

Statistical approaches in QTL mapping and molecular breeding for complex traits

XU HaiMing & ZHU Jun*

Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China

Received October 29, 2011; accepted February 27, 2012; published online April 5, 2012

Most of the important agronomic traits in crops, such as yield and quality, are complex traits affected by multiple genes with gene \times gene interaction as well as gene \times environment interaction. Understanding the genetic architecture of complex traits is a long-term task for quantitative geneticists and plant breeders who wish to design efficient breeding programs. Conventionally, the genetic properties of traits can be revealed by partitioning the total variation into variation components caused by specific genetic effects. With recent advances in molecular genotyping and high-throughput technology, the unraveling of the genetic architecture of complex traits by analyzing quantitative trait locus (QTL) has become possible. The improvement of complex traits has also been achieved by pyramiding individual QTL. In this review, we describe some statistical methods for QTL mapping that can be used to analyze QTL \times QTL interaction and QTL \times environment interaction, and discuss their applications in crop breeding for complex traits.

complex traits, QTL mapping, epistasis, environment interaction, crop breeding

Citation: Xu H M, Zhu J. Statistical approaches in QTL mapping and molecular breeding for complex traits. *Chin Sci Bull*, 2012, 57: 2637–2644, doi: 10.1007/s11434-012-5107-1

The most important economic traits in crops are quantitative in nature. The genetic variation of quantitative traits is usually controlled by minor-effect genes and the environments as well as by epistatic effects between different genes and by interactions between genes and environments. Because of their complicated genetic architecture, quantitative traits are usually referred to as complex trait. The genetic architecture includes knowledge about the numbers and positions of the genes in the chromosomes, the magnitude of their effects, and the contributions of additive, dominance and epistatic effects [1].

Most plant breeders and quantitative geneticists have long been interested in uncovering the genetic architecture of yield and other complex traits [2]. Understanding genetic architecture can provide insights into how the architecture is translated into genetic variation and selection response [3]. Based on diallele cross experiments, the genetic architecture of a trait has been analyzed by decomposing the trait varia-

tion into additive, dominance, and epistasis variances and studying their interaction variances with the environments [4–7]. The total genetic covariance between two traits can also be dissected into genetic covariate components. With the increasing availability of molecular markers, the quantitative genetics theory of complex traits and marker-assisted selection has undergone major advances. It is now possible to connect marker variations with phenotypic variations, and to dissect genetic architecture of traits into individual genetic variants. The selection of traits can also be achieved using individual quantitative trait loci (QTLs). In this review, we discuss some of the statistical methods used for QTL mapping of experimental populations and examine their applications in marker-assisted selection (MAS) for improving complex traits.

1 QTL mapping

In the late 1980s, the discovery of abundant molecular

*Corresponding author (email: jzhu@zju.edu.cn)

markers and advances in rapid genotyping technology led to the development of statistical methods for use in the QTL mapping of complex quantitative traits. One of the most well-known methods is the interval mapping (IM) approach developed by Lander and Botstein [8]. This landmark method for QTL mapping has established a statistical framework for most methods that are currently used to analyze the QTLs of complex traits. The IM approach can be used to search for QTLs in whole genomes using a log likelihood (LOD) profile. A QTL associated with a quantitative trait is assumed to be located on the genome under the peak where the LOD is higher than a specified threshold. Zeng [9,10] proposed the composite interval mapping (CIM) method which improved upon the precision of QTL mapping by including associated markers as covariant variables. The CIM approach can provide unbiased estimations of QTL positions and effects, under the assumption of no QTL \times environment interaction. Similarly, Jansen [11] suggested that markers outside of the scanned interval in the regression model could be included as covariant variables to control type I and type II errors in QTL mapping. The IM and CIM methods have been widely applied in experimental populations, such as doubled haploid (DH) population, recombinant inbred line (RIL) population, backcross (BC) population, and F_2 population.

Epistasis refers to the masking of genotypic effects at one locus by the genotypes of another locus [3] and also to any statistical interaction between the genotypes at two or more loci [12]. It is believed that epistasis and interaction between genes and environments are involved in the genetic variations that underlie quantitative traits. Because multi-environmental data cannot be integratively analyzed by the IM and CIM methods, mapping analyses are usually conducted separately in individual environments [13–15]. Separate analysis for the different environmental data cannot provide reliable estimates either for the main QTL effects or for the QTL \times environment interaction effects across multiple environments. Kao and Zeng [16,17] developed the multiple interval mapping (MIM) method to analyze the epistatic effects of QTLs. This method first screens for significant QTLs using the CIM approach; then, a multiple QTL model is constructed by including the epistatic effects of paired significant QTLs. The MIM method can only detect epistasis between two QTLs that are first detected using CIM. Often, epistatic effects can be as large as the main QTL effects, and epistasis can also occur in opposite directions between different pairs of interacting loci and even between loci that have no effect on the trait [12,18–20].

A mixed linear model approach that includes the main QTL effects, environment effects and QTL \times environment interaction effects has been proposed for mapping QTLs [21,22]. The main QTL effects are treated as the fixed effects and the random effects are the environment effects, the QTL \times environment interaction effects, and the associated markers included in the model to control for background

genetic effects. Similar to the procedure used in the CIM method, the profile of the log likelihood ratio (LOD) is constructed for the entire genome, the significance of the main QTL effect is tested by the jackknife resampling method, and the position of a significant QTL is determined by the peak of the LOD profile. The method proposed by Zhu [21] is abbreviated to MCIM, denoting mixed-model-based composite interval mapping. Based on the analysis framework of the MCIM method, Wang et al. [23] proposed a genetic model of QTL mapping for DH and RIL populations and developed the QTLMapper software (<http://ibi.zju.edu.cn/software/qtlmapper/>) to map QTLs with additive and epistatic effects and QTL \times environment interactions. This model can analyze the additive effects of two QTLs (a_i, a_j), the additive \times additive interaction effect (aa_{ij}) of paired QTLs, and the interaction effects of the QTLs with environments ($ae_{ih}, ae_{jh}, aae_{ijh}$). Gao and Zhu [24] proposed a genetic model for analyzing the immortalized F_2 population from random mating among individuals of either DH or RIL populations. This model can simultaneously analyze additive effects (a_i, a_j), dominance effects (d_i, d_j), and epistatic effects for paired QTLs ($aa_{ij}, ad_{ij}, da_{ij}, dd_{ij}$) along with their interaction effects with environments ($ae_{ih}, ae_{jh}, aae_{ijh}, dae_{ijh}, dde_{ijh}$). An advantage of MCIM over the other methods is that it provides unbiased estimations of the positions and effects of QTLs, and also predicts the random QTL \times environment interaction effects. The MCIM method is very useful for analyzing complex traits that are affected by epistasis and QTL \times environment interactions.

The MCIM approach adopts a two-dimensional search strategy to scan paired QTLs involved in epistasis and the jackknife resampling technique is used to test the significance of QTL effects. Although digenic epistasis could be analyzed by MCIM, the two loci model cannot reflect the intricate architecture of quantitative traits where the interaction of more than two loci may be involved. To calculate the inverse of the matrix that is generated using the MCIM method a great deal of computation time is required. Therefore, the MCIM method cannot appropriately handle the large amount of experimental data that are generated when multiple environments with replications are considered. To address this problem, Yang et al. [25] proposed a new QTL mapping approach for the complex traits in the form of an optimized MCIM that involves four steps: (i) the use of an F -statistic based on Henderson method III for hypothesis tests. This method is less computationally greedy than the corresponding likelihood ratio test; (ii) permutation testing to control the genome-wide false positive rate, and model selection to reduce ghost peaks in the F -statistic profile; (iii) construction of a full-QTL model in which the significance of each QTL is re-tested; (iv) application of the Bayesian method via Gibbs sampling for mixed linear model to estimate the QTL effects and to conduct statistical inference. Based on this new approach, the software of QTLNetwork 2.0 was developed and can be downloaded from <http://ibi>.

zju.edu.cn/software/qtlnetwork/ [26,27]. Notably, QTLNetwork can be used to analyze data from DH, RIL and IF2 populations, as well as data from advanced populations (B \times F $_y$) derived from the two parents by backcrossing x times and selfing y times. The software also provides the potentially superior QTL design for improving complex traits using the main QTL effects and the QTL \times environments interaction effects [28]. Because of the advantages of using MCIM to analyze digenic epistasis and QTL \times environment interactions, and the user-friendly software that is now available, the MCIM approach has been widely applied to the analyses of the QTLs for complex traits in various plant experimental populations such as wheat [29–34], barley [35], soybean [36–39], cucumber [40], cabbage [41], cotton [42], rapeseed [43,44] and rice [45–48], as well as in *Drosophila* [49].

Most QTL mapping methods can only analyze the QTLs in one genome at a time. To completely understand the genetic basis of, for example, host-parasite interactions, it is necessary to study not only the resistance/infection genes, but also to investigate gene interactions between the two species. A genetic model that integrates genetic information from the host and parasite genomes has been proposed [50]. This model can map the QTLs that are involved in host-parasite interactions and can be used to detect interactions between these QTLs.

The accurate estimation of QTL positions and effects requires the normal distribution of residuals in the experimental data. However, the data often contain outliers and influential observations that may seriously affect the estimation of model parameters and can lead to errors in the detection of QTL positions and their predicted effects [51]. A method for detecting outliers and influential observations prior to QTL mapping analysis has been described [52]. The statistical analysis software QTModel (<http://ibi.zju.edu.cn/software/qtmodel/>) can be used to identify outliers in experimental data. Simulations have shown that the presence of a small proportion of outliers can increase false QTL discovery rates and hence decrease the detection power of QTLs.

2 Dynamic mapping for developmental traits

The development of quantitative traits occurs through the actions and interactions of many genes that may behave differentially during growth periods [53]. Conventional genetic analyses of quantitative traits have shown that gene effects and their mode of action are different at different developmental stages [54,55]. Most QTL mapping studies have focused on analyzing the performance traits at the final growth stage; however, understanding dynamic changes in the genetic architecture of quantitative traits is crucial in developmental genetics. Currently, several approaches for analyzing the dynamics of QTL expression are available. The phenotypic values measured at different growth stages

are usually analyzed directly by QTL mapping, and the genetic architecture of a trait is dynamically inferred by comparing of the results of the mapping [56–58]. Alternatively, the increment of traits measured at sequential time points can be analyzed to predict the genetic architecture [56,58]. A third method includes an appropriate growth function in the QTL model and then fits the model before analyzing the model parameters using all the phenotypic values [58–60]. A fourth method treats the phenotypic values at different developing time points as different traits, and use the methods of the multi-trait CIM method [61] to analyze the QTLs [58]. None of these methods take account of either the epistasis of genes or gene \times environment interactions, nor do they include the possible correlation between different phenotypes at different growth stages. A conditional mapping approach has been proposed to first predict the conditional effects $y_{(t|t-1)}$ for the phenotypic values at time (t) given the phenotype at time ($t-1$); $\hat{y}_{(t|t-1)}$ is then used to analyze the QTLs using an MCIM approach [21,22]. The predicted $\hat{y}_{(t|t-1)}$ can be obtained by running the sub-menu “QTL Data” in the software QGASStation (<http://ibi.zju.edu.cn/software/qga/>). Because the conditional effects $y_{(t|t-1)}$ are independent of the phenotypes at time ($t-1$), they can effectively measure the net genetic effects of gene expression from the time ($t-1$) to the time (t). Therefore, this conditional QTL mapping strategy can identify new QTLs for which the accumulated expression effects are just large enough to be detected, as well as QTLs that begin to appear at an early specific growth stage and new genetic effects occur in the period.

Yan et al. [53] conducted conditional QTL mapping on the plant height of rice in a DH population. The results showed that the number of QTLs detected at various measuring stages was different. Some QTLs were detected at all stages and some were seen only at one or several stages. Cao et al. [62] studied the impact of epistasis and QTL \times environment interactions on the developmental behavior of plant height in rice by conventional and conditional QTL mapping. Their results showed that most epistasis effects could be detected by conditional QTL mapping alone, and that QTL \times environment interaction effects were detected more often than the main QTL effects for plant-height behavior. Conditional QTL mapping was widely used as a valid way of revealing dynamic gene expression for the development of quantitative traits. Recently, conditional QTL mapping, or a combination of unconditional and conditional QTL mapping has been used to study the dynamics of developmental traits. Sun et al. [63] used conditional and unconditional mapping to study the developmental behavior of soybean. Yang et al. [45] used conditional mapping to dissect the developmental behavior of tiller number and plant height, and found that conditional mapping was superior to conventional mapping for studying developmental traits. Han et al. [36] studied the impact of epistasis and QTL \times environment interactions on the accumulation of seed mass in soybean using conditional mapping. Wang et

al. [48] employed conditional and unconditional QTL mapping to analyze rice fat content and fat index. Jiang et al. [47] used a combined conditional and unconditional mapping strategy to find QTLs underlying plant height and tiller number in rice exposed to two different nitrogen levels. Cheng et al. [41] used conditional mapping to identify QTLs associated with plant height at various stages of non-heading Chinese cabbage. Han et al. [64] conducted conditional QTL analysis of the linolenic acid content in different developmental stages of soybean seed.

3 QTL dissection of genetic correlation between traits

Complex traits are sometimes determined by their component traits; for example, rice yield may be measured by panicle number per plant, filled grain number per panicle and/or weight per grain. The relationship at the molecular level between traits of interest can yield insights that may be of importance for the improvement of multiple traits or for designing indirect selection schemes. Studies of this kind can be supported by a combined conditional and unconditional QTL mapping approach. Such an approach can be used to estimate the genetic contribution of one component trait (cause-trait) to the resultant trait of interest. For example, Guo et al. [65] studied rice yield and its component traits using this kind of approach and a similar study has been reported by Liu et al. [66]. Zhao et al. [67] analyzed the relationship between oil content in rapeseed with protein content and other traits related to plant development. They found clear evidence for a strong genetic link between protein content and oil content, and also discovered two QTLs that controlled oil content but were independent of the protein content. It was proposed that these two QTLs could be used in practical breeding programs to increase oil content without affecting seed protein content [67]. Li et al. [68] identified three QTLs in popcorn that controlled popping expansion volume (PEV) independent from grain weight per plant (GWP)/100-grain weight (100 GW). These QTLs may be potential candidates to increase PEV in popcorn without decreasing the GWP/100 GW value. Ye et al. [69] studied the genetic relationships in the panicle characteristics of rice using a similar approach.

4 Marker-assisted selection based on QTLs

The principle of the Marker-assisted selection (MAS) uses molecular markers in linkage disequilibrium with the quantitative trait of interest, to predict the genetic value of individuals and to improve the selection efficiency of a trait. In early 1923, the improvement of a quantitative trait by selection of a single marker gene, the colour of the seed, was first exploited by Sax [70]. Since then, MAS studies of single

loci have been carried out [71–73]. However, the extensive theoretical and practical studies have been conducted only in the last 15 years. This growth owes much to the technique for the development of molecular markers [74] and the revolution in statistical methodology for mapping the QTLs of quantitative traits. Lande and Thompson [75] proposed a marker-index selection method based on multiple variables regression, and this is the theory that was used in initial studies of MAS.

Currently, the main MAS breeding methods include: (i) Marker-assisted introgression (MAI) or marker-assisted backcross (MAB); (ii) the simple screening of populations; (iii) gene pyramiding schemes; (iv) Marker-based recurrent selection; and (v) selection on an index combining molecular and phenotypic scores [76]. Recently, using information from QTL analyses, MAS has been successfully applied to improve quantitative traits. For example, Tanksley and McCouch [77] successfully introgressed QTLs from the wide variety of *Lycopersicon hirsutum* into a cultivar variety (*Lycopersicon esculencum*) to create a line that outperformed the original elite variety for yield, soluble solids content, and fruit color. Ashikari et al. [78] simultaneously improved grain number and plant height by pyramiding one QTL for grain number and another QTL for plant height; however, the trait values of the generated line were slightly lower than expected based on single introgression lines. Barloy et al. [79] pyramided two cereal cyst nematode resistance genes from *Aegilops variabilis* into wheat by MAS. Many other successful applications of MAS in numerous species have been reported [80,81].

Despite its successful use, MAS is still not as efficient as was anticipated for most quantitative traits, probably because of the epistasis of two or more QTLs, QTL \times genetic background interactions and/or QTL \times environment interactions that may reduce QTL effects compared to what is expected from single-gene effects [76]. The efficiency of MAS also depends on the magnitude of heritability and the accuracy of the QTL information, including their positions and effects. Traditionally, theoretical and simulation studies on MAS have focused only on the additive quantitative traits and the non-additive effects, such as epistatic effects, have not been included. Gimelfarb and Lande [82] showed that selection using marker scores that contain both additive and non-additive genetic effects have a higher efficiency as compared to selection using the marker scores that contain only the additive genetic effects. Liu et al. [83] conducted a simulation study on a F_2 population derived from two inbred lines to investigate the efficiency of MAS when the epistatic effects were considered. They showed that MAS generally yielded a longer persistence response when the epistatic effects were considered, as compared with the response that was based exclusively on additive or additive and dominance effects. Thus, neglecting epistasis can result in a considerable loss in response which becomes more pronounced in later generations; the precision of QTL de-

tection and the magnitude of its effects are imperative to realize the potentials of MAS [83]. Liu et al. [84] included the additive, dominance and epistatic effects in MAS and proposed a potential index to measure the probability of an individual showing superior genotypes under selfing. They showed that MAS not only revealed larger genetic responses but also dramatically increased the frequencies of superior genotypes compared with phenotypic selection. However, the advantages of MAS over phenotypic selection were considerably reduced when conducting selection in later generations. Hu [74] extended Lande and Thompson's theory [75] to incorporate both additive and non-additive effects into MAS with reference to the mass selection case. There were four different indices proposed, which were phenotype-, general combining ability (GCA)-, and GCA with reciprocal effects-based marker scores. Hu found that these indices had higher selection efficiencies than the index with only additive effects-associated markers as long as detectable transient non-additive effects are present; the improvement of selection efficiency depends on the magnitude of non-additive variances and the proportion of them that are explained by the markers. The index with the phenotype-based marker scores that are applicable to any population of non-random mating has the highest selection efficiency [74].

Although some successful examples in improvement of quantitative traits by MAS have been reported, some quantitative traits were changed from complex traits to Mendelian traits after discovering that they were controlled by a surprisingly low number of QTLs with large effects [76]. It is still challenging to improve the efficiency of MAS for the majority of quantitative traits that are generally controlled by a large number of QTLs with small effects, and that are affected by their interaction with environments. It is well known that the efficiency of MAS is essentially determined by the genetic architecture of the trait. Yang et al. [28] developed new methods for predicting superior genotypes (superior line and superior hybrid) based on QTL effects including epistasis and QTL \times environment interactions. When QTL mapping is conducted using QTLNetwork, the predicted genotypic values (G) can be obtained for two parents (P_1 and P_2) and F_1 with G in all environments and G + environment interaction in the h th environment (GEh), and also the superior genotypes of QTLs are also presented. The predicted superior lines and superior hybrids can be found with superiorities over parents and F_1 hybrid, indicating the vast breeding potential that can still be further improved. Apparently, the dissection of genetic architecture, screening of novel QTLs with large effects, prediction of QTL effects and their interaction effects with environments, are imperative for effectively improving complex traits of interest by the MAS.

5 Conclusions

Most plant segregating populations for QTL mapping are

generated via crosses of two inbred lines with large difference in the traits of interest. Thus, only the QTLs with genotypes that differ in the two parents can be detected and only limited genetic architecture of the traits can be revealed by using most of the QTL mapping methods. Further, the QTL effects will only apply in the context of this population. Recently, new mapping populations derived from crosses of multiple inbred lines have been proposed, and corresponding analysis methods have been designed [85]. Nevertheless, only a few statistical methods of QTL mapping are available for conducting integrated analysis for multiple segregating populations. Multiple population integrated analysis can explicitly incorporate genetic heterogeneity into the QTL models, and this can help identify QTL alleles carried by specific parental lines. For populations with pedigree information or for the integration of multiple experimental populations, the mixed model approach based on identical by descent could be a choice for QTL analysis [86]. However, this framework needs further extension to be able to predict QTL effects and QTL by environment interaction effects.

The development of a quantitative trait and its variations usually involves a series of metabolism pathways that are regulated by a gene network system; this could generate higher-order epistatic interactions. Epistatic interactions among loci contribute substantially to the genetic variation of complex traits [87]. Recent studies indicated that statistical epistasis was a generic feature of gene regulatory networks, and the proportion of genetic variation that could be attributed to statistical epistasis varied from zero to very high [88]. Apparently, the incorporation of two-order or higher-order interactions of genes in QTL models can objectively describe the context of complex traits, and accurate results from QTL mapping could be anticipated using this approach. Stich et al. [89] investigated the power of various methods to detect three-way epistatic interactions among QTLs involved in metabolic pathway. Although there is more interest in analyzing gene networks or higher-order epistatic interactions, an effective methodology for examining the quantitative genetics of gene networks has not yet been developed.

QTLs often cover large regions of the chromosome region where many genes reside. How to finely resolve QTLs to be able to distinguish candidate genes or variants among many genes in QTL regions is still a challenge in quantitative genetics. More large mapping populations or specifically designed fine-mapping populations, such as the advanced intercross or backcross populations, recurrent selection and backcross populations, are required to develop sufficient recombination within the QTLs. Association mapping has been applied to analyze plant populations, although there are many limitations like the influence of population structure and non-random mating. To improve the accuracy of QTL mapping, some methods that combine interval mapping with association mapping have been proposed

[90,91]. Currently, profiting from the rapid advance in high-throughput genotyping, transcriptional profiling and decreasing costs, it has become possible to measure many intermediate phenotypes in the chain of causation from genetic perturbation to phenotypic variation [12] for large mapping population, and even to sequence each individual in a population. Genetic data of this kind can provide an unprecedented opportunity to dissect the genetic architecture of complex traits by integrated analysis of gene expression differences, eQTLs, networks of gene regulation, and/or protein interactions. We are now developing a new methodology for mapping quantitative trait SNPs (QTS) and quantitative trait transcripts (QTT) based on genome-wide association studies (GWAS) by which real gene-network can be detected using gene-by-gene interactions and gene-by-environment interactions; QTLNetwork that incorporates this method will be released soon for mapping two-way and three-way epistasis and their environmental interactions. This development will further progress the analysis of the genetic architecture of complex traits. A better knowledge on the genetic architecture will help in the development of efficient strategies for direct gene selection designed to improve complex traits by enhancing the use of additive and non-additive genetic effects.

This work was supported by the National Basic Research Program of China (2011CB109306 and 2010CB126006), the National Special Program for Breeding New Transgenic Variety (2009ZX08009-004B), the CNTC (110200701023) and the YNTC (08A05).

- 1 Mackay T F. The genetic architecture of quantitative traits. *Annu Rev Genet*, 2001, 35: 303–309
- 2 Cooper M, van Eeuwijk F A, Hammer G L, et al. Modeling QTL for complex traits: detection and context for plant breeding. *Curr Opin Plant Biol*, 2009, 12: 231–240
- 3 Phillips P C. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet*, 2008, 9: 855–867
- 4 Griffing B. Concept of general and specific combining ability in relation to diallel crossing system. *Aust J Biol Sci*, 1956, 9: 463–493
- 5 Zhu J. Mixed model approaches for estimating genetic variances and covariances. *J Biomath*, 1992, 7: 1–11
- 6 Zhu J. Methods of predicting genotype value and heterosis for offspring of hybrids. *J Biomath*, 1993, 8: 32–44
- 7 Zhu J. *Analysis Methods for Genetic Models* (in Chinese). Beijing: Agriculture Publication House of China, 1997
- 8 Lander E S, Botstein D. Mapping mendelian factors underlying quantitative traits using rflp linkage maps. *Genetics*, 1989, 121: 185–199
- 9 Zeng Z B. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc Natl Acad Sci USA*, 1993, 90: 10972–10976
- 10 Zeng Z B. Precision mapping of quantitative trait loci. *Genetics*, 1994, 136: 1457–68
- 11 Jansen R C. Interval mapping of multiple quantitative trait loci. *Genetics*, 1993, 135: 205–211
- 12 Mackay T F, Stone E A, Ayroles J F. The genetics of quantitative traits: Challenges and prospects. *Nat Rev Genet*, 2009, 10: 565–77
- 13 Jansen R C, van Ooijen J W, Stam P, et al. Genotype-by-environment interaction in genetic-mapping of multiple quantitative trait loci. *Theor Appl Genet*, 1995, 91: 33–37
- 14 Tinker N A, Mather D E. Methods for QTL analysis with progeny replicated in multiple environments. *J Quant Trait Loci*, 1995, 1: 1
- 15 Sari-Gorla M, Calinski T, Kaczmarek Z, et al. Detection of QTL x environment interaction in maize by a least squares interval mapping method. *Heredity*, 1997, 78: 146–157
- 16 Kao C H, Zeng Z B, Teasdale R D. Multiple interval mapping for quantitative trait loci. *Genetics*, 1999, 152: 1203–1216
- 17 Kao C H, Zeng Z B. Modeling epistasis of quantitative trait loci using Cockerham's model. *Genetics*, 2002, 160: 1243–1261
- 18 Kroymann J, Mitchell-Olds T. Epistasis and balanced polymorphism influencing complex trait variation. *Nature*, 2005, 435: 95–98
- 19 Steinmetz L M, Sinha H, Richards D R, et al. Dissecting the architecture of a quantitative trait locus in yeast. *Nature*, 2002, 416: 326–330
- 20 Sinha H, David L, Pascon R C, et al. Sequential elimination of major-effect contributors identifies additional quantitative trait loci conditioning high-temperature growth in yeast. *Genetics*, 2008, 180: 1661–1670
- 21 Zhu J. Mixed model approaches for mapping quantitative trait loci. *Heredity (Beijing)*, 1998, 20(suppl): 137–138
- 22 Zhu J, Weir B S. Mixed model approaches for genetic analysis of quantitative traits. In: Chen L S, Ruan S G, Zhu J, eds. *Advanced Topics in Biomathematics: Proceedings of International Conference on Mathematical Biology*. Singapore: World Scientific Publishing Co., 1998. 321–330
- 23 Zhang K, Tian J, Zhao L, et al. Mapping QTLs with epistatic effects and QTLxenvironment interactions by mixed linear model approaches. *Theor Appl Genet*, 1999, 99: 1255–1264
- 24 Gao Y M, Zhu J. Mapping QTLs with digenic epistasis under multiple environments and predicting heterosis based on QTL effects. *Theor Appl Genet*, 2007, 115: 325–333
- 25 Yang J, Zhu J, Williams R W. Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics*, 2007, 23: 1527–1536
- 26 Hu C C, Ye X Z, Zhang Y, et al. 3D graphical visualization of the genetic architectures underlying complex traits in multiple environments. *J Zhejiang Univ Sci A*, 2007, 8: 563–567
- 27 Yang J, Hu C C, Hu H, et al. QTLNetwork: Mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics*, 2008, 24: 721–723
- 28 Yang J, Zhu J. Methods for predicting superior genotypes under multiple environments based on QTL effects. *Theor Appl Genet*, 2005, 110: 1268–1274
- 29 Huynh B L, Wallwork H, Stangoulis J C, et al. Quantitative trait loci for grain fructan concentration in wheat (*Triticum aestivum* L.). *Theor Appl Genet*, 2008, 117: 701–709
- 30 Mohan A, Kulwal P L, Singh R, et al. Genome-wide QTL analysis for pre-harvest sprouting tolerance in bread wheat. *Euphytica*, 2009, 168: 319–329
- 31 Kumar A, Kumar J, Singh R, et al. QTL analysis for grain colour and pre-harvest sprouting in bread wheat. *Plant Sci*, 2009, 177: 114–122
- 32 Sun X Y, Wu K, Zhao Y, et al. QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica*, 2009, 165: 615–624
- 33 Conti V, Roncallo P F, Beaufort V, et al. Mapping of main and epistatic effect QTLs associated to grain protein and gluten strength using a RIL population of durum wheat. *J Appl Genet*, 2011, 52: 287–298
- 34 Bennett D, Izanloo A, Edwards J, et al. Identification of novel quantitative trait loci for days to ear emergence and flag leaf glaucousness in a bread wheat (*Triticum aestivum* L.) population adapted to southern Australian conditions. *Theor Appl Genet*, 2012. 124: 697–711
- 35 Walker K R, Able J A, Mather D, et al. Black point formation in barley: Environmental influences and quantitative trait loci. *Aus J Agric Res*, 2008, 59: 1021–1029
- 36 Han Y, Teng W, Sun D, et al. Impact of epistasis and QTLx environment interaction on the accumulation of seed mass of soybean (*Glycine max* L. Merr.). *Genet Res (Camb)*, 2008, 90: 481–491
- 37 Zhang D, Cheng H, Geng L Y, et al. Detection of quantitative trait

- loci for phosphorus deficiency tolerance at soybean seedling stage. *Euphytica*, 2009, 167: 313–322
- 38 Gai J Y, Chen L, Zhang Y H, et al. Genome-wide genetic dissection of germplasm resources and implications for breeding by design in soybean. *Breed Sci*, 2012, 61: 495–510
- 39 Wu X, Blake S, Slepner D A, et al. QTL, additive and epistatic effects for SCN resistance in PI 437654. *Theor Appl Genet*, 2009, 118: 1093–1105
- 40 Li X Z, Yuan X J, Jiang S, et al. Detecting QTLs for plant architecture traits in cucumber (*Cucumis sativus* L.). *Breed Sci*, 2008, 58: 453–460
- 41 Cheng Y, Wang Q, Ban Q Y, et al. Unconditional and conditional quantitative trait loci mapping for plant height in nonheading Chinese cabbage. *Hortscience*, 2009, 44: 268–273
- 42 Wu J, Gutierrez O, Jenkins J, et al. Quantitative analysis and QTL mapping for agronomic and fiber traits in an RI population of upland cotton. *Euphytica*, 2009, 165: 231–245
- 43 Radoev M, Becker H C, Ecker W. Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by quantitative trait locus mapping. *Genetics*, 2008, 179: 1547–1558
- 44 Zhao J Y, Dimov Z, Becker H C, et al. Mapping QTL controlling fatty acid composition in a doubled haploid rapeseed population segregating for oil content. *Mol Breed*, 2008, 21: 115–125
- 45 Yang G, Xing Y, Li S, et al. Molecular dissection of developmental behavior of tiller number and plant height and their relationship in rice (*Oryza sativa* L.). *Hereditas*, 2006, 143: 236–245
- 46 Wan X, Weng J, Zhai H, et al. Quantitative trait loci (QTL) analysis for rice grain width and fine mapping of an identified QTL allele gw-5 in a recombination hotspot region on chromosome 5. *Genetics*, 2008, 179: 2239–2252
- 47 Hua J, Jiang L, Guo L B, et al. Conditional and unconditional mapping of quantitative trait loci underlying plant height and tiller number in rice (*Oryza sativa* L.) grown at two nitrogen levels. *Prog Nat Sci*, 2008, 18: 1539–1547
- 48 Wang H L, Zhang W W, Liu L L, et al. Dynamic QTL analysis on rice fat content and fat index using recombinant inbred lines. *Cereal Chem*, 2008, 85: 769–775
- 49 Ng C S, Hamilton A M, Frank A, et al. Genetic basis of sex-specific color pattern variation in *Drosophila melanogaster*. *Genetics*, 2008, 180: 421–429
- 50 Yang J, Wu W, Zhu J. Mapping interspecific genetic architecture in a host-parasite interaction system. *Genetics*, 2008, 178: 1737–1743
- 51 Fernandes E, Pacheco A, Penha-Goncalves C. Mapping of quantitative trait loci using the skew-normal distribution. *J Zhejiang Univ Sci B*, 2007, 8: 792–801
- 52 Hayat Y, Yang J, Xu H M, et al. Influence of outliers on QTL mapping for complex traits. *J Zhejiang Univ Sci B*, 2008, 9: 931–937
- 53 Yan J Q, Zhu J, He C, et al. Molecular dissection of developmental behavior of plant height in rice (*Oryza sativa* L.). *Genetics*, 1998, 150: 1257–1265
- 54 Zhu J. Analysis of conditional genetic effects and variance components in developmental genetics. *Genetics*, 1995, 141: 1633–1639
- 55 Atchley W R, Zhu J. Developmental quantitative genetics, conditional epigenetic variability and growth in mice. *Genetics*, 1997, 147: 765–776
- 56 Verhaegen D, Plomion C, Gion J M, et al. Quantitative trait dissection analysis in Eucalyptus using RAPD markers. 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theor Appl Genet*, 1997, 95: 597–608
- 57 Price A H, Tomos A D. Genetic dissection of root growth in rice (*Oryza sativa* L.). 2. Mapping quantitative trait loci using molecular markers. *Theor Appl Genet*, 1997, 95: 143–152
- 58 Wu W R, Lia W M, Tanga D Z, et al. Time-related mapping of quantitative trait loci underlying tiller number in rice. *Genetics*, 1999, 151: 297–303
- 59 Wu R L, Ma C X, Lin M, et al. A general framework for analyzing the genetic architecture of developmental characteristics. *Genetics*, 2004, 166: 1541–1551
- 60 Wu R L, Lin M. Opinion—Functional mapping—how to map and study the genetic architecture of dynamic complex traits. *Nat Rev Genet*, 2006, 7: 229–237
- 61 Jiang C J, Zeng Z B. Multiple-trait analysis of genetic-mapping for quantitative trait loci. *Genetics*, 1995, 140: 1111–1127
- 62 Cao G Q, Zhu J, He C, et al. Impacts of epistasis and QTL × environment interaction for developmental behavior of plant height in rice (*Oryza sativa* L.). *Theor Appl Genet*, 2001, 103: 153–160
- 63 Sun D H, Li W B, Zhang Z C, et al. Quantitative trait loci analysis for the developmental behavior of Soybean (*Glycine max* L. Merr.). *Theor Appl Genet*, 2006, 112: 665–673
- 64 Han Y, Xie D, Teng W, et al. Dynamic QTL analysis of linolenic acid content in different developmental stages of soybean seed. *Theor Appl Genet*, 2011, 122: 1481–1488
- 65 Guo L B, Xing Y Z, Mei H W, et al. Dissection of component QTL expression in yield formation in rice. *Plant Breed*, 2005, 124: 127–132
- 66 Liu G F, Yang J, Xu H M, et al. Genetic analysis of grain yield conditioned on its component traits in rice (*Oryza sativa* L.). *Aus J Agric Res*, 2008, 59: 189–195
- 67 Zhao J Y, Becker H C, Zhang D Q, et al. Conditional QTL mapping of oil content in rapeseed with respect to protein content and traits related to plant development and grain yield. *Theor Appl Genet*, 2006, 113: 33–38
- 68 Li Y L, Dong Y B, Cui D Q, et al. The genetic relationship between popping expansion volume and two yield components in popcorn using unconditional and conditional QTL analysis. *Euphytica*, 2008, 162: 345–351
- 69 Ye Z H, Wang J M, Liu Q, et al. Genetic relationships among panicle characteristics of rice (*Oryza sativa* L.) using unconditional and conditional QTL analyses. *J Plant Biol*, 2009, 52: 259–267
- 70 Sax K. The association of size difference with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, 1923, 8: 552–560
- 71 Neimann-Sorenson A, Robertson A. The association between blood groups and several production characteristics in three Danish cattle breeds. *Acta Agric Scand*, 1961, 11: 163–196
- 72 Soller M, Plotkinhazan J. Use marker alleles for introgression of linked quantitative alleles. *Theor Appl Genet*, 1977, 51: 133–137
- 73 Stuber C W, Goodman M M, Moll R H. Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. *Crop Sci*, 1982, 22: 737–740
- 74 Hu X S. A general framework for marker-assisted selection. *Theor Popul Biol*, 2007, 71: 524–542
- 75 Lande R, Thompson R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, 1990, 124: 743–756
- 76 Hospital F. Challenges for effective marker-assisted selection in plants. *Genetica*, 2009, 136: 303–310
- 77 Tanksley S D, McCouch S R. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science*, 1997, 277: 1063–1066
- 78 Ashikari M, Sakakibara H, Lin S Y, et al. Cytokinin oxidase regulates rice grain production. *Science*, 2005, 309: 741–745
- 79 Barloy D, Lemoine J, Abelard P, et al. Marker-assisted pyramiding of two cereal cyst nematode resistance genes from *Aegilops variabilis* in wheat. *Mol Breed*, 2007, 20: 31–40
- 80 Park S, Yoon M K, Lee S S, et al. Development of uniform double-crossed varieties using near-isogenic lines produced by marker-assisted selection in radish (*Raphanus sativus* L.). *Hortscience*, 2007, 42(suppl): 856
- 81 Nocente F, Gazza L, Pasquini M. Evaluation of leaf rust resistance genes Lr1, Lr9, Lr24, Lr47 and their introgression into common wheat cultivars by marker-assisted selection. *Euphytica*, 2007, 155: 329–336
- 82 Gimelfarb A, Lande R. Simulation of marker assisted selection for non-additive traits. *Genet Res*, 1994, 64: 127–36
- 83 Liu P, Zhu J, Lou X, et al. A method for marker-assisted selection based on QTLs with epistatic effects. *Genetica*, 2003, 119: 75–86
- 84 Liu P Y, Zhu J, Lu Y. Marker-assisted selection in segregating generations of self-fertilizing crops. *Theor Appl Genet*, 2004, 109: 370–376
- 85 Paulo M J, Boer M, Huang X Q, et al. A mixed model QTL analysis

- for a complex cross population consisting of a half diallel of two-way hybrids in *Arabidopsis thaliana*: Analysis of simulated data. *Euphytica*, 2008, 161: 107–114
- 86 Xu S Z, Gessler D D G. Multipoint genetic mapping of quantitative trait loci using a variable number of sibs per family. *Genet Res*, 1998, 71: 73–83
- 87 Marchini J, Donnelly P, Cardon L R. Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet*, 2005, 37: 413–417
- 88 Gjuvsland A B, Hayes B J, Omholt S W, et al. Statistical epistasis is a generic feature of gene regulatory networks. *Genetics*, 2007, 175: 411–420
- 89 Stich B, Yu J, Melchinger A E, et al. Power to detect higher-order epistatic interactions in a metabolic pathway using a new mapping strategy. *Genetics*, 2007, 176: 563–570
- 90 Lou X Y, Casella G, Toddhunter R J, et al. A general statistical framework for unifying interval and linkage disequilibrium mapping: Toward high-resolution mapping of quantitative traits. *J Am Statist Assoc*, 2005, 100: 158–171
- 91 Lou X Y, Ma J Z, Yang M C, et al. Improvement of mapping accuracy by unifying linkage and association analysis. *Genetics*, 2006, 172: 647–661

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.