agriculture

MDPI

*Article*

# Genetic Analysis of Maternal Haploid Inducibility for In Vivo Haploid Induction in Maize

**Jiaojiao Ren, Xiaoyu Zhang, Zongze Li and Penghao Wu ***

College of Agronomy, Xinjiang Agricultural University, Urumqi 830052, China; renjiaojiao789@sina.com (J.R.); zhangxiaoyu5581@163.com (X.Z.); lizongze0517@163.com (Z.L.)
* Correspondence: craie788@126.com

**Abstract:** Doubled haploid (DH) technology based on in vivo haploid induction has gradually become the key technology in modern maize breeding. The ability of maternal germplasm to be induced into haploids, inducibility, varies among genotypes. To dissect the genetic basis of maternal haploid inducibility (MHI), an $F_2$ population derived from inbred lines B73 and Zheng58 was used for single environment QTL analysis and QTL by environment interaction analysis. The mapping population was genotyped by the 48K liquid-phase hybridization probe capture technique and phenotyped in multi-environment trials for MHI. A total of ten QTLs located on chromosome bins 4.05, 4.09, 5.05/5.06, 6.07, 7.00, 7.01, 7.02, 7.03, 9.02, and 10.06 were identified for MHI. The PVE value of each QTL ranged from 4.79% to 10.01%. The QTL *qMHI5* is a stable QTL identified in JSH, HN, and across environments with the highest PVE value of 10.01%. Three QTLs, *qMHI4-1*, *qMHI5*, and *qMHI 9-1*, were detected by both methods. Three genes, Zm00001d017366, Zm00001d017420, and Zm00001d017432, involved in seed development were the most likely candidate genes. This study provides valuable information for the genetic basis of MHI.

**Keywords:** maize; haploid induction rate; maternal haploid inducibility; QTL mapping



**Citation:** Ren, J.; Zhang, X.; Li, Z.; Wu, P. Genetic Analysis of Maternal Haploid Inducibility for In Vivo Haploid Induction in Maize. *Agriculture* **2022**, *12*, 845. https://doi.org/10.3390/agriculture12060845

Academic Editor: Panagiotis Madesis

Received: 12 May 2022
Accepted: 8 June 2022
Published: 11 June 2022



## 1. Introduction

By successfully taking advantage of hybridization technology, maize is the most widely used commercial hybrid crop in the world. To comprise the hybrid, inbred line development is the first step that requires seven or more generations through traditional breeding technology. Therefore, shorten the breeding cycle and speeding up the breeding process is gradually being realized by lots of breeders. Doubled haploid (DH) breeding technology has become an important way to breed homozygous lines quickly and efficiently. It could produce highly homozygous DH lines in only two generations, with the process of haploid induction, identification, and doubling [1–3]. Because of its advantage in greatly reducing the time and cost, DH technology is able to improve the overall breeding process. Haploid breeding has been applied by many foreign companies at a large scale [4].

Haploid inducer line is a special germplasm to produce haploid by crossing with other materials. This way, the haploid production efficiency is much higher than that obtained by natural hybridization [5]. In 1956, the first haploid inducer Stock6 with haploid induction rate (HIR) of 2–3% was discovered [6]. After years of haploid inducer line improvement by breeders, the HIR has been continuously improved from ~2% to over 10% [7–9]. RWS [7] and UH400 [5] developed from KEMS (HIR of ~6%) have the HIR of 8.65–13.39% and 8–15%, respectively. Tropical inducer candidates (TIC) developed from crosses between temperate inducers (HIR of 8–10%) and tropical CIMMYT maize lines have the HIR up to 10% [10]. A high oil inducer line, CHOI4, derived from CAU2 (HIR of ~10%) and Beijing High Oil population has a HIR of ~10.59–20.84% and average kernel oil content of ~11% [11]. These high HIR inducers have been widely used in DH breeding programs, and significantly improved the efficiency of DH line production.

However, haploid inducer is not the only genetic factor impact HIR. The other key factor is the genotype of maternal germplasm [12,13], which could be found in many previous research studies. In the study of Prigge [12], there were significant genotypic differences among inducers and maternal germplasm for HIR, but no interactions between the two factors. A diallel analysis of a maize donor population was used to detect the maternal influence on HIR by De la Fuente et al. [13]. The HIR ranged from 7.3% to 14.6% by pollinate 30 $F_1$ with RWS × RWK-76 inducer.

Inducibility, first proposed by Wu et al. [14], was defined as the ability of maternal germplasm to be induced into haploids. In an $F_2$ population, Wu et al. [14] identified two QTLs for MHI on chromosomes 1 and 3, explaining 14.7% and 23.12% of the total phenotypic variation, respectively. Benjamin et al. [15] reported four QTLs for MHI on chromosomes 2, 4, 5, and 8 with a total phenotypic variation explained (PVE) value of 37.4%. Li [16] detected 13 QTLs related to MHI, of which *qMHI9-3* and *qMHI10-3* with PVE values of 22.14 and 17.44% were major QTLs. The genome-wide association study (GWAS) for MHI was conducted using 671 tropical inbreed lines in two association mapping panels [17]. A total of 29 significant SNPs distributed over all 10 chromosomes were identified.

Only several genetic analyses had been performed on maternal haploid inducibility (MHI) in different genetic backgrounds [14–16]. Most of the QTLs were reported in a special population or in a special environment, there was no stable QTL identified in different environments. Therefore, the stable QTL region identified in a cross environment needs to be individually explored for molecular marker-assisted selection. In this study, an $F_2$ population including 200 $F_{2:3}$ families was used to perform QTL mapping. The population was phenotyped for MHI in three different environments and genotyped by the 48K liquid-phase hybridization probe capture technique. The objectives of this study were to: (1) evaluate the phenotypic variation for MHI; (2) identify QTL conferring MHI; (3) explore the stable QTL region in different environments; and (4) predict candidate genes for major QTLs of MHI.

## 2. Materials and Methods

### *2.1. Materials and Experimental Design*

In this study, 200 $F_{2:3}$ families developed from a bi-parental cross between maize inbred lines B73 (high MHI) and Zheng58 (low MHI) were used. The high-oil maize haploid inducer line CHOI3 [18] with HIR of 8.72% was used as male parent to induce haploids.

The $F_{2:3}$ families were planted at the Jiushenghe Experimental Station (N 43°57′3.2″ E 87°13′19.3″, JSH) in Changji, Xinjiang and the Erliugong Town Experimental Station (N 44°3′10.0″ E 87°10′23.1″, ELG) in Urumqi, Xinjiang in the summer of 2020, and at the Ledong Experimental Station, Hainan (N 18°45′0.2″ E 109°10′24.1″, HN) in the winter of 2020. A randomized complete block design with two replications at each site was used. In each block, plants were planted in a single row with a row length of 2.5 m, a plant distance of 0.25 m, and a row distance of 0.60 m. CHOI3 was planted on three different planting dates to guarantee sufficient pollen for haploid induction. The $F_{2:3}$ family was pollinated by CHOI3.

### *2.2. Phenotypic Evaluation*

The R1-nj seed-based color marker system was used for haploid selection. All the kernels were classified into two main types: diploid kernels with purple endosperm and embryo, and haploid kernels with purple endosperm and colorless embryo (Figure 1) [19]. MHI was evaluated using the following formula:

$$\text{MHI}(\%) = \frac{\text{number of haploid kernels}}{\text{total number of kernels}} \times 100\%$$

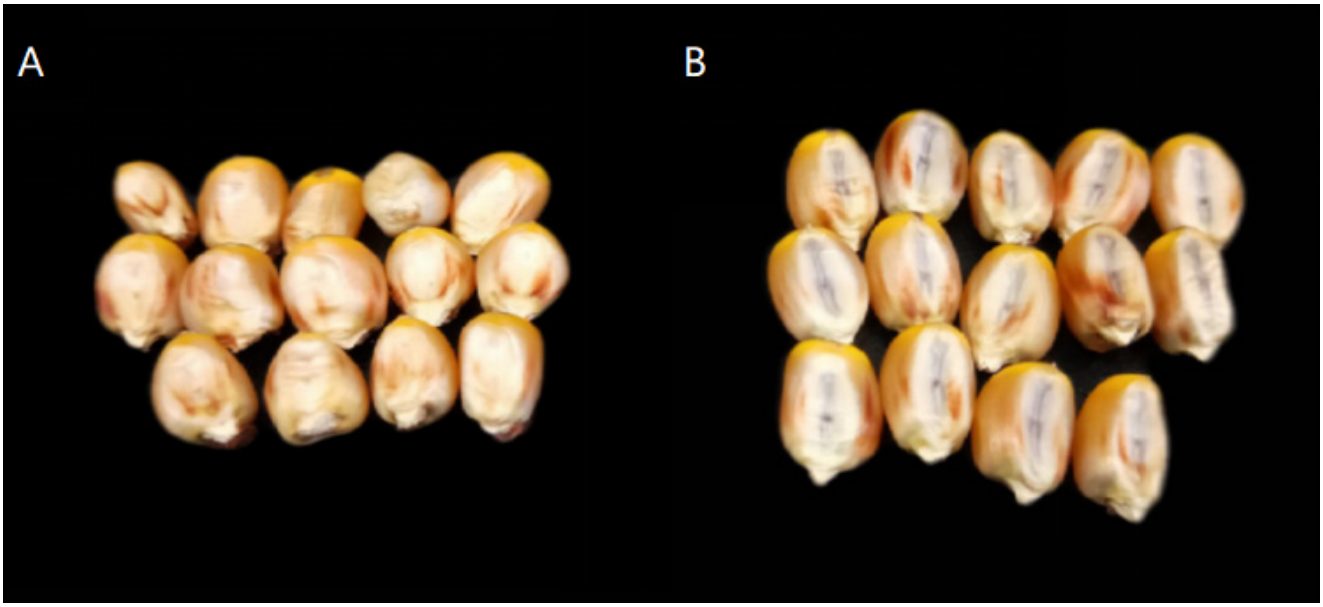


**Figure 1.** Two different types of hybrid induction kernels: (**A**) haploid kernels; (**B**) diploid kernels.

### *2.3. Data Analysis*

The phenotypic data were analyzed using a mixed linear model:

$$Y_{ijk} = \mu + G_i + E_j + R_{k(j)} + GE_{ij} + \varepsilon_{ijk}$$

where $Y_{ijk}$ refers to the phenotypic value of the *i*th genotype in the *j*th environment in the jkth replicate, μ refers to the overall phenotypic mean, $G_i$ refers to the effect of *i*th genotype, $E_j$ refers to the effect of the *j*th environment, $R_{k(j)}$ refers to the effect of the kth replication nested on the *j*th environment, $GE_{ij}$ refers to the interaction effect of the *i*th genotype with the *j*th environment, and $\varepsilon_{ijk}$ refers to the effect of the experimental error. Analysis of variance was performed using the AOV (ANOVA of multi-environmental trials) function of the QTL IciMapping software [20]. The broad-sense heritability ($h^2$) was calculated as follows:

$$h^2 = \sigma_G^2 / \left(\sigma_G^2 + \sigma_{GE}^2 / E + \sigma_\varepsilon^2 / ER\right)$$

where $\sigma_G^2$ refers to the genotypic variance, $\sigma_{GE}^2$ refers to the genotype × environment interaction variance, $\sigma_\varepsilon^2$ refers to the error variance, *E* refers to the number of environments, and *R* refers to the number of replications. The Best Linear Unbiased Estimator (BLUE) values for single and combined environments were evaluated by the META-R Version 6.04 software [21]. The violin plots were plotted by R software [22].

### *2.4. Genotyping*

Young leaves of the $F_2$ population and parental lines were sampled. DNA was extracted according to the CTAB method [23] and genotyped using the 48K liquid-phase hybridization probe capture technique (China golden marker, Beijing, China). BWA 0.7.17 software [24] was used to anchor reads to the Maize B73_RefGen_v4 reference genome. For each line, 62,504 SNPs evenly distributed on 10 maize chromosomes were obtained. The SNP data were strictly filtered using vcftools software. The filtering parameters were set as follows: MAF = 0.05, minDP = 10, maxDP = 300, minGQ = 30, minQ = 30, and max-missing = 0.8. A total of 1855 polymorphic SNPs were used for further analysis.

### *2.5. Linkage Map Construction and QTL Analysis*

QTL iciMapping V4.2 software was used for linkage map construction and QTL analysis. Linkage map construction was performed using the "MAP" function. Kosambi

mapping function was used to convert recombination frequencies into centi Morgans (cM). The SNPs were first ordered by physical position and then rippled by the sum of adjacent distances (DIS) with a window size of 5. QTL in single environment were identified using the "BIP" function with the inclusive composite interval mapping of additive and dominant (ICIM-ADD) method. QTLs by environment interactions for multi-environment trials were detected by the "MET" function with the "ICIM-ADD" method. The LOD thresholds were 2.5 and 5.5 for the single environment QTL analysis and QTL by environment interaction analysis, respectively. The dominant (D) effect, additive (A) effect, and phenotypic variation explained (PVE) by each QTL were estimated. According to the ration of D and A (D/A), gene action was classified into four types: A = 0 to 0.20, partial dominance (PD) = 0.21 to 0.80, D = 0.81 to 1.20, and over-dominance (OD) > 1.20. The candidate gene identification and annotation were based on MaizeGDB [25].

## 3. Results

### *3.1. Phenotypic Analysis of Maternal Haploid Inducibility*

The phenotypic variation for MHI in the three environments is shown in Table 1 and Figure 2. MHI varied across environments and the means in different environments were similar, ranging from 9.01% in ELG to 11.10% in HN. For each environment, there were abundant phenotypic variations. The greatest variation occurred in ELG, where the MHI ranged from 2.56 to 16.62. The smallest variation was detected in the combined environment, where MHI ranged from 5.64 to 14.12. All the values of skewness were negative, indicating that some lines had quite low MHI.

**Table 1.** Descriptive statistics of maternal haploid inducibility (%) in the $F_2$ population in three environments.

| Environment | Mean | Min | Max | Median | SD | Kurtosis | Skewness |
|---|---|---|---|---|---|---|---|
| JSH | 10.46 | 2.94 | 14.45 | 10.73 | 2.04 | 0.26 | −0.61 |
| ELG | 9.01 | 2.26 | 16.62 | 8.85 | 2.68 | −0.32 | −0.02 |
| HN | 11.10 | 3.11 | 15.75 | 11.31 | 2.41 | −0.16 | −0.47 |
| Combine | 10.19 | 5.64 | 14.12 | 10.29 | 1.76 | −0.26 | −0.32 |

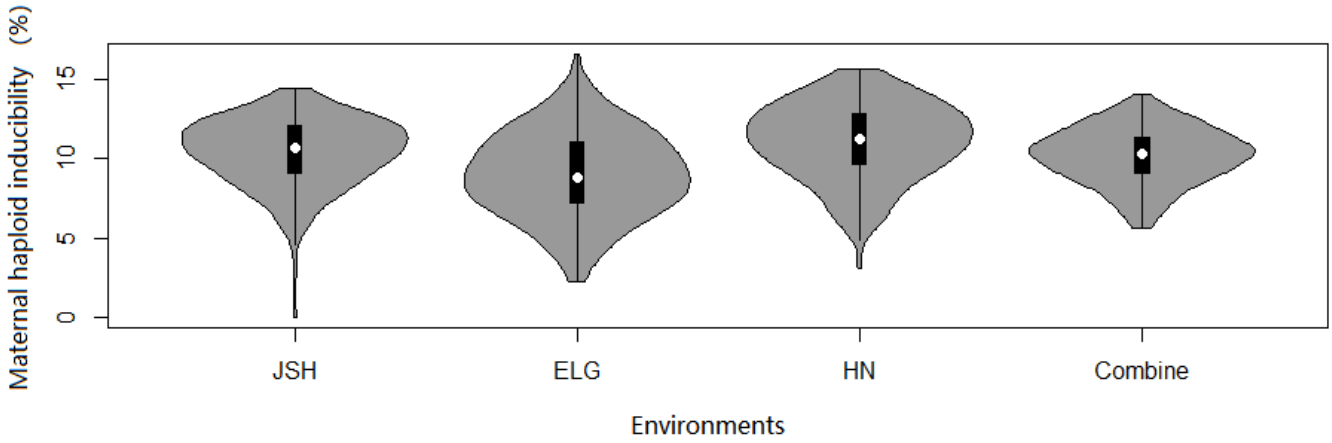


**Figure 2.** Violin plots of maternal haploid inducibility in the $F_2$ population.

ANOVA analysis results revealed significant differences among genotypes, environments, and genotype by environment interaction (Table 2). The heritability of MHI was 0.40 indicating that MHI was a complex quantitative trait easily affected by environments.

**Table 2.** Analysis of variance (ANOVA) and broad-sense heritability ($H^2$) for maternal haploid inducibility (%) in the $F_2$ population.

| Source of Variation | Degree of Freedom | Sum of Square | Mean of Square | F-Value | *p*-Value | $H^2$ |
|---|---|---|---|---|---|---|
| Environment | 2 | 20.46 | 10.23 | 3.25 | <0.01 ** | |
| Genotype | 199 | 4008.93 | 20.14 | 2.74 | <0.01 ** | |
| GXE | 398 | 8983.60 | 22.57 | 3.12 | <0.01 ** | 0.40 |
| Residual | 597 | 3117.01 | 5.21 | | | |
| Total | 1195 | 16,130.00 | | | | |

** significance at $p < 0.01$.

### 3.2. Construction of Linkage Map

The results of linkage map construction are shown in Table 3 and Figure 3. The map length was 3085.72 cM with an average distance between adjacent markers of 1.66 cM. The number of SNPs on each chromosome ranged from 85 to 346 on chromosomes 6 and 3, respectively. Chromosome 1 was the longest with 444.11 cM, while chromosome 6 was the shortest with 178.57 cM. The genetic gap interval ranged from 6.34 cM on chromosome 3 to 22.70 cM on chromosome 6. Although the average distance between markers was small, there were still some relatively large gaps.

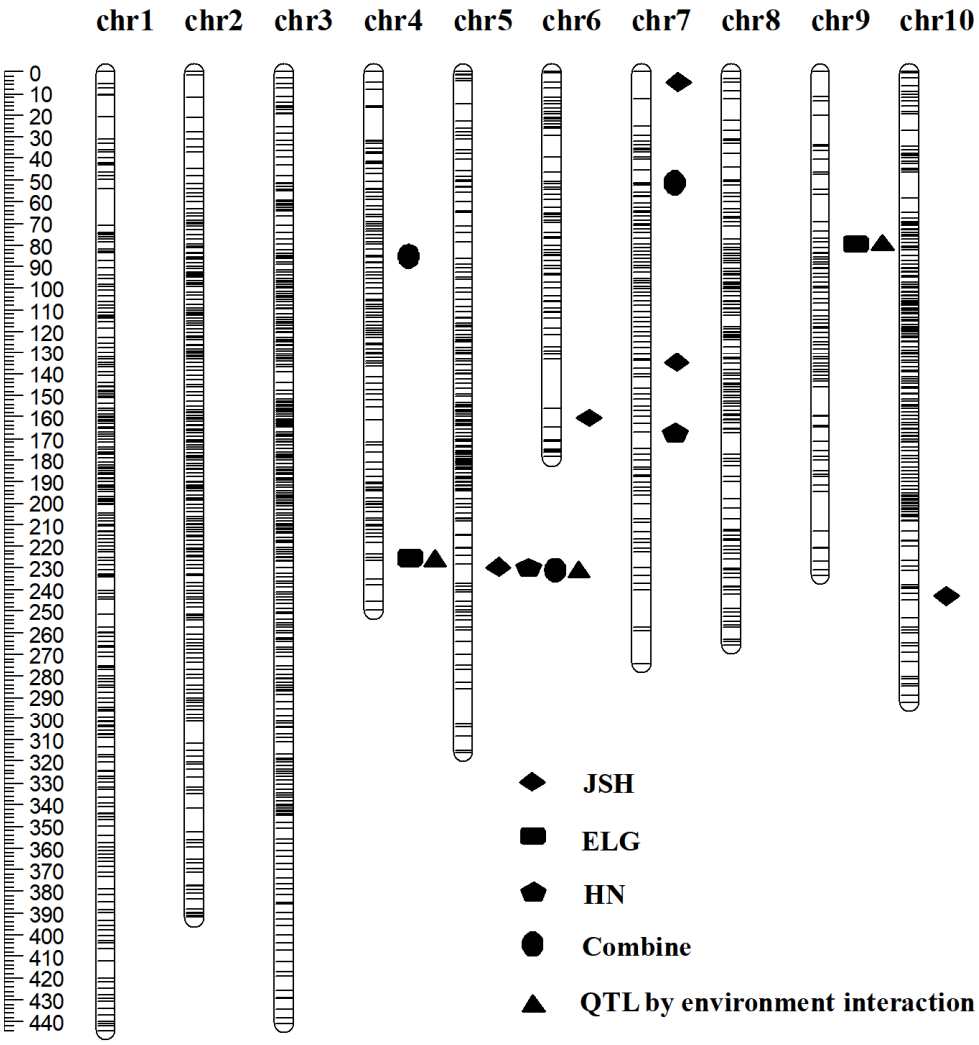


**Figure 3.** QTL analysis for maternal haploid inducibility in the $F_2$ population.

**Table 3.** Basic information of genetic linkage map for the $F_2$ population.

| Chromosomes | Number of SNPs | Chromosome Length (cM) | Average SNP Distance (cM) | Maximum Gap (cM) |
|---|---|---|---|---|
| 1 | 270 | 444.11 | 1.65 | 17.01 |
| 2 | 257 | 392.01 | 1.53 | 10.93 |
| 3 | 346 | 440.67 | 1.28 | 6.34 |
| 4 | 129 | 249.25 | 1.95 | 14.91 |
| 5 | 193 | 315.80 | 1.64 | 17.14 |
| 6 | 85 | 178.57 | 2.13 | 22.70 |
| 7 | 111 | 274.18 | 2.49 | 16.95 |
| 8 | 150 | 265.54 | 1.78 | 10.19 |
| 9 | 89 | 233.36 | 2.65 | 17.95 |
| 10 | 225 | 292.23 | 1.30 | 11.51 |
| Total | 1855 | 3085.72 | | |

### *3.3. Single Environment QTL Analysis*

The single environment QTL analysis results are presented in Table 4 and Figure 3. In JSH, five QTLs were detected on chromosome bins 5.05/5.06, 6.07, 7.00, 7.02, and 10.06 explaining 6.84%, 5.24%, 5.65%, 5.26%, and 4.81% of the total phenotypic variation, respectively. The total PVE value was 27.8%. For QTLs *qMHI5* and *qMHI6*, the MHI increasing alleles were derived from the inbred line B73. For the other three QTLs, *qMHI7-1*, *qMHI7-2*, and *qMHI10*, the favorite alleles were provided by Zheng58. In ELG, two QTLs located on chromosome bins 10.06 and 4.09 were detected. The PVE value of each QTL was 5.97% and 5.43%. The inducibility increasing alleles were derived from B73. In HN, two QTLs were detected on chromosome bins 5.05/5.06 and 7.03 explaining 6.04% and 5.97% of the phenotypic variation, respectively. For *qMHI7-3*, the inducibility increasing allele were derived from Zheng58.

**Table 4.** Single environment QTL mapping of maternal inducibility in the $F_2$ population.

| QTL | Env | Chr | Bin | Genetic Position (cM) | Left Marker [a] | Right Marker | LOD | PVE [b] (%) | Add [c] | Dom | Gene Effect [d] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *qMHI5* | JSH | 5 | 5.05/5.06 | 229 | S5_192539104 | S5_198930944 | 3.67 | 6.84 | 0.86 | −0.14 | A |
| *qMHI6* | JSH | 6 | 6.07 | 160 | S6_168195277 | S6_169331347 | 2.64 | 5.24 | 0.68 | 0.65 | D |
| *qMHI7-1* | JSH | 7 | 7.00 | 4 | S7_2687664 | S7_5084568 | 2.63 | 5.65 | −0.4 | 0.85 | OD |
| *qMHI7-2* | JSH | 7 | 7.02 | 133 | S7_118697494 | S7_119793590 | 3.09 | 5.26 | −0.45 | 0.73 | OD |
| *qMHI10* | JSH | 10 | 10.06 | 243 | S10_139299755 | S10_139367146 | 2.63 | 4.81 | −0.70 | 0.12 | A |
| *qMHI4-1* | ELG | 4 | 4.09 | 226 | S4_238620990 | S4_238810734 | 3.53 | 5.97 | −1.12 | −0.05 | A |
| *qMHI9* | ELG | 9 | 9.02 | 79 | S9_20780866 | S9_21925680 | 2.99 | 5.43 | −0.05 | 1.40 | OD |
| *qMHI5* | HN | 5 | 5.05/5.06 | 229 | S5_192539104 | S5_198930944 | 2.62 | 6.04 | 0.89 | −0.06 | A |
| *qMHI7-3* | HN | 7 | 7.03 | 167 | S7_144352936 | S7_149327555 | 2.83 | 5.97 | −0.39 | 1.01 | OD |
| *qMHI4-2* | Combine | 4 | 4.05 | 85 | S4_63834477 | S4_67727362 | 3.17 | 4.79 | 0.28 | −0.78 | OD |
| *qMHI5* | Combine | 5 | 5.05/5.06 | 231 | S5_192539104 | S5_198930944 | 5.72 | 10.10 | 0.87 | −0.41 | PD |
| *qMHI7-4* | Combine | 7 | 7.01 | 50 | S7_9620150 | S7_10978116 | 3.33 | 5.55 | −0.24 | 0.84 | OD |

[a] Marker name, chromosome_position, for example, S5_192539104 refers that the SNP is located at the physical position of 192,539,104 bp on chromosome 5. The physical position was based on the Maize B73_RefGen_v4 reference genome. [b] PVE, phenotypic variation explained. [c] Additive effect, positive values refer that the favorite alleles came from B73 and negative values refer that the favorite alleles came from Zheng58. [d] Gene action was classified to: additive (A), partially dominant (PD), dominant (D), and overdominant (OD).

Three QTLs located on chromosome bins 4.05, 5.05/5.06, and 7.01 were identified across environments. The QTL *qMHI5* with the PVE value of 10.1% was a major QTL for MHI. The favorite alleles were derived from B73, except for the QTL *qMHI7-3*.

The QTL *qMHI5* was detected in JSH, HN, and across environments, indicating that it is an environmentally stable QTL for MHI. Most of the QTLs showed additive and overdominant gene action.

### *3.4. QTL by Environment Interaction Analysis*

A total of three QTLs were detected on chromosome bins 4.09, 5.05/5.06, and 9.02 explaining 2.28%, 2.47%, and 2.04% of the total phenotypic variation, respectively (Figure 3).

All the three QTLs were detected in the single environment QTL analysis. Two QTLs, *qMHI4-1* and *qMHI 9-1* (Table 5), had higher PVE by additive × environment effect than PVE by additive effects, indicating strong QTL × environment interactions. The QTL *qMHI5* showed strong additive effects (Table 5).

**Table 5.** QTL by environment interaction analysis of maternal inducibility in the $F_2$ population.

| QTL | Chr | Bin | Genetic Position (cM) | Left Marker [a] | Right Marker | LOD | PVE (%) [b] | PVE (A) (%) [c] | PVE (A by E) (%) [d] | Add [e] | Dom | Gene Effect [f] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *qMHI4-1* | 4 | 4.09 | 227 | S4_238810734 | S4_241403945 | 5.97 | 2.28 | 1.06 | 1.22 | −0.39 | 0.45 | OD |
| *qMHI5* | 5 | 5.05/5.06 | 229 | S5_192539104 | S5_198930944 | 7.56 | 2.47 | 2.30 | 0.17 | 0.74 | −0.23 | PD |
| *qMHI9-1* | 9 | 9.02 | 79 | S9_20780866 | S9_21925680 | 5.78 | 2.04 | 0.17 | 1.87 | 0.20 | 0.03 | A |

[a] Marker name, chromosome_position. The physical position was based on the Maize B73_RefGen_v4 reference genome. [b] PVE, phenotypic variation explained. [c] PVE (A), phenotypic variation explained by additive effects. [d] PVE (A by E), phenotypic variation explained by additive × environment effects. [e] Additive effect, positive values refer that the favorite alleles came from B73 and negative values refer that the favorite alleles came from Zheng58. [f] Gene action was classified to: additive (A), partially dominant (PD), dominant (D), and overdominant (OD).

### *3.5. Prediction of Candidate Genes*

The QTL *qMHI5* was detected by both methods with the biggest PVE value of 10.10%. Therefore, candidate gene prediction was performed for *qMHI5*. Three candidate genes Zm00001d017366, Zm00001d017420, and Zm00001d017432 encoding AP2-EREBP-transcription factor 76, trihelix-transcription factor 20, and U6 biogenesis like1 (UBL1) were annotated (Table 6). All the three candidate genes may play important roles in seed development.

**Table 6.** Putative candidate genes for maternal haploid inducibility.

| Chr | Physical Location (Mb) | Candidate Gene | Annotation |
|---|---|---|---|
| 5 | 193,683,447–193,687,796 | Zm00001d017366 | AP2-EREBP-transcription factor 76 |
| 5 | 195,390,564–195,397,434 | Zm00001d017420 | Trihelix-transcription factor 20 |
| 5 | 195,843,468–195,848,738 | Zm00001d017432 | U6 biogenesis like1 |

## 4. Discussion

In vivo haploid induction is a backbone of doubled haploid technology in maize. HIR is not only affected by haploid inducers, but also by maternal lines [14,16]. In the present study, MHI was significantly affected by genotypes, environments, and genotype by environment interactions. Selecting sites with suitable climatic conditions and appropriate germplasm can increase the MHI. In this study, materials planting in Hainan showed slightly higher MHI than that of Jiushenghe and Erliugong in Xinjiang. This may be due to the low pollen viability of the inducer caused by the high temperature during anther shedding pollen stage in Xinjiang. Similar results were also observed by Li [16], where the MHI of Hainan was higher than that of Beijing, Shijianzhuang, and Jinan. Therefore, carrying out haploid induction in Hainan can obtain a higher MHI and improve the efficiency of haploid breeding.

The heritability of MHI was 0.40, slightly lower than previous studies, but within an acceptable range. The heritability of MHI was 049, 0.58, and ranged from 0.44 to 0.59 in the study of Benjamin et al. [15], Wu et al. [14], and Li [16]. In our study, the lower heritability may be due to the misclassification of haploids. Because of high cost and time-consuming in manual work, correction for HIR was not conducted.

To better understand the genetic mechanism of MHI, an $F_2$ population was used to identify QTL related to MHI. High-density markers for QTL mapping can not only improve the resolution of QTL mapping, but also identify more minor QTLs [26]. Li et al. [27] identified a major QTL with a confidence interval of 3.1 Mb using 3K SNP chips in a RIL population. In our study, a total of 1855 high-quality SNPs were used for linkage mapping construction. The total length of the linkage map was longer than previous studies, possibly

due to the potential genotyping errors. The average distance between adjacent markers was 1.66 cM, which effectively improved the resolution of QTL mapping. The interval of each QTL ranged from 0.07 Mb of *qMHI10* to 6.39 Mb of *qMHI5*, with an average interval of 2.88 cM.

A total of ten QTLs were detected for MHI. The QTLs *qMHI4-1*, *qMHI5*, and *qMHI9* were reliable QTLs detected by both single environment QTLs and by environment interaction analysis. The QTL *qMHI5* was also an environment stable QTL detected in multiple environments. All loci were minor QTLs except *qMHI5*.

QTLs identified in different studies may be different due to the use of different populations, environments, and markers [28,29]. Four QTLs detected in our study were reported in previous studies [16,17]. The QTLs *qMHI9* and *qMHI10* coincided with the QTLs identified by Li [16]. In the study of Li, the two QTLs were stable QTLs identified in different environments. They were minor QTLs in our study, but major QTLs in Li's study. The QTLs *qMHI5* and *qMHI7-3* contain the significant SNPs S5_191373611, S7_147071275, and S7_147071295 identified by Nair et al. [17] in tropical and subtropical maize. These four common loci were consistent in different studies, indicating that they were stable QTLs across different genetic backgrounds. The remaining five loci may be population-dependent QTLs, providing new genetic sources for MHI improvement. The stable QTLs identified in different environments and genetic backgrounds need to be explored further for fine-mapping and functional marker development for maker-assisted selection (MAS).

Three candidate genes related to MHI were annotated. *Zm00001d017366* encodes a AP2-EREBP-transcription factor 76 protein, which controls seed mass and seed yield [30]. *Zm00001d017420* encodes a trihelix-transcription factor 20 protein, which plays crucial roles in seed development and seed filling. In Arabidopsis, *ASIL1*, a member of the plant-specific trihelix family of transcription factor, controls embryonic gene expression [31]. The mutation of *ASIL1* leads to early embryo development. *Zm00001d017432* encodes a UBL1 protein, which is involved in kernel and seedling development. The mutation of *ubl1* showed small kernel in Maize [32,33]. These results provide valuable information on fine-mapping and candidate genes cloning.

Overall, MHI is a complex trait controlled by a few large effect QTLs with multiple small effect QTLs. Both MAS and genomic prediction can be used to improve the inducibility in maize. The average prediction of genomic prediction was 41% and 33% in different GWAS panels, which increased to 49% and 44% by the inclusion of significant makers in the prediction model [17]. Thus, the loci detected in our study can be used to develop functional markers for MAS or included in the genomic selection model to improve the prediction ability.

## 5. Conclusions

In this study, an $F_2$ population, phenotyped in multi-environment trials for MHI and genotyped by the 48K liquid-phase hybridization probe capture technique, was used for single environment QTL analysis and QTL by environment analysis. The heritability of MHI was 0.40, which was significantly affected by environment, genotypes, and environment × genotype interactions. A total of ten QTLs distributed on chromosomes 4, 5, 6, 7, 9, and 10, explaining 4.79% to 10.01% of the total phenotypic variation, were identified by single environment QTL analysis. The QTL *qMHI5* is a major QTL identified in JSH, HN, and across environments. Three QTLs, *qMHI4-1*, *qMHI5*, and *qMHI 9-1*, were also detected using QTL by environment interaction analysis. Zm00001d017366, Zm00001d017420, and Zm00001d017432 were candidate genes for MHI that play important roles in seed development. Our study contributes to a better understanding of the genetic mechanism of MHI.

**Author Contributions:** Conceptualization, P.W. and J.R.; formal analysis, X.Z. and Z.L.; funding acquisition, P.W. and J.R.; investigation, X.Z., Z.L. and P.W.; project administration, P.W., Z.L. and X.Z.; writing—original draft, X.Z., Z.L. and P.W.; writing—review and editing, J.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Foundation of China, grant numbers: U2003304, 32001561, 32060484; Tianshan innovation team funding, grant number: 2022D14017; Xinjiang Key R&D funding, grant number: 2021A02001-2, 2021B02002-2; the China Postdoctoral Science Foundation, grant number: 2018M643774; and Xinjiang Youth Foundation, grant number: 2019D01A41.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in the graphs and tables provided in the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Szarejko, L.; Forster, B.P. Doubled haploidy and induced mutation. *Euphytica* **2007**, *158*, 359–370. [CrossRef]
2. Chang, M.T.; Coe, E.H. Doubled haploids. In *Molecular Genetic Approaches to Maize Improvement*; Kriz, A.L., Larkins, B.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 63, pp. 127–142. [CrossRef]
3. Geiger, H.H. Doubled haploids. In *Handbook of Maize, Genetics and Genomics*; Bennetzen, J.L., Hake, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 2, pp. 641–657. [CrossRef]
4. Liu, Z.; Wang, Y.; Ren, J.; Mei, M.; Frei, U.K.; Trampe, B.; Lübberstedt, T. Maize doubled haploids. In *Plant Breeding Reviews*; Janick, J., Ed.; Wiley-Blackwell: Hoboken, NJ, USA, 2016; Volume 40, pp. 123–166. [CrossRef]
5. Prigge, V.; Xu, X.X.; Li, L.; Babu, R.; Chen, S.J.; Atlin, G.N.; Melchinger, A.E. New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. *Genetics* **2012**, *190*, 781–793. [CrossRef] [PubMed]
6. Coe, E.H. A Line of Maize with High Haploid Frequency. *Am. Nat.* **1959**, *93*, 381–382. [CrossRef]
7. Röber, F.K.; Gordillo, G.A.; Geiger, H.H. In vivo haploid induction in maize-performance of new inducers and significance of doubled haploid lines in hybrid breeding. *Maydica* **2005**, *50*, 275–283.
8. Zhang, Z.; Qiu, F.Z.; Liu, Y.Z.; Ma, K.J.; Li, Z.Y.; Xu, S.Z. Chromosome elimination and in vivo haploid production induced by Stock 6-derived inducer line in maize (*Zea mays* L.). *Plant Cell Rep.* **2008**, *27*, 1851–1860. [CrossRef] [PubMed]
9. Rotarenco, V.A.; Dicu, G.; State, D.; Fuia, S. New inducers of maternal haploids in maize. *Maize Genet. Newsl.* **2010**, *84*, 15.
10. Prigge, V.; Schipprack, W.; Mahuku, G.; Atlin, G.N.; Melchinger, A.E. Development of in vivo haploid inducers for tropical maize breeding programs. *Euphytica* **2012**, *185*, 481–490. [CrossRef]
11. Liu, C.; Li, J.; Chen, M.; Li, W.; Zhong, Y.; Dong, X.; Xu, X.; Chen, C.; Tian, X.; Chen, S. Development of high-oil maize haploid inducer with a novel phenotyping strategy. *Crop J.* **2022**, *10*, 524–531. [CrossRef]
12. Prigge, V.; Sánchez, C.; Dhillon, B.S.; Schipprack, W.; Araus, J.L.; Bänziger, M.; Melchinger, A.E. Doubled haploids in tropical maize. I. Effects of inducers and source germplasm on in vivo haploid induction rates. *Crop Sci.* **2011**, *51*, 1498–1506. [CrossRef]
13. De La Fuente, G.N.; Frei, U.K.; Trampe, B.; Nettleton, D.; Zhang, W.; Lübberstedt, T. A diallel analysis of a maize donor population response to in vivo maternal haploid induction I: Inducibility. *Crop Sci.* **2018**, *58*, 1830–1837. [CrossRef]
14. Wu, P.; Li, H.; Ren, J.; Chen, S. Mapping of maternal QTLs for in vivo haploid induction rate in maize (*Zea mays* L.). *Euphytica* **2014**, *196*, 413–421. [CrossRef]
15. Trampe, B.; Batîru, G.; Pereira da Silva, A.; Frei, U.K.; Lübberstedt, T. QTL Mapping for Haploid Inducibility Using Genotyping by Sequencing in Maize. *Plants* **2022**, *11*, 878. [CrossRef] [PubMed]
16. Li, W. *Effects of In Vivo Haploid Induction in Maize and Application of Machine Learning in Doubled Haploid Technology*; China Agricultural University: Beijing, China, 2017.
17. Nair, S.K.; Chaikam, V.; Cowda, M.; Hindu, V.; Melchinger, A.E.; Boddupalli, P.M. Genetic dissection of maternal influence on in vivo haploid induction in maize. *Crop J.* **2022**, *8*, 287–298. [CrossRef]
18. Dong, X.; Xu, X.; Li, L.; Liu, C.; Tian, X.; Li, W.; Chen, S. Marker-assisted selection and evaluation of high oil in vivo haploid inducers in maize. *Mol. Breed.* **2014**, *34*, 1147–1158. [CrossRef]
19. Li, L.; Xu, X.; Jin, W.; Chen, S.J. Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. *Planta* **2009**, *230*, 367–376. [CrossRef]
20. Meng, L.; Li, H.; Zhang, L.; Wang, J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* **2015**, *3*, 269–283. [CrossRef]
21. Alvarado, G.; Rodríguez, F.M.; Pacheco, A.; Burgueño, J.; Crossa, J.; Vargas, M.; Pérez-Rodríguez, P.; Lopez-Cruz, M.A. META-R: A software to analyze data from multi-environment plant breeding trials. *Crop J.* **2020**, *8*, 745–756. [CrossRef]
22. R core Team. *R: A Language and Environment for Statistical Computing Version 3.6.2*; Springer: Vienna, Austria, 2019.
23. Murray, M.G.; Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucl. Acids Res.* **1980**, *8*, 4321–4326. [CrossRef]
24. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef]
25. Portwood, J.C.L.; Woodhouse, M.R.; Cannon, E.K.; Gardiner, J.M.; Harper, L.C.; Schaeffer, M.L.; Walsh, J.R.; Sen, T.Z.; Cho, K.T.; Schott, D.A.; et al. MaizeGDB 2018: The maize multi-genome genetics and genomics database. *Nucleic Acids Res.* **2019**, *47*, D1146–D1154. [CrossRef]

26. Stuber, C.W.; Edwards, M.D.; Wendel, J.F. Molecular Marker-Facilitated Investigations of Quantitative Trait Loci in Maize. II. Factors Influencing Yield and its Component Trait. *Crop Sci.* **1987**, *27*, 639–648. [CrossRef]
27. Wang, L.; Wang, A.; Huang, X.; Zhao, Q.; Dong, G.; Qian, Q.; Sang, T.; Han, B. Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombinant inbred lines. *Theor. Appl. Genet.* **2011**, *122*, 327–340. [CrossRef] [PubMed]
28. Li, K.; Yan, J.; Li, J.; Yang, X. Genetic architecture of rind penetrometer resistance in two maize recombinant inbred line populations. *BMC Plant Biol.* **2014**, *14*, 152. [CrossRef] [PubMed]
29. Ren, J.; Boerman, N.A.; Liu, R.; Wu, P.; Trampe, B.; Vanous, K.; Frei, U.K.; Chen, S.; Lübberstedt, T. Mapping of QTL and identifification of candidate genes conferring spontaneous haploid genome doubling in maize (*Zea mays* L.). *Plant Sci.* **2020**, *293*, 110337. [CrossRef]
30. Ren, J.; Li, Z.; Wu, P.; Zhang, A.; Liu, Y.; Hu, G.; Cao, S.; Qu, J.; Dhliwayo, T.; Zheng, H.; et al. Genetic dissection of quantitative resistance to common rust (*Puccinia sorghi*) in tropical Maize (*Zea mays* L.) by combined genome-wide association study, linkage mapping, and genomic prediction. *Frontiers* **2021**, *12*, 692205. [CrossRef]
31. Goldberg, R.B.; Nakamura, K.; Harada, J.J.; Ohto, M.A.; Fischer, R.L. Control of seed mass by APETALA2. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3123–3128.
32. Gao, M.J.; Li, X.; Lui, H.; Gropp, G.M.; Lydiate, D.D.; Wei, S.; Hegedus, D.D. Asil1 is required for proper timing of seed filling in arabidopsis. *Plant Signal. Behav.* **2011**, *6*, 1885–1888. [CrossRef]
33. Li, J.; Fu, J.; Chen, Y.; Fan, K.; He, C.; Zhang, Z.; Li, L.; Liu, Y.; Zheng, J.; Ren, D.; et al. The U6 biogenesis-Like 1 plays an important role in maize kernel and seedling development by affecting the 3′ end processing of U6 snRNA. *Mol. Plant* **2017**, *10*, 470–482. [CrossRef]