Dissecting Genetic Architecture Underlying Seed Traits in Multiple Environments

Ting Qi,* Yujie Cao,* Liyong Cao,[†] Yongming Gao,[†] Shuijin Zhu,* Xiangyang Lou,[§] and Haiming Xu^{*,1}

^{*}Institute of Crop Science and Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, 310058, People's Republic of China, [†]State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, 311402, People's Republic of China, [‡]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, 100081, People's Republic of China, and [§]Department of Biostatistics, University of Alabama, Birmingham, Alabama 35294

ABSTRACT The seeds of flowering plants develop from double fertilization and play a vital role in reproduction and supplying human and animal food. The genetic variation of seed traits is influenced by multiple genetic systems, e.g., maternal, embryo, and/or endosperm genomes. Understanding the genetic architecture of seed traits is a major challenge because of this complex mechanism of multiple genetic systems, especially the epistasis within or between different genomes and their interactions with the environment. In this study, a statistical model was proposed for mapping QTL with epistasis and QTL-by-environment (QE) interactions underlying endosperm and embryo traits. Our model integrates the maternal and the offspring genomes into one mapping framework and can accurately analyze maternal additive and dominant effects, endosperm/embryo additive and dominant effects, and epistatic effects of two loci in the same or two different genomes, as well as interaction effects of each genetic component of QTL with environment. Intensive simulations under different sampling strategies, heritabilities, and model parameters were performed to investigate the statistical properties of the model. A set of real cottonseed data was analyzed to demonstrate our methods. A software package, QTLNetwork-Seed-1.0.exe, was developed for QTL analysis of seed traits.

CROP seeds are the primary source of human food, animal feed, and industrial raw materials, hence understanding how seed traits are genetically determined is of great significance. The embryo and/or endosperm are the principal components of mature cereal grains. Endosperm in monocots such as rice, wheat, and maize or cotyledon in species such as *Arabidopsis thaliana* function as storage organs for nutrients such as carbohydrates, proteins, lipids, and minerals (Mazur *et al.* 1999), which are used in embryo development and maturation (Olsen 1998).

Most crop seed traits are quantitative. Quantitative trait locus (QTL) mapping has become an effective way to dissect the genetic architecture of quantitative traits. So far, several genetic models and statistical methods for seed trait mapping have been developed and applied in analysis of breeding data (Wu *et al.* 2002; Tsilo *et al.* 2011; Walker *et al.* 2013). However,

some of these approaches ignored the difference between diploid and triploid inheritance. Because endosperm is triploid, it exhibits more complex genetics than diploid organisms. For example, in a locus with two alleles (Q and q), endosperm has four genotypes (QQQ, QQq, Qqq, and qqq) whereas there are only three genotypes for the embryo (QQ, Qq, and qq). Thus, endosperm usually exhibits reciprocal effects. Furthermore, the endosperm is the next generation of maternal plants. Based on such a genetic feature, Hu and Xu (2005) proposed a statistical method for characterizing the genetic effects of the maternal genome on offspring traits, incorporating both a quantitative genetic model for diploid maternal traits and one for triploid endosperm traits into a unified QTL mapping framework. Wen and Wu (2007) further proposed a method of interval mapping endosperm traits based on a two-stage hierarchical design, which could estimate all kinds of main effects and further improved the mapping accuracy. However, these methods ignored epistatic interactions and QTL-by-environment (QE) interactions, potentially leading to biased estimation of parameters.

Epistasis has long been recognized as fundamentally important to understanding the structure and function of genetic pathways and the evolutionary dynamics of complex

Copyright © 2015 by the Genetics Society of America
doi: 10.1534/genetics.114.168203

Manuscript received July 18, 2014; accepted for publication October 8, 2014;
published Early Online October 21, 2014.

¹Corresponding author: Institute of Crop Science and Institute of Bioinformatics,
College of Agriculture and Biotechnology, Zhejiang University, 866 Yu-Hang Tang Rd.,
Hangzhou, 310058, People's Republic of China. E-mail: hmxu@zju.edu.cn

Table 1 Mapping power and estimates of QTL position and main effects

QTL	Chr.	Pos.		a^m		d^m		a^e		d^e		Power
		Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	
Q ₁	3	26.0	25.87 (1.90)	1.8	1.73 (0.55)	2.4	2.31 (0.40)	0.8	0.82 (0.52)	2.4	2.31 (0.82)	99.8
Q ₂	4	63.0	63.60 (1.27)	1.8	1.75 (0.53)	3.6	3.49 (0.36)	2.5	2.45 (0.51)	0.8	0.78 (0.79)	100.0
Q ₃	5	51.0	50.98 (1.85)	2.5	2.48 (0.48)	1.8	1.76 (0.35)	0.9	0.91 (0.46)	2.6	2.53 (0.69)	100.0

QTL, the name of the simulated QTL; Chr., the order number of simulated chromosomes; Pos., the distance (centimorgans) between the QTL and the first marker on the same chromosome; a^m (d^m) and a^e (d^e), the maternal additive (dominance) effect and the endosperm additive (dominance) effect of one QTL, respectively; Power, the percentage of the QTL detected correctly at the significance of 0.05; Par., the true value of the parameter in simulation; Est., the estimate of the parameter; SD, the standard deviation of the estimate of the parameter.

genetic systems. It describes the general situation in which the phenotype of a quantitative trait cannot be predicted by the sum of all single-locus effects (Lark *et al.* 1995; Xing *et al.* 2002; Fan *et al.* 2005). Intensive research has illustrated that epistasis is an important factor in complex traits such as grain yield and its components, as well as heterosis and inbreeding depression (Yu *et al.* 1997; Mackay *et al.* 2009). With increasing recognition of the importance of epistasis in complex traits, it is appealing to develop a statistical model that integrates maternal and offspring genome into one mapping framework for dissecting the genetic networks of seed traits. General epistasis is defined as the interaction of two genes at different loci of the same genome; however, some studies have indicated that two epistatic genes can reside in different genomes (Cui *et al.* 2004; Cui and Wu 2005a, b). Zhang *et al.* (2004) proposed a statistical model for testing and estimating the effects of maternal–offspring genome interactions on embryo and endosperm traits during seed development in autogamous plants. This approach can detect epistasis from different genomes, but at the price of ignoring maternal effects of a single QTL. Cui *et al.* (2004; Cui and Wu 2005a,b) proposed a statistical method to detect the genetic mechanism of maternal–offspring genome interactions. Under the assumption that genetic expression of endosperm traits is controlled by both the maternal and the offspring genomes, the method can estimate the direct (offspring) effect, the indirect (maternal) effect, and the effects of maternal–offspring genome interaction simultaneously. But this method is based on the arguable assumption that the two interacting QTL cannot have maternal and offspring effects simultaneously. In other words, the methods proposed by Cui *et al.* (2004; Cui and Wu 2005a,b) can handle only an interaction of paired QTL in which one QTL has only a maternal effect while the other has only an endosperm effect. As the endosperm develops from the maternal plant, it is natural to consider that both maternal and offspring QTL can have phenotypic manifestations. Consequently, assuming the QTL potentially act in both the maternal and the endosperm genomes, the maternal effects and the endosperm effects, respectively, of each QTL can be tested.

GE interaction is also a very important component of phenotypic variation. GE interaction occurs when the environment effect depends on a genotype or, equivalently, when the genotypic effect depends on the environment. Identification

of GE interactions has been a hot topic in mapping QTL for crop complex traits and human disease susceptibility genes (Zheng *et al.* 2008; Renz *et al.* 2011; Nickels *et al.* 2013). Qi *et al.* (2014) proposed two statistical models for mapping QTL of crop seed traits with inclusion of maternal effects, embryo or endosperm effects of QTL, environment effects, and QE interactions. All the above methods do not integrate epistatic effects either from one genome or from two different genomes and their interaction effects with environment, which are of great importance for marker-assisted selection in crop improvement (He *et al.* 1999; Bao *et al.* 2004).

To deal with the above issues and further improve the mapping power and estimation accuracy of QTL effects for seed traits, we propose a new methodology for systematically mapping QTL based on a mixed linear model approach. We integrated maternal and offspring genomes into one model. The QTL effects, including epistatic effects between two loci in the same genome and different genomes, and their interaction effects with environment were estimated and tested using the Markov chain Monte Carlo (MCMC) algorithm for Gaussian mixed linear model via Gibbs sampling (Smith and Roberts 1993). Monte Carlo simulations were conducted to investigate the reliability and efficiency of the method. A worked example was provided to demonstrate the effectiveness of our method. A computer program (QTLNetwork-Seed-1.0.exe) was developed for implementing QTL analysis of seed traits.

Materials and Methods

Genetic model

Consider a population derived from random mating of immortalized F₂ (IF₂) of which genetically identical individuals can be regenerated for replication of experiments and phenotyping in different environments for detection of GE interactions. Marker information for maternal plants and the phenotypes of offspring are employed to perform QTL mapping. The genetic experiment is conducted in p different environments, each with b blocks. Suppose a seed trait is controlled by both the maternal genetic effects and endosperm (embryo) genetic effects of s segregating QTL (Q₁, Q₂, ..., Q_s) in which t pairs of QTL are involved in epistatic interactions. Let a random variable E_{ki} be the genotype of Q_k

Table 2 Estimates of interaction effects between QTL and environment

QTL	Chr.	a_1^m		ae_2^m		de_1^m		de_2^m		ae_1^e		ae_2^e		de_1^e		de_2^e	
		Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)
Q ₁	3	-1.28	-0.97 (0.46)	1.28	1.00 (0.45)	-1.34	-1.24 (0.21)	1.34	1.20 (0.23)	1.38	1.09 (0.41)	-1.38	-1.13 (0.40)	-1.09	-0.71 (0.59)	1.09	0.74 (0.59)
Q ₂	4	1.18	0.92 (0.43)	-1.18	-0.89 (0.44)	1.20	1.06 (0.23)	-1.20	-1.09 (0.22)	-1.64	-1.36 (0.40)	1.64	1.42 (0.38)	1.53	1.26 (0.56)	-1.53	-1.27 (0.56)
Q ₃	5	0.00	0.00 (0.05)	0.00	0.00 (0.05)	0.00	0.00 (0.06)	0.00	0.00 (0.06)	0.00	0.00 (0.05)	0.00	0.00 (0.05)	0.00	0.00 (0.11)	0.00	0.00 (0.11)

$ae_1^m, ae_2^m, de_1^m, de_2^m, ae_1^e, ae_2^e, de_1^e, de_2^e$ are the interaction effects of the maternal additive, the maternal dominance, the endosperm additive, and the endosperm dominance with the first (subscript 1) and the second (subscript 2) environments, respectively. QTL, Chr., Par., Est., and SD have the same definitions as those in Table 1.

from the i th line. If QTL genotypes are available, we can easily get the genotypic value (Hu and Xu 2005). In reality, we do not know such information, but we can infer the conditional probabilities based on flanking markers of QTL. Let $x_{ki} = E(E_{ki} | \text{the genotypes of flanking markers})$, which can be estimated by a general algorithm with dominant, co-dominant, or missing markers (Zeng 1994). For the i th strain in the j th block from the h th environment, its phenotype of seed trait y_{hij} can be expressed by the following mixed linear model:

$$\begin{aligned}
 y_{hij} = & \mu + \sum_k^s (a_k^m x_{ik}^{Am} + d_k^m x_{ik}^{Dm} + a_k^e x_{ik}^{Ae} + d_k^e x_{ik}^{De}) + e_h \\
 & + \sum_{k,l \in \{1,2,\dots,s\}, k \neq l}^t (aa_{kl}^{mm} x_{ik}^{Am} x_{il}^{Am} + aa_{kl}^{ee} x_{ik}^{Ae} x_{il}^{Ae} \\
 & + aa_{kl}^{me} x_{ik}^{Am} x_{il}^{Ae} + aa_{kl}^{em} x_{ik}^{Ae} x_{il}^{Am}) \\
 & + \sum_k^s (ae_{kh}^m x_{ik}^{Am} + de_{kh}^m x_{ik}^{Dm} + ae_{kh}^e x_{ik}^{Ae} + de_{kh}^e x_{ik}^{De}) \\
 & + \sum_{k,l \in \{1,2,\dots,s\}, k \neq l}^t (aae_{klh}^{mm} x_{ik}^{Am} x_{il}^{Am} + aae_{klh}^{ee} x_{ik}^{Ae} x_{il}^{Ae} \\
 & + aae_{klh}^{me} x_{ik}^{Am} x_{il}^{Ae} + aae_{klh}^{em} x_{ik}^{Ae} x_{il}^{Am}) + B_{j(h)} + \varepsilon_{hij}, \quad (1)
 \end{aligned}$$

where μ is the population mean; a_k^m and d_k^m are the maternal additive and dominance effects of Q_k , with coefficients x_{ik}^{Am} and x_{ik}^{Dm} , respectively; a_k^e and d_k^e are the endosperm (embryo) additive and dominance effects of Q_k , with coefficients x_{ik}^{Ae} and x_{ik}^{De} ; e_h is the random effect of the h th environment, $e_h \sim (0, \sigma_e^2)$; ae_{kh}^m and de_{kh}^m are the interaction effects of the maternal additive and the maternal dominance with environment, $ae_{kh}^m \sim (0, \sigma_{AmE}^2)$, $de_{kh}^m \sim (0, \sigma_{DmE}^2)$; ae_{kh}^e and de_{kh}^e are the endosperm (embryo) additive- and dominance-by-environment interaction effects, $ae_{kh}^e \sim (0, \sigma_{AeE}^2)$, $de_{kh}^e \sim (0, \sigma_{DeE}^2)$; aa_{kl}^{mm} is the maternal additive-additive epistatic effect between Q_k and Q_l , with coefficient $x_{ik}^{Am} x_{il}^{Am}$; aa_{kl}^{ee} is the endosperm (embryo) additive-additive epistatic effect between Q_k and Q_l , with coefficient $x_{ik}^{Ae} x_{il}^{Ae}$; aa_{kl}^{me} is the maternal-endosperm (embryo) additive-additive epistatic effect between Q_k and Q_l , with coefficient $x_{ik}^{Am} x_{il}^{Ae}$; aa_{kl}^{em} is the endosperm (embryo)-maternal additive-additive epistatic effect between Q_k and Q_l , with coefficient $x_{ik}^{Ae} x_{il}^{Am}$; aae_{klh}^{mm} is the interaction effect between aa_{kl}^{mm} and the h th environment; aae_{klh}^{ee} is the interaction effect between aa_{kl}^{ee} and the h th environment; aae_{klh}^{me} is the interaction effect between aa_{kl}^{me} and the h th environment; aae_{klh}^{em} is the interaction effect between aa_{kl}^{em} and the h th environment; $B_{j(h)}$ is the block effect in the h th environment; and ε_{hij} is the residual random effect, $\varepsilon_{hij} \sim (0, \sigma_e^2)$.

Mapping strategy

In model (1), we assume that the position and genotype of each QTL are known. In reality, such information is unavailable before mapping. We first need to specify the positions

Table 3 Mapping power and estimates of QTL position and epistatic effect

Epi.	Pos. of QTL1				Pos. of QTL2				aa^{mm}		aa^{ee}		aa^{me}		aa^{em}		Power
	Chr.	Par.	Est. (SD)		Chr.	Par.	Est. (SD)		Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	
EQ ₁	1	12.0	12.58 (1.30)		2	31.0	31.08 (1.45)		1.50	1.50 (1.72)	1.20	1.23 (1.44)	1.50	1.46 (1.54)	1.60	1.52 (1.56)	100
EQ ₂	3	26.0	25.78 (1.66)		4	63.0	63.62 (1.24)		1.80	1.61 (1.91)	1.30	1.15 (1.70)	1.60	1.62 (1.78)	1.80	1.81 (1.81)	100
EQ ₃	1	12.0	12.67 (1.34)		4	63.0	63.65 (1.18)		1.60	1.57 (1.89)	1.30	1.24 (1.47)	0.80	0.76 (1.65)	1.50	1.45 (1.66)	99.8

aa^{mm} , aa^{ee} , aa^{me} , and aa^{em} are the maternal additive-additive epistatic effect, the endosperm additive-additive epistatic effect, the maternal-endosperm additive-additive epistatic effect, and the endosperm-maternal additive-additive epistatic effect, respectively. Epi. (epistasis), the name of the detected paired QTL (QTL1 and QTL2) with epistasis; Pos. of QTL1 (or -2), the distance (centimorgans) between the first marker and QTL1 (or -2) of the paired QTL with epistasis; Chr., Par., Est., and SD have same definitions as those in Table 1.

of all QTL, and then the coefficients of each genetic effect can be determined by the conditional probability of the QTL genotype on the observed flanking marker genotype. Then, all the effects can be estimated through solving the mixed linear model. To search for s segregating QTL and t pairs of epistases, we adopt a whole-genome one-dimensional scanning strategy to detect QTL with individual effects and two-dimensional scanning to detect paired QTL with epistatic effects (Yang *et al.* 2007).

Mapping QTL and epistasis: Candidate marker intervals potentially harboring QTL are first selected through one-dimensional scanning across the whole genome. Then the F -test can be conducted with a walking step, such as 1 cM, along the whole genome, while the selected marker intervals are included in the model as the cofactors to control background genetic effects. Without loss of generality, we use the endosperm model to illustrate the proposed analysis. The following model is used to identify the significant marker intervals on the whole genome,

$$y_{hij} = \mu_h + \zeta_{it}^{+Am} a_{th}^{+m} + \zeta_{it}^{+Dm} d_{th}^{+m} + \zeta_{it}^{+Ae} a_{th}^{+e} + \zeta_{it}^{+De} d_{th}^{+e} + \zeta_{it}^{-Am} a_{th}^{-m} + \zeta_{it}^{-Dm} d_{th}^{-m} + \zeta_{it}^{-Ae} a_{th}^{-e} + \zeta_{it}^{-De} d_{th}^{-e} + B_{j(h)} + \varepsilon_{hij}, \quad (2)$$

where μ_h is the population mean in the h th environment; t ($t = 1, \dots, T$) indexes the t th marker interval to be tested in T intervals; a_{th}^{+m} (a_{th}^{-m}) and d_{th}^{+m} (d_{th}^{-m}) stand for the maternal additive and dominance effects for the right marker (left marker) of the t th marker interval in the h th environment, with coefficients ζ_{it}^{+Am} (ζ_{it}^{-Am}) and ζ_{it}^{+Dm} (ζ_{it}^{-Dm}), respectively; a_{th}^{+e} (a_{th}^{-e}) and d_{th}^{+e} (d_{th}^{-e}) stand for the endosperm additive effect and the dominance effect for the right marker (left marker) of the t th marker interval in the h th environment with coefficients ζ_{it}^{+Ae} (ζ_{it}^{-Ae}) and ζ_{it}^{+De} (ζ_{it}^{-De}), respectively; and the other parameters have the same definitions as model (1).

All ζ 's corresponding to different marker effects can be determined the same way as for the coefficients of QTL effects.

Once the significant intervals are selected, the QTL mapping model for testing a locus k within a particular genomic region can be written as

$$y_{hij} = \mu_h + x_{ik}^{Am} a_{kh}^m + x_{ik}^{Dm} d_{kh}^m + x_{ik}^{Ae} a_{kh}^e + x_{ik}^{De} d_{kh}^e + \sum_{t=1}^s (\zeta_{it}^{+Am} a_{th}^{+m} + \zeta_{it}^{+Dm} d_{th}^{+m} + \zeta_{it}^{+Ae} a_{th}^{+e} + \zeta_{it}^{+De} d_{th}^{+e}) + \sum_{t=1}^s (\zeta_{it}^{-Am} a_{th}^{-m} + \zeta_{it}^{-Dm} d_{th}^{-m} + \zeta_{it}^{-Ae} a_{th}^{-e} + \zeta_{it}^{-De} d_{th}^{-e}) + B_{j(h)} + \varepsilon_{hij}, \quad (3)$$

where a_{kh}^m and d_{kh}^m are the maternal additive and dominance effects of the k th putative QTL in the h th environment; a_{kh}^e and d_{kh}^e are the endosperm additive and dominance effects of the k th putative QTL in the h th environment, respectively; and all the other parameters have the same definitions as those in models (1) and (2).

When the F values for a region exceed the critical threshold determined by permutation testing (Doerge and Churchill 1996), a QTL is indicated at that position with the regional maximum F value.

Suppose that s QTL have been mapped by a one-dimensional genome scan. To find all possible epistasis, the s QTL as well as significant marker interval pairs identified by a marker pair selection (Piepho and Gauch 2001) approach are included in the model as cofactors. For a pair of marker intervals consisting of the l th and the r th intervals, we can use the following model to test its significance,

$$y_{hij} = \mu_h + \zeta_{il}^{+Am} \zeta_{ir}^{+Am} aa_{th}^{+mm} + \zeta_{il}^{+Am} \zeta_{ir}^{+Ae} aa_{th}^{+me} + \zeta_{il}^{+Ae} \zeta_{ir}^{+Am} aa_{th}^{+em} + \zeta_{il}^{+Ae} \zeta_{ir}^{+Ae} aa_{th}^{+ee} + \zeta_{il}^{-Am} \zeta_{ir}^{-Am} aa_{th}^{-mm} + \zeta_{il}^{-Am} \zeta_{ir}^{-Ae} aa_{th}^{-me} + \zeta_{il}^{-Ae} \zeta_{ir}^{-Am} aa_{th}^{-em} + \zeta_{il}^{-Ae} \zeta_{ir}^{-Ae} aa_{th}^{-ee} + \sum_{k=1}^s (\zeta_{ik}^{+Am} a_{kh}^{+m} + \zeta_{ik}^{+Dm} d_{kh}^{+m} + \zeta_{ik}^{+Ae} a_{kh}^{+e} + \zeta_{ik}^{+De} d_{kh}^{+e}) + \sum_{k=1}^s (\zeta_{ik}^{-Am} a_{kh}^{-m} + \zeta_{ik}^{-Dm} d_{kh}^{-m} + \zeta_{ik}^{-Ae} a_{kh}^{-e} + \zeta_{ik}^{-De} d_{kh}^{-e}) + B_{j(h)} + \varepsilon_{hij}, \quad (4)$$

where aa_{th}^{+mm} , aa_{th}^{+me} , aa_{th}^{+em} , and aa_{th}^{+ee} are the maternal, the maternal-endosperm, the endosperm-maternal, and the endosperm additive-additive interaction effects between two

Table 4 Estimates of interaction effects between QTL epistasis and environment

Epi.	aae_1^{mm}			aae_2^{mm}			aae_1^{ee}			aae_2^{ee}			aae_1^{me}			aae_2^{me}			aae_1^{em}			aae_2^{em}		
	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)
EQ ₁	0.00	0.15 (0.35)	0.00	-0.15 (0.35)	-1.32	-0.67 (0.83)	1.32	0.65 (0.84)	0.00	-0.12 (0.33)	0.00	0.13 (0.34)	1.34	0.61 (0.92)	-1.34	-0.59 (0.92)	0.00	0.13 (0.30)	0.00	0.13 (0.30)	0.00	-0.13 (0.30)	0.00	-0.13 (0.30)
EQ ₂	1.52	0.52 (1.11)	-1.52	-0.51 (1.11)	0.00	-0.09 (0.24)	0.00	0.09 (0.23)	-1.14	-0.28 (0.95)	1.14	0.29 (0.95)	0.00	0.13 (0.30)	0.00	-0.13 (0.30)	0.00	0.13 (0.30)	0.00	0.13 (0.30)	0.00	-0.13 (0.30)	0.00	-0.13 (0.30)
EQ ₃	0.00	-0.38 (0.62)	0.00	0.38 (0.63)	0.00	-0.31 (0.53)	0.00	0.30 (0.52)	-1.45	-1.04 (0.63)	1.44	1.08 (0.62)	-1.34	-0.93 (0.63)	1.34	0.97 (0.62)	0.00	0.13 (0.30)	0.00	0.13 (0.30)	0.00	-0.13 (0.30)	0.00	-0.13 (0.30)

aae_1^{mm} , aae_2^{mm} , aae_1^{ee} , aae_2^{ee} , aae_1^{me} , aae_2^{me} , aae_1^{em} , and aae_2^{em} are the maternal additive-additive $\times$ environment interaction effects in the first (subscript 1) and the second (subscript 2) environments, the endosperm additive-additive $\times$ environment interaction effects in the first (subscript 1) and the second (subscript 2) environments, the epistasis $\times$ environment interaction effects between maternal additive and endosperm additive in the first (subscript 1) and the second (subscript 2) environments, and the epistasis $\times$ environment interaction effects between endosperm additive and maternal additive in the first (subscript 1) and the second (subscript 2) environments, respectively. Epi. has the same definition as that in Table 3. Par., Est., and SD have the same definitions as those in Table 1.

right markers of the t th paired interval in the h th environment, respectively; and similarly, aa_{th}^{-mm} , aa_{th}^{-me} , aa_{th}^{-em} , and aa_{th}^{-ee} are the maternal, the maternal-endosperm, the endosperm-maternal, and the endosperm additive-additive interaction effects between two left markers of the t th paired interval in the h th environment and the remaining parameters are the same as in models (1) and (2).

To reduce computational complexity, we scan only the regions in which significant marker interval interactions were detected by the interval interaction analysis to find all paired QTL with potential epistatic effects. Suppose p candidate QTL are identified in one-dimensional QTL scanning and f pairs of candidate marker intervals are selected in marker pair selection; we can use the following model to test the epistatic effects between two putative QTL,

$$\begin{aligned}
 y_{hij} = & \mu_h + x_{i1}^{Am} x_{i2}^{Am} aa_h^{mm} + x_{i1}^{Am} x_{i2}^{Ae} aa_h^{me} + x_{i1}^{Ae} x_{i2}^{Am} aa_h^{em} \\
 & + x_{i1}^{Ae} x_{i2}^{Ae} aa_h^{ee} + \sum_k^f (\zeta_{il}^{+Am} \zeta_{ir}^{+Am} aa_{kh}^{+mm} \\
 & + \zeta_{it}^{+Am} \zeta_{ir}^{+Ae} aa_{kh}^{+me} + \zeta_{il}^{+Ae} \zeta_{ir}^{+Am} aa_{kh}^{+em} \\
 & + \zeta_{it}^{+Ae} \zeta_{ir}^{+Ae} aa_{kh}^{+ee}) + \sum_k^f (\zeta_{il}^{-Am} \zeta_{ir}^{-Am} aa_{kh}^{-mm} \\
 & + \zeta_{it}^{-Am} \zeta_{ir}^{-Ae} aa_{kh}^{-me} + \zeta_{il}^{-Ae} \zeta_{ir}^{-Am} aa_{kh}^{-em} \\
 & + \zeta_{it}^{-Ae} \zeta_{ir}^{-Ae} aa_{kh}^{-ee}) + \sum_s^p (x_{is}^{Am} a_{sh}^m + x_{is}^{Dm} d_{sh}^m + x_{is}^{Ae} a_{sh}^e \\
 & + x_{is}^{De} d_{sh}^e) + B_{j(h)} + \varepsilon_{hij},
 \end{aligned} \tag{5}$$

where all parameters and variables have the same definitions as those described above. Two-dimensional scanning is conducted and permutation testing is employed to determine an empirical threshold value of the F statistic. All significant paired epistatic QTL are distinguished. To remove false positive QTL, stepwise regression on all distinguished QTL is performed; as a result, a QTL full model could be established based on all significant QTL.

Genetic effect estimation: After obtaining the positions of the QTL and their epistatic interactions, a QTL full model is constructed to estimate all the parameters in model (1). To fit the mixed linear model, we first obtain a set of initial estimations of parameters including variances of random effects by the minimum norm quadric unbiased estimation (MINQUE) (Rao 1971), the fixed effects by the ordinary least-squares estimation (OLSE), and the random effects by the adjusted unbiased prediction (AUP). These estimates are then used as the prior values of parameters in the MCMC procedure for the mixed linear model implemented with Gibbs sampling (Smith and Roberts 1993). Parameter estimation and statistical inference are conducted by summarizing the Gibbs samplers (Wang *et al.* 1994; Yang *et al.* 2007). In the full model, we use MCMC to fit the parameters, implying an assumption of normality for the random effects. When the

Table 5 Estimates of QTL position and power values under three different simulation cases

QTL	Chr.	Pos.	Est. (SD)				Power	
Q ₁	3	26.0	^I 25.87 (1.90)	^{II} 25.59 (0.92)	^{III} 25.87 (1.90)	^I 99.80	^{II} 100.00	^{III} 99.80
Q ₂	4	63.0	63.60 (1.27)	63.63 (0.92)	63.60 (1.27)	100.00	100.00	100.00
Q ₃	5	51.0	50.98 (1.85)	51.14 (0.90)	50.98 (1.85)	100.00	100.00	100.00

I, II, and III stand for three different simulation cases, in which (I) the epistatic model is applied when epistasis exists, (II) the epistatic model is applied when epistasis does not exist, and (III) the reduced model without epistatic effects is applied when epistasis exists, respectively, under the same heritability of 70% and the same population structure of 600:1 (600 IF₂ lines and 1 offspring each IF₂ line). QTL, Chr., Pos., Est., SD, and Power have same definitions as those in Table 1.

assumption of normality does not hold true, the MINQUE can be used to estimate the variance components, OLSE for estimation of fixed effects, and AUP for prediction of random effects.

Results

Monte Carlo simulation

Simulation scenarios: We performed a series of Monte Carlo simulations to verify the unbiasedness and robustness of our method. The simulation study was conducted under different heritabilities, sampling strategies, and model parameters. The experiment design was random mating of IF₂, with 600 genotypes and two environments. The genetic map consisted of five chromosomes, each with 11 markers evenly distributed. The distance between 2 consecutive markers was set as 10 cM. Three QTL (Q₁, Q₂, and Q₃) and three epistatic interactions (EQ₁, EQ₂, and EQ₃) controlling an endosperm trait were scattered throughout the five chromosomes. EQ₁ indicated the interaction of Q₁ and Q₂; EQ₂ indicated the interaction of Q₂ and a locus on chromosome 1 without main effects; EQ₃ was the interaction between two loci on chromosomes 1 and 2, respectively, both of which had no main effects. Factors simulated include (I) three QTL heritabilities of 30%, 50%, and 70%; (II) 600 and 300 maternal plant lines in the segregating population, each line having one or two offspring; and (III) different models with or without inclusion of epistatic effects. Such scenarios not only allow us to examine the performance of parameter estimation but also provide useful guidance for practical experiment design. Each scenario repeated 500 simulations.

Estimation of QTL parameter: Since the simulation results had similar patterns for all the scenarios, we chose to report one simulation result under the heritability of 70% with population structure of 600:1 (*i.e.*, 600 IF₂ lines, each with one offspring). The results are summarized in Table 1, Table 2, Table 3, and Table 4. It was shown that all the parameter estimates were not significantly different from their true values in all the cases. All the statistical powers of QTL detection were extremely high and the lowest still reached 99.80% for Q₁. For the estimation of QTL effects, the estimation accuracy of main effects was slightly better than that of QE interaction effects.

Influence of different models: The influence of different models was investigated and three scenarios were simulated

under the same heritability (70%) with a population structure of 600:1 in the simulation procedure: (I) the simulated trait was controlled by epistatic effects and the epistatic model was used, (II) the simulated trait was controlled only by QTL main effects but the epistatic model was used, and (III) the simulated trait was controlled by epistatic effects but the reduced model without epistatic effects was used. The simulation results are summarized in Table 5, Table 6, Table 7, and Table 8. Unbiased estimates of QTL positions and close estimates of all effects (Table 6, Table 7, and Table 8) could be observed under the three situations, consistent with the results at different levels of heritability. Under the three simulations, we found that the estimation of main effects was slightly more stable than for the QE interaction effects. In scenario III, the QTL positions and main effects were well estimated by the reduced model but with higher SD. Moreover, this model could not detect epistasis. Hence, it could be concluded that in the presence of epistatic interaction, the model without epistatic interactions tended to result in unstable estimates of QTL parameters. Although the QTL full model was employed, statistical power and SD in scenario II were slightly better than in the other scenarios, which may be due to the absence of epistatic effects on variance of the simulated trait. Compared with scenario III, scenario I, which was based on the QTL full model, exhibited relatively lower SD and could detect epistasis, indicating the full model is more suitable for dissecting the genetic basis of complex traits. Hence, in practice, regardless of whether epistasis exists or not, we suggest using the QTL full model with inclusion of epistatic effects.

Analysis of cottonseed data

A set of cottonseed data from an IF₂ population was analyzed by the proposed method. A total of 188 recombinant inbred lines (RILs) were developed from intraspecific crossing between HS46 and WARKCBUCAG8US-1-88 parental strains, which have wide genetic differences in yield, fiber quality, disease resistance, and seed quality traits. In this study, every 2 lines randomly sampled from 188 RILs were crossed during flowering to produce 376 IF₂ lines in 2009 and 2010, which were used for QTL analysis. Seeds of the IF₂ population and two parents were manually harvested at maturity. Fiber percentage (FP) was measured for all seeds. The genetic map consists of 388 molecular markers across 30 linkage groups. It covers a total length of 1946.22 cM,

Table 6 Estimates of main effects under three different simulation cases

QTL	Chr.	a^m		d^m		a^e		d^e	
		Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)	Par.	Est. (SD)
Q ₁	3	1.80	^I 1.73 (0.55)	^{III} 1.79 (0.83)	^{II} 2.31 (0.40)	0.80	^{II} 0.82 (0.52)	^{III} 0.81 (0.80)	^{II} 2.34 (0.47)
Q ₂	4	1.80	^I 1.75 (0.53)	^{II} 1.96 (0.80)	^{III} 3.49 (0.36)	2.50	^I 2.45 (0.51)	^{II} 2.57 (0.76)	^{III} 0.78 (1.41)
Q ₃	5	2.50	^I 2.48 (0.48)	^{II} 2.47 (0.79)	^{III} 1.80 (0.35)	0.90	^I 0.91 (0.46)	^{II} 0.93 (0.74)	^{III} 2.53 (1.09)

^I, ^{II}, ^{III} stand for the same simulation cases as those in Table 5; QTL, Chr., Par., Est., SD, a^m , d^m , a^e , and d^e have the same definitions as those in Table 1.

accounting for 41.55% of the whole genome, with an average distance of 5.03 cM between adjacent markers.

The critical F value to declare QTL was calculated by permutation test at the 0.05 genome-wide significance level. Our method detected two significant QTL with sharp and narrow F -statistic peaks located on chromosomes 19 ($F = 5.20$) and 21 ($F = 9.93$) by a one-dimensional genome scan (Figure 1A). Moreover, a total of four QTL with two pairs of epistatic interactions ($F = 4.83$ and 5.47) were detected by a two-dimensional genome scan (Figure 1B). All the QTL and epistases seemed to be sensitive to environment. The genetic effects and the corresponding P -values of all the QTL and epistases are presented in Table 9. The estimated relative contribution (RC) of QTL 21-18 is 14.05%, which means that ~14.05% of phenotypic variation could be explained by this QTL. The suffix 21-18 indicates the QTL is located on the 18th marker interval on chromosome 21. Moreover, this study revealed that nearly 16.72% of the phenotypic variance is attributed to the epistatic interactions, with 4.01% from between-genome interactions and 12.71% from the interactions within one genome. Hence, in practical applications, we should take account of the epistasis from both QTL within the same genome and those in different genomes. More attention should be paid to QTL 4-2, which is involved in two epistases but has no significant marginal effects.

Discussion

In this study, we proposed a statistical method for mapping endosperm or embryo traits of crops that integrated into one mapping framework the maternal and offspring effects of multiple QTL, epistatic interactions either within one genome or between two genomes, and QE interactions. Epistasis is a topic of current interest in molecular and quantitative genetics and has been proposed to be the source of missing heritability (Zuk *et al.* 2012). Uncovering genetic interaction networks from epistatic interactions between loci will improve our understanding of biological systems that give rise to quantitative trait variation and help reveal the mechanisms that underlie genetic homeostasis and speciation. Knowledge of interaction loci will improve our ability to predict individual disease risk factors in humans and response to natural and artificial selection, as well as inbreeding depression and heterosis in agricultural species. In addition, GE interactions are also an important issue for breeders and quantitative geneticists. It has been well documented that environmental conditions, especially temperature during grain filling, significantly affect cooking and eating quality (Zheng *et al.* 2008). Therefore, detection of GE interactions is very valuable for breeding environment-specific or broadly adapted varieties in genetic improvement of crop traits by marker-assisted selection.

However, despite being important genetic components of seed trait variation, epistasis and GE interactions have not been integrated into a single-QTL mapping model, mainly due to lack of an appropriate methodology. One major difficulty in developing a powerful statistical approach for

Table 7 Estimates of the maternal interaction effects with environments under three different simulation cases

QTL Chr.	Par.	ae_1^m			ae_2^m			de_1^m			de_2^m		
		Est.	SD	Par.	Est.	SD	Par.	Est.	SD	Par.	Est.	SD	Par.
Q ₁	3	-1.28	1.07(0.46)	1.28	1.07(0.46)	1.28	1.07(0.46)	1.28	1.07(0.46)	1.28	1.07(0.46)	1.28	1.07(0.46)
Q ₂	4	1.18	0.92(0.43)	-1.18	0.92(0.43)	-1.18	0.92(0.43)	1.18	0.92(0.43)	-1.18	0.92(0.43)	-1.18	0.92(0.43)
Q ₃	5	0.00	-0.00(0.05)	0.00	-0.00(0.05)	0.00	-0.00(0.05)	0.00	-0.00(0.05)	0.00	-0.00(0.05)	0.00	-0.00(0.05)

I, II, III stand for the same simulation cases as those in Table 5; QTL, Chr., Par., Est., SD, ae_1^m , ae_2^m , de_1^m , and de_2^m have the same definitions as those in Table 1 and Table 2.

mapping seed trait QTL with epistasis and QE interaction is fitting many parameters of multiple QTL in the full model, especially for epistasis from multiple genomes. Unlike previous one-QTL models that produce bias when there is more than one linked QTL, we use a QTL full model and conduct model selection to remove false positive QTL. Thus our approach is more effective for the precise estimation of all QTL positions and main effects, as well as the QE interaction effects, although the number of parameters in the model increases rapidly compared with that in the other methods. In QTL genome-wide scanning and model selection on QTL, we use the *F*-test to test significance of QTL based on Henderson's method III, which avoids inversion of large matrices and therefore requires less computation.

Permanent genetic resources such as RILs or IF₂ possess many advantages in QTL studies. One is that they can be used to conduct multiple-environment experiments, which are required for studying GE interaction. Another advantage is that the population derived from them need not be genotyped and can be inferred from their genotypes. In this study, we utilize a one-generation design that needs only the maternal marker information and the offspring phenotypes and thus can reduce a large amount of labor and costs, making it easier to implement for breeders in practical studies. Certainly, the work of hybridization in the random-mating design is laborious, but it can produce a very large population of hybrid lines without additional genotyping costs and effectively increases the statistical power of QTL analysis.

As shown by simulation studies, our model is quite robust and reliable in estimation of parameters for a modest sample size (200) and low heritability (0.3). Mapping power for both QTL and epistasis increased when the number of maternal plant lines increased from 100 to 600, while the false discovery rate (FDR) decreased (data not shown). Based on our simulations, we suggest that at least 200 maternal plant lines should be used for sufficient detection power for epistasis and GE interactions, while keeping a reasonable hybridization workload. It is noteworthy that increasing maternal line numbers may be more important than increasing offspring numbers to increase estimation precision. The simulation results from different models in the current study indicate that when the epistatic effect is small and a reduced model is employed, the estimations of positions and effects of QTL are relatively precise. However, when the epistatic effect is large and we still use a reduced model, the position estimations also reach better precision, but the estimated effects of QTL will be biased severely. This is because when we search the significant QTL by a one-dimensional scan, the reduced model does not include interaction effects of paired markers to control background genetic effects. As a result, when we estimate all kinds of effects in the reduced model, the main effects of QTL will be confounded with the epistatic effects. However, the *F*-values profile still exhibits a very similar pattern, and therefore the estimation of QTL position is less influenced by the epistatic effects. Since epistasis is mostly

Table 8 Estimates of the endosperm interaction effects with environments under three different simulation cases

QTL Chr.	Par.	ae_1^e			ae_2^e			de_1^e			de_2^e		
		Est.	(SD)	Par.	Est.	(SD)	Par.	Est.	(SD)	Par.	Est.	(SD)	Par.
Q ₁	3	1.38	1.09 (0.41)	1.22 (0.24)	1.09 (0.41)	1.38 (0.69)	1.09 (0.41)	1.32 (0.21)	1.09 (0.41)	1.32 (0.21)	1.09 (0.41)	1.32 (0.21)	1.09 (0.41)
Q ₂	4	-1.64	-1.36 (0.40)	-1.55 (0.22)	-1.64 (0.61)	-1.16 (0.04)	-1.64 (0.61)	1.42 (0.38)	1.49 (0.24)	1.14 (0.62)	1.42 (0.38)	1.49 (0.24)	1.14 (0.62)
Q ₃	5	0.00	-0.00 (0.05)	-0.00 (0.05)	-0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.05)	0.00 (0.05)	0.00 (0.04)	0.00 (0.05)	0.00 (0.04)	0.00 (0.04)

I, II, III stand for the same simulation cases as those in Table 5; QTL, Chr., Est., SD, ae_1^e , ae_2^e , de_1^e , and de_2^e have the same definitions as those in Table 1 and Table 2.

involved in genetic variation of quantitative traits and our method will give accurate estimation of QTL effects no matter whether the epistasis and QE interaction exist or not, we suggest using the QTL full model with inclusion of epistasis and QE interactions for seed trait QTL studies where the genetic architecture is unknown. In the cottonseed data analysis, we found some epistatic QTL without significant individual effects detected. It is noteworthy that these types of epistasis are also very important in genetic buffering and can respond to natural and artificial selection. Moreover, the epistases from different genomes also contribute to phenotypic variation; this kind of epistasis should not be ignored in mapping seed trait QTL.

Although the mechanisms for genetic imprinting are not totally understood, this phenomenon is thought to offer an evolutionary advantage through the maintenance of greater genetic variation. An imprinting QTL can function in a coordinated network of gene-gene and gene-environment interactions. It is worthwhile to develop new statistical models for the detection of interacting regulatory genes that affect the imprinting expression of any QTL involved in a genetic network composed of maternal and zygotic genomes. Under our framework, it is in principle feasible to extend the model to accommodate imprinting effects.

The mixed linear model (1) developed in the present study provides a basic framework that can easily be extended to cover more complex experimental designs, such as double backcross of IF₂ or selfing of IF₂ based on a one- or two-generation design. In the two-generation designs, such as double backcross of IF₂ and selfing of IF₂, that include maternal and offspring marker information as well as the phenotypes of the offspring, our model can give accurate position estimations, all the QTL main effects, and the interaction effects, including GE interactions and epistatic interactions. In one-generation designs, which use only maternal marker information and phenotypes of offspring, for double backcross of IF₂ lines linear dependency will occur between genetic effects of any pair of QTL when the QTL full model (1) is applied. To tackle this problem, a potential strategy is to keep the endosperm dominance effect of the QTL while merging the dominance effect of all other QTL, so that the linear dependency could be removed; however, this procedure will increase programming complexity and require much more computation time because the model has to be adjusted for different QTL. Given that the effectiveness of this strategy still needs to be investigated, we do not currently implement it in our software.

Based on the models and methods proposed in the present study, the QTLNetwork-Seed-1.0.exe software was developed in the C++ programming language. This software can be run on commonly used operation systems and can analyze data for 1000 individuals and 120 markers in ~2 min with 45 MB memory (data not shown). The simulations on median running time and memory usage suggested that, as expected, the computing time and the memory usage increased with increasing maternal plant lines and marker density. It would

Table 9 Estimated position and effects of QTL for cottonseed fiber percentage in an IF₂ population

QTL	Chr.	Pos.	a^m	a^o	d^o	ae_1^m	ae_2^m	ae_1^o	ae_2^o	de_1^o	de_2^o
19-5	19	21.3	1.51**	-0.49**	1.61**	1.09**	-1.10**	-0.16	0.16	1.44*	-1.47**
21-18	21	47.9	-0.91**	1.93**	0.71**	-0.44	0.46	1.58**	-1.52**	0.40	-0.41

Epi.	QTL1		QTL2		aa^{mm}	aa^{oo}	aa^{mo}	aa^{om}	aae_1^{mm}	aae_2^{mm}	aae_1^{oo}	aae_2^{oo}	aae_1^{mo}	aae_2^{mo}	aae_1^{om}	aae_2^{om}
	Chr.	Pos.	Chr.	Pos.												
(4-2, 20-2)	4	4.1	20	7.8	-0.92**	1.14**	-0.42	0.33	-0.31	0.32	1.16*	-1.19*	-0.51	0.51	0.00	0.00
(4-2, 25-13)	4	4.1	25	25.3	0.76**	0.01	-0.31	0.63*	0.28	-0.30	0.00	0.00	0.00	0.00	0.79*	-0.79*

The QTL are designed with the serial numbers of chromosome and marker interval; a^o (d^o) is the embryo additive (dominance) effect of one QTL; ae_1^o , ae_2^o , de_1^o , and de_2^o are the interaction effects of the embryo additive and the embryo dominance with the first (subscript 1) and the second (subscript 2) environments, respectively; aa^{oo} , aa^{mo} , and aa^{om} are the embryo additive-additive epistatic effect, the maternal-embryo additive-additive epistatic effect, and the embryo-embryo additive-additive epistatic effect, respectively; aae_1^{oo} , aae_2^{oo} , aae_1^{mo} , aae_2^{mo} , aae_1^{om} , and aae_2^{om} are the embryo additive-additive $\times$ environment interaction effects in the first (subscript 1) and the second (subscript 2) environments, the epistasis $\times$ environment interaction effects between maternal additive and embryo additive in the first (subscript 1) and the second (subscript 2) environments, and the epistasis $\times$ environment interaction effects between embryo additive and maternal additive in the first (subscript 1) and the second (subscript 2) environments, respectively; * and ** indicate reaching significance at a level of 0.05 and 0.01, respectively. Chr., Pos., a^m , ae_1^m , ae_2^m , aa^{mm} , aae_1^{mm} , and aae_2^{mm} have the same definitions as those in Table 1, Table 2, Table 3 and Table 4.

take 2.28 hr and 131 MB memory for a set of samples consisting of 600 maternal lines, 1198 individuals, and 1005 markers, suggesting that our software can handle much

larger data than those commonly used. We noted that running time and memory usage comparisons might vary as a function of computing environment. Different experimental

Figure 1 *F*-statistic plots from (A) one-dimensional genome scan for QTL with individual effects and (B) two-dimensional genome scan for epistasis of fiber percentage (FP) in cottonseed. (A) Two peaks exceed the threshold *F* value (4.55) calculated by permutation tests on chromosomes 19 and 21, respectively. (B) Two significant peaks are detected that have a much larger *F* value than the threshold (4.24).

designs, such as double backcross of IF₂; selfing of IF₂ based on one- or two-generation designs; and random mating for IF₂, double haploid, or RILs can be handled by this software for detecting genetic architecture of seed endosperm or embryo traits.

Acknowledgments

This work was supported in part by the National Basic Research Program of China (973 programs 2011CB109306 and 2010CB126006), the National Natural Science Foundation grants 31271608 and 31470083, the National Institutes of Health grant DA025095, the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences, and the Bill and Melinda Gates Foundation Project. The authors declare no conflict of interest.

Literature Cited

- Bao, J., M. Sun, L. Zhu, and H. Corke, 2004 Analysis of quantitative trait loci for some starch properties of rice (*Oryza sativa* L.): thermal properties, gel texture and swelling volume. *J. Cereal Sci.* 39: 379–385.
- Cui, Y., and R. Wu, 2005a Mapping genome-genome epistasis: a high-dimensional model. *Bioinformatics* 21: 2447–2455.
- Cui, Y., and R. Wu, 2005b Statistical model for characterizing epistatic control of triploid endosperm triggered by maternal and offspring QTLs. *Genet. Res.* 86: 65–76.
- Cui, Y., G. Casella, and R. Wu, 2004 Mapping quantitative trait loci interactions from the maternal and offspring genomes. *Genetics* 167: 1017–1026.
- Doerge, R. W., and G. A. Churchill, 1996 Permutation tests for multiple loci affecting a quantitative character. *Genetics* 142: 285–294.
- Fan, C. C., X. Q. Yu, Y. Z. Xing, C. G. Xu, L. J. Luo *et al.*, 2005 The main effects, epistatic effects, and environmental interactions of QTLs on the cooking and eating quality of rice in a doubled-haploid line population. *Theor. Appl. Genet.* 110: 1445–1452.
- He, P., S. Li, Q. Qian, Y. Ma, J. Li *et al.*, 1999 Genetic analysis of rice grain quality. *Theor. Appl. Genet.* 98: 502–508.
- Hu, Z., and C. Xu, 2005 A new statistical method for mapping QTLs underlying endosperm traits. *Chin. Sci. Bull.* 50: 1470–1476.
- Lark, K. G., K. Chase, F. Adler, L. M. Mansur, and J. H. Orf, 1995 Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. *Proc. Natl. Acad. Sci. USA* 92: 4656–4660.
- Mackay, T. F., E. A. Stone, and J. F. Ayroles, 2009 The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10: 565–577.
- Mazur, B., E. Krebbers, and S. Tingey, 1999 Gene discovery and product development for grain quality traits. *Science* 285: 372–375.
- Nickels, S., T. Truong, R. Hein, K. Stevens, K. Buck *et al.*, 2013 Evidence of gene–environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genet.* 9: e1003284.
- Olsen, O. A., 1998 Endosperm developments. *Plant Cell* 10: 485–488.
- Piepho, H.-P., and H. G. Gauch, 2001 Marker pair selection for mapping quantitative trait loci. *Genetics* 157: 433–444.
- Qi, T., B. Jiang, Z. Zhu, C. Wei, Y. Gao *et al.*, 2014 Mixed linear model approach for mapping quantitative trait loci underlying crop seed traits. *Heredity* 113: 224–232.
- Rao, C. R., 1971 Minimum variance quadratic unbiased estimation of variance components. *J. Multivariate Anal.* 1: 445–456.
- Renz, H., E. von Mutius, P. Brandtzaeg, W. O. Cookson, I. B. Autenrieth *et al.*, 2011 Gene–environment interactions in chronic inflammatory disease. *Nat. Immunol.* 12: 273–277.
- Smith, A. F., and G. O. Roberts, 1993 Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *J. R. Stat. Soc. B* 55: 3–23.
- Tsilo, T. J., S. Simsek, J.-B. Ohm, G. A. Hareland, S. Chao *et al.*, 2011 Quantitative trait loci influencing endosperm texture, dough-mixing strength, and bread-making properties of the hard red spring wheat breeding lines. *Genome* 54: 460–470.
- Walker, C. K., R. Ford, M. Muñoz-Amatriaín, and J. F. Panozzo, 2013 The detection of QTLs in barley associated with endosperm hardness, grain density, grain size and malting quality using rapid phenotyping tools. *Theor. Appl. Genet.* 126: 2533–2551.
- Wang, C., J. Rutledge, and D. Gianola, 1994 Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genet. Sel. Evol.* 26: 91–115.
- Wen, Y., and W. Wu, 2007 Interval mapping of quantitative trait loci underlying triploid endosperm traits using F3 seeds. *J. Genet. Genomics* 34: 429–436.
- Wu, R., X. Y. Lou, C. X. Ma, X. Wang, B. A. Larkins *et al.*, 2002 An improved genetic model generates high-resolution mapping of QTL for protein quality in maize endosperm. *Proc. Natl. Acad. Sci. USA* 99: 11281–11286.
- Xing, Z., F. Tan, P. Hua, L. Sun, G. Xu *et al.*, 2002 Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. *Theor. Appl. Genet.* 105: 248–257.
- Yang, J., J. Zhu, and R. W. Williams, 2007 Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics* 23: 1527–1536.
- Yu, S., J. Li, C. Xu, Y. Tan, Y. Gao *et al.*, 1997 Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA* 94: 9226–9231.
- Zeng, Z.-B., 1994 Precision mapping of quantitative trait loci. *Genetics* 136: 1457–1468.
- Zhang, L., M. C. Yang, X. Wang, B. A. Larkins, M. Gallo-Meagher *et al.*, 2004 A model for estimating joint maternal-offspring effects on seed development in autogamous plants. *Physiol. Genomics* 19: 262–269.
- Zheng, X., J. Wu, X. Lou, H. Xu, and C. Shi, 2008 The QTL analysis on maternal and endosperm genome and their environmental interactions for characters of cooking quality in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 116: 335–342.
- Zuk, O., E. Hechter, S. R. Sunyaev, and E. S. Lander, 2012 The mystery of missing heritability: genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. USA* 109: 1193–1198.

Communicating editor: E. A. Stone