames Agronomy Farm

[b] A: GF1 allele

B: GF5 allele

[c] Adjusted *P* value was obtained by Bonferroni corrections for multiple testing

Table 6 Significant candidates at the gene level and P = 0.05 following Benjamini-Hochberg adjustment within the chromosome 5 region of 84,495,519 – 130,789,375 bases. Also included are the Ensembl gene IDs, descriptive statistics, and ploidy/genotype combinations for each candidate.

Description	Gene ID	Adjusted P value	Log 2 fold change	Std. error	Biological variance	Technical variance	Ploidy
Uncharacterized Protein	Zm00001d015502	0.001	-3.611	0.837	4.478	0.111	Diploid
Probable cinnamyl alcohol dehydrogenase	Zm00001d015618	0.009	0.665	0.178	0.168	0.001	Diploid
Chlorophyll a-b binding protein 6 chloroplastic	Zm00001d015385	0.009	0.706	0.189	0.193	0.0001	Diploid
Uncharacterized Protein	Zm00001d015532	0.009	-3.42	0.922	3.517	0.022	Diploid
Protein RRC1	Zm00001d015495	0.015	-0.704	0.198	0.19	0.004	Diploid
Myb family transcription factor APL	Zm00001d015407	0.027	-1.01	0.307	0.382	0.01	Diploid
Cytochrome P450 CYP71K14	Zm00001d015592	0.035	2.405	0.752	2.414	0.003	Diploid
dentin sialophosphoprotein-related	Zm00001d015824	0.04	-2.972	0.949	3.53	0.347	Diploid
Clathrin heavy chain 2	Zm00001d015676	0.046	2.205	0.718	2.078	0.091	Diploid
Uncharacterized Protein	Zm00001d015880	0.01	0.822	0.224	0.234	0.003	Diploid

Table 7 Significant candidates at the isoform level and P = 0.05 following Benjamini-Hochberg adjustment within the chromosome 5 region of 84,495,519 – 130,789,375 bases. Also included are the Ensembl Transcript IDs, descriptive statistics, and ploidy/genotype combinations for each candidate.

Description	Gene ID	Adjusted P value	Log 2 fold change	Std. error	Biological variance	Technical variance	Ploidy
Aconitate hydratase 3 mitochondrial	Zm00001d015497_T008	0	-6.498	1.327	13.51	0.853	Diploid
Formin-like protein 5	Zm00001d015326_T002	0.001	-6.102	1.342	11.182	0.665	Diploid
Glyceraldehyde-3-phosphate dehydrogenase%2C cytosolic	Zm00001d015383_T001	0.01	5.255	1.386	9.5	0.721	Diploid
Core-2/I-branching beta-16-N-acetylglucosaminyltransferase family protein	Zm00001d015429_T001	0.006	-4.814	1.1	6.958	0.018	Diploid
Protein phosphatase 2C isoform gamma	Zm00001d015504_T007	0.013	-4.949	1.341	9.507	1.331	Diploid
Pectin lyase-like superfamily protein	Zm00001d015825_T001	0.02	-1.28	0.361	0.547	0.01	Diploid
Cytochrome P450 CYP71K14	Zm00001d015592_T001	0.048	2.405	0.752	2.414	0.003	Diploid
Chlorophyll a-b binding protein 6 chloroplastic	Zm00001d015385_T001	0.045	0.706	0.219	0.193	0.001	Diploid