

# Bayesian mapping of multiple traits in maize: the importance of pleiotropic effects in studying the inheritance of quantitative traits

Marcio Balestre · Renzo Garcia Von Pinho ·  
Claudio Lopes de Souza Junior ·  
Júlio Sílvio de Sousa Bueno Filho

Received: 7 October 2011 / Accepted: 5 March 2012 / Published online: 22 March 2012  
© Springer-Verlag 2012

**Abstract** Pleiotropy has played an important role in understanding quantitative traits. However, the extensiveness of this effect in the genome and its consequences for plant improvement have not been fully elucidated. The aim of this study was to identify pleiotropic quantitative trait loci (QTLs) in maize using Bayesian multiple interval mapping. Additionally, we sought to obtain a better understanding of the inheritance, extent and distribution of pleiotropic effects of several components in maize production. The design III procedure was used from a population derived from the cross of the inbred lines L-14-04B and L-08-05F. Two hundred and fifty plants were genotyped with 177 microsatellite markers and backcrossed to both parents giving rise to 500 backcrossed progenies, which were evaluated in six environments for grain yield and its components. The results of this study suggest that mapping isolated traits limits our understanding of the genetic architecture of quantitative traits. This architecture

can be better understood by using pleiotropic networks that facilitate the visualization of the complexity of quantitative inheritance, and this characterization will help to develop new selection strategies. It was also possible to confront the idea that it is feasible to identify QTLs for complex traits such as grain yield, as pleiotropy acts prominently on its subtraits and as this “trait” can be broken down and predicted almost completely by the QTLs of its components. Additionally, pleiotropic QTLs do not necessarily signify pleiotropy of allelic interactions, and this indicates that the pervasive pleiotropy does not limit the genetic adaptability of plants.

## Introduction

Understanding the genetic architecture of complex traits remains a great challenge for geneticists. Phenomena such as epistasis, pleiotropy and heterosis make understanding the genetic control of quantitative traits extremely complex, a veritable black box. Recent studies have sought to shed light on these phenomena (Melchinger et al. 2007; Flint and Mackay 2009; Stearns 2010). However, dissecting these elements is only possible using complex molecular techniques and sophisticated statistical models.

One of these highly complex architectural elements is pleiotropy. Understanding pleiotropy is important for plant breeding because pleiotropy helps link genetic correlations to economically and agriculturally important traits that can also be observed through linkage disequilibrium. However, the correlations obtained from linkage disequilibrium are temporary, and the rate at which the disequilibrium dissipates depends primarily on the distance between the genes. In contrast, because pleiotropy is a phenomenon where the same locus controls different traits, its effect is more stable.

---

Communicated by E. Carbonell.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-012-1847-1) contains supplementary material, which is available to authorized users.

---

M. Balestre (✉) · J. S. de Sousa Bueno Filho  
Departamento de Ciências Exatas, Universidade Federal de Lavras, CP 3037, Lavras, MG 37200-000, Brazil  
e-mail: marciobalestre@dex.ufla.com

R. G. Von Pinho  
Departamento de Agricultura, Universidade Federal de Lavras, CP 3037, Lavras, MG 37200-000, Brazil

C. L. de Souza Junior  
Departamento de Genética, Escola de Agricultura Luiz de Queiroz, Universidade de São Paulo (ESALQ/USP), CP 83, Piracicaba, SP 13400-970, Brazil

Therefore, the identification of pleiotropic quantitative trait loci (QTLs) is important. To successfully perform multi-trait selections and breeding strategies, distinguishing between genetic correlations that arise from the effects of linkage and correlations that originate from pleiotropic genes is essential. The phenomena that are involved in the architecture of pleiotropic genes also need further characterization (Flint and Mackay 2009; Mackay et al. 2009).

The distribution of pleiotropic framework across the genome is, until now, unknown. In the “*universal pleiotropy*” hypothesis, all of the genes affect (at high or low intensity) all of the traits that constitute an individual, and a mutation at a locus potentially affects (directly or indirectly) the whole phenotype of the individual (Wright 1968). In the universal pleiotropy hypothesis, there is not a specific gene “for” a trait but rather an extremely complex network where all genes contribute with differing magnitudes to the trait in question. In the context of the universal hypothesis, the identification of QTLs “for” a given trait is misleading (Buchanan et al. 2009; Mackay et al. 2009). Although this hypothesis is still supported by some authors (Carbone et al. 2006; Buchanan et al. 2009; Mackay et al. 2009), this phenomenon has not been commonly observed in recent studies where “*modular pleiotropy*” has been observed (Wagner et al. 2008; Su et al. 2010; Wagner and Zhang 2011). Modular pleiotropy is restricted to a defined group of genes that control particular traits, rather than all characteristics of an individual.

In addition to the current understanding of the function and distribution of pleiotropic genes, there is evidence to suggest that epistasis also shows a correlation between traits (Pleiotropic epistasis: Wolf et al. 2005, 2006). In addition, epistasis between loci can influence the pleiotropic effects of an individual locus (differential epistasis: Cheverud et al. 2004; Pavlicev et al. 2008). These authors raise the possibility that pleiotropy could be greatly influenced by epistasis, which means that there is no certainty whether pleiotropic genes can induce stable covariances under any circumstance. The magnitude of this effect can be variable when different epistatic interactions occur between loci that affect multiple traits (Mackay et al. 2009).

Taking the above information together, we suggest that pleiotropy is an extremely complex phenomenon and has a fundamental role in defining the genetic outcomes in all species. In the field of genetic improvement, few studies have paid sufficient attention to this phenomenon, and when they do, they do not account for the diverse implications of pleiotropy on the improvement process. In the context of selecting for multiple traits, pleiotropy is often ascribed a limited role in relation to genetic correlation, where one seeks merely to distinguish it from the effects of linkage between loci, which is a temporary event.

Therefore, in addition to distinguishing pleiotropy from linkage, it is necessary to also verify the extent of this effect across the genome for various traits to better understand the inheritance of quantitative traits and to subsequently develop new selection strategies. The identification and distribution of pleiotropic QTLs across the genome can be achieved by Bayesian shrinkage multiple interval analysis (Wang et al. 2005; Xu et al. 2009). This approach is not restricted by the number of parameters, and it assumes that QTL must be spaced some distance apart, depending on the sample size and marker density. It is also assumed that each QTL possess individual variance forcing marginal QTLs to values close to zero and that the inclusion of these marginal QTLs in the model has little effect on other intervals. Conversely, QTLs with outstanding effects tend to show a variance that is negligibly penalized by the model error. These properties describe Bayesian shrinkage analysis (Wang et al. 2005). Other Bayesian multi-trait and multiple interval approaches such as composite model space, using the so-called Stochastic Source Variable Selection SSVS (George and McCulloch 1993), are also suggested to depict the genetic architecture of pleiotropic QTLs (Banerjee et al. 2008). In Bayesian shrinkage analysis, the magnitude of the QTL effects are conditional on its variance or its heritability while SSVS uses binary variables that determine if a putative QTL effect may or not be shrunk to zero.

Although Bayesian shrinkage analysis appears to violate the principle of model parsimony, it is the most representative of the possible architectures of an oligogenic/polygenic trait where key genes have a moderate to large influence on a trait and when combined with various other genes with small effects may determine the genetic variation of a quantitative trait (Mackay 2001; Flint and Mackay 2009). Thus, it is important to maintain QTLs with negligible effects in the model, for they can perform an important role in the genetic variance of a trait (Wang et al. 2005).

In this study, we mapped multiple traits using a Bayesian approach to identify pleiotropic QTLs related to grain yield and its components in maize. In addition, we sought to obtain a better understanding of the inheritance, extent and distribution of the pleiotropic effects related to these traits.

## Materials and methods

### Genetic material

To obtain the F<sub>2</sub> population, two inbred lines genetically divergent for several traits were crossed (L-14-04 B and L-08-05F). These lines belong to different heterotic groups:

L-14-04 B was derived from the BR-106 population (yellow dent kernels) and L-08-05 F from the IG-1 population (flint and orange kernels). The BR-106 population was developed at EMBRAPA Maize and Sorghum Research Center and the IG-1 population was developed from the ESALQ/USP Department of Genetics by crossing Brazilian and Thai populations.

Two hundred and fifty progenies from the  $F_2$  generation were self-fertilized to yield the  $F_{2:3}$  progenies. These progenies were backcrossed to the L-14-04 B parental line to create the  $RC_1$  progenies and also backcrossed to L-08-05 F to create the  $RC_2$  progenies. These crosses were performed in isolated, detasseled blocks and yielded 500 backcross progenies following the Design III procedure.

### Experimental procedures

The 500 backcrossed progenies were evaluated in six environments with two replications per environments. The experimental design was  $10 \times 10$  lattices, and in each lattice 50 pairs of progenies backcrossed to the two parental inbreds were allocated; then there were five lattices per environment. In all environments, the experiments and their replications were randomized within the experimental area. Each row in the plots was 4.0 m long, spaced 0.8 m between rows, and they were overplanted and thinned to 20 plants  $\text{plot}^{-1}$  (62,500 plants  $\text{ha}^{-1}$ ). Data for grain yield (kg  $\text{plot}^{-1}$ ) adjusted to 15 % grain moisture and to average stand (GY), ear diameter (ED) (cm  $\text{ear}^{-1}$ ), ear length (cm  $\text{ear}^{-1}$ ) (EL), kernel row number (number  $\text{ear}^{-1}$ ) (NKR), and row number per ear (number  $\text{ear}^{-1}$ ) (NR) were recorded in all environments. Except GY, the other traits were recorded in five plants  $\text{plot}^{-1}$  and the plot means were used for the analyses. A complete description of experimental procedures can be obtained in Aguiar (2003)—PhD Thesis—data not published.

### Genetic map

The genetic map and mapping procedure used to obtain it were previously described by Sibov et al. (2003). The molecular analysis was performed at CEBEMEG/Unicamp. Briefly, the  $F_2$  plants that generated the  $F_{2:3}$  progenies were genotyped using microsatellite markers. The genetic map was developed using the program MAP-MAKER/EXP version 3.0b (Lincoln et al. 1992) with a LOD of 3.0 and a maximum distance between adjacent markers of 50 cM to form the linkage groups. Sixty new microsatellite markers were added to the Sibov et al. (2003) map, resulting in a total of 177 markers distributed across the 10 maize chromosomes. The genetic map covered 2,052 cM of the maize genome, with an average distance of 11.6 cM between markers.

### QTL analysis

The Bayesian analysis of multiple intervals and traits was adapted from Xu et al. (2009) for Design III. In this approach, it is assumed that each interval contains a potential QTL that produces random effects and individual variance. The data obtained in this method are the markers ( $m$ ) and the phenotype ( $y$ ). The unobserved variables are the QTL genotypes ( $x$  and  $w$ ), their effects ( $a$  and  $d$ ), the additive, dominant and residual (co)variance matrices ( $A_k$ ,  $D_k$  and  $\Sigma$ ) and QTL position ( $\lambda$ ).

Therefore, utilizing the phenotypic averages obtained from the analysis groups in six environments and the 177 microsatellite markers, we adopted the following linear model:

$$y_{ij} = b_j + \sum_{k=1}^p x_{ijk} a_k + \sum_{k=1}^p w_{ijk} d_k + e_i \quad (1)$$

where  $y_{ij}$  is the vector of the phenotypic observations in the  $i$ th progeny in the  $j$ th backcross  $[y_{1ij} \dots y_{qij}]^T$ ,  $q$  is the number of traits evaluated;  $b_j = [b_{1j} \dots b_{qj}]^T$  is the vector of the population averages for the trait  $q$  in the  $j$ th backcross; and  $a_k = [a_{1k} \dots a_{qk}]^T$  and  $d_k = [d_{1k} \dots d_{qk}]^T$  are the additive and dominant effects, respectively, for the  $k$ th locus. The residual is a vector  $q \times 1$  assumed to be multivariate normal  $MVN(0, \Sigma)$ , where  $\Sigma$  is a  $q \times q$  matrix. The variables  $x_{ijk}$  and  $w_{ijk}$  were obtained from the  $F_\infty$  metric and were therefore orthogonal and similar to the  $F_2$  metric when the epistatic effects are not adjusted for in the model (Yang 2004; Zeng et al. 2005). Thus, we obtain the following:

$$x_{ijk} = \begin{cases} 1 & \text{if } QQ \\ 0 & \text{if } Qq \\ -1 & \text{if } qq \end{cases} \text{ and } w_{ijk} = \begin{cases} 0 & \text{if } QQ \\ 1 & \text{if } Qq \\ 0 & \text{if } qq \end{cases} \quad (2)$$

The variables  $x_{ijk}$  and  $w_{ijk}$  are not observed but can be inferred from the genotype information of the markers flanking the QTL. This model (1) is clearly a pleiotropic model and describes only one QTL simultaneously affecting two or more traits with one QTL per interval and defined variances–covariances matrices.

### Likelihood function

To simplify, the following vector equivalence will be used:  $y = y_{ij}$ ,  $b = b_j$ ,  $x = x_{ij}$  and  $w = w_{ij}$ . Under the assumption of normality of the residuals, one can assume that the conditional probability of  $y$  is the following:

$$p(y|b, a, d, x, w, \Sigma) = N(\mu, \Sigma) \quad (3)$$

$$\mu = \sum_{k=1}^p x_k a_k + \sum_{k=1}^p w_{ik} d_k$$

Thus, the likelihood function is given by:

$$p(y|\lambda, b, a, d, x, w, \Sigma, ) \propto |\Sigma \otimes I|^{-1/2} x \exp\left\{-\frac{1}{2}(y - \mu)(\Sigma \otimes I)^{-1}(y - \mu)\right\} \tag{4}$$

The parameters set is:  $\theta = \{\lambda, b, a, d, \Sigma, A_k, D_k\}$ . The number of QTLs is usually considered a parameter of interest in classical mapping, but in Bayesian shrinkage analysis, the number of QTLs is a constant that is dependent on the number of markers or intervals. If an interval does not contain a QTL, its value shrinks to zero, which is equivalent to a QTL being excluded from the model.

*A priori distribution*

Each parameter of the model possesses an a priori distribution. Therefore,  $p(b) \propto 1$ ,  $p(a_k) \propto N(0, A_k)$ ,  $p(d_k) \propto N(0, D_k)$ . The additive ( $A_k$ ) and dominant ( $D_k$ ) variance and covariance matrices have the dimensions  $q \times q$ . These matrices have values that vary for different loci and demonstrate a specific effect. The prior distribution of these matrices follows an inverted Wishart distribution and can be represented by  $p(A_k) = \text{inv - Wishart}(\tau, \Gamma)$  and  $p(D_k) = \text{inv - Wishart}(\tau, \Gamma)$  with  $\tau > q$  and  $\Gamma > 0$ . In this work, minimum values were assumed for the hyperparameters in order to prevent a bias in the a posteriori inference (Xu et al. 2009). For the residual matrix ( $\Sigma$ ), the same priori  $p(\Sigma) = \text{inv - Wishart}(\tau, \Gamma)$  is assumed, and in this case, due to the number of degrees of freedom contained in the data, the hyperparameters will have little influence on the estimated values of  $\Sigma$ . The prior distribution relative to the QTL position may be uniform. In other words, given that we assume that  $M_k^L$  and  $M_k^R$  are two markers that flank the genotype of the QTL  $Q_k$  and that  $L_k$  and  $U_k$  are the distances between  $M_k^L \leftrightarrow Q_k$  and  $Q_k \leftrightarrow M_k^R$ , respectively, a uniform prior for each interval is given by a group of ordered numbers of the same probability and that vary from  $[L_k \text{ to } U_k]$ .

Thus, the group a priori distribution is:

$$p(\theta) = p(b)p(\Sigma) \prod_{k=1}^p p(\lambda_k)p(a_k)p(d_k)p(A_k)p(D_k) \tag{5}$$

*A posteriori distribution*

The posterior distribution of the join parameters is given by:

$$p(\theta, x, w|y) \propto p(y|b, a, d, x, w, \Sigma)p(x, w|\lambda)p(\theta) \tag{6}$$

where  $p(x, w|\lambda)$  is the probability of the genotype of the QTL, given its position.

Although it is difficult to obtain an analytic solution with this distribution, the Monte Carlo Markov Chain (MCMC) method can be used to obtain samples of the joint posterior distribution. To facilitate the sampling process, each parameter is sampled individually, and the sampling is dependent on all of the other parameters. If the conditional probability distribution of the parameters has an explicit form, the Gibbs sampler can be applied, removing samples directly from known distributions. The conditional distributions of the parameters are given below.

The conditional posterior distribution of the population averages from the two backcrosses is multivariate normal with the mean and variance given by:

$$b_j = \frac{1}{n_j} \sum_{i=1}^{n_j} (y_{ij} - \mu) \tag{7}$$

and

$$\text{var}(b_j) = \left(\frac{1}{n_j}\right)\Sigma \tag{8}$$

Thus, a distinct mean is assumed for each backcross.

The conditional posterior distribution of the additive effects ( $a_k$ ) is multivariate normal with the following mean and variance:

$$a_k = \left(\sum_{j=1}^2 \sum_{i=1}^n x_{ijk}^2 \Sigma^{-1} + A_k^{-1}\right)^{-1} \sum_{j=1}^2 \sum_{i=1}^n x_{ijk}^2 \Sigma^{-1} \left(y_{ij} - b_j - \sum_{k' \neq k}^p x_{ijk'} a_{k'} - \sum_{k=1}^p w_{ijk} d_k\right) \tag{9}$$

and

$$\text{var}(a_k) = \left(\sum_{j=1}^2 \sum_{i=1}^n x_{ijk}^2 \Sigma^{-1} + A_k^{-1}\right)^{-1} \tag{10}$$

Similarly, the conditional posterior distribution for dominance effects is multivariate normal with the following mean and variance:

$$d_k = \left(\sum_{j=1}^2 \sum_{i=1}^n w_{ijk}^2 \Sigma^{-1} + D_k^{-1}\right)^{-1} \sum_{j=1}^2 \sum_{i=1}^n w_{ijk}^2 \Sigma^{-1} \left(y_{ij} - b_j - \sum_{k=1}^p x_{ijk} a_k - \sum_{k' \neq k}^p w_{ijk'} d_{k'}\right) \tag{11}$$

and

$$\text{var}(d_k) = \left(\sum_{j=1}^2 \sum_{i=1}^n w_{ijk}^2 \Sigma^{-1} + D_k^{-1}\right)^{-1} \tag{12}$$

The a priori (co)variances matrices are conjugated, thus yielding the following posterior conditionals of  $A_k$  and  $D_k$ :

$$p(A_k|\dots) = \text{Inv - Wishart}(\tau + 1, \Gamma + a_k a_k^T) \tag{13}$$

$$p(D_k|\dots) = \text{Inv - Wishart}(\tau + 1, \Gamma + d_k d_k^T) \tag{14}$$

Similarly, the a posteriori conditional of the residual (co)variance matrix is:

$$p(\Sigma|\dots) = \text{Inv - Wishart}(\tau + n, \Gamma + SS) \tag{15}$$

where

$$SS = \sum_{i=1}^n \left( y_{ij} - b_j - \sum_{k=1}^p x_{ijk} a_k - \sum_{k=1}^p w_{ijk'} d_k \right) \left( y_{ij} - b_j - \sum_{k=1}^p x_{ijk} a_k - \sum_{k=1}^p w_{ijk'} d_k \right)^T \tag{16}$$

### Design III

In Design III, each backcross creates a distinct population. Although the backcrosses are analyzed together, the conditional probability of the QTL genotype based on the markers is obtained for each population separately (Table 1). Assuming the equation given in (2) and considering  $g = (1, 0, -1)$  for  $x_{ijk}$  and  $h = (0, 1, 0)$  for  $w_{ijk}$ , the posterior conditional probability for the QTL genotype is calculated utilizing the Bayes theorem as follows:

$$p(x_{ijk} = g_z|\dots) = \frac{p(x_{ijk} = g_z) H_{kL}(g_z, m_l) H_{kR}(g_z, m_r) p(y_{ij}|b_j, a_k, d_k, x_{ijk}, w_{ijk}, \Sigma)}{\sum_{z=1}^3 p(x_{ijk} = g_z) H_{kL}(g_z, m_l) H_{kR}(g_z, m_r) p(y_{ij}|b_j, a_k, d_k, x_{ijk}, w_{ijk}, \Sigma)} \tag{17}$$

where  $p(x_{ijk} = g_z)$  is the prior probability of the QTL genotype given the expected segregation of the backcross. As an example,  $p(x_{ijk} = g_1) = p(x_{ijk} = g_2) = \frac{1}{2}$  for RC1 and  $p(x_{ijk} = g_2) = p(x_{ijk} = g_3) = \frac{1}{2}$  for RC2. The variables  $H_{kL}(g, m_l)$  and  $H_{kR}(g, m_r)$  are the matrices for the transition between the markers  $M_k^L$  and  $M_k^R$  and the QTL  $Q_k$ . These  $3 \times 2$  matrices are constructed based on the conditional probabilities expressed in Table 1, and the conditional probability of the interval is given by the Kronecker

**Table 1** Conditional probability of segregation of a QTL with the adjacent marker (M)

	RC <sub>1</sub>			RC <sub>2</sub>		
	Q <sub>1</sub> Q <sub>1</sub>	Q <sub>1</sub> Q <sub>2</sub>	Q <sub>2</sub> Q <sub>2</sub>	Q <sub>1</sub> Q <sub>1</sub>	Q <sub>1</sub> Q <sub>2</sub>	Q <sub>2</sub> Q <sub>2</sub>
M <sub>1</sub> M <sub>1</sub>	(1 - r)	r	0	0	(1 - r)	r
M <sub>1</sub> M <sub>2</sub>	1/2	1/2	0	0	1/2	1/2
M <sub>2</sub> M <sub>2</sub>	r	(1 - r)	0	0	r	(1 - r)

The conditional probability of the interval is given by the Kronecker product of the backcross matrices (RC<sub>1</sub> ⊗ RC<sub>1</sub> e RC<sub>2</sub> ⊗ RC<sub>2</sub>)

product of  $H_{kL}(g, m_l)$  and  $H_{kR}(g, m_r)$ , resulting in a modified  $9 \times 2$  matrix.

The full conditional distribution for parameter  $\lambda$  does not have a closed form and Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970) should be used for sampling. This algorithm makes use of an auxiliary sampling function and candidate values are accepted with probability  $\alpha$ . A uniform distribution was used to generate samples to each interval  $[\max(\lambda_{j-1}, \lambda_j - d), \min(\lambda_{j+1}, \lambda_j + d)]$ , where  $d$  is a tuning parameter for the interval  $j$ , normally fixed between 1 and 2 cM. This function is denoted by  $u(\lambda^*, \lambda)$ , and the new position will be accepted in the  $k$ th iteration with of probability  $\min(1, \alpha)$ , where  $\alpha$  is given by:

$$\alpha = \frac{p(\lambda_j^*|y, b, x, w, a, d, \Sigma) u(\lambda_j^*, \lambda_j)}{p(\lambda_j^0|y, b, x, w, a, d, \Sigma) u(\lambda_j, \lambda_j^*)} \tag{18}$$

Finally, the lost genotypes of the markers are sampled from their posterior conditional distribution, which can be calculated using the Bayes theorem:

$$p(m_{ijk} = g_z|\dots) = \frac{p(m_{ijk} = g_z) M_{k-1}(g_z, m_{k-1}) M_{k+1}(g_z, m_{k+1})}{\sum_{z=1}^3 p(m_{ijk} = g_z) M_{k-1}(g_z, x_{k-1}) M_{k+1}(g_z, x_{k+1})} \tag{19}$$

where  $p(m_{ijk} = g_1) = p(m_{ijk} = g_3) = \frac{1}{4}$  and  $p(m_{ijk} = g_2) = \frac{1}{2}$  and  $M_{k-1}$  and  $M_{k+1}$  are the matrices for the genotypes of the markers that flank the lost marker and their distances.

These expressions clearly show that the markers follow an F<sub>2</sub> segregation pattern, whereas QTLs segregate following the expected segregation of the backcross populations. These observations suggest that the prediction of the lost marker based on the QTL genotypes becomes tedious and requires matrices that describe the probabilities of the two backcrosses. Thus, the high density of the utilized markers limits the effects that are predicted by lost markers based on flanking markers.

### Post-MCMC analysis

The QTL position was described following (Yang and Xu 2007). The conditional posterior for the position  $f(\lambda)$  was weighted by the quadratic effect of the QTL  $t(\lambda)$ . Thus, the profile of these quadratic effects weighted by their frequencies in the posterior distribution is given by  $g(\lambda) =$

$W(\lambda)f(\lambda)$ , where  $W(\lambda) = a^T V_a^{-1} a + d^T V_d^{-1} d$  follows a  $\chi^2$  distribution with two degrees of freedom and  $V_a^{-1}$  and  $V_d^{-1}$  are the inverses of the variances of the QTL effects given in (10, 12).

### Pleiotropy versus linkage

The pleiotropy versus linkage test can be performed by comparing the original pleiotropy model ( $\text{model}_{\text{plei.}}$ ) with the linkage model ( $\text{model}_{\text{link.}}$ ) using the Bayes factor (Varona et al. 2004; Liu et al. 2007).

$$\text{BF} = \frac{p(y|\text{mod}_{\text{pleiotropic}})}{p(y|\text{mod}_{\text{linkage}})} \quad (20)$$

where  $p(y|\text{mod} \dots)$  is the probability of the conditional phenotypic observations relative to all of the other parameters. The model  $p(y|\text{mod}_{\text{linkage}}$  can be easily adapted to the pleiotropy model (1) assuming two or more QTL positions per interval and a null covariance across traits (Varona et al. 2004; Liu et al. 2007). During the MCMC process,  $p(y|\text{mod} \dots)$  was obtained, following which the BF was calculated, thus resulting in a chain of BF values.

The harmonic mean of this chain corresponds to the probability of the observations, given the parameters (Liu et al. 2007; Raftery et al. 2007). The criteria adopted to infer the most favorable model was given by the log of the Bayes factor as follows:  $\log(\text{BF}) > 10$ , decisive in favor of the pleiotropy model;  $10 > \log(\text{BF}) > 5$ , strong evidence of pleiotropy; and  $5 > \log(\text{BF}) > 0$ , moderate evidence of pleiotropy. Negative values for  $\log(\text{BF})$  provide evidence that favors the linkage model with the same scale described for pleiotropy. A detailed derivation of Bayes factor for linkage and pleiotropic models can be obtained in Varona et al. (2004). The pleiotropy test was performed in pairs of two traits, which is equivalent to the decomposition of the group analysis with five traits.

A study was carried out in order to verify the power of our method to detect linkage effects as opposed to pleiotropic effects. Two correlated traits were simulated presenting one pleiotropic QTL plus two linked QTLs sited in one linkage group with 15 molecular markers. Average distance between adjacent markers was 10 cM. Simulated heritability was 0.5 and population size ranged from 200 to 500 individuals. Results from using our methods were compared with the analysis suggested by Jiang and Zeng (1995) using Qgene program (Joehanes and Nelson 2008). All of the analyses were performed using SAS/IML code.

### Distribution of pleiotropic effects

The direction and magnitude of the pleiotropic effects was represented using a biplot. This biplot was constructed using a singular value decomposition of a double entry

table with the effects of the QTLs (significant for at least one trait) over different evaluated traits. This table was centered on the traits, and the effects were corrected to the same scale, generating a pattern of QTL interaction magnitudes across traits. The biplot was generated using the SAS computational package.

The extent of the pleiotropic effects and the identification of possible modules were analyzed by bipartite hierarchy networks (Buchanan et al. 2009; Wang et al. 2010). The analyses of the networks were performed using the program Cytoscape v. 2.8.1 (Shannon et al. 2003).

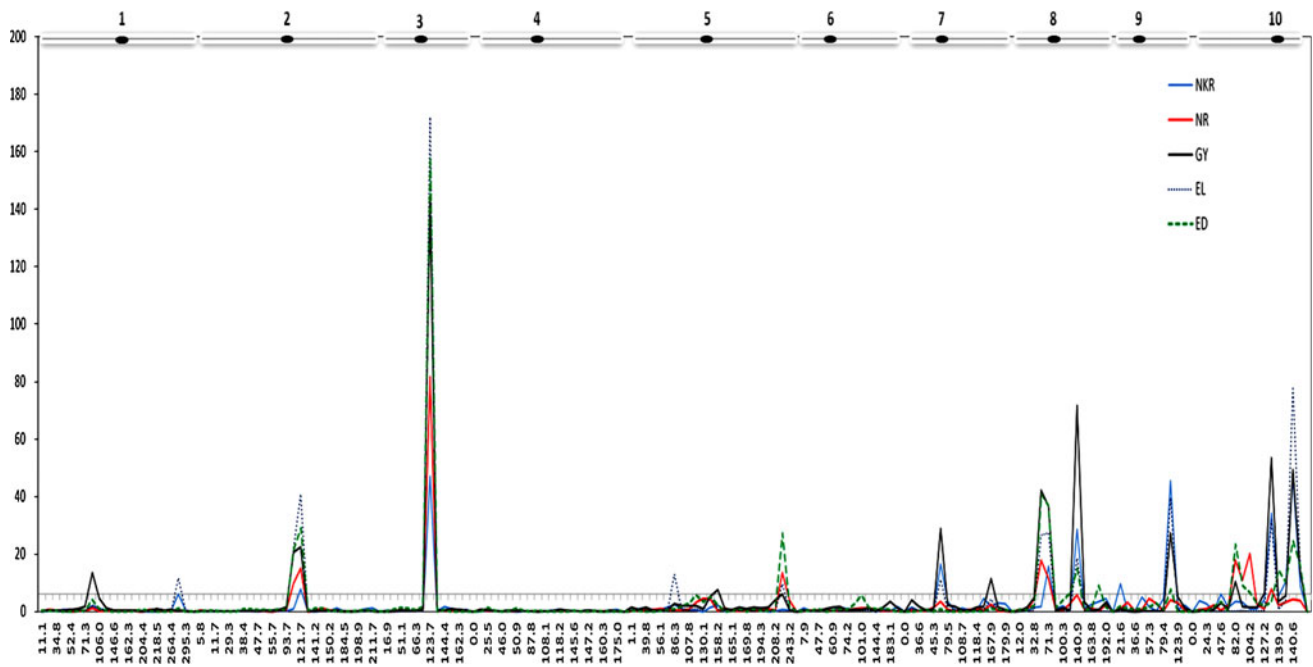
## Results

### Genetic control of grain yield and yield components

In this study, 27 QTLs that affect the grain yield (GY), the ear diameter (ED), the ear length (EL), the number of kernels per row (NKR) and the number of rows per ear (NR) were identified. Figure 1 shows the distribution of these 27 QTLs across the genome. QTLs with a large effect were not identified on chromosomes 4 and 6, but a high density of QTLs was found on chromosomes 8 and 10. The sharpest peaks suggest the presence of only one QTL in the region, whereas the broader peaks suggest linked QTLs, but with  $g(\lambda)$  that do not overlap (Fig. 1). In the multiple interval analysis, the 10 linkage groups were joined to simulate a chromosome with extensive intervals between the groups. However, the scale in Fig. 1 is given in cM and specified for linkage groups.

Upon examining NKR, 13 QTLs were identified, and the sum of their effects suggests that the trait is completely dominant ( $LD = 0.88$ ). However, QTLs with additive effects (6), partial dominance (3), complete dominance (1) and overdominance (4) were found by analyzing their individual effects. A large portion of the QTLs that affect the NKR (85 %) also had a significant effect on EL, and 14 QTLs were identified that together showed partial dominance. Allelic interaction, or the degree of dominance, exerted by these QTLs on NKR and EL were very distinct, and seven additive, six overdominant and only one partially dominant QTL were found to specifically affect EL. The posterior distribution of the effects of the QTLs on NKR and EL is shown in Fig. 2.

The QTLs acting on the number of rows (10) were nearly identical to those acting on ED (17), with a coincidence rate of 90 %. Contrary to what was observed for EL and NKR, one can infer that the genetic controls of ED and NR are virtually additive. A dominant or overdominant QTL acting on NR was not found, as 60 % of these QTLs showed additive inheritance and the others exhibited only partial dominance. Only one overdominant QTL was found



**Fig. 1** Profile of the location of QTLs in the maize genome (chromosomes 1–10) responsible for grain yield (GY) and its components, defined by the number of kernels per row (NKR), the number of rows (NR), the ear length (EL) and the ear diameter (ED).

The horizontal dotted line represents the critical value for the univariate Wald test ( $\chi^2_{2,1-0.05} = 5.99$ )

for ED, and the rest demonstrated additive effects (9) or were partially dominant (7). The additive inheritance of these QTLs is shown in Fig. 2, where few dominance peaks are found in the posterior distribution.

The genetic control of GY demonstrated partial dominance when the sum of the 17 QTLs that act on this trait were considered together. Generically, the inheritance of this trait was equivalent to a combination of the effects of the QTLs present in its components (Fig. 2). In this figure, it is possible to note that the trend of the additive effects for grain yield were similar to those observed for ED and NR, and their dominant effects were similar to those observed for EL and NKR. Among the 17 QTLs that influence grain yield, there was an almost equal number of additive effects (6) and partial dominant (6) and overdominant (5) loci. These results indicate that the trait GY is a result of mixed allelic interactions of its components.

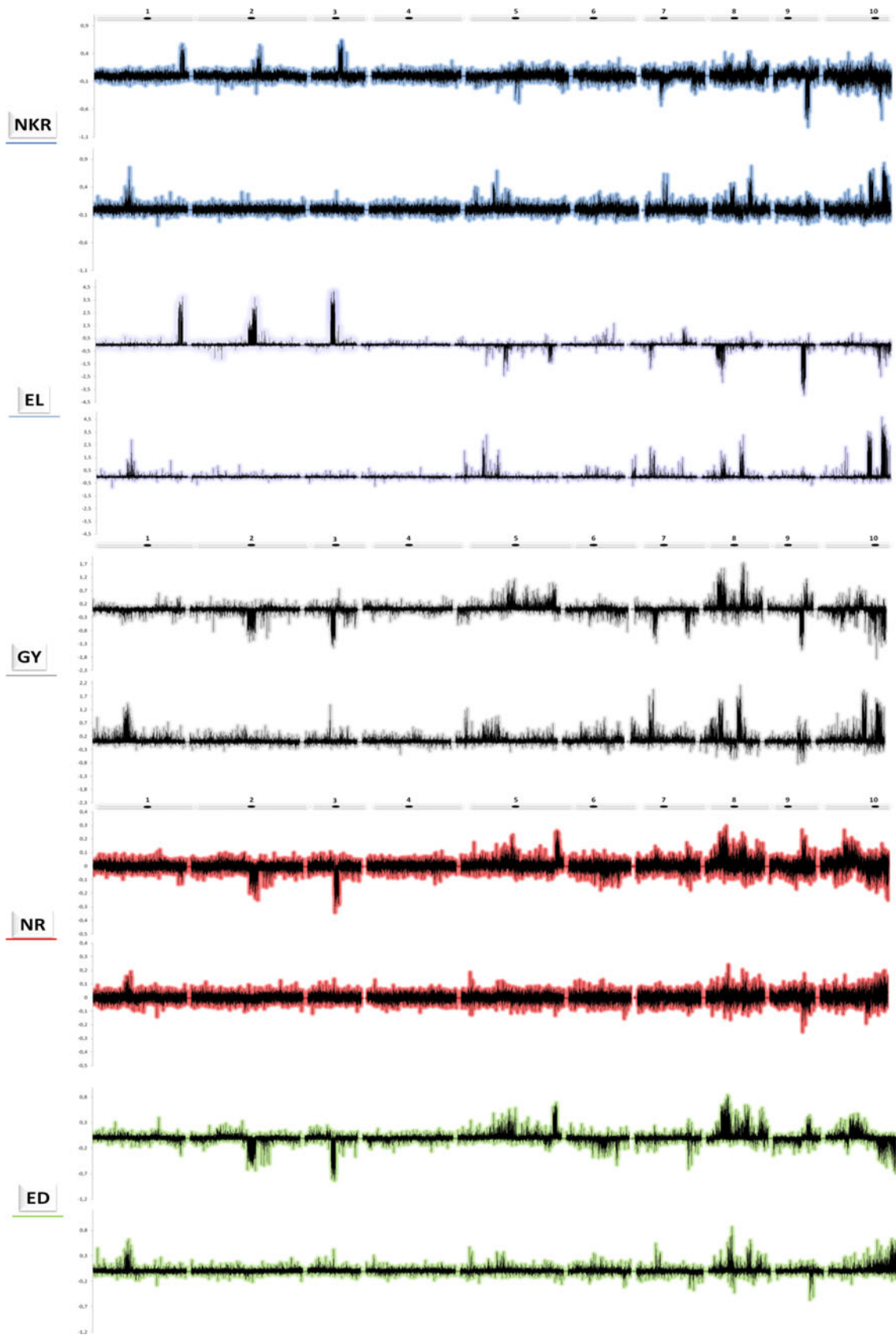
#### Pleiotropy and its distribution

The majority of the QTLs identified in this study (63 %) had pleiotropic effects. These pleiotropic QTLs were also the most important for the genetic control of the five traits we evaluated (Fig. 1). The regions of the genome that have the highest  $W(\lambda)$  also show more than one peak above the critical value. The most important QTLs were the following: QTL55, which is located 110.8 cM from the beginning of chromosome 3 and acts on all five traits, QTL 175,

which is located 140.5 cM from the beginning of chromosome 10 and acts on four traits and QTL 141, which is located 58.8 cM from the beginning of chromosome 8 and acts on all five traits.

The comparative analyses between the linkage and pleiotropy models provide evidence for the pleiotropic model according to Bayes factor criterion (Table 2). The reduction of the residual variance in the pleiotropy model was as much as four times compared to the linkage model when analyzing the EL and ED traits together. The results in Table 2 show that it is unlikely that two QTLs that are very close together act indirectly on the traits, i.e., where the covariance of traits would just be a function of the recombination frequency.

Although our results provide evidence in support of pleiotropic model, the simulation experiments revealed that close linkage is difficult to sort from pleiotropic effects in multiple interval mapping, specially on populations in linkage disequilibrium when the QTLs are in the same interval (Fig. 1 and 2—Supplementary Materials). On the other hand, our analysis was more powerful in detecting the pattern of linked QTLs than conventional multiple interval mapping (MIM) methods. Two linked QTLs spaced by 11 cM were found on neighborhood of true linked QTLs spaced by 8 cM. In conventional MIM, these linked QTLs were described as pleiotropic effects (just one QTL estimated at 68 cM ahead from starting point) and the true pleiotropic QTL was wrongly estimated as three linked





◀ **Fig. 2** Posterior distribution for additive effects (*upper lines*) and dominant effects (*lower lines*) of the QTLs in the maize genome (chromosomes 1–10) responsible for the trait grain yield (GY) and its components defined by number of kernels per row (NKR), number of rows (NR), ear length (EL) and ear diameter (ED)

**Table 2** Evidence supporting the pleiotropy model obtained from the logarithm of the Bayes factor, log (BF), against the linkage model

	ED	EL	GY	NR	NKR
ED		615.61	550.93	453.54	562.22
EL			424.21	453.54	440.32
GY				362.08	552.91
NR					557.01

ED ear diameter, EL ear length, GY grain yield, NR number of rows, NKR number of kernels per row

QTLs at 14, 24 and 26 cM (Fig. 3—Supplementary Materials). Thus, our pleiotropic model was more powerful than Jiang and Zeng approach in both detecting pleiotropic and true linkage patterns, since the simulated pleiotropic QTL was detected with larger resolution in Bayesian analysis than in likelihood approach. However, if there are tightly linked QTL, both methods will be challenged.

The pleiotropic model was efficient in detecting closely linked QTLs distributed in different intervals. For instance, at distance of 139.6 cM in the chromosome 10, a significant QTL was found for NR, but not for ED. On the other hand, at 140.2 cM, on the same linkage group, an inverted situation was obtained (Fig. 1). We suggest that our framework is useful to sort out patterns of pleiotropy and linkage in multiple QTLs mapping; although may fail to distinguish close linkage from pleiotropy in QTLs located in the same interval (Supplementary Materials).

In Fig. 2, there is a superimposition of the posterior distributions for EL and NKR, suggesting that there is no distinction between these traits but rather between measurements of the same trait. Similarly, there was no distinction between ED and NR in the position and the direction of the effects of the QTLs (Fig. 2). It is also evident that GY is a mixture of these four components, which may suggest, a priori, that QTLs for grain yield are very restrictive since GY is a function of the QTLs that act on its subtraits.

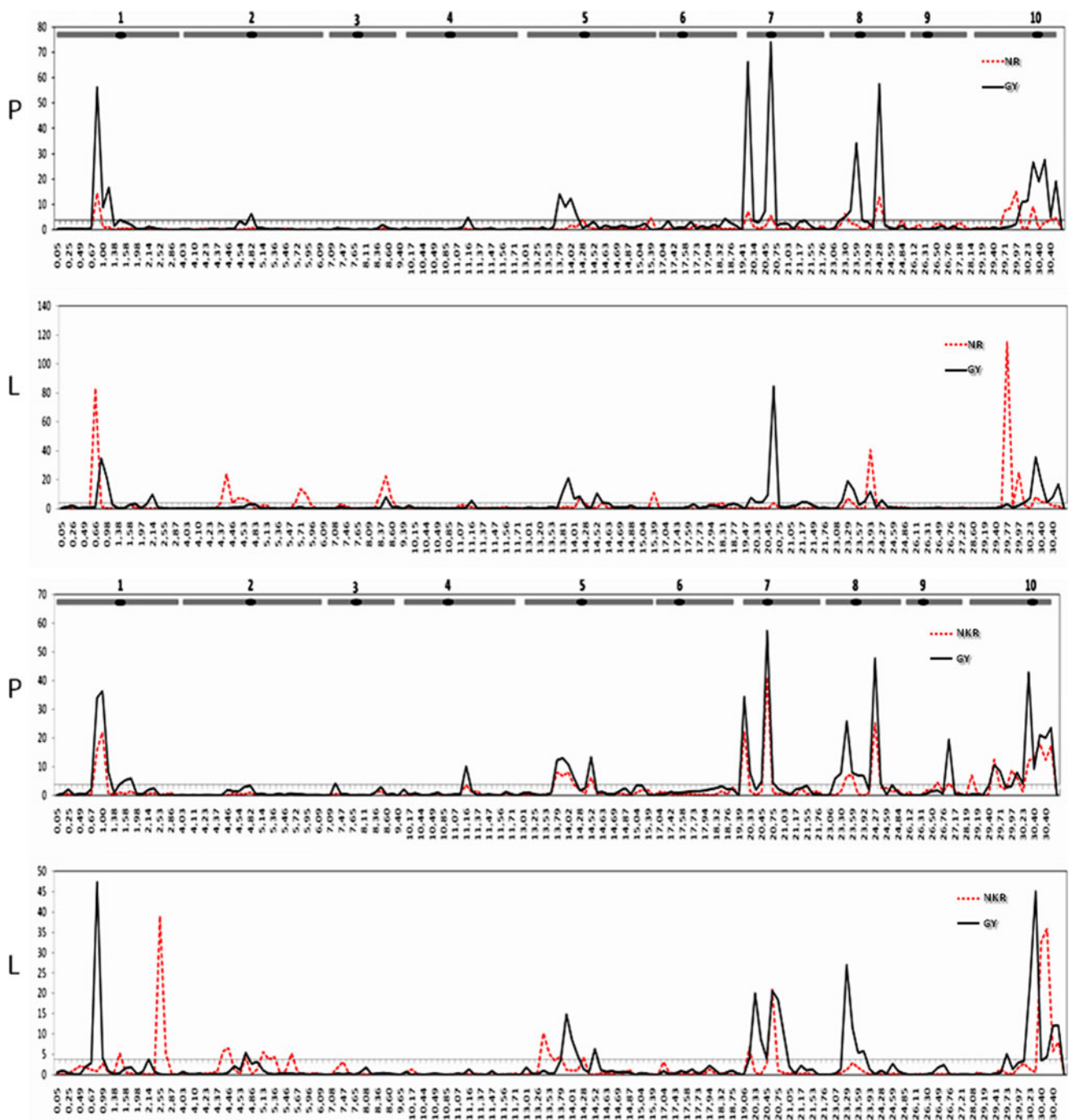
Evidence for the pleiotropy model can also be seen in Fig. 3. In this figure, the pleiotropy model adjusts the effects of very close QTLs and concentrates these effects in just one region of the genome and this feature of the model results in lower residual variance (Table 2). For example, the two-trait linkage analysis for GY and NR demonstrates that the first linkage group contains two peaks in neighboring intervals that were later corrected to two peaks within the same interval in the pleiotropy model (Fig. 3).

The same situation was observed for the traits GY and NKR, where nearly all of the neighboring peaks in the linkage model were placed in just one region of the genome. This pattern repeated itself in all of the comparative analyses between the linkage and pleiotropy models.

The results of this study suggest that grain yield and yield components do not possess modules of pleiotropic genes, or in other words, it was not possible to separate the QTLs according to their importance and amplitude of action (Fig. 4b). Although both the 168 and 169 QTLs act on ED and NR and are located in group 10, they were outside the module due to their low relative importance, and specific traits could not be separated by their particular QTL groups. In this module, we have identified that the most important QTLs were closely located to the five traits, and those presenting minor importance were placed at the edge of the module. This grouping takes both the amplitude of the action of the QTL and its magnitude into consideration. Figure 4a shows a hierarchical network where the QTLs were grouped based on the magnitude of their influence on the traits (but not their amplitude) and where the traits were grouped based on the influence of these QTLs. In this case, the QTLs located most closely to the traits are those with a lower average influence, and those more distant are the QTLs with greater influence, not accounting for the amplitude of their action. This hierarchical system also suggests that the genetic control of these five traits is extremely complex and that pleiotropy is very dispersed. It also suggests that EL and NKR are highly correlated, as are ED and NR.

The high correlation between the five traits can also be observed through the use of biplots, where the normalized effects of the 27 QTLs on the five traits were decomposed into the principal components. Figure 5 depicts a group of QTLs that controls EL and NKR and another group that controls ED and NR. In addition, these groups have effects with opposing signals. The QTLs with low additive pleiotropy are found near the coordinates (0,0) in the biplot.

It is also possible to estimate the additive genetic correlation between the traits from Fig. 5 by taking the cosine of the lesser angle between the vectors for EL, ED, PROD, NR and NKR. The correlation was estimated from the biplot and was representative of the values obtained by variance components (Table 3—Aguiar 2003—PhD Thesis—data not published). For example, the correlation between ED and NR in the biplot was 0.982 and was 0.63 via variance components (VC). The correlation between EL and NKR was 0.72 in the biplot and was 0.46 via CV. The trait with the largest additive correlation with GY was ED (0.89 via biplot and 0.52 via CV) and the lowest correlation (in scale) was EL (−0.50 via biplot and 0.12 via CV). This result suggests that the additive biplot illustrates the importance of pleiotropy in determining the additive

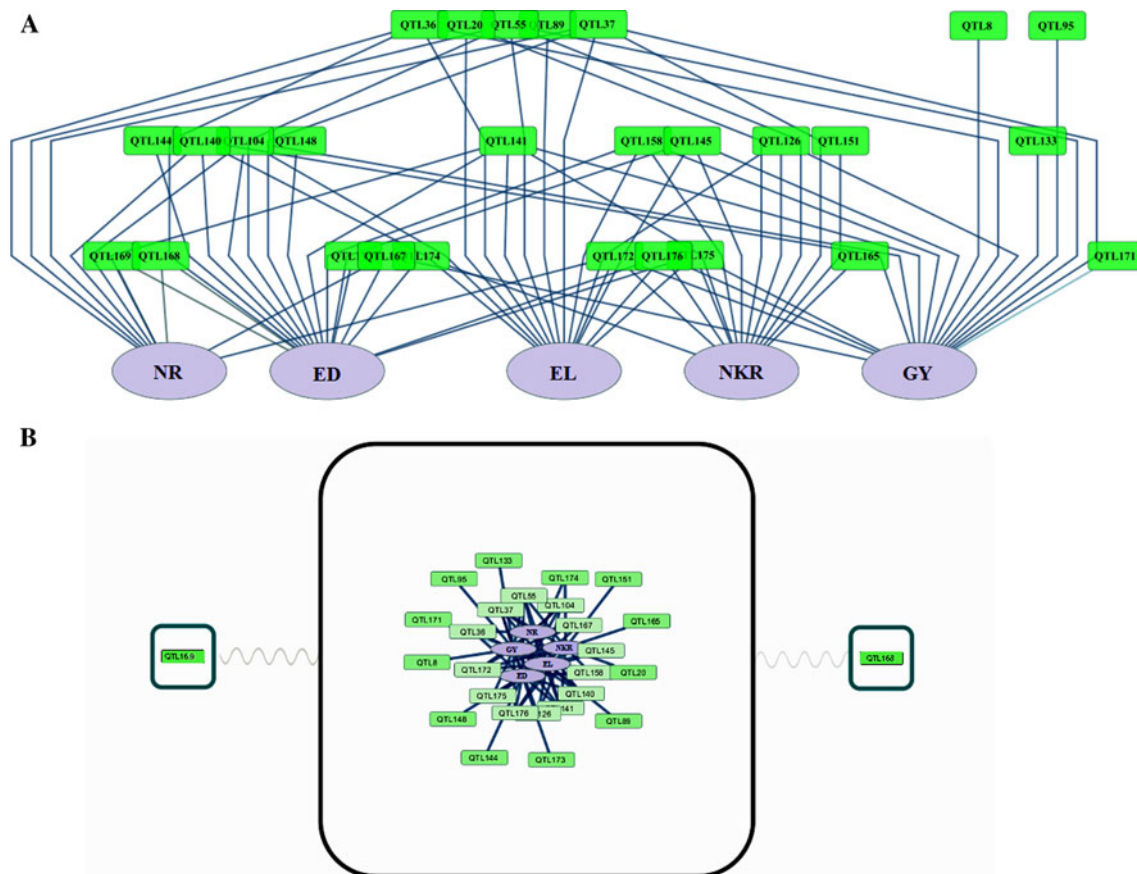


**Fig. 3** Profile of the QTL location in the maize genome by applying the pleiotropic model (P) and the linkage model (L) for the traits grain yield (GY), the number of rows per ear (NR) and the number of kernels per row (NKR)

genetic correlation between these traits, or in other words, the 17 pleiotropic QTLs found were highly descriptive of the genetic correlation estimated by conventional methods using VC. The heritability of these traits ranged from 0.70 to 0.89 for GY and NR, respectively.

Figure 5 also demonstrates the dominant effects of the 27 QTLs and their importance in the genetic control of GY and its components. In the additive biplot, GY was more

influenced by the traits ED and NR, but the biplot of the dominance effects showed that GY was more influenced by the traits EL and NKR. Additionally, it was observed that the QTLs with greater dominant relevance were somewhat distinct from the QTLs with greater additive relevance. As we know, the parental line with higher grain yield per se has smaller ears, but a wider diameter and, consequently, more rows per ears. Thus, QTL evidence suggests that in



**Fig. 4** Hierarchical pleiotropic network (a) and modular analysis (b) of QTLs acting on grain yield (PROD) in maize and its components, defined as the number of kernels per row (NKR), the number of rows (NR), the ear length (EL) and the ear diameter (ED)

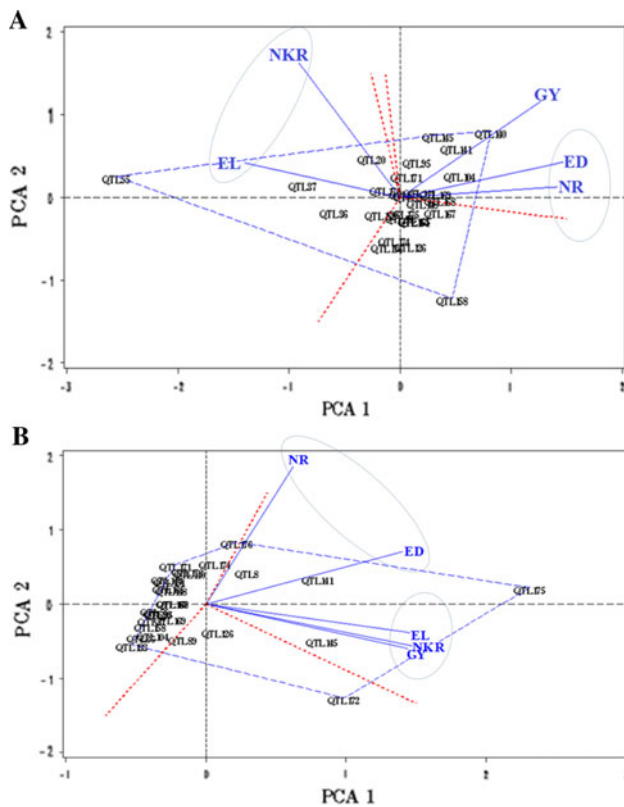
the hybrid, the allelic interaction and complementation of the pleiotropic QTLs allows selection to increase ER and NR. One could, for instance, simultaneously obtain dominant QTLs that increase the EL and NKR. This property may compensate for the pleiotropic QTLs that have different signs of additive effects across the characters as shown in Fig. 2. For example, QTL 55 (located in linkage group 3) has a positive additive effect for NKR and EL and a negative additive effect for ED and NR, but null dominance effect. This allows for hybrid combinations of contrasting lines increasing both characters. Thus, the pleiotropy of the allelic interactions may minimize the possible losses from negative correlations from the pleiotropy of the QTL because the signs of these interactions are not always the same and the degree of dominance that the pleiotropic QTLs have over certain traits is also quite varied.

## Discussion

The complexity of the inheritance of quantitative traits has become an area of intense focus in recent reviews (Mackay

2001; Flint and Mackay 2009; Mackay et al. 2009; Buchanan et al. 2009; Wagner and Zhang 2011). These reviews have shown that the “complexity cost” of an organism is directly linked to the distribution of pleiotropy in the genome. Our results suggest that the genetic control of grain yield and its components may be inseparable from the action of QTLs and that pleiotropy has a fundamental role in the genetic architecture of these traits. In these analyses, it was difficult to separate the genetic control of these five traits based on the action of the QTLs (Figs. 1, 2, 3 and 4). It is apparent from Fig. 2 that grain yield emerges from the interaction of the QTLs present in ED, EL, NR and NKR, which may lead to the development of a multiplicative model acting on these traits (Schnell and Cockerham 1992).

Schnell and Cockerham (1992) evaluated the possibility that heterosis is a product of a multiplicative interaction between subtraits. Because the posterior distribution for grain yield can be superimposed over the additive distributions of ED and NR and over the dominance distributions for EL and NKR (Fig. 2) and because pleiotropic QTLs that act on these traits possess different allelic interactions (Fig. 5), we can predict the genetic control of



**Fig. 5** Principal component analysis of additive (a) and dominant (b) QTL effects on maize grain yield (GY) and its components, defined as the number of kernels per row (NKR), the number of rows (NR), the ear length (EL) and the ear diameter (ED)

**Table 3** Additive genetic correlation obtained by components of variance

	GY	ED	EL	NR	NKR
GY	1.00	0.52	0.12	0.13	0.52
ED		1.00	-0.12	0.63	0.05
EL			1.00	-0.19	0.46
NR				1.00	-0.07
NKR					1.00

Aguiar 2003, unpublished data

ED ear diameter, EL ear length, GY grain yield, NR number of rows, NKR number of kernels per row

GY by analyzing the QTLs of its components (Schnell and Cockerham 1992). For example, if we consider QTL 20 (chromosome 1—first peak in EL and NKR), we can see its additive specificity on EL and NKR, but a null effect on ED and NR (Fig. 2). Thus, in the multiplicative model by Schnell and Cockerham (1992), the prediction of this QTL would also have a null effect on GY as observed in Fig. 2, that is, its predicted effect would be  $QTL_{20} = \mu_{el}\mu_{ed} + \alpha_{el}\mu_{ed} + \alpha_{el}\mu_{ed} + d_{el}\mu_{ed} + d_{el}\mu_{ed} + \alpha_{el}\alpha_{ed} + d_{el}d_{ed}$ , where  $\mu_{ed} = 0$ ,  $\alpha_{ed} = 0$ ,  $d_{ed} = 0$   $e d_{ed} = 0$ . The same prediction

can be made for the behavior of QTL 145 (second peak on chromosome 8), which had an overdominant effect on EL and NKR and an additive effect on ED and NR but had a completely dominant effect on GY. Thus, it is possible to predict the behavior of the majority of the pleiotropic QTLs acting on GY but not the behavior of QTLs specifically for GY. Although four QTLs were found to specifically act on PROD, this does not mean that they are not related to other traits that were not evaluated in this study.

The universality of pleiotropic effects is speculated to be an obstacle for individual adaptation due to the “cost of complexity” or “pleiotropic cost,” where the adaptive network can be described in an  $n$ -dimensional space described by the Fisher Geometric Model (Wagner and Zhang 2011). According to Wagner and Zhang (2011) and several other authors (Wagner et al. 2008; Su et al. 2010; Stearns 2010), modular pleiotropy would be favorable from an adaptive/evolutionary standpoint, whereas universal pleiotropy would be an obstacle to the process.

From the genetic improvement viewpoint, the division of GY into modules of pleiotropic genes could facilitate genetic improvement because changing the modules of pleiotropic genes to arrive at the ideal genotype would be easier than changing several pleiotropic QTLs within one module. It would be similar to putting together a complex puzzle with part of it pre-assembled. However, our results show that although there was widely distributed pleiotropy among the QTLs (but not universally due to the number of traits studied), the allelic interaction within these QTLs does not necessarily show strong pleiotropy (Fig. 5). In this Figure, it is clear that the QTLs that show prevalent additive pleiotropic effect (QTL 55, QTL 37, QTL 36, QTL 104, QTL 141, QTL 145 and QTL 140) do not necessarily have prevalent dominant pleiotropy (QTL 175, QTL 172, QTL 145, QTL 141, QTL 8, QTL 176, QTL 174 and QTL 126). Furthermore, it was shown that QTL 145 had a strong additive effect on ED and NR but significant dominant control over EL and NKR, suggesting that pleiotropic QTLs can have different allelic interactions across traits.

These results suggest that even in widely dispersed pleiotropy, it is possible to obtain highly productive hybrids or highly adapted genotypes resulting from the various pleiotropic effects exerted by the allelic interactions within of the QTLs. These observations are similar to those described by Carbone et al. (2006), who showed that the sites in the pleiotropic gene *Catsup* in *Drosophila*, described by their quantitative trait nucleotides (QTNs), acted independently for some traits. Their results suggested that mutations in the *Catsup* gene do not necessarily act universally across all the traits, and therefore, the allelic effect of the gene corresponded to varying degrees of pleiotropy.

It is also important to note that in pericentromeric regions in maize, suppression of the recombination ratio may occur interfering with the linkage/pleiotropy relationship (Gore et al. 2009). In addition, it is difficult to characterize pleiotropic genes in linkage disequilibrium mapping studies because a quantitative locus can contain several genes (Mackay et al. 2009). If so, mostly linkage disequilibrium mapping might fail in detecting close linkage in genome analysis and thus pleiotropic models might be favored. We have found in simulation that inconclusive decision can be obtained based on BF values and eventually pleiotropy and close linkage can be unraveled (Fig. 1—Supplementary Materials). This is a function of sample size, marker density and strength of effects. Furthermore, estimates of strongly linked QTLs localized in different interval were obtained in our mapping using pleiotropic model (for instance in chromosome 10).

Another simulation has been carried out using several approaches and, analyzing initial results, we can infer that selection of linkage models may be conditional on the number of phenotyped/genotyped individuals, heritability of the traits, level of genetic correlation, number of intervals and its length and number of QTLs retained in the model. However, due to strong genetical relationship expected between traits in this study (grain yield and its components) and to the potential of pleiotropic effects to be more informative than linkage in the models considered, our results suggest that pleiotropy is prevalent on grain yield and its components (just 37 % of the QTLs had linkage). However, further studies adding non-correlated biological traits such as grain density and plant height may eventually show prevalence of linkage QTLs.

It should be emphasized that our linkage model did not account for two pleiotropic QTLs in one interval but two QTLs with independent effects in an interval. In this case, our inference about the evidence supporting a pleiotropic model is limited by the small probability of observing QTLs that are very near each other and have independent effects on the traits. Buckler et al. (2009) has tried to identify pleiotropic and linkage effects from QTLs using correlation of QTL effects in 25 families. We think as correlations would be affected by both effects, such type of studies could be done in advance to access simpler models (candidate QTL regions) to further detection linkage or pleiotropic effects.

This study supports the power of Bayesian shrinkage analysis to describe the genetic architecture of quantitative traits, as has also been observed by Xu (2003), Wang et al. (2005) and Xu et al. (2009). However, some authors have suggested other approaches for mapping multiple traits (Jiang and Zeng 1995; Weller et al. 1996; Gilbert and Le Roy 2003; Malosetti et al. 2008; Banerjee et al. 2008). Some of these methods have large computational

advantages but are not as genetically interpretable as our approach. Some others are heavily dependent on model selection or are more computationally intensive.

Recent studies have suggested a joint modeling that combines multiple traits and multiple environments into one single model (Malosetti et al. 2008; Van Eeuwijk et al. 2010). These authors consider the QTL effect to be fixed and the marginal effects of background as random. Although joint modeling of several environments and traits is desirable, using simple interval analysis may counterbalance this possible gain. Additionally, in these approaches, the covariance of the QTLs is confused with the residual covariance, for both direction and magnitude, because they do not directly explore the specific (co)variances of each QTL across the traits. Furthermore, treating the QTL as fixed appears to contradict the uncertainty present in the genetic architecture of a complex trait (Buchanan et al. 2009).

Recent studies have attempted to describe the genetic control of heterosis by mapping complex traits in maize (Lu et al. 2003; Frascaroli et al. 2007; Garcia et al. 2008; Schön et al. 2010). The majority of these studies have observed a strong dominant/overdominant effect on the genetic control of heterosis and a marginal effect from epistasis. Our results suggest that the control of heterosis is partially dominant but has an equal distribution of QTLs with additive effects, partial dominance and overdominance. Although epistasis was not investigated in this work, the literature shows that it has marginal effects on the genetic control of heterosis (Garcia et al. 2008; Schön et al. 2010). Two points require further attention: (1) reconciling epistatic effects only in QTLs with a large effect may not describe the real contribution of this phenomenon in the genetic control of heterosis, thus requiring the application of complex models for the study of the distribution of epistasis across the genome, and (2) epistasis may not be trait specific and may be better studied within the concept of multiple pleiotropic QTLs because these two phenomena (pleiotropy and epistasis) may be inseparable (Schnell and Cockerham 1992; Cheverud et al. 2004; Wolf et al. 2005; Wolf et al. 2006; Pavlicev et al. 2008; Mackay et al. 2009).

Our results reinforce the discussion raised by Mackay (2001), Mackay et al. (2009) and Buchanan et al. (2009) and suggest that the inheritance of traits is limitedly described if analyzed individually but may be better understood through pleiotropic networks that allow us to visualize the complexity of inheritance of quantitative traits and, from this, develop new selection strategies. It was also possible to confront the idea that it is possible to identify QTLs for complex traits such as grain yield, as pleiotropy acts prominently on its subtraits and as this “trait” can be broken down and predicted almost completely by the QTLs

of its components. Furthermore, the pleiotropy of QTLs does not necessarily describe the pleiotropy of the allelic interactions, and therefore, the wide distribution of the pleiotropic effects does not limit the genetic adaptability of plants.

**Acknowledgments** Authors would like to thank Professor Anete Pereira de Souza of the Department of Biology and Evolution of the State University of Campinas for genotyping the population. Criticisms of both referees were very insightful and made this a better paper. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) has supported this research.

## References

- Banerjee S, Yandell BS, Yi N (2008) Bayesian quantitative trait loci mapping for multiple traits. *Genetics* 179:2275–2289
- Buchanan AV, Sholtis S, Richtsmeier J, Weiss KM (2009) What are genes “for” or where are traits “from”? What is the question? *Bioessays* 31:198–208
- Buckler ES, Holland JB, Bradbury PJ et al (2009) The genetic architecture of maize flowering time. *Science* 325:714–718
- Carbone MA, Jordan KW, Lyman RF et al (2006) Phenotypic variation and natural selection at *Catsup*, a pleiotropic quantitative trait gene in *Drosophila*. *Curr Biol* 16:912–919
- Cheverud JT, Ehrich H, Vaughn TT et al (2004) Pleiotropic effects on mandibular morphology II: differential epistasis and genetic variation in morphological integration. *J Exp Zool* 302:424–435
- Flint J, Mackay TFC (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res* 19:723–773
- Frascaroli E, Canè MA, Landi PG et al (2007) Classical genetic and quantitative trait loci analyses of heterosis in a maize hybrid between two elite inbred lines. *Genetics* 176:625–644
- García AA, Wang S, Melchinger AE, Zeng ZB (2008) Quantitative trait mapping and the genetic basis of heterosis in maize and rice. *Genetics* 180:1707–1724
- George EI, McCulloch RE (1993) Variable selection via Gibbs Sampling. *J Am Stat Assoc* 88:881–889
- Gilbert H, Le Roy P (2003) Comparison of three multitrait methods for QTL detection. *Genet Sel Evol* 35:281–304
- Gore MA, Chia JM, Elshire R et al (2009) A first-generation haplotype map of maize. *Science* 326:1115–1117
- Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109
- Jiang C, Zeng Z (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140:1111–1127
- Joehanes R, Nelson JC (2008) QGene 4.0, an extensible Java QTL-analysis platform. *Bioinformatics* 24:2788–2789
- Lincoln SE, Daly MJ, Lander ES (1992) Constructing genetic maps with Mapmaker Exp 3.0, 3rd edn. Whitehead Institute for Biometrical Research, Cambridge
- Liu J, Liu Y, Liu X, Deng HW (2007) Bayesian mapping of quantitative trait loci for multiple complex traits with the use of variance components. *Am J Hum Genet* 81:304–320
- Lu H, Romero-Severson J, Bernardo R (2003) Genetics basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. *Theor Appl Genet* 107:494–502
- Mackay TFC (2001) The genetic architecture of quantitative traits. *Annu Rev Genet* 35:336–339
- Mackay TFC, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* 10:565–577
- Malosetti M, Ribaut JM, Vargas M, Crossa J, Van Eeuwijk FAA (2008) Multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* 161:241–257
- Melchinger AE, Utz HF, Piepho HP, Zeng ZB, Schön CC (2007) The role of epistasis in the manifestation of heterosis: a systems-oriented approach. *Genetics* 177:1815–1825
- Metropolis N, Rosenblut AW, Rosenblut MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092
- Pavlicev M, Kenney-Hunt JP, Norgard EA et al (2008) Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution* 62:199–213
- Raftery AE, Newton MA, Satagopan J, Krivitsky P (2007) Estimating the integrated likelihood via posterior simulation using the harmonic mean identity (with discussion). *Bayesian Stat* 8:1–45
- Schnell FW, Cockerham CC (1992) Multiplicative vs. arbitrary gene action in heterosis. *Genetics* 131:461–469
- Schön CC, Dhillon BS, Utz HF, Melchinger AE (2010) High congruency of QTL positions for heterosis of grain yield in three crosses of maize. *Theor Appl Genet* 120:321–332
- Shannon P, Markiel A, Ozier O et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
- Sibov ST, Souza-Júnior CL, Garcia AAF et al (2003) Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 1. Map construction and localization of loci showing distorted segregation. *Hereditas* 139:96–106
- Stearns FW (2010) One hundred years of pleiotropy: a retrospective. *Genetics* 186:767–773
- Su Z, Zeng Y, Gu X (2010) A preliminary analysis of gene pleiotropy estimated from protein sequences. *J Exp Zool* 314:115–122
- Van Eeuwijk FA, Boer M, Totir RL et al (2010) Mixed model approaches for the identification of QTLs within a maize hybrid breeding program. *Theor Appl Genet* 120:429–440
- Varona L, Gomez-Raya L, Rauw WM, Clop A, Noguera JL (2004) Derivation of a Bayes factor to distinguish between linked or pleiotropic quantitative trait loci. *Genetics* 166:1025–1035
- Wagner GP, Zhang J (2011) The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat Rev Genet* 12:204–213
- Wagner GP, Kenney-Hunt JP, Pavlicev M et al (2008) Pleiotropic scaling of gene effects and the ‘cost of complexity’. *Nature* 27:470–472
- Wang H, Zhang YM, Li X et al (2005) Bayesian shrinkage estimation of quantitative trait loci parameters. *Genetics* 170:465–480
- Wang Z, Liao BY, Zhang J (2010) Genomic patterns of pleiotropy and the evolution of complexity. *Proc Natl Acad Sci* 107:18034–18039
- Weller JI, Wiggans GR, Vanraden PM, Ron M (1996) Application of a canonical transformation to detection of quantitative trait loci with the aid of genetic markers in a multitrait experiment. *Theor Appl Genet* 92:998–1002
- Wolf JB, Leamy LL, Routman EJ, Cheverud JM (2005) Epistatic pleiotropy and the genetic architecture of covariation within early and late-developing skull trait complexes in mice. *Genetics* 171:1–12
- Wolf JB, Pomp D, Eisen EJ, Cheverud JM, Leamy LL (2006) The contribution of epistatic pleiotropy to the genetic architecture of covariation among polygenic traits in mice. *Evol Dev* 8:468–476

- Wright S (1968) Evolution and the genetics of populations. University of Chicago Press, Chicago
- Xu S (2003) Estimating polygenic effects using markers of the entire genome. *Genetics* 163:789–801
- Xu C, Wang X, Li Z, Xu S (2009) Mapping QTL for multiple traits using Bayesian statistics. *Genet Res* 91:23–37
- Yang RC (2004) Epistasis of quantitative trait loci under different gene action models. *Genetics* 167:1493–1505
- Yang R, Xu S (2007) Bayesian shrinkage analysis of quantitative trait loci for dynamic traits. *Genetics* 176:1169–1175
- Zeng ZB, Wang T, Zou W (2005) Modeling quantitative trait loci and interpretation of models. *Genetics* 169:1711–1725