The proof of Theorem 4 is complete.

4 Concluding Remarks

In this paper we have generalized the idea on the net reproductive number and used it to discuss the asymptotic behavior and the periodic solutions of some age structured population models. Some very natural conclusions have been presented and proved. The results and ideas are on the non-autonomous and non-linear age structured population models. Many criteria can be obtained based on the average net reproductive number.

Acknowledgments

This research is supported by the State Education Commission Foundation for those who studied abroad.

References

1. L. Chen, J. Chen, *Nonlinear Dynamics of Biological Systems*, Science Press, Beijing, 1993.
2. G. F. Webb, *Theory of Age Dependent Population Danymics*, Marcel Dekker, 1985.
3. J. M. Cushing, Bifurcations of time periodic solutions of the McKendrick equations with applications to population dynamics, *Comp. Math. Appl.* **9(3)**, 459–478 (1983).
4. J. H. Swart, Hopf bifurcation and the stability of the a nonlinear age-dependent population models, *Comp. Math. Appl.* **15**, 555–564 (1988).
5. J. Yuan, Periodic solutions of a nonlinear autonomous age-dependent model of single species population dynamics, *Nonlin. Anal.* **12(10)**, 1079–1086 (1988).
6. N. E. Chipev and T. Kostova, A model of the dynamics of intramolluscan trematode populations: some problems concerning oscillatory behavior, *Comp. Math. Appl.* **21(5)**, 1–13 (1991).
7. Y. Zhou and X. Li, The continous solutions and stability of the age-structured disabled population model, *The proceedings of the 4th CSIAM conference*, Fudan University Press, Shanghai, 1996.

MIXED MODEL APPROACHES FOR GENETIC ANALYSIS OF QUANTITATIVE TRAITS

JUN ZHU
Department of Agronomy, Zhejiang Agricultural University, Hangzhou 310029, China

BRUCE S. WEIR
Department of Statistics, NC State University, Raleigh, NC 27695, USA

Cockerham's general genetic model was extended to seed models including direct, cytoplasm and maternal effects as well as genotype×environment interaction effects. The mixed linear model approaches for estimating variances and covariances and for predicting random effects were presented. A mixed model, with quantitative trait loci (QTL) effects being fixed and molecular marker effects being random, was suggested for searching QTLs. The appropriate mixed model approaches were proposed for searching QTLs with genetic main effects and GE interaction effects.

1 Introduction

After Fisher (1925) proposed methods for analysis of variance (ANOVA), many genetic models have been developed based on the ANOVA approaches. Some of these models, e.g. NC design I and II (Comstock *et al.*, 1952; Hallauer and Miranda, 1981), diallel models (Griffing, 1956; Gardner and Eberhart, 1966), are still widely used by plant and animal breeders. But ANOVA approaches have some deficiencies in analyzing genetic models with unbalanced data, or non-integer values of coefficients, or correlated random factors. The further development of quantitative genetics has been restrained in some ways by its prevailing dependency on ANOVA approaches.

In 1970s statisticians developed some new methods for analyzing mixed linear models which can be applied in quantitative genetics. Mixed linear model approaches overcome the shortcomings of ANOVA methods for handling unbalanced data and complicated models. Development of mixed linear model approaches and its application in quantitative genetics will create enormous challenges for quantitative geneticists in dealing with complicated genetic problems. In this paper we will present some of our recent work in extending Cockerham's general genetic model methodology (1980) and the mixed linear model approaches for quantitative genetics. Several genetic models, which can not be analyzed by ANOVA, will be presented. Recently developed methods for mixed linear models with their applications will be illustrated to show the

ways of solving the real complicated problems in quantitative genetics.

2 General Genetic Models and Its Extensions

When a genetic experiment is conducted in one environment, the average phenotypic performance (y) of a genetic population can be expressed in a simple linear model,

$$y = \mu + G + e,$$

where μ is population mean, G is total genotypic effect, and e is residual effect.

Cockerham (1980) proposed a general genetic model for partitioning total genetic effect G. If there are only additive and dominance effects, G can be partitioned into two components

$$G = \sum_i \alpha_i A_i + \sum_i \sum_{j \geq i} \delta_{ij} D_{ij}, \tag{1}$$

where A_i = additive effect with coefficient $\alpha_i (\sum_i \alpha_i = 2)$, D_{ij} = dominance effect with coefficient $\delta_{ij} (\sum_i \sum_{j \geq i} \delta_{ij} = 1)$.

Zhu and Weir (1994a) extended Cockerham's general genetic model by including seed direct gene effect (G_{O}), cytoplasm gene effect (G_{C}), and maternal nuclear gene effect (G_{M})($G = G_{\mathrm{O}} + G_{\mathrm{C}} + G_{\mathrm{M}}$). Further partitioning was also proposed:

$$\begin{aligned} G_{\mathrm{O}} &= \textstyle\sum_i \tau_i A_i + \sum_i \sum_{j \geq i} \delta_{ij} D_{ij}, \\ G_{\mathrm{C}} &= \textstyle\sum_i \gamma_i C_i, \\ G_{\mathrm{M}} &= \textstyle\sum_i \tau_{m_i} A_{m_i} + \sum_i \sum_{j \geq i} \delta_{m_{ij}} D_{m_{ij}}, \end{aligned} \tag{2}$$

where A_i = direct additive effect, D_{ij} = direct dominance effect, C_i =cytoplasm gene effect, A_{m_i} = maternal additive effect, $D_{m_{ij}}$ = maternal dominance effect.

Genetic models were proposed for quantitative traits of diploid seeds and animals (Zhu and Weir, 1994a) and of triploid endosperm (Zhu and Weir, 1994b). Usually means of only three generations (P's, F_1's and F_2's) are required for analyzing seed traits. In these models, some coefficients are non-integer, and direct genetic effects are correlated with maternal genetic effects. Therefore they can be appropriately handled only by mixed linear model approaches.

Genotype × environment (GE) interactions have been detected for quantitative traits of many plants and animals. With genetic experiments conducted in multiple environments, the average phenotypic performance of a genetic population in one environment can be expressed by the following genetic model,

$$y = \mu + E + G + GE + e,$$

where μ = population mean, E = environment effect, G = total genotypic effect, GE = genotype×environment interaction effect, e = residual effect.

Cockerham's (1980) general genetic model can be extended by including $G + GE$ interaction (Zhu, 1994). The total genotypic effect G is defined the same as in Equation (1), the partitioning of GE interaction for interaction of additive and dominance effects is

$$GE = \sum_h \sum_i \alpha_{hi} AE_{hi} + \sum_h \sum_i \sum_{j \geq i} \beta_{hij} DE_{hij}, \tag{3}$$

where $AE_{hi} = A_i \times E_h$ interaction effect, $DE_{hij} = D_{ih} \times E_h$ interaction effect.

The general genetic model for seeds (Zhu and Weir, 1994a) can also be extended by including genotype by environment terms (Zhu, 1994),

$$GE = G_{\mathrm{O}} + G_{\mathrm{C}} E + G_{\mathrm{M}} E. \tag{4}$$

The GE interaction terms can be further partitioned into its components:

direct interaction

$$G_{\mathrm{O}} E = \sum_h \sum_i \alpha_{hi} AE_{hi} + \sum_h \sum_i \sum_{j \geq i} \beta_{hij} DE_{hij},$$

cytoplasm interaction

$$G_{\mathrm{C}} E = \sum_h \sum_i \lambda_{hi} CE_{hi}, \tag{5}$$

maternal interaction

$$G_{\mathrm{M}} E = \sum_h \sum_i \alpha_{m_{hi}} A_m E_{hi} + \sum_h \sum_i \sum_{j \geq i} \beta_{m_{hij}} D_m E_{hij}.$$

Based on this extension of general genetic model for seeds, experiments of a diallel cross with three generations (P's, F_1's and F_2's) in multiple environments can be appropriately analyzed (Zhu, 1996). For dicot seeds the partition of total $G + GE$ effect for three generations is,

$$\begin{aligned} G(P_i) + GE(P_i) &= 2A_i + D_{ii} + C_i + 2A_{m_i} + D_{m_{ii}} \\ &\quad +2AE_{hi} + DE_{hii} + CE_{hi} \\ &\quad +2A_m E_{hi} + D_m E_{hii}, \\ G(F_{1ij}) + GE(F_{1ij}) &= A_i + A_j + D_{ij} + C_i + 2A_{m_i} + D_{m_{ii}} \\ &\quad +AE_{hi} + AE_{hj} + DE_{hij} \\ &\quad +CE_{hi} + 2A_m E_{hi} + D_m E_{hii}, \\ G(F_{2ij}) + GE(F_{2ij}) &= A_i + A_j + \tfrac{1}{4} D_{ii} + \tfrac{1}{4} D_{jj} + \tfrac{1}{2} D_{ij} + C_i \\ &\quad +A_{m_i} + A_{m_j} + D_{m_{ij}} + AE_{hi} + AE_{hj} \\ &\quad +\tfrac{1}{4} DE_{hii} + \tfrac{1}{4} DE_{hjj} + \tfrac{1}{2} DE_{hij} + CE_{hi} \\ &\quad +A_m E_{hi} + A_m E_{hj} + D_m E_{hij}. \end{aligned} \tag{6}$$

For endosperm the partition of total GE interaction effect for three generations is,

$$\begin{aligned}
G(P_i)+GE(P_i) &= 3A_i+3D_{ii}+C_i+2A_{m_i}+D_{m_{ii}} \\
&\quad +3AE_{hi}+3DE_{hii}+CE_{hi} \\
&\quad +2A_mE_{hi}+D_mE_{hii}, \\
G(F_{1ij})+GE(F_{1ij}) &= A_i+A_j+D_{ij}+C_i+2A_{m_i}+D_{m_{ii}} \\
&\quad +2AE_{hi}+AE_{hj}+DE_{hii}+2DE_{hij} \\
&\quad +CE_{hi}+2A_mE_{hi}+D_mE_{hii}, \\
G(F_{2ij})+GE(F_{2ij}) &= \tfrac{3}{2}A_i+\tfrac{3}{2}A_j+D_{ii}+D_{jj}+D_{ij}+C_i \\
&\quad +A_{m_i}+A_{m_j}+D_{m_{ij}}+\tfrac{3}{2}AE_{hi}+\tfrac{3}{2}AE_{hj} \\
&\quad +DE_{hii}+DE_{hjj}+DE_{hij}+CE_{hi} \\
&\quad +A_mE_{hi}+A_mE_{hj}+D_mE_{hij}.
\end{aligned} \tag{7}$$

In the general genetic model and its extensions, genetic effects can be defined as random effects or fixed effects. Since genetic experiments are usually conducted in small number of environments (e.g. years, locations, or treatments), environment effects could be treated as fixed. If experiments use genetic parents as a sample from a reference population and need to infer the genetic variation for the population, genetic effects and GE interaction effects are all defined as random effects. If genetic entries are selected specifically for evaluating their merit, genetic effects could be treated as fixed.

3 Mixed Linear Model Approaches

Any general genetic models can be expressed by a form of mixed linear model,

$$\begin{aligned}
\mathbf{y} &= \mathbf{X}_1\mathbf{b}_1+\mathbf{X}_2\mathbf{b}_2+\cdots+\mathbf{X}_n\mathbf{b}_n+\mathbf{U}_1\mathbf{e}_1+\mathbf{U}_2\mathbf{e}_2+\cdots+\mathbf{U}_m\mathbf{e}_m \\
&= \mathbf{Xb}+\sum_{u=1}^{m}\mathbf{U}_u\mathbf{e}_u,
\end{aligned} \tag{8}$$

where $\mathbf{y}$ is the vector of phenotypic mean for all entries of the mating design; $\mathbf{b}$ is the vector of fixed environment effects; $\mathbf{X}$ is the known incidence matrix with coefficients 1 or 0 relating to the fixed environment effects; $\mathbf{e}_u$ is the vector of random effects; $\mathbf{U}_u$ is the known coefficient matrix relating to the random vector $\mathbf{e}_u$.

The parameters in Equation (8) can be analyzed by mixed linear model approaches such as maximum likelihood (ML) method (Hartley and Rao, 1967), restricted maximum likelihood (REML) method (Patterson and Thompsom, 1971), or minimum norm quadratic unbiased estimation (MINQUE) method

(Rao, 1971). Estimated variances obtained by ML method tend to be influenced by the fixed effects, therefore ML is rare in use. Monte Carlo simulations (Zhu and Weir, 1994a, 1994b, 1996) showed that MINQUE has advantages over REML for its (1) simple computation without iterations, (2) no requirement for normality distribution, and (3) unbiased estimation.

For seed models (6) and (7), the phenotypic mean can be expressed as

$$\begin{aligned}
\mathbf{y} &= \mathbf{Xb}+\mathbf{U}_\mathrm{A}\mathbf{e}_\mathrm{A}+\mathbf{U}_\mathrm{D}\mathbf{e}_\mathrm{D}+\mathbf{U}_\mathrm{C}\mathbf{e}_\mathrm{C}+\mathbf{U}_{\mathrm{A_m}}\mathbf{e}_{\mathrm{A_m}}+\mathbf{U}_{\mathrm{D_m}}\mathbf{e}_{\mathrm{D_m}} \\
&\quad +\mathbf{U}_\mathrm{AE}\mathbf{e}_\mathrm{AE}+\mathbf{U}_\mathrm{DE}\mathbf{e}_\mathrm{DE}+\mathbf{U}_\mathrm{CE}\mathbf{e}_\mathrm{CE} \\
&\quad +\mathbf{U}_{\mathrm{A_mE}}\mathbf{e}_{\mathrm{A_mE}}+\mathbf{U}_{\mathrm{D_mE}}\mathbf{e}_{\mathrm{D_mE}}+\mathbf{U}_\mathrm{B}\mathbf{e}_\mathrm{B}+\mathbf{e}_\mathrm{e} \\
&= \mathbf{Xb}+\sum_{u=1}^{12}\mathbf{U}_u\mathbf{e}_u
\end{aligned}$$

with variance-covariance matrix

$$\begin{aligned}
\mathrm{Var}(\mathbf{y}) &= \sigma^2_\mathrm{A}\mathbf{V}_1+\sigma^2_\mathrm{D}\mathbf{V}_2+\sigma^2_\mathrm{C}\mathbf{V}_3+\sigma^2_{\mathrm{A_m}}\mathbf{V}_4+\sigma^2_{\mathrm{D_m}}\mathbf{V}_5+ \\
&\quad \sigma^2_\mathrm{AE}\mathbf{V}_6+\sigma^2_\mathrm{DE}\mathbf{V}_7+\sigma^2_\mathrm{CE}\mathbf{V}_8+\sigma^2_{\mathrm{A_mE}}\mathbf{V}_9+\sigma^2_{\mathrm{D_mE}}\mathbf{V}_{10}+\sigma^2_\mathrm{B}\mathbf{V}_{11}+ \\
&\quad \sigma_{\mathrm{A.A_m}}\mathbf{V}_{12}+\sigma_{\mathrm{D.D_m}}\mathbf{V}_{13}+\sigma_{\mathrm{AE.A_mE}}\mathbf{V}_{14}+\sigma_{\mathrm{DE.D_mE}}\mathbf{V}_{15}+\sigma^2_\mathrm{e}\mathbf{V}_{16} \\
&= \sum_{u=1}^{16}\theta_u\mathbf{V}_u,
\end{aligned}$$

where $\mathbf{V}_u=\mathbf{U}_u\mathbf{U}_u^T(u=1,2,\cdots,11)$, $\mathbf{V}_{12}=(\mathbf{U}_1\mathbf{U}_4^T+\mathbf{U}_4\mathbf{U}_1^T)$, $\mathbf{V}_{13}=(\mathbf{U}_2\mathbf{U}_5^T+\mathbf{U}_5\mathbf{U}_2^T)$, $\mathbf{V}_{14}=(\mathbf{U}_6\mathbf{U}_9^T+\mathbf{U}_9\mathbf{U}_6^T)$, $\mathbf{V}_{15}=(\mathbf{U}_7\mathbf{U}_{10}^T+\mathbf{U}_{10}\mathbf{U}_7^T)$, $\mathbf{V}_{16}=\mathbf{I}$; $\theta_1=\sigma^2_\mathrm{A}$, $\theta_2=\sigma^2_\mathrm{D}$, $\theta_3=\sigma^2_\mathrm{C}$, $\theta_4=\sigma^2_{\mathrm{A_m}}$, $\theta_5=\sigma^2_{\mathrm{D_m}}$, $\theta_6=\sigma^2_\mathrm{AE}$, $\theta_8=\sigma^2_\mathrm{DE}$, $\theta_9=\sigma^2_\mathrm{CE}$, $\theta_{10}=\sigma^2_{\mathrm{A_mE}}$, $\theta_{11}=\sigma^2_{\mathrm{D_mE}}$, $\theta_{12}=\sigma^2_{\mathrm{A.A_m}}$, $\theta_{13}=\sigma^2_{\mathrm{D.D_m}}$, $\theta_{14}=\sigma^2_{\mathrm{AE.A_mE}}$, $\theta_{15}=\sigma^2_{\mathrm{DE.D_mE}}$, $\theta_{16}=\sigma^2_\mathrm{e}$.

MINQUE(0/1), which is a MINQUE method with 0 for all the prior covariances and 1 for all the prior variances, was suggested by Zhu and Weir (1994a) for unbiased estimation of variances for one traits and covariances between two traits. variances and covariances for one trait ($\mathbf{y}_a=\mathbf{y}_b$), and covariances between two traits ($\mathbf{y}_a\neq\mathbf{y}_b$) can be estimated by the following MINQUE(0/1) equations,

$$[\mathrm{tr}(\mathbf{Q}_{(0/1)}\mathbf{V}_u\mathbf{Q}_{(0/1)}\mathbf{V}_v)]\,[\hat{\theta}_u]=[\mathbf{y}_a^T\mathbf{Q}_{(0/1)}\mathbf{V}_u\mathbf{Q}_{(0/1)}\mathbf{y}_b], \tag{9}$$

where

$$\begin{aligned}
\mathbf{Q}_{(0/1)} &= \mathbf{V}^{-1}_{(0/1)}-\mathbf{V}^{-1}_{(0/1)}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}_{(0/1)}X)^{+}\mathbf{X}^T\mathbf{V}^{-1}_{(0/1)} \\
\mathbf{V}_{(0/1)} &= \sum_{u=1}^{15}\mathbf{U}_u\mathbf{U}_u^T+\mathbf{I}.
\end{aligned}$$

Monte Carlo simulation showed that both variances and covariances of seed traits can be unbiasedly estimated by the MINQUE(0/1) procedure (Zhu and Weir, 1994a, 1994b).

Prediction of genetic merits in the mixed linear models can be obtained by: (1) the best linear unbiased prediction (BLUP) (Henderson, 1963), (2) linear unbiased prediction (LUP) (Zhu and Weir, 1994a), and (3) adjusted unbiased prediction (AUP) (Zhu, 1993; Zhu and Weir, 1996).

For seed models having multivariate normal distribution with correlated random variables $\mathrm{Cov}(\mathbf{e}_u,\ \mathbf{e}_v^T) = \sigma_{uv}\mathbf{I}$, the BLUP for random effects $\mathbf{e}_u$ is given by,

$$\hat{\mathbf{e}}_{u(\theta)} = (\sigma_u^2\mathbf{U}_u^T + \sigma_{u.v}\mathbf{U}_v^T)\mathbf{Q}\mathbf{y}, \tag{10}$$

where $\mathbf{Q} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{+}\mathbf{X}^T\mathbf{V}^{-1}$.

Since the true variances and covariances are always unknown in practice, estimated variances are usually used in prediction:

$$\hat{\mathbf{e}}_{u(\hat{\theta})} = (\hat{\sigma}_u^2\mathbf{U}_u^T + \hat{\sigma}_{u.v}\mathbf{U}_v^T)\hat{\mathbf{Q}}\mathbf{y}, \tag{11}$$

where $\hat{\mathbf{Q}} = \hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1}\mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{X})^{+}\mathbf{X}^T\hat{\mathbf{V}}^{-1}$.

With such prediction by using estimates, only a so-called "BLUP" is obtainable, and the linearity and unbiasedness of BLUP may be lost. Instead of using parameters or their estimates for predicting random effects, Zhu and Weir (1994a, 1996) suggested choosing prior values α_u as in the case of MINQUE method. When MINQUE(0/1) is used for estimating variances and covariances for seed traits, LUP can be used for predicting random genetic effects $\mathbf{e}_u$.

$$\hat{\mathbf{e}}_{u(0/1)} = \mathbf{U}_u^T\mathbf{Q}_{(0/1)}\mathbf{y}, \tag{12}$$

where $\mathbf{Q}_{(0/1)} = \mathbf{V}_{(0/1)}^{-1} - \mathbf{V}_{(0/1)}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}_{(0/1)}^{-1}\mathbf{X})^{+}\mathbf{X}^T\mathbf{V}_{(0/1)}^{-1}$. And AUP is obtainable by

$$\hat{\mathbf{e}}_{u(0/1)} = \kappa\mathbf{U}_u^T\mathbf{Q}_{(0/1)}\mathbf{y}, \tag{13}$$

where $\kappa = \sqrt{(n_u - 1)\hat{\sigma}_u^2/(\mathbf{y}^T\mathbf{Q}_{(0/1)}\mathbf{V}_u\mathbf{Q}_{(0/1)}\mathbf{y}}$.

Monte Carlo simulation revealed that both BLUP and LUP will give prediction with unbiased mean but under estimated variance ($\mathrm{E}(\hat{\mathbf{e}}^T\hat{\mathbf{e}}) < \sigma^2$) for random variables (Zhu and Weir, 1996), and that AUP can give both unbiased mean and estimated variance (Zhu, 1993; Zhu and Weir, 1996).

4 Mixed Model Approaches for Mapping QTLs

Quantitative trait loci (QTLs) have been searched by several methods, such as interval mapping method (Lander and Botstein, 1989) and composite interval mapping method (Zeng, 1994), which are based on the regression approaches. In the present paper we proposed a new method for mapping QTLs based on the mixed model approaches. When a genetic analysis is conducted to search QTLs using information of genetic markers, only QTL position and effects are especially interested and markers can be treated as a random sample from all potential markers. Therefore a mixed model can be employed with effects of the searching QTL being fixed and marker effects being random.

If a putative QTL is within two flanking codominant genetic markers M_{i-} and M_{i+}, the phenotypic value of quantitative trait measured on the jth individual can be expressed as a mixed linear model,

$$\begin{aligned} y_j &= \mu + ax_{A_j} + dx_{D_j} + \textstyle\sum_{i\neq i-,i+} e_{M_k}z_{M_{kj}} + \varepsilon_j \\ &= \mathbf{e}_j^T\mathbf{b} + \mathbf{z}_{M_j}^T\mathbf{e}_M + \varepsilon_j, \end{aligned} \tag{14}$$

where μ is the population mean; a and d are the additive and dominance effects for the QTL searched; x_{A_j} and x_{D_j} are coefficients for genetic effects; e_{M_k} is the random effect for the kth marker genotype with its coefficient $z_{M_{kj}}$ taking the value of 1 for $M_{k1}M_{k1}$, 0 for $M_{k1}M_{k2}$, or -1 for $M_{k2}M_{k2}$; and ε_j is the random residual effect; $\mathbf{b}$ is a vector of fixed parameters including u, a and d; $\mathbf{e}_M$ is a random vector of marker effects; $\mathbf{z}_{M_j}^T$ is a row vector of the coefficients for $\mathbf{e}_M$ of the jth individual; $\mathbf{x}_j^T$ is a row vector of the coefficients for $\mathbf{b}$ of the jth individual.

Since QTL genotype of the jth individual is unknown and model (14) is a mixture model, the second and third elements of $\mathbf{x}_j^T = [1,\ x_{A_j},\ x_{D_j}]$ can only be inferred by the probability of QTL genotype given observed flanking marker genotype.

If QTL mapping is conducted in several environments (years or locations) for individuals sampled from the same reference population, QTL genetic main effects as well as GE interaction effects can be evaluated. The phenotypic value of quantitative trait measured on the jth individual in the hth environment can be expressed as

$$\begin{aligned} y_{hj} &= \mu + ax_{A_j} + dx_{D_j} + \sum_{k\neq i-,i+} e_{M_k}z_{M_k} + e_{E_h}z_{E_{hj}} \\ &\quad + e_{AE_h}z_{AE_{hj}} + e_{DE_h}z_{DE_{hj}} + \sum_{k\neq i-,i+} e_{ME_{hk}}z_{ME_{hkj}} + \varepsilon_{hj}, \end{aligned} \tag{15}$$

where μ is the population mean; a and d are the additive and dominance main effects for the searching QTL; x_{A_j} and x_{D_j} are coefficients for genetic main effects; e_{M_k} is the random main effect over environments for the kth marker genotype with its coefficient $z_{M_{kj}}$; e_{E_h} is the hth environment effect with its

coefficient $z_{E_{hj}}$; e_{AE_h} is the additive by environment interaction effect with its coefficient $z_{AE_{hj}}$; e_{DE_h} is the dominance by environment interaction effect with its coefficient $z_{DE_{hj}}$; $e_{ME_{hk}}$ is the marker by environment interaction effect with its coefficient $z_{ME_{hkj}}$; and ε_j is the random residual effect.

Model (14) and (15) can be expressed as a matrix form of general mixed model,

$$\begin{aligned} \mathbf{y} &= \mathbf{Xb} + \sum_u \mathbf{Z}_u \mathbf{e}_u \\ &\sim N(\mathbf{Xb},\ \mathbf{V} = \sum_u \sigma_u^2 \mathbf{Z}_u \mathbf{F}_u \mathbf{Z}_u^T), \end{aligned} \tag{16}$$

where $\mathbf{y}$ is a vector of phenotypic values of quantitative trait studied; $\mathbf{b}$ is a vector of the fixed effects; $\mathbf{X}$ is the coefficient matrix with row vectors $\mathbf{x}_j^T$; $\mathbf{e}_u \sim N(\mathbf{0},\ \sigma_u^2 \mathbf{Z}_u \mathbf{F}_u \mathbf{Z}_u^T)$ is a vector of random effects, $\mathbf{F}_u$ is a constant matrix describing the relationship of $\mathbf{e}_u$; $\mathbf{Z}_u$ is the coefficient matrix for $\mathbf{e}_u$, and $\mathbf{Z}_u^T$ is the transpose matrix of $\mathbf{Z}_u$. Phenotypic vector $\mathbf{y}$ has mean $\mathbf{Xb}$ and variance $\mathbf{V} = \sum_u \sigma^2 \mathbf{Z}_u \mathbf{F}_u \mathbf{Z}_u^T$.

The likelihood function (L) for the parameters of fixed effects $\mathbf{b}$ and variance components $[\sigma_u^2]$ is

$$L(\mathbf{b},\ \mathbf{V}) = (2\pi)^{-\frac{n}{2}} |\mathbf{V}|^{-\frac{1}{2}} \exp[-\frac{1}{2}(\mathbf{y} - \mathbf{Xb})^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{Xb})] \tag{17}$$

with the log of the likelihood function (L) is

$$l(\mathbf{b},\ \mathbf{V}) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln |\mathbf{V}| - \frac{1}{2}(\mathbf{y} - \mathbf{Xb})^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{Xb}). \tag{18}$$

If variance components of the model are known, the maximum likelihood estimates of QTL effects in $\mathbf{b}$ can be obtained by

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}$$

with sampling variance matrix

$$\mathrm{Var}(\hat{\mathbf{b}}) = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1}.$$

We can search QTL within two flanking markers M_{i-} and M_{i+} for the whole genome by setting a prior value for recombination fraction $r_{M_{i-}Q}$ between mater M_{i-} and locus Q. For each prior value $r_{M_{i-}Q}$, the likelihood ratio statistic (LR) can be calculated by

$$LR = 2l_1(\hat{\mathbf{b}},\ \hat{\mathbf{V}},\ \hat{r}_{M_{i-}Q}) - 2l_0(\hat{\mathbf{b}},\ \hat{\mathbf{V}},\ r_{M_{i-}Q} = 0.5),$$

where variance components in $\mathbf{V}$ can be replaced by their unbiased estimates,

$$\hat{\mathbf{V}} = \sum_u \hat{\sigma}_u^2 \mathbf{Z}_u \mathbf{F}_u \mathbf{Z}.$$

The likelihood ratio statistic can also be used for testing the null hypothesis $\mathrm{H_0}$: $r_{M_{i-}Q} = 0.5$ vs. the alternative hypothesis $\mathrm{H_1}$: $r_{M_{i-}Q} < 0.5$. LR approximately has a χ^2 distribution with $df = 1$.

When the null presentation of non QTL within two flanking markers M_{i-} and M_{i+} is rejected, $r_{M_{i-}Q}$ infers the position of the QTL while $\mathbf{b}$ gives the estimates of additive and dominance effects of this QTL. Hypothesis for additive and dominance effects can be conducted by a t-test in a general way for

$$\mathrm{H_0}:\ \mathbf{c}^T\mathbf{b} = m \text{ vs.} \qquad \mathrm{H_1}:\ \mathbf{c}^T\mathbf{b} \neq m.$$

If the statistic $|\mathbf{c}^T(\hat{\mathbf{b}} - \mathbf{b}) / \sqrt{\mathbf{c}^T(\mathbf{X}^T\hat{\mathbf{v}}^{-1}\mathbf{X})\mathbf{c}}| > t_\alpha$, the null hypothesis is then rejected.

Acknowledgments

This research was supported in part by the China Natural Science Foundation 39670390 and by the NIH Grant GM32518.

References

1. Cockerham C. C., 1980. Random and fixed effects in plant genetics. *Theor. Appl. Genet.*, **56**, 119–131.
2. Comstock R. E. and Robinson H. F., 1952. Estimation of average dominance of genes. *In Heterosis.* Gowan J. W. (editor), Iowa State University Press, Ames, Iowa.
3. Fisher R. A., 1925. *Statistical Methods for Research Workers.* 1st ed. Oliver & Boyd. Edinburgh and London.
4. Griffing B., 1956. Concept of general and specific combining ability in relation to diallel crossing systems. *Aust. J. Biol. Sci.*, **9**, 463–493.
5. Hallauer A. R. and Miranda J. B., 1981. *Quantitative Genetics in Maize Breeding.* Iowa State University Press, Ames, Iowa.
6. Hartley H. D. and Rao J. N. K., 1967. Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika*, **54**, 93–108.
7. Henderson C. R., 1963. Selection index and expected genetic advance. *In Statistical Genetics and Plant Breeding.* Hanson W.D. and Robinson H.F. (eds.). Publication 982, pp 141–163. Washington: National Academy of Science, National Research Council.

8. Lander E. S. and Botstein D., 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185–189.
9. Patterson H.D. and Thompson R., 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika*, **58**, 545–554.
10. Rao C. R., 1971. Estimation of variance and covariance components MINQUE theory. *Journal of Multivariate Analysis*, **1**, 257–275.
11. Zeng Z.-B., 1994. Precision mapping of quantitative trait loci. *Genetics*, **136**, 1457–1468.
12. Zhu J., 1993. Methods of predicting genotype value and heterosis for offspring of hybrids. (Chinese). *Journal of Biomathematics* (Chinese), **8(1)**, 32–44.
13. Zhu J., 1994. General genetic models and new analysis methods for quantitative traits. (Chinese). *Journal of Zhejiang Agricultural University*, **20(6)**, 551–559.
14. Zhu J., 1996. Analysis methods for seed models with genotype×environment interactions. (Chinese). *Acta Genetica Sinica*, **23(1)**, 56–68.
15. Zhu J. and Weir B. S., 1994a. Analysis of cytoplasmic and maternal effects: I. a genetic model for diploid plant seeds and animals. *Theor. Appl. Genet.*, **89**, 153–159.
16. Zhu J. and Weir B. S., 1994b. Analysis of cytoplasmic and maternal effects II. genetic models for triploid endosperm. *Theor. Appl. Genet.*, **89**, 160–166.
17. Zhu J. and Weir B. S., 1996. Diallel analysis for sex-linked and maternal effects. *Theor. Appl. Genet.*, **92(1)**, 1–9.