**MATERIALS AND METHODS**

**Plant materials**

H3 is a high oil inducer selected by Dong et al (2014)([Dong et al. 2014](#_ENREF_2)), its HIR is ~8%, and the kernel oil content is ~8.5%. A series of maize inbred lines were induced by H3 as pollen parent in a preliminary experiment. After two years comparison, we found out two lines with large difference in haploid inducibility (see Figure S1). Q319 shows a higher HIR of ~16%, C7-2 ~8%. Then a cross was made between Q319 and C7-2, and induced by haploid inducer to generate haploids. After chromosome doubling, a doubled haploid (DH) population with 135 lines was constructed.

**Field experiment**

The population of 135 DH lines and their two parentas were planted across four location environments, including Shunyi in spring, Jinan and Shijiazhuang (SJZ) in summer and Sanya in winter nursery in the year of 2015. For all trials at each location, experimental units consisted of single-row plots arranged in randomized block design with two replications. Three blocks were used in every replication, two parents were set randomly as checks in each block.

H3 was used as pollen father to pollinate with all the DH lines and their parents. To ensure flowering time encountered, two more copies of H3 were planted every 5 days after sowing at the same time with DHs.

**Phenotype and data analysis**

To discriminate with HIR of inducer, maternal haploid induction rate (MHIR) were used to represent the haploid inducibility of source germplasm. MHIR of each line was determined ear by ear after harvest. The calculation method was already described by Dong ([Dong et al. 2013](#_ENREF_3)).

A mixed linear model was used to estimate the entry effect of each line by using lmer function in the R package ‘lme4’ ([Mahuku et al. 2016](#_ENREF_6)). Through the model a best linear unbiased prediction (BLUP) was obtained for single and combined environments separately.

Single Environment BLUP：MHIR~1 + Block + Replication + (1|Geno) + (1|Replication:Geno)

Combined Environments BLUP：MHIR~1+ Replication + Block + (1|Location) + (1|Geno) + (1|Location:Geno)

where MHIR was trait data, replication of each environment and blocks in every replication were used as fixed factors which including all possible levels in the study, Geno refers to DH lines, Location to the four environments, The parentheses indicate random effects, the vertical bar character “|” separates an expression for a model matrix

and a grouping factor, and “:” refers to interactions.

An analysis of variance (ANOVA) was performed in SAS 9.4 (SAS Institute). Variance components were estimated using the restricted maximum likelihood (REML) method with PROC VARCOMP in SAS. Broad sense heritability () for single and combined environments were calculated using the following formulae ([Knapp et al. 1985](#_ENREF_4))：

Single Environment:

Combined Environments:

represents genetic variance, genotype × environment variance, error variance. And l is number of environments, r is number of replication in each environment, which is 4 and 2 in our study.

**DNA extraction, genotyping and genetic map construction**

Young leaf tissue was harvested in the field for genomic DNA extraction according to the method described by Murray and Thompson([Murray and Thompson 1980](#_ENREF_8)). Then the purified DNA was genotyped with MaizeSNP6K chip at the genomic platform of National Maize Improvement Center of China, China Agricultural University. 5124 qualified SNPs were checked out according to the method described by Yang et al. ([Yang et al. 2011](#_ENREF_9)).

After that, we have screened 2471 polymorphic markers between Q319 and C7-2. Then the missing rate, minor allele frequency (MAF) and heterozygosity of each SNP in this DH population were calculated as well as heterozygosity for each line. The DH lines with heterozygosity ≤ 0.1, and the SNPs with a missing rate ≤ 0.2 and MAF ≥ 0.05 were used for constructing genetic linkage map with software package JoinMap4.0([Meng et al. 2016](#_ENREF_7)). The linkage groups were determined by a minimum Logarithm of odds (LOD) of 7, and regression-mapping algorithm were used for calculating map distances. During this process the kosambi mapping function was used to adjust the map distance.

**QTL analysis**

Composite interval mapping (CIM) method was performed in the software Windows QTL Cartographer 2.5 ([Zeng 1994](#_ENREF_10)) to detect QTLs for single environment and combined environments BLUP trait values. Model 6 of the Zmapqtl module in the CIM analysis with forward and backward stepwise regression was implemented in the whole genome scan with a window size of 10 cM, and a step size of 1 cM. Permutation test was performed with 1000 times to determine the threshold logarithm (base 10) of odds (LOD) values at a threshold p value equals 0.05 ([Churchill and Doerge 1994](#_ENREF_1)). The LOD peak location of a QTL was described, and the surrounding region with 95% confidence interval was calculated in the WinQTLCart to determine the confidence interval of a QTL ([Lander and Botstein 1989](#_ENREF_5)). If the position of two QTLs was less than 20cM, they would be treated as one. The QTLs were named after trait name and chromosome number as described by

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. **Genetics** 138: 963-971

Dong X, Xu X, Li L, Liu C, Tian X, Li W, Chen S (2014) Marker-assisted selection and evaluation of high oil in vivo haploid inducers in maize. **Molecular Breeding** 34: 1147-1158

Dong X, Xu X, Miao J, Li L, Zhang D, Mi X, Liu C, Tian X, Melchinger AE, Chen S (2013) Fine mapping of qhir1 influencing in vivo haploid induction in maize. **Theoretical and Applied Genetics** 126: 1713-1720

Knapp S, Stroup W, Ross W (1985) Exact confidence intervals for heritability on a progeny mean basis. **Crop Science** 25: 192-194

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics** 121: 185

Mahuku G, Chen J, Shrestha R, Narro LA, Guerrero KVO, Arcos AL, Xu Y (2016) Combined linkage and association mapping identifies a major QTL (qRtsc8-1), conferring tar spot complex resistance in maize. **Theoretical and Applied Genetics** 129: 1217-1229

Meng Y, Li J, Liu J, Hu H, Li W, Liu W, Chen S (2016) Ploidy effect and genetic architecture exploration of stalk traits using DH and its corresponding haploid populations in maize. **BMC plant biology** 16: 1

Murray M, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. **Nucleic acids research** 8: 4321-4326

Yang X, Xu Y, Shah T, Li H, Han Z, Li J, Yan J (2011) Comparison of SSRs and SNPs in assessment of genetic relatedness in maize. **Genetica** 139: 1045-1054

Zeng ZB (1994) Precision mapping of quantitative trait loci. **Genetics** 136: 1457