seedling forestry. Eds. GIBSON, G. L., GRIFFIN, A. R. and MATHESON, A. C. OFI, Oxford, UK and Winrock, Arlington. pp. 39–51 (1988). — ELDRIDGE, K. G.: Southern Tree Breeding Association 1987 Working Plan. Unpublished report. CSIRO (1987). — HALLAUER, A. R. and MIRANDA, J. B.: Quantitative Genetics in Maize Breeding. Iowa State Univ. Press, Ames (1981). — HOPKINS, I. R.: Some optimum age structures and selection methods in open nucleus breeding schemes with overlapping generations. Anim. Prod. **26**: 267–276 (1978). — JAMES, J. W.: Open nucleus breeding systems. Anim. Prod. **24**: 287–305 (1977). — JAMES, J. W.: Effective population size in open nucleus breeding schemes. Acta. Agric. Scand. **28**: 387–392 (1978). — JAYAWICKRAMA, K. J. S. and BALOCCHI, C.: Growth and form of provenances of *Pinus radiata* in Chile. Aust. For. **56**: 172–178 (1993). — JOHNSON, G. R.: STBA's Nucleus Population: A Proposal for the Breeding Population Structure. Tech Rep. 90–01. STBA, Mt. Gambier, Australia. 22 p. (1990). — JOHNSON, I. G.: The Breeding Strategy for Radiata Pine In New South Wales. Tech. Pap. 47. New South Wales Forestry Commission, Sydney, Australia. 58 p. (1989). — KANG, H.: Long-term tree breeding. In: Proc. 15th South. For. Tree Improve. Conf. June 19 to 21, Starkeville, Miss. pp. 66–72 (1979). — LINDGREN, D.: Genetic gain by progeny testing as a function of mating design and cost. In: Third World Consultation on Forest Tree Breeding. March 21 to 26, Canberra, Australia. pp. 1223–1235 (1977). — LINDGREN, D.: How should breeders respond to breeding values? In: Proc. IUFRO Conf. Breeding Theory, Progeny Testing and Seed Orchards. October 13 to 17, Williamsburg, VA. pp. 361–372 (1986). — LINDGREN, D. and GREGORIUS, H. R.: Inbreeding and coancestry. In: Proc. IUFRO Adv. Gen. Breeding. June 1976, Bordeaux, France. pp. 49–65 (1976). — LINDGREN, D. and MATHESON, A. C.: An algorithm for increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. Silvae Genet. **35**: 173–177 (1986). — LOWE, W. J. and VAN BUIJTENEN, J. P.: The development of a sublining system in an operational tree improvement program. In: Proc. IUFRO Conf. on Breeding Theory, Progeny Testing and Seed Orchards. Oct. 1986, Williamsburg, VA. pp. 98–106 (1986). — MATHESON, A. C. and BROWN, A. G.: Radiata pine breeding Manual (CSIRO Division of Forest Research, Canberra). (1983). — MCKEAND, S. and BEINEKE, F.: Sublining for half-sib breeding populations of forest trees. Silvae Genet. **29**: 14–17 (1980). — MCKEAND, S.E. and BRIDGEWATER, F.E.: Third-generation breeding strategy for the North Carolina State University-Industry Cooperative Tree Improvement Program. In: Proc. IUFRO, Resolving Trop. For. Resource Concerns Through Tree Improv., Gene Conserv., and Domestication of New Species. October, Cali, Colombia. pp. 234–240 (1992). — NAMKOONG, G.: A multiple-index selection strategy. Silvae Genet. **25**: 199–201 (1976). — NAMKOONG, G. H., BURLEY, J. and BARNES, R. D.: A philosophy of breeding strategy for tropical forest trees. Tropical Forestry Paper No. 16. Commonwealth Forestry Institute, Oxford. 67 pp. (1980). — NAMKOONG, G., KANG, H. and BROUARD, J. S.: Tree Breeding: Principles and Strategies. Springer-Verlag, New York, NY. 175 p. (1989). — NEWTON TURNER, H. and YOUNG, S. S. Y.: Quantitative Genetics in Sheep Breeding. Macmillan, Melbourne, Australia (1969). — PEDERICK, L. A.: A Plan for Breeding Radiata Pine by the Department of Conservation, Forests and Lands. Res. Rep 38. Dept. Cons., Forests and Lands, Melbourne, Australia. 50 p. (1987). — PURNELL, R. C. and KELLISON, R. C.: A tree improvement program for southern hardwoods. In: Proc. 17th South. For. Tree Improve. Conf. June 7 to 9, Athens, GA. pp. 90–98 (1983). — SHELBOURNE, C. J. A.: Tree breeding methods. In: For. Res. Inst. Tech. Pap. 55. New Zealand For. Serv., Wellington, NZ. 45 p. (1969). — SHELBOURNE, C. J. A.: Genetic gains from different kinds of breeding population and seed or plant production population. South African Forestry Journal **160**: 49–66 (1992). — SHELBOURNE, C. J. A., BURDON, R. D., CARSON, S. D., FIRTH, A. and VINCENT, T. G.: Development plan for radiata pine breeding. New Zealand Forest Serv., Rotorua, NZ. 142 p. (1986). — SKRØPPA, T.: Breeding strategies with Norway spruce in southeastern Norway. In: Proc. IUFRO Conf. Breeding Strategies Including Multiclonal Varieties. Sensenstein, West Germany (1982). — STRICKBERGER, M. W.: Genetics. The Macmillan Co., London, GB. 868 p. (1968). — TALBERT, J. T.: An advanced-generation breeding plan for the N.C. State University-Industry pine tree improvement cooperative. Silvae Genet. **28**: 72–75 (1979). — VAN BUIJTENEN, J. P.: Mating designs. In: Proc. IUFRO Joint Meeting on Advanced Generation Breeding. June, Bordeaux, France. pp. 11–27 (1976). — VAN BUIJTENEN, J. P. and BRIDGWATER, F.: Mating and genetic test designs. In: Advanced Generation Breeding of Forest Trees. Southern Coop. Series Bull. 309. Louisiana Ag. Exp. Stn., Baton Rouge, LA. pp. 5–10 (1986). — VAN BUIJTENEN, J. P. and LOWE, W. J.: The use of breeding groups in advanced generation breeding. In: Proc. 15th South. For. Tree Improve. Conf. June 19 to 21, Starkeville, Miss. pp. 59–65 (1979). — VAN BUIJTENEN, J. P. and NAMKOONG, G.: Mating designs. In: Progeny Testing of Forest Trees. South. Coop. Series Bull. No. 275. Texas A&M Univ., College Station, TX. pp. 7–13 (1983). — VAN VLECK, L. D., POLLAK, E. J. and OLTENACU, E. A.: Genetics for the Animal Sciences. W. H. Freeman and Co., New York, NY. 391 p. (1987). — WHITE, T. L.: Advanced-generation breeding populations: Size and structure. In: Proc. IUFRO Conf. Breeding Tropical Trees. Oct., Cali, Colombia (1992). — WHITE, T. L. and HODGE, G. R.: Predicting Breeding Values with Applications in Forest Tree Improvement. Kluwer Academic Pub., Dordrecht, The Netherlands. 367 p. (1989). — WHITE, T. L. and HODGE, G. R.: Test designs and optimal age for parental selection in advanced-generation progeny tests of slash pine. Silvae Genet. **41**: 293–302 (1992). — WHITE, T. L., HODGE, G. R. and POWELL, G. L.: Advanced-generation breeding strategy for slash pine in the southeastern United States. Silvae Genet. **42**: 359–371 (1993). — WHITE, T. L., MATHESON, A. C., BOOMSMA, D. B. and ROUT, A. F.: Logistics, costs and genetic gains of five options of nucleus breeding strategies. STBA Tech. Rep. 92–01. Southern Tree Breeding Association, Mt. Gambier, South Australia. 38 p. (1992a). — WHITE, T. L., ROUT, A. F., BOOMSMA, D. B. and DUTKOWSKI, G. W.: Predicted breeding values for 1213 first-generation parents. STBA Tech. Rep. 92–02. Southern Tree Breeding Association, Mt. Gambier, South Australia. 27 p. (1992b).

Genetic Mapping of Quantitative Trait Loci Underlying Complex Genotype-Phenotype Relationships in Forest Trees

By R. WU[1)2)3)] and Y. HAN[2)]

(Received 25th September 1998)

Abstract

The complex relationship between genotype and phenotype can be attributed to individual quantitative trait loci (QTLs). These underling QTLs are often complex in terms of their statistical and biological properties. They could be pleiotropic, linked, developmentally or environmentally plastic, and interacting. Statistical methods have been developed to map QTLs and proven to be successful in detecting QTLs of large effects on the phenotype. Yet, these are not vastly adequate, because the application of these methods is limited by the complex nature of QTLs. In this review, we outline recent developments of statistical methods for mapping these complex QTLs within the framework of composite interval mapping. In each section, we discuss the statistical model for dealing with a key topic, followed by computational algorithms. The topics discussed include mapping pleiotropic QTLs for multiple quantitative traits, QTLs linked on the same chromosome, development-

[1)] Forest Biotechnology Group, Department of Forestry, Box 8008, North Carolina State University, Raleigh, NC 27695-8008, USA

[2)] Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China

[3)] Corresponding to the current address:
RONGLING WU, Program in Statistical Genetics, Department of Statistics, Box 8203, North Carolina State University, Raleigh, NC 27695, USA
Tel: (919)-515-1932
Fax: (919)-515-7315
E-mail: rwu@statgen.ncsu.edu

dependent QTLs during a growth process, environmentally plastic QTLs causing significant genotype x environment interaction, and epistatic QTLs between which gene effects are not linear. Further considerations for QTL mapping are discussed.

Key words: Composite interval mapping, development, epistasis, linkage, pleiotropy, quantitative trait loci (QTLs), QTL x environment interaction.

1. Introduction

The relationship between genotype and phenotype is very complex for a quantitatively-inherited trait. This relationship is often controlled by specific quantitative trait loci (QTLs), which could be different in the magnitude of their effects, inheritance mode, and genomic locations, as well as in the degree to which the expression of these QTLs is affected by environmental and developmental signals. Based on molecular markers, many statistical methods have been developed to map individual QTLs throughout the genome (LANDER and BOTSTEIN, 1989; JANSEN, 1993; JANSEN and STAM, 1994; ZENG, 1994). These methods have been successful in detecting a few QTLs of large effects on the phenotype in a number of organisms (JACOB et al., 1991; PATERSON et al., 1991; STUBER et al., 1992; GROOVER et al., 1994; BRADSHAW and STETTLER, 1995; GRATTAPAGLIA et al., 1995, 1996; CHEVERUD et al., 1996; WU, 1998a). However, for a particular organism, the influences of QTLs on phenotypes are never an isolated event. A particular QTL may not only exert its influence on a single trait, but also be responsible for variation in other traits (pleiotropy; WU et al., 1997); through such pleiotropic effects, the organism can respond or adapt to changing environments in an integrated manner (CHAPIN, 1991). Among the functional loci affecting a trait, some may be stable regardless of environmental alterations, yet the others may be plastic in gene expression, which is the major cause of genotype x environment interactions (FALCONER, 1989; WU, 1998b). The expression of QTLs may also be developmental-dependent. The morphogenesis, growth, development and senescence of an organ can be attributed to the turn-on or turn-off of certain QTLs (e.g., DOEBLEY et al., 1995). The expression of a QTL may be mediated, regulated or suppressed by other QTLs on the same or different chromosomes (epistasis; PHILLIPS, 1998). The significance of gene interactions in quantitative variation has been debated for 80 years.

Attention to detecting QTLs underlying a specific phenomenon has been paid by QTL-methodological geneticists. For example, JIANG and ZENG (1995) developed a multivariate statistical method to map pleiotropic or linked QTLs, which are traditionally viewed as two different mechanisms for trait correlations. Similar methods have been reported by KOROL et al. (1995, 1997). KOROL et al. (1998) proposed statistical methods for mapping QTL x environment interactions. Epigenetic factors are considered to exist during growth and development for a species. Such factors can be identified by a mapping approach developed by YAN et al. (1998) and WU (1999). Despite these significant progresses, no comprehensive treatment of these problems has been reported. Without such a treatment, however, readers may still not be clear about a whole picture of the relationship between genotype and phenotype.

In this article, we review new developments of statistical methods for mapping QTLs that display complex relationships with quantitative traits affected by external or internal environments. In genetics, the concept of "complex" is often used to reflect multifactorial causes of the variation of a quantitative trait (LANDER and KRUGLYAK, 1995). It is contrast to a "simple" MENDELian trait, such as disease resistance, in which variation is controlled by a single gene. However, such a traditional usage of "complex" cannot reveal dynamic properties of an underling gene. From a statistical perspective, those genes responsible for a simple trait, but whose expression changes with development or environment, should be regarded as complex. Statistical methods reviewed in this article are all developed within the framework of composite interval mapping developed by ZENG (1994).

2. Mapping Pleiotropic QTLs

Using composite interval mapping, we can map a QTL that affects v quantitative traits simultaneously. Suppose that there are m markers on a chromosome. For the F_2 progeny of n individuals, the statistical model for mapping multiple traits on a marker interval $\mathbf{M}_i - \mathbf{M}_{i+1}$ is the extension of single-trait mapping and can be expressed in matrix notation:

$$\mathbf{Y} = \mathbf{x}^*\mathbf{a}^* + \mathbf{z}^*\mathbf{d}^* + \mathbf{XB} + \mathbf{E} \qquad (1)$$

where $\mathbf{Y}$ is the (n x v) matrix of trait values $y_{j\xi}$ (j = 1, 2, ..., n, ξ = 1, 2, ..., v), $\mathbf{x}^*$ and $\mathbf{z}^*$ are the (n x 1) column vectors of indicator variables x^*_j and z^*_j describing the QTL genotypes, respectively, whose values are defined as:

$$x^*_j = \begin{cases} 2 & \text{if the QTL genotype is } QQ \\ 1 & \text{if the QTL genotype is } Qq \\ 0 & \text{if the QTL genotype is } pp \end{cases}$$

$$z^*_j = \begin{cases} 1 & \text{if the QTL genotype is } Qq \\ 0 & \text{if the QTL genotype is } QQ \text{ or } qq \end{cases}$$

$\mathbf{a}^*$ and $\mathbf{d}^*$ are the (1 x v) row vectors of additive and dominant effects of the QTL on the v traits, respectively, $\mathbf{X}$ is the [n x (2m–3)] matrix of data on m – 2 markers, x_{jk}'s (k = 1 ,2 , ..., m, $k \neq i$, $k \neq i$+1) fitted with additive and dominant effects in the model for background control, including also the mean effect, $\mathbf{B}$ is the [2(m+1) x v] matrix of additive and dominant effects for the control markers and the mean effect, and $\mathbf{E}$ is the (n x v) matrix of residual errors, $\varepsilon_{j\xi}$. It is assumed that the residual effects are correlated among traits over individuals with covariance Cov($e_{j\xi}$, $e_{j\zeta}$) = $\sigma_{\xi\zeta} = \rho_{\xi\zeta}\sigma_\xi\sigma_\zeta$ ($\rho_{\xi\zeta}$, σ_ξ and σ_ζ are the corresponding correlation and standard deviations), but are independent among individuals. Further, they are assumed to follow a multivariate normal distribution with means zero and covariance matrix:

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix} \qquad (2)$$

Given model (1) and the assumption of multivariate normal distribution of error terms, the likelihood function of the v traits is defined as:

$$L_1 = \prod_{j=1}^{n} \left[p_{2j} f_2(\mathbf{y}_j) + p_{1j} f_1(\mathbf{y}_j) + p_{0j} f_0(\mathbf{y}_j) \right] \qquad (3)$$

where p_{2j}, p_{1j} and p_{0j} denote the prior probabilities of individual j taking values 2, 1, and 0 for the three QTL genotypes, QQ, Qq and qq, respectively, and $f_2(\mathbf{y}_j)$, $f_1(\mathbf{y}_j)$, and $f_0(\mathbf{y}_j)$ represent the multivariate normal density functions of the vector variable $\mathbf{y}_j$ within the three QTL genotypes with $\mathbf{u}_{2j} = \mathbf{x}_j\mathbf{B} + 2\mathbf{a}^*$, $\mathbf{u}_{1j} = \mathbf{x}_j\mathbf{B} + \mathbf{a}^* + \mathbf{d}^*$, and $\mathbf{u}_{0j} = \mathbf{x}_j\mathbf{B}$, respectively, and the covariance matrix (2).

The ML estimates of unknown parameters can be obtained by iteration through the ECM algorithm (MENG and RUBIN, 1993). In the (τ+1)th iteration, the E-step calculates the poster-

ior probabilities of individual j being a particular genotype, *QQ*, *Qq* and *qq*, at the putative QTL as:

$$q_{2j}^{(\tau+1)} = p_{2j}f_2^{(\tau)}(\mathbf{y}_j) / \left[p_{2j}f_2^{(\tau)}(\mathbf{y}_j) + p_{1j}f_1^{(\tau)}(\mathbf{y}_j) + p_{0j}f_0^{(\tau)}(\mathbf{y}_j)\right] \quad (4a)$$

$$q_{1j}^{(\tau+1)} = p_{2j}f_1^{(\tau)}(\mathbf{y}_j) / \left[p_{2j}f_2^{(\tau)}(\mathbf{y}_j) + p_{1j}f_1^{(\tau)}(\mathbf{y}_j) + p_{0j}f_0^{(\tau)}(\mathbf{y}_j)\right] \quad (4b)$$

$$q_{0j}^{(\tau+1)} = p_{2j}f_0^{(\tau)}(\mathbf{y}_j) / \left[p_{2j}f_2^{(\tau)}(\mathbf{y}_j) + p_{1j}f_1^{(\tau)}(\mathbf{y}_j) + p_{0j}f_0^{(\tau)}(\mathbf{y}_j)\right] \quad (4c)$$

where $f_2^{(\tau)}(\mathbf{y}_j)$, $f_1^{(\tau)}(\mathbf{y}_j)$ and $f_0^{(\tau)}(\mathbf{y}_j)$ are the corresponding normal density functions with parameters replaced by estimates in the τth iteration. In the CM-step, parameters in $f_2(\mathbf{y}_j)$, $f_1(\mathbf{y}_j)$ and $f_0(\mathbf{y}_j)$ are divided into three groups ($\mathbf{a}^*$, $\mathbf{d}^*$), $\mathbf{B}$ and $\mathbf{V}$, and estimated consecutively between groups, but simultaneously within each group. These estimators can be shown as:

$$\mathbf{a}^{*(\tau+1)} = \mathbf{q}_2^{(\tau+1)\prime}(\mathbf{Y} - \mathbf{X}\mathbf{B}^{(\tau)}) / (2\,\mathbf{q}_2^{(\tau+1)\prime}\mathbf{1}) \quad (5a)$$

$$\mathbf{d}^{*(\tau+1)} = \left[\mathbf{q}_1^{(\tau+1)\prime}/(\mathbf{q}_1^{(\tau+1)\prime}\mathbf{1}) - \mathbf{q}_2^{(\tau+1)\prime}/(2\mathbf{q}_2^{(\tau+1)\prime}\mathbf{1})\right](\mathbf{Y} - \mathbf{X}\mathbf{B}^{(\tau)}) \quad (5b)$$

$$\mathbf{B}^{(\tau+1)} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\left[\mathbf{Y} - (2\mathbf{q}_2^{(\tau+1)} + \mathbf{q}_1^{(\tau+1)})\mathbf{a}^{*(\tau+1)} - \mathbf{q}_1^{(\tau+1)}\mathbf{d}^{*(\tau+1)}\right] \quad (5c)$$

$$\mathbf{V}^{(\tau+1)} = \frac{1}{n}\Big[(\mathbf{Y} - \mathbf{X}\mathbf{B}^{(\tau+1)})'(\mathbf{Y} - \mathbf{X}\mathbf{B}^{(\tau+1)}) - 4(\mathbf{q}_2^{(\tau+1)\prime}\mathbf{1})\mathbf{a}^{*(\tau+1)}\mathbf{a}^{*(\tau+1)} - (\mathbf{q}_1^{(\tau+1)\prime}\mathbf{1})(\mathbf{a}^{*(\tau+1)} + \mathbf{d}^{*(\tau+1)})(\mathbf{a}^{*(\tau+1)} + \mathbf{d}^{*(\tau+1)})\Big] \quad (5d)$$

where $\mathbf{q}_2^{(\tau+1)}$ and $\mathbf{q}_1^{(\tau+1)}$ are the (n x 1) vectors of $q_{2j}^{(\tau+1)}$ and $q_{1j}^{(\tau+1)}$, and $\mathbf{1}$ is a column vector of ones. A prime represents the transpose of a matrix or a vector.

The calculations begins with $q_{2j}^{(0)} = p_{2j}$, $q_{1j}^{(0)} = p_{1j}$, and $q_{0j}^{(0)} = p_{0j}$, and some starting values, e.g., zero, for $\mathbf{a}^{*(0)}$ and $\mathbf{d}^{*(0)}$. Iterations are then made between (4) to (5), and terminated when a predetermined criterion is satisfied. The criterion for termination is set to be that the changes of the parameter estimates, or the increment of the log-likelihood value, at each iteration become less than ε (a small positive number, say, 10^{-8}). The final estimates are denoted as $\hat{\mathbf{a}}^*$,$\hat{\mathbf{d}}^*$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{V}}$, which will then be used for the calculation of the maximum likelihood value for hypothesis testing.

The log-likelihood of (3) is calculated, with the parameters replaced by the estimates, as

$$\ln(L_1) = K - \tfrac{1}{2}n\ln(|\hat{\mathbf{V}}|) - \tfrac{1}{2}\sum_{j=1}^{n}(\mathbf{y}_j - \mathbf{x}_j\hat{\mathbf{B}})\hat{\mathbf{V}}^{-1}(\mathbf{y}_j - \mathbf{x}_j\hat{\mathbf{B}})' + \sum_{j=1}^{n}\ln\Big\{p_{2j}\exp\left[2\hat{\mathbf{a}}^*\hat{\mathbf{V}}^{-1}(\mathbf{y}_j - \hat{\mathbf{a}}^* - \mathbf{x}_j\hat{\mathbf{B}})'\right] + p_{1j}\exp\left[(\hat{\mathbf{a}}^* + \hat{\mathbf{d}}^*)\hat{\mathbf{V}}^{-1}(\mathbf{y}_j - \tfrac{1}{2}\hat{\mathbf{a}}^* - \tfrac{1}{2}\hat{\mathbf{d}}^* - \mathbf{x}_j\hat{\mathbf{B}})'\right] + p_{0j}\Big\} \quad (6)$$

where $|\hat{\mathbf{V}}|$, is the determinant of the covariance matrix, and $K = -\frac{1}{2}nv\ln(2\pi)$.

The test for the existence of the putative QTL can be performed by testing the hypotheses:

H_0: $\mathbf{a} = \mathbf{O}$ and $\mathbf{d} = \mathbf{O}$ vs. H_1: at least one element $\neq 0$ (7)

The log-likelihood under H_0 is:

$$\ln(L_0) = \ln\left[\prod_{j=1}^{n} f_0(\mathbf{y}_j)\right] = K - \tfrac{n}{2}\ln(|\hat{\mathbf{V}}_0|) - \tfrac{1}{2}nv \quad (8)$$

where $\hat{\mathbf{V}}_0 = \frac{1}{n}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_0)'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_0)$ and $\mathbf{B}_0 = (\mathbf{X'X})^{-1}\mathbf{X'Y}$. The test is performed with a likelihood ratio statistic:

$$LR_1 = -2\ln\left[\frac{L_0}{L_1}\right] \quad (9)$$

Under H_0, LR_1 is approximately chi-square distributed with a degree of freedom $2v$ +1. The critical values for the test can be determined based on the permutation test proposed by CHURCHILL and DOERGE (1994).

The significance of pleiotropic effects on any pair (ξ, ζ) of the v traits can be tested using

H_{10}: $a_\xi = 0$, $d_\xi = 0$, $a_\zeta \neq 0$, $d_\zeta \neq 0$ *vs.* H_{11}: $a_\xi \neq 0$, $d_\xi \neq 0$, $a_\zeta \neq 0$, $d_\zeta \neq 0$ (10a)

H_{20}: $a_\xi \neq 0$, $d_\xi \neq 0$, $a_\zeta = 0$, $d_\zeta = 0$ *vs.* H_{21}: $a_\xi \neq 0$, $d_\xi \neq 0$, $a_\zeta \neq 0$, $d_\zeta \neq 0$ (10b)

The estimates of model parameters under H_{10} and H_{20} can be obtained as in joint mapping of (4) to (5) except that some estimates in (5a) and (5b) are set to zero. The likelihood ratio test statistics under the null hypothesis of (10a) and (10b) will each be asymptotically chi-square distributed with two degrees of freedom. The permutation test of CHURCHILL and DOERGE (1994) can be used to determine the critical values for the test. Only when both of the null hypotheses are rejected, one can suggest the significance of pleiotropic effects.

3. Mapping Linked QTLs

We have considered the situation in which the v traits are assumed to be under the control of a single QTL. However, it is possible that multiple QTLs may be responsible for the correlations among these traits. For simplicity, we only assume two such QTLs that may be in the same or different marker intervals. Under the two-QTL model, Eq. (1) should be changed as:

$$\mathbf{Y} = \mathbf{x}_1^*\mathbf{a}_1^* + \mathbf{z}_1^*\mathbf{d}_1^* + \mathbf{x}_2^*\mathbf{a}_2^* + \mathbf{z}_2^*\mathbf{d}_2^* + \mathbf{X}\mathbf{B} + \mathbf{E} \quad (11)$$

where the subscripts 1 and 2 denote the two different QTLs under consideration. For example, $\mathbf{a}^*_1$ is the (1 x v) vector containing $a^*_{1\xi}$ which is the additive effect of QTL 1 on the ζth trait. Other variables can be similarly defined using the same notations as (1). Yet, $\mathbf{X}$ is the [n x ($2m$+1)] matrix when the two QTLs are assumed in the same interval and [n x ($2m$-1)] matrix when the two QTLs are assumed in different intervals. A similar difference is held for $\mathbf{B}$. For an F_2 population, the two QTLs will lead to nine possible QTL genotypes. Let $p_{t_1t_2j}$ (t_1, t_2 = 0, 1, and 2) be the probability of individual j having genotype t_1 for QTL 1 and t_2 for QTL 2. The likelihood function for the multivariate data assuming two QTLs can be given by:

$$L_2 = \prod_{j=1}^{n} \sum_{t_1=0}^{2} \sum_{t_2=0}^{2} p_{t_1 t_2 j} f_{t_1 t_2}(\mathbf{y}_j) \qquad (12)$$

where $f_{t_1 t_2}(\mathbf{y}_j)$ is a multivariate normal density function for $\mathbf{y}_j$ with a mean vector

$$\mathbf{u}'_{t_1 t_2 j} = \begin{pmatrix} \mathbf{x}_j \mathbf{b}_1 + t_1 a^*_{11} + \eta(t_1) d^*_{11} + t_2 a^*_{21} + \eta(t_2) d^*_{21} \\ \mathbf{x}_j \mathbf{b}_2 + t_1 a^*_{12} + \eta(t_1) d^*_{12} + t_2 a^*_{22} + \eta(t_2) d^*_{22} \\ \vdots \\ \mathbf{x}_j \mathbf{b}_\nu + t_1 a^*_{1\nu} + \eta(t_1) d^*_{1\nu} + t_2 a^*_{2\nu} + \eta(t_2) d^*_{2\nu} \end{pmatrix} \qquad (13)$$

and covariance matrix (2), where the indicator function $\eta(t_1)$ [or $\eta(t_2)$] = 1 if t_1 (or t_2) = 1 and 0 otherwise. Specifically, $a^*_{1\xi}$ and $a^*_{2\xi}$ are the additive effects of QTLs 1 and 2, respectively, on the ξth trait, whereas $d^*_{1\xi}$ and $d^*_{2\xi}$ are the dominant effects of these two QTLs, respectively, on the ξth trait.

The probability $p_{t_1 t_2 j}$ can be inferred from the observed genotypes of the flanking markers. If the two QTLs are tested in different marker intervals, the conditional probability of QTL genotypes can be calculated by assuming independence of the two QTLs, i.e., $p_{t_1 t_2 j} = p_{t_1 j} p_{t_2 j}$ (assuming no crossover interference), where $p_{t_1 j}$ and $p_{t_2 j}$ are the conditional probabilities of the QTL genotypes at a single QTL for individual j (see *Table 1* of Wu et al., 1999). If the two QTLs are tested in the same marker interval, the conditional probability of two-QTL genotypes can be calculated from *table 1.*

The ECM algorithm can be used to solve the unknown parameters. The E-step is to calculate the posterior probabilities of individual j having genotype t_1 for QTL 1 at position $p(1)$ and t_2 for QTL 2 at position $p(2)$:

Table. 1. – Probability of QTL genotype given flanking marker genotype for two QTLs within a marker interval.

Marker genotype	$Q_1Q_1Q_2Q_2$ (22)	$Q_1Q_1Q_2q_2$ (21)	$Q_1Q_1q_2q_2$ (20)	$Q_1q_1Q_2Q_2$ (12)	$Q_1q_1Q_2q_2$ (11)	$Q_1q_1q_2q_2$ (10)	$q_1q_1Q_2Q_2$ (02)	$q_1q_1Q_2q_2$ (01)	$q_1q_1q_2q_2$ (00)
$M_1M_1M_2M_2$	1	0	0	0	0	0	0	0	0
$M_1M_1M_2m_2$	p_3	p_2	0	0	p_1	0	0	0	0
$M_1M_1m_2m_2$	p_3^2	$2p_2p_3$	p_2^2	0	$2p_1p_3$	$2p_1p_2$	0	0	p_1^2
$M_1m_1M_2M_2$	p_1	0	0	p_2	p_3	0	0	0	0
$M_1m_1M_2m_2$	δp_1p_3	δp_1p_2	0	δp_2p_3	$\delta(p_1^2 + p_2^2 + p_3^2) + (1-\delta)$	δp_2p_3	0	δp_1p_2	δp_1p_3
$M_1m_1m_2m_2$	0	0	0	0	p_1	p_2	0	0	p_1
$m_1m_1M_2M_2$	p_1^2	$2p_1p_2$	0	$2p_2p_3$	$2p_1p_3$	0	p_2^2	0	p_3^2
$m_1m_1M_2m_2$	0	0	0	0	p_1	0	0	p_2	p_3
$m_1m_1m_2m_2$	0	0	0	0	0	0	0	0	1

It is assumed that the order of markers and QTL are $M_1Q_1Q_2M_2$ and the recombination fractions between M_1Q_1, Q_1Q_2, Q_2M_1 and M_1M_2 and r_1, r_2, r_3 and r, respectively. Double recombination is ignored. $p_1 = r_1/r$, $p_2 = r_2/r$, $p_3 = r_3/r$, and $\delta = r^2/[(1-r)^2+r^2]$.

$$q^{(\tau+1)}_{t_1 t_2 j} = p_{t_1 t_2 j} \hat{f}^{(\tau)}_{t_1 t_2 j}(\mathbf{y}_j) \Big/ \sum_{t_1=0}^{2} \sum_{t_2=0}^{2} p_{t_1 t_2 j} \hat{f}^{(\tau)}_{t_1 t_2}(\mathbf{y}_j)$$

$$\text{for } t_1, t_2 = 0, 1, 2 \qquad (14)$$

The CM-step is to calculate:

$$a^{*(\tau+1)}_{1\xi} = \qquad (15a)$$

$$\left\{ \mathbf{q}_{2\cdot}^{(\tau+1)\prime} \left[(\mathbf{y}_\xi - \mathbf{X}\mathbf{b}^{(\tau)}_\xi) - (\rho^{(\tau)}_{12} \sigma^{(\tau)}_1 / \sigma^{(\tau)}_2)(\mathbf{y}_1 - \mathbf{X}\mathbf{b}^{(\tau)}_1) - \ldots - (\rho^{(\tau)}_{n(n-1)} \sigma^{(\tau)}_{n-1} / \sigma^{(\tau)}_n)(\mathbf{y}_n - \mathbf{X}\mathbf{b}^{(\tau)}_n) \right] \right.$$

$$+ (\rho^{(\tau)}_{12} \sigma^{(\tau)}_1 / \sigma^{(\tau)}_2) \left[2\mathbf{q}^{(\tau+1)}_{22} + \mathbf{q}^{(\tau+1)}_{21} \right]' \mathbf{1} a^{*(\tau)}_{21} + \mathbf{q}^{(\tau+1)\prime}_{21} \mathbf{1} d^{*(\tau)}_{21} \Big]$$

$$+ \ldots$$

$$\left. + (\rho^{(\tau)}_{(n-1)n} \sigma^{(\tau)}_{n-1} / \sigma^{(\tau)}_n) \left[2\mathbf{q}^{(\tau+1)}_{22} + \mathbf{q}^{(\tau+1)}_{21} \right]' \mathbf{1} a^{*(\tau)}_{2n} + \mathbf{q}^{(\tau+1)\prime}_{21} \mathbf{1} d^{*(\tau)}_{2n} \right] \Big\} / (2\mathbf{q}^{(\tau+1)\prime}_{2\cdot} \mathbf{1})$$

$$d_{1\xi}^{*(\tau+1)} =$$

(15b)

$$\left\{\mathbf{q}_{1\cdot}^{(\tau+1)\prime}\left[(\mathbf{y}_\xi - \mathbf{X}\mathbf{b}_\xi^{(\tau)}) - (\rho_{12}^{(\tau)}\sigma_1^{(\tau)}/\sigma_2^{(\tau)})(\mathbf{y}_1 - \mathbf{X}\mathbf{b}_1^{(\tau)}) - \ldots - (\rho_{n(n-1)}^{(\tau)}\sigma_{n-1}^{(\tau)}/\sigma_n^{(\tau)})(\mathbf{y}_n - \mathbf{X}\mathbf{b}_n^{(\tau)})\right]\right.$$
$$+ (\rho_{12}^{(\tau)}\sigma_1^{(\tau)}/\sigma_2^{(\tau)})\left[2\mathbf{q}_{12}^{(\tau+1)} + \mathbf{q}_{11}^{(\tau+1)})'\mathbf{1}a_{21}^{*(\tau)} + \mathbf{q}_{11}^{(\tau+1)\prime}\mathbf{1}d_{21}^{*(\tau)}\right]$$
$$+ \ldots$$
$$\left. + (\rho_{(n-1)n}^{(\tau)}\sigma_{n-1}^{(\tau)}/\sigma_n^{(\tau)})\left[2\mathbf{q}_{12}^{(\tau+1)} + \mathbf{q}_{11}^{(\tau+1)})'\mathbf{1}a_{2n}^{*(\tau)} + \mathbf{q}_{11}^{(\tau+1)\prime}\mathbf{1}d_{2n}^{*(\tau)}\right]\right\} / (\mathbf{q}_{1\cdot}^{(\tau+1)\prime}\mathbf{1})$$

$$a_{2\xi}^{*(\tau+1)} =$$

(15c)

$$\left\{\mathbf{q}_{\cdot 2}^{(\tau+1)\prime}\left[(\mathbf{y}_\xi - \mathbf{X}\mathbf{b}_\xi^{(\tau)}) - (\rho_{21}^{(\tau)}\sigma_2^{(\tau)}/\sigma_1^{(\tau)})(\mathbf{y}_1 - \mathbf{X}\mathbf{b}_1^{(\tau)}) - \ldots - (\rho_{n(n-1)}^{(\tau)}\sigma_{n-1}^{(\tau)}/\sigma_n^{(\tau)})(\mathbf{y}_n - \mathbf{X}\mathbf{b}_n^{(\tau)})\right]\right.$$
$$+ (\rho_{12}^{(\tau)}\sigma_1^{(\tau)}/\sigma_2^{(\tau)})\left[2\mathbf{q}_{22}^{(\tau+1)} + \mathbf{q}_{12}^{(\tau+1)})'\mathbf{1}a_{11}^{*(\tau)} + \mathbf{q}_{12}^{(\tau+1)\prime}\mathbf{1}d_{11}^{*(\tau)}\right]$$
$$+ \ldots$$
$$\left. + (\rho_{(n-1)n}^{(\tau)}\sigma_{n-1}^{(\tau)}/\sigma_n^{(\tau)})\left[2\mathbf{q}_{22}^{(\tau+1)} + \mathbf{q}_{12}^{(\tau+1)})'\mathbf{1}a_{1n}^{*(\tau)} + \mathbf{q}_{12}^{(\tau+1)\prime}\mathbf{1}d_{1n}^{*(\tau)}\right]\right\} / (2\mathbf{q}_{\cdot 2}^{(\tau+1)\prime}\mathbf{1})$$

$$d_{2\xi}^{*(\tau+1)} =$$

(15d)

$$\left\{\mathbf{q}_{\cdot 1}^{(\tau+1)\prime}\left[(\mathbf{y}_\xi - \mathbf{X}\mathbf{b}_\xi^{(\tau)}) - (\rho_{12}^{(\tau)}\sigma_2^{(\tau)}/\sigma_1^{(\tau)})(\mathbf{y}_1 - \mathbf{X}\mathbf{b}_1^{(\tau)}) - \ldots - (\rho_{n(n-1)}^{(\tau)}\sigma_{n-1}^{(\tau)}/\sigma_n^{(\tau)})(\mathbf{y}_n - \mathbf{X}\mathbf{b}_n^{(\tau)})\right]\right.$$
$$+ (\rho_{12}^{(\tau)}\sigma_1^{(\tau)}/\sigma_2^{(\tau)})\left[2\mathbf{q}_{12}^{(\tau+1)} + \mathbf{q}_{11}^{(\tau+1)})'\mathbf{1}a_{11}^{*(\tau)} + \mathbf{q}_{11}^{(\tau+1)\prime}\mathbf{1}d_{11}^{*(\tau)}\right]$$
$$+ \ldots$$
$$\left. + (\rho_{(n-1)n}^{(\tau)}\sigma_{n-1}^{(\tau)}/\sigma_n^{(\tau)})\left[2\mathbf{q}_{22}^{(\tau+1)} + \mathbf{q}_{21}^{(\tau+1)})'\mathbf{1}a_{1n}^{*(\tau)} + \mathbf{q}_{21}^{(\tau+1)\prime}\mathbf{1}d_{1n}^{*(\tau)}\right]\right\} / (2\mathbf{q}_{1\cdot}^{(\tau+1)\prime}\mathbf{1})$$

$$\mathbf{B}^{(\tau+1)} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W})^{(\tau+1)} \qquad (15e)$$

$$\mathbf{V}^{(\tau+1)} = \tfrac{1}{n}(\mathbf{W}^{(\tau+1)} - \mathbf{X}\mathbf{B}^{(\tau+1)})'(\mathbf{W}^{(\tau+1)} - \mathbf{X}\mathbf{B}^{(\tau+1)}) \qquad (15f)$$

where

$$\mathbf{B}^{(\tau+1)} = (\mathbf{b}_1^{(\tau+1)} \;\; \mathbf{b}_2^{(\tau+1)} \;\; \ldots \;\; \mathbf{b}_v^{(\tau+1)})$$

Formel A

$$\mathbf{W}^{(\tau+1)\prime} = \begin{pmatrix} \mathbf{y}_1' - (2\mathbf{q}_{2\cdot}^{(\tau+1)} + \mathbf{q}_{1\cdot}^{(\tau+1)})'a_{11}^{*(\tau+1)} - \mathbf{q}_{1\cdot}^{(\tau+1)\prime}d_{11}^{*(\tau+1)} - (2\mathbf{q}_{\cdot 2}^{(\tau+1)} + \mathbf{q}_{\cdot 1}^{(\tau+1)})'a_{21}^{*(\tau+1)} - \mathbf{q}_{\cdot 1}^{(\tau+1)\prime}d_{21}^{*(\tau+1)} \\ \mathbf{y}_2' - (2\mathbf{q}_{2\cdot}^{(\tau+1)} + \mathbf{q}_{1\cdot}^{(\tau+1)})'a_{12}^{*(\tau+1)} - \mathbf{q}_{1\cdot}^{(\tau+1)\prime}d_{12}^{*(\tau+1)} - (2\mathbf{q}_{\cdot 2}^{(\tau+1)} + \mathbf{q}_{\cdot 1}^{(\tau+1)})'a_{22}^{*(\tau+1)} - \mathbf{q}_{\cdot 1}^{(\tau+1)\prime}d_{22}^{*(\tau+1)} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \mathbf{y}_v' - (2\mathbf{q}_{2\cdot}^{(\tau+1)} + \mathbf{q}_{1\cdot}^{(\tau+1)})'a_{1v}^{*(\tau+1)} - \mathbf{q}_{1\cdot}^{(\tau+1)\prime}d_{1v}^{*(\tau+1)} - (2\mathbf{q}_{\cdot 2}^{(\tau+1)} + \mathbf{q}_{\cdot 1}^{(\tau+1)})'a_{2v}^{*(\tau+1)} - \mathbf{q}_{\cdot 1}^{(\tau+1)\prime}d_{2v}^{*(\tau+1)} \end{pmatrix}$$

$$\mathbf{q}_{2\cdot}^{(\tau+1)} = \mathbf{q}_{22}^{(\tau+1)} + \mathbf{q}_{21}^{(\tau+1)} + \mathbf{q}_{20}^{(\tau+1)}$$
$$\mathbf{q}_{1\cdot}^{(\tau+1)} = \mathbf{q}_{12}^{(\tau+1)} + \mathbf{q}_{11}^{(\tau+1)} + \mathbf{q}_{10}^{(\tau+1)}$$
$$\mathbf{q}_{\cdot 2}^{(\tau+1)} = \mathbf{q}_{22}^{(\tau+1)} + \mathbf{q}_{12}^{(\tau+1)} + \mathbf{q}_{02}^{(\tau+1)}$$
$$\mathbf{q}_{\cdot 1}^{(\tau+1)} = \mathbf{q}_{21}^{(\tau+1)} + \mathbf{q}_{11}^{(\tau+1)} + \mathbf{q}_{01}^{(\tau+1)}$$

After the convergence of the ECM algorithm, log-likelihood value of (12) can be calculated on which several hypotheses can be tested. Three of the most important hypotheses include testing (1) if there are two different QTLs that affect the v traits at the same time, (2) if there two QTLs, if any, are linked, and (3) if these two linked QTLs each have a pleiotropic effect on the v traits. The test for the first hypothesis can be formulated by:

$$H_0: p(1) = p(2) \text{ vs. } H_1: p(1) \neq p(2) \qquad (16)$$

where $p(1)$ and $p(2)$ describe the positions of these two QTLs, respectively. If L_{20} is the maximum of the likelihood values which correspond to the H_0: $p(1) = p(2)$, then, under the null hypothesis, the test statistic

$$LR_2 = -2\ln\left[\frac{L_{20}}{L_2}\right] \qquad (17)$$

will be asymptotically chi-square distributed with 1 degree of freedom. The critical values for the test can be determined by the permutation test (CHURCHILL and DOERGE, 1994). If an analysis is based on the same interval or different chromosomes, the second hypothesis is not necessarily tested. However, if the two different intervals of interest are on the same chromosome, this hypothesis should be tested. The likelihood value for its null hypothesis is calculated by replacing the ML estimate for the position of the first QTL and letting the position of the second QTL away from the first one with the recombination fraction of 0.5. The third hypothesis can be formulated as (10a) and (10b).

4. Mapping Epigenetic QTLs

If the expression of a trait at different ages is regarded as a different "trait", composite interval mapping can be used to analyze the molecular genetic basis of trait development over age. In a developmental genetic study, the same genotypes with molecular information are measured phenotypically at different ages and, thus, the principle to map multi-trait correlations can be directly used. Our main interests in tree developmental genetics include: (1) are there some QTLs that are responsible for all developmental stages from seed germination to tree mature or for a particular developmental period? (2) between what ages do QTLs change their function so abruptly that a distinct transition occurs? and (3) is the QTL x age interaction effect significant between a certain two ages? The first question is addressed by combining all measurements at different ages into a multivariate dataset. The multi-trait QTL analysis method is used to test whether there exist one or two QTLs that can significantly account for the phenotypic co-variation in this multivariate system. If evidence shows the effects of two QTLs, a hypothesis about their linkage should be tested (see above). To answer the second question, one should construct and analyze a bivariate system composed of measurements taken at each pair of two successive ages. Thus, the timing of maximum genetic differentiation is suggested as that at which a QTL takes a significant change in its expression for trait development. Assuming a QTL in a marker interval, the third question can be solved by testing the following hypotheses

$$H_0: \ a_1^* = a_2^*, d_1^* = d_2^* \text{ vs. } H_1: a_1^* \neq a_2^*, d_1^* \neq d_2^* \qquad (18)$$

where the subscripts denote two different ages. If H_0 is rejected, then it means that there is a significant additive and dominant QTL x age interaction. This test should be performed after a significant QTL is suggested. For all these three questions, the relative importance of pleiotropic vs. linked QTLs in trait development can be estimated by performing the corresponding hypothesis test similarly to the procedures described for the multi-trait analysis.

One of the most important issues in developmental genetics is to detect new genes during a particular developmental period. Different from those which function during previous periods, these new genes are activated by some developmental (internal) or environmental (external) signals and further result in the formation of new morphological features or physiological processes. For example, a hormone or mitogen may activate certain genes to induce a population of cells to undergo differentiation or proliferation (VOGL et al., 1993). Supplied with nutrients, a growing plant may turn off genes for lateral roots to invest more energy on above-ground growth (ZHANG and FORDE, 1998). The influence of an inducive signal on final structures is defined as epigenetic effects (ATCHLEY and HALL, 1991; COWLEY and ATCHLEY, 1992; ATCHLEY et al., 1994). The genes that control morphological changes through epigenetic effects are defined as "epigenetic genes". WU (1999) developed a statistical model to map epigenetic QTLs using molecular markers. WU's basic idea is to combine a conditional quantitative genetic model and composite interval mapping. The conditional quantitative genetics model can be used to eliminate the confounding effect of early development on morphology or structure measured at subsequent times. The QTLs identified based on phenotypic means at time 2 conditional on phenotypic means measured at time 1 represent epigenetic QTLs.

5. Mapping Plasticity QTLs

The QTLs whose effects display plastic changes across various environments are defined as plasticity genes (WU, 1998b). Differential expression of such plasticity QTLs is the molecular genetic basis for genotype x environment interaction, a phenomenon commonly observed in tree breeding (WU and STETTLER, 1997). In several molecular experiments, efforts have been made to identify plasticity QTLs for a few organisms, such as tomatoes (PATERSON et al., 1991) and maize (STUBER et al., 1992). However, QTL analyses in these experiments were conducted separately for individual environments and, thus, only a few plasticity QTLs could be detected, in a sharp contrast to considerable genotype x environment interactions. From a statistical perspective, joint mapping for multiple environments would have more power to detect plasticity QTLs. QTL mapping for genotype x environment interactions can be performed by viewing different performance of the same trait in different environments as a distinct trait or trait state, a concept originally introduced by FALCONER (1952). If trees can be cloned, then the same genotypes recorded on markers can be measured in various environments. In this case, the method will be the same as the multi-trait genetic mapping. To test QTL x environment interaction for a putative QTL, we just formulate and test the hypotheses like (7), (10a) and (10b).

If the null hypotheses are rejected, this indicates that the QTL exists and is also environment-dependent, contributing significantly to genotype x environment interaction. Similar tests can be also formulated under a two-QTL model (17).

In practice, cloning may be expensive or impossible, especially in pines. When genotypes cannot be repeated in different environments, one should construct a genetic map for each environment using the same markers. Because the sample for

each environment independently represents the same overall population, there should be a great similarity in arrangement orders for most markers among the maps. However, errors due to genetic drifts cannot assure marker orders and marker distances to be exactly identical for all maps. Assume a marker interval bracketed by the same markers, $\mathbf{M}_i - \mathbf{M}_{i+1}$, but with possibly different genetic distances among the maps. The corresponding statistical model incorporating a pleiotropic QTL within this interval is given in matrix notation as:

$$\mathbf{y}_1 = a_1^* \mathbf{x}_1^* + d_1^* \mathbf{z}_1^* + \mathbf{b}_1 \mathbf{X}_1 + \mathbf{e}_1$$

$$\mathbf{y}_2 = a_2^* \mathbf{x}_2^* + d_2^* \mathbf{z}_2^* + \mathbf{b}_2 \mathbf{X}_2 + \mathbf{e}_2$$

(19)

...

$$\mathbf{y}_v = a_v^* \mathbf{x}_v^* + d_v^* \mathbf{z}_v^* + \mathbf{b}_v \mathbf{X}_v + \mathbf{e}_v$$

It is assumed that $\mathbf{e}_\xi (\xi = 1, 2, ..., v)$ is independently normally distributed with means zero and variance σ_ξ^2. Thus, parameters in each environment can be estimated separately. The likelihood function across all environments is the product of the likelihoods in each environment, i.e.,

$$L_3 = \prod_{j=1}^{n_1} \left[\sum_{t=0}^{2} p_{1tj} f_t(y_{1j}) \right] \prod_{j=1}^{n_2} \left[\sum_{t=0}^{2} p_{2tj} f_t(y_{2j}) \right] \cdots \prod_{j=1}^{n_v} \left[\sum_{t=0}^{2} p_{vtj} f_t(y_{vj}) \right]$$

(20)

where n_ξ is the sample size of the F_2 progeny in the ξth environment, and $p_{\xi tj}$ is the prior probability of individual j taking $x^*_\xi = t$ in the ξth environment, and $f_t(y_{\xi j})$ is the density function of the phenotype of individual j with QTL genotype t in the ξth environment. Differences in the effects and/or positions of the QTL among environments reflect the strength of QTL x environment interaction on a quantitative trait, which can be tested by constructing the related hypotheses. Only after the hypotheses in both QTL positions and effects are tested, QTL x environment interactions can be determined.

The parameters can be estimated jointly via the ECM algorithm. In each ECM iteration, the E-step constitutes:

$$q_{\xi tj}^{(\tau+1)} = p_{\xi tj} f_t^{(\tau)}(y_{\xi j}) / \sum_{t=0}^{2} \left[p_{\xi tj} f_t^{(\tau)}(y_{\xi j}) \right]$$

(21)

for the ξth environment, the tth genotype and the jth individual. The CM-step constitutes:

$$a_\xi^{*(\tau+1)} = \sum_{\xi=1}^{v} \left[\mathbf{q}_{\xi 2}^{(\tau+1)\prime} (\mathbf{y}_\xi - \mathbf{X}_\xi \mathbf{b}_\xi^{(\tau)}) / \sigma_\xi^{2(\tau)} \right] / \sum_{\xi=1}^{v} \left[\mathbf{q}_{\xi 2}^{(\tau+1)\prime} \mathbf{1} / \sigma_\xi^{2(\tau)} \right]$$

(22a)

$$d_\xi^{*(\tau+1)} = \sum_{\xi=1}^{v} \left[\mathbf{q}_{\xi 1}^{(\tau+1)\prime} (\mathbf{y}_\xi - \mathbf{X}_\xi \mathbf{b}_\xi^{(\tau)}) / \sigma_\xi^{2(\tau)} \right] / \sum_{\xi=1}^{v} \left[\mathbf{q}_{\xi 1}^{(\tau+1)\prime} \mathbf{1} / \sigma_\xi^{2(\tau)} \right] - a_\xi^{*(\tau+1)}$$

(22b)

$$\mathbf{b}_\xi^{(\tau+1)} = (\mathbf{X}_\xi' \mathbf{X}_\xi)^{-1} \mathbf{X}_\xi' \left[\mathbf{y}_\xi - (2\mathbf{q}_{\xi 2}^{(\tau+1)} + \mathbf{q}_{\xi 1}^{(\tau+1)}) a_\xi^{*(\tau+1)} - \mathbf{q}_{\xi 1}^{(\tau+1)} d_\xi^{*(\tau+1)} \right]$$

(22c)

$$\sigma_\xi^{2(\tau+1)} = \frac{1}{n_\xi} \left[(\mathbf{y}_\xi - \mathbf{X}_\xi \mathbf{b}_\xi^{(\tau+1)})' (\mathbf{y}_\xi - \mathbf{X}_\xi \mathbf{b}_\xi^{(\tau+1)}) - 4\mathbf{q}_{\xi 2}^{(\tau+1)\prime} \mathbf{1} a_\xi^{*2(\tau+1)} - \mathbf{q}_{\xi 1}^{(\tau+1)\prime} \mathbf{1} (a_\xi^{*2(\tau+1)} + d_\xi^{*2(\tau+1)})^2 \right]$$

(22d)

The test for QTL x environment interaction includes two steps: (1) testing the position and (2) effect of the QTL in each environment. The hypotheses for these two steps are formulated, respectively, by:

H_{10}: $p(1) = p(2) = ... = p(v)$ vs. H_{11}: ≥ 2 QTL positions are not equal (23a)

H_{20}: $\hat{a}_1^* = \hat{a}_2^* = ... = \hat{a}_v^*$ and $\hat{d}_1 = \hat{d}_2 = ... \hat{d}_v$ vs. H_{20}: ≥ 2 variables are not equal (23b)

The log-likelihood values, $\ln(L_{31})$ and $\ln(L_{32})$, under the H_0's of (23) will be similar to (20) in form but with the corresponding constraints. The likelihood ratio test statistics for (23a) and (23b) are given, respectively, by:

$$LR_{31} = -2\ln\left[\frac{L_{31}}{L_3}\right] \qquad (24a)$$

$$LR_{32} = -2\ln\left[\frac{L_{32}}{L_3}\right] \qquad (24b)$$

both of which are asymptotically chi-square distributed under the null hypothesis, with one and two degrees of freedom, respectively. The critical values for these tests can be determined based on the permutation test proposed by CHURCHILL and DOERGE (1994).

If the H_{10} in step 1 (23a) is rejected, this indicates the existence of QTL x environment interaction, irrespective of the testing result for step 2 (23b). Only after the H_{10} of (23a) is accepted, it is necessary to conduct the test for step 2.

6. Mapping Epistatic QTLs

Consider two epistatic QTLs located in the same or different marker intervals, $\mathbf{M}_{i_1} - \mathbf{M}_{i_1+1}$ and $\mathbf{M}_{i_2} - \mathbf{M}_{i_2+2}$. Under the epistasis model, the phenotypic value of a quantitative trait for the jth individual in the F_2 progeny:

$$y_j = \mu + a_1 x_{1j}^* + d_1 z_{1j}^* + a_2 x_{2j}^* + d_2 z_{2j}^* + e_{aa} \omega_{aaj}^* + e_{ad} \omega_{adj}^* + e_{da} \omega_{daj}^* + e_{dd} \omega_{ddj}^* + \sum_{k \neq i_1, i_1+1}^{m} b_k x_{kj} + \varepsilon_j$$

(25)

where a_1 and d_1 are the additive and dominant effect of QTL 1, a_2 and d_2 are the additive and dominant effects of QTL 2, e_{aa}, e_{ad}, e_{da}, and e_{dd} are the additive x additive, additive x dominant, dominant x additive, and dominant x dominant epistatic effects between the two QTLs, and all variables with asterisks are the indicator variables and they are defined as:

$$x^*_{1j} = \begin{cases} 1 & \text{if the first QTL is } Q_1Q_1 \\ 0 & \text{if the first QTL is } Q_1q_1 \\ -1 & \text{if the first QTL is } q_1q_1 \end{cases} \qquad x^*_{2j} = \begin{cases} 1 & \text{if the second QTL is } Q_2Q_2 \\ 0 & \text{if the second QTL is } Q_2q_2 \\ -1 & \text{if the second QTL is } q_2q_2 \end{cases}$$

$$z^*_{1j} = \begin{cases} 1/2 & \text{if the first QTL is } Q_1Q_1 \\ -1/2 & \text{otherwise} \end{cases} \qquad z^*_{2j} = \begin{cases} 1/2 & \text{if the second QTL is } Q_2Q_2 \\ -1/2 & \text{otherwise} \end{cases}$$

$$w^*_{aaj} = \begin{cases} 1 & \text{if the QTLs are } Q_1Q_1Q_2Q_2 \text{ or } q_1q_1q_2q_2 \\ -1 & \text{if the QTL are } Q_1Q_1q_2q_2 \text{ or } q_1q_1Q_2Q_2 \\ 0 & \text{otherwise} \end{cases}$$

$$w^*_{adj} = \begin{cases} 1/2 & \text{if the QTLs are } Q_1Q_1Q_2q_2,\ q_1q_1Q_2Q_2 \text{ or } q_1q_1q_2q_2 \\ -1/2 & \text{if the QTLs are } q_1q_1Q_2q_2 \text{ or } Q_1Q_1Q_2Q_2 \text{ or } Q_1Q_1q_2q_2 \\ 0 & \text{otherwise} \end{cases}$$

$$w^*_{adj} = \begin{cases} 1/2 & \text{if the QTLs are } Q_1q_1Q_2Q_2,\ Q_1Q_1q_2q_2 \text{ or } q_1q_1q_2q_2 \\ -1/2 & \text{if the QTLs are } Q_1q_1q_2q_2 \text{ or } Q_1Q_1Q_2Q_2 \text{ or } q_1q_1Q_2Q_2 \\ 0 & \text{otherwise} \end{cases}$$

$$w^*_{adj} = \begin{cases} 1/4 & \text{if the QTLs are } Q_1q_1Q_2q_2,\ Q_1Q_1Q_2Q_2,\ Q_1Q_1q_2q_2,\ q_1q_1Q_2Q_2 \text{ or } q_1q_1q_2q_2 \\ -1/4 & \text{if the QTLs are } Q_1q_1Q_2Q_2 \text{ or } Q_1q_1q_2q_2,\ Q_1Q_1Q_2q_2 \text{ or } q_1q_1Q_2q_2 \end{cases}$$

The nine unknown genotypes at the two putative QTLs can be inferred from two pairs of flanking markers, with a total of 81 different marker genotypes. If the two QTLs are in two different intervals, the joint probabilities of genotypes at the two QTLs conditional on the two intervals is the products of the probabilities of QTL genotypes conditional on individual intervals, assuming no crossover.

The likelihood function of the data under the two-locus interaction epistasis model is written as:

$$L_4 = \prod_{j=1}^{n}\left[\sum_{t_1=0}^{2}\sum_{t_2=0}^{2} p_{t_1t_2j} f_{t_1t_2}(y_j)\right] \qquad (26)$$

where $f_{t_1t_2}(y_j)$ is the density function of the phenotype of individual j with QTL genotype t_1t_2 at the two QTLs:

Formel B

$$f_{22}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j - a_1 + \frac{1}{2}d_1 - a_2 + \frac{1}{2}d_2 - e_{aa} + \frac{1}{2}e_{ad} + \frac{1}{2}e_{da} - \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{12}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j - \frac{1}{2}d_1 - a_2 + \frac{1}{2}d_2 - \frac{1}{2}e_{da} + \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{02}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j + a_1 + \frac{1}{2}d_1 - a_2 + \frac{1}{2}d_2 + e_{aa} - \frac{1}{2}e_{ad} + \frac{1}{2}e_{da} - \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{21}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\left[-\frac{(y_j - a_1 + \frac{1}{2}d_1 - d_2 - \frac{1}{2}e_{ad} + \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{11}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j - \frac{1}{2}d_1 - \frac{1}{2}d_2 - \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{01}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j + a_1 + \frac{1}{2}d_1 - \frac{1}{2}d_2 + e_{aa} + \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{20}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j - a_1 + \frac{1}{2}d_1 + a_2 + \frac{1}{2}d_2 + e_{aa} + \frac{1}{2}e_{ad} - \frac{1}{2}e_{da} - \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{10}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j - \frac{1}{2}d_1 + a_2 + \frac{1}{2}d_2 + \frac{1}{2}e_{da} + \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

$$f_{00}(y_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(y_j + a_1 + \frac{1}{2}d_1 + a_2 + \frac{1}{2}d_2 - e_{aa} - \frac{1}{2}e_{ad} - \frac{1}{2}e_{da} - \frac{1}{4}e_{dd} - \mathbf{X}_j\mathbf{B})^2}{2\sigma^2}\right]$$

The maximum likelihood estimates of the unknown parameters can be obtained via the ECM algorithm (KAO, 1995). In each ECM iteration, the E-step is constituted by the posterior probability of individual j with different QTL genotypes at the epistatic QTLs, with the same form as (14). The M-step constitutes:

$$a_1^{(\tau+1)} = 1/\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} + q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)})\times$$

(27a)

$$\left[\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})(\mathbf{y}_j - \mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+\tfrac{1}{2}d_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+a_2^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)}) + \tfrac{1}{2}d_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} + q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+e_{aa}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$\left.+\tfrac{1}{2}e_{da}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)})\right]$$

$$d_1^{(\tau+1)} = 2/n\times$$

(27b)

$$\left[\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} + q_{11j}^{(\tau+1)} - q_{01j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} - q_{00j}^{(\tau+1)})(\mathbf{y}_j - \mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+\ a_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+\ a_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+\ \tfrac{1}{2}d_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} + q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+\ e_{aa}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+\ \tfrac{1}{2}e_{da}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$\left.+\ \tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} - q_{11j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})\right]$$

$$a_2^{(\tau+1)} = 1/\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{21j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})\times$$

(27c)

$$\left[\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} - q_{00j}^{(\tau+1)})(\mathbf{y}_j - \mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+a_1^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)}) + \tfrac{1}{2}d_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+\tfrac{1}{2}d_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$+e_{aa}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{2}e_{da}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$\left.+\tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})\right]$$

$$d_2^{(\tau+1)} = = 2/\mathrm{n}\times \tag{27d}$$

$$\left[\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}-q_{12j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}+q_{11j}^{(\tau+1)}+q_{01j}^{(\tau+1)}-q_{20j}^{(\tau+1)}-q_{10j}^{(\tau+1)}-q_{00j}^{(\tau+1)})(\mathbf{y}_j-\mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+\ a_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}-q_{21j}^{(\tau+1)}+q_{01j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})$$

$$+\ a_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{12j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}+q_{10j}^{(\tau+1)}-q_{00j}^{(\tau+1)})$$

$$+\ \tfrac{1}{2}d_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})$$

$$+\ e_{aa}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})+\tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})$$

$$+\ \tfrac{1}{2}e_{da}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{12j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{10j}^{(\tau+1)}+q_{00j}^{(\tau+1)})$$

$$\left.+\ \tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{12j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}-q_{11j}^{(\tau+1)}+q_{01j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{10j}^{(\tau+1)}+q_{00j}^{(\tau+1)})\right]$$

$$e_{aa}^{(\tau+1)}=1/\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})\times\left[\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})(\mathbf{y}_j-\mathbf{X}_j\mathbf{b}^{(\tau)})\right. \tag{27e}$$

$$+a_1^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})+\tfrac{1}{2}d_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})$$

$$+a_2^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})+\tfrac{1}{2}d_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})$$

$$+\tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})+\tfrac{1}{2}e_{da}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})$$

$$\left.+\tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})\right]$$

$$e_{ad}^{(\tau+1)} = 1/\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}+q_{01j}^{(\tau+1)}+q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})\times \tag{27f}$$

$$\left[\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})(\mathbf{y}_j-\mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+\ a_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}-q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}+q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})$$

$$+\ \tfrac{1}{2}d_1^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}-q_{20j}^{(\tau+1)}+q_{00j}^{(\tau+1)})$$

$$+\ a_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}+q_{00j}^{(\tau+1)})+\tfrac{1}{2}d_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}+q_{12j}^{(\tau+1)}+q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}-q_{10j}^{(\tau+1)}-q_{00j}^{(\tau+1)})$$

$$+\ e_{aa}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}-q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})+\tfrac{1}{2}e_{da}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)}+q_{02j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})$$

$$\left.+\ \tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)}-q_{02j}^{(\tau+1)}+q_{21j}^{(\tau+1)}-q_{01j}^{(\tau+1)}+q_{20j}^{(\tau+1)}-q_{00j}^{(\tau+1)})\right]$$

$$e_{da}^{(\tau+1)} = 1/\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})\times \tag{27g}$$

$$\left[\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})(\mathbf{y}_j - \mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+ a_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{2}d_1^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ a_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ \tfrac{1}{2}d_2^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ e_{aa}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)}) + \tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$\left. + \tfrac{1}{4}e_{dd}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} - q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})\right]$$

$$e_{dd}^{(\tau+1)} = 4/n\times \tag{27h}$$

$$\left[\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} + q_{11j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})(\mathbf{y}_j - \mathbf{X}_j\mathbf{b}^{(\tau)})\right.$$

$$+ a_1^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} + q_{01j}^{(\tau+1)} - q_{20j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ d_1^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{21j}^{(\tau+1)} - q_{11j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} + q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ a_2^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ d_2^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{11j}^{(\tau+1)} + q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} + q_{00j}^{(\tau+1)})$$

$$+ e_{aa}^{(\tau)}\sum_{j=1}^{n}(-q_{22j}^{(\tau+1)} + q_{02j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)}) + \tfrac{1}{2}e_{ad}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} - q_{02j}^{(\tau+1)} + q_{21j}^{(\tau+1)} - q_{01j}^{(\tau+1)} + q_{20j}^{(\tau+1)} - q_{00j}^{(\tau+1)})$$

$$\left. + \tfrac{1}{4}e_{da}^{(\tau)}\sum_{j=1}^{n}(q_{22j}^{(\tau+1)} + q_{12j}^{(\tau+1)} + q_{02j}^{(\tau+1)} - q_{20j}^{(\tau+1)} - q_{10j}^{(\tau+1)} - q_{00j}^{(\tau+1)})\right]$$

$$\mathbf{b}^{(\tau+1)} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{Y} + \mathbf{q}\mathbf{D}\beta^{(\tau)}] \tag{27i}$$

$$\sigma^{2(\tau+1)} = \tfrac{1}{2}[(\mathbf{Y} - \mathbf{X}\mathbf{b}^{(\tau+1)})'(\mathbf{Y} - \mathbf{X}\mathbf{b}^{(\tau+1)}) + 2(\mathbf{Y} - \mathbf{X}\mathbf{b}^{(\tau+1)})'\mathbf{q}\mathbf{D}\beta^{(\tau+1)} + \beta^{(\tau+1)}\mathbf{V}\beta^{(\tau+1)}] \tag{27j}$$

where

$$\beta' = [a_1 \ \tfrac{1}{2}d_1 \ a_2 \ \tfrac{1}{2}d_2 \ e_{aa} \ \tfrac{1}{2}e_{ad} \ \tfrac{1}{2}e_{da} \ \tfrac{1}{2}e_{dd}]$$

Formel C

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 0 & -1 & -1 & 1 & 0 & 0 & -1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & -1 \\ 1 & 1 & 0 & -1 & 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 0 & -1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

$$= [\mathbf{D}_1 \ \mathbf{D}_2 \ \mathbf{D}_3 \ \mathbf{D}_4 \ \mathbf{D}_5 \ \mathbf{D}_6 \ \mathbf{D}_7 \ \mathbf{D}_8]$$

$$\mathbf{q} = \begin{bmatrix} q_{221} & q_{121} & q_{021} & q_{211} & q_{111} & q_{011} & q_{201} & q_{101} & q_{001} \\ q_{222} & q_{122} & q_{022} & q_{212} & q_{112} & q_{012} & q_{202} & q_{102} & q_{002} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ q_{22n} & q_{12n} & q_{02n} & q_{21n} & q_{11n} & q_{01n} & q_{20n} & q_{10n} & q_{00n} \end{bmatrix}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{1}'\mathbf{q}\mathbf{D}_1^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_1\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_2^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_2\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_3^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_3\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_4^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_4\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_5^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_5\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_6^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_6\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_7^2 & \mathbf{1}'\mathbf{q}\mathbf{D}_7\mathbf{D}_8 \\ \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_1 & \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_2 & \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_3 & \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_4 & \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_5 & \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_6 & \mathbf{1}'\mathbf{q}\mathbf{D}_8\mathbf{D}_7 & \mathbf{1}'\mathbf{q}\mathbf{D}_8^2 \end{bmatrix}$$

The putative epistatic QTLs can be analyzed using two separate steps. The first step is to test for the existence of the interacting QTLs using a two-dimensional search; the second step is to test for existence of epistasis between the two QTLs. Different hypothesis tests are used between these two steps. The hypotheses to be tested for the first step are

$$H_0: a_1 = d_1 = a_2 = d_2 = e_{aa} = e_{ad} = e_{da} = e_{\mathrm{dd}} = 0 \ \textit{vs.} \text{ at least one variable} \neq 0 \qquad (28)$$

By calculating the log-likelihood ratio under these two hypotheses, the existence of the two QTLs can be tested for each combination of positions in the two intervals (two-dimensional search). The coordinate with the highest ratio can be viewed as the likely positions of the two QTLs. The hypothesis for the second step is set as:

$$H_0: e_{aa} = e_{ad} = e_{da} = e_{dd} = 0 \ \textit{vs.} \text{ at least one variable} \neq 0 \qquad (29)$$

The log-likelihood ratio for testing the significance of epistasis can be similarly calculated, yet the two hypotheses of (29) should be tested only at the two estimated positions from the first step.

If the two QTLs are tested in the same marker interval, the conditional probability of two-QTL genotypes given this interval can be obtained from *table 1.*

7. Discussion

Given the complexity of the genetic architecture of a quantitative trait, sophisticated statistical methods should be developed to detect the underling QTLs. To develop a robust statistical method, however, one should first understand the genetic causes of complex genetic architecture. Of all these causes, some may play a more important role in determining the phenotype of a quantitative trait than others. Based on the existing knowledge about genetic variation, we suggest that pleiotropy, linkage, QTL x environment interaction and epistasis may be among the most important causes. In this article, we review recent developments of statistical methods for mapping QTLs in relation to these aspects. We present a statistical model behind each aspect, followed by computational algorithms.

Pleiotropy and linkage among QTLs are considered two major genetic mechanisms for genetic correlations of quantitative traits (FALCONER, 1989). Within the framework of a multivariate analysis, statistical models have been proposed to detect their effects on trait correlations (JIANG and ZENG, 1995;

KOROL et al., 1995, 1998; RONIN et al., 1995). JIANG and ZENG' (1995) simulation experiments indicated that the statistical power to detect QTLs would be increased when multiple traits are mapped simultaneously than when they are mapped separately. Thus, in spite of an increased number of parameters to be estimated, compared to the single-trait formulation, the multi-trait mapping allows for an improvement of detection power and estimation accuracy of linked QTLs.

The statistical methods described to estimate pleiotropic and linked QTLs in this paper are developed within the framework of ZENG's (1994) composite interval mapping. However, increased complexity in model manipulation and computation due to an increased number of parameters to be estimated simultaneously may restrain their application when the number of traits is large. MANGIN et al. (1998) proposed an alternative method based on two separate steps, one being to estimate the canonical variables associated with the traits and the other to map QTLs for each canonical variable using maximum likelihood. Because each canonical variable represents a complex of multiple related characteristics, the QTLs detected can be considered as pleiotropic to these characteristics. Other methods have also been developed to explore issues related to pleiotropy and linkage. For example, a residual maximum likelihood method based on a deterministic, derivative-free algorithm has been proposed by GRIGNOLA et al. (1997) to test one vs. two QTLs linked to a group of markers. RONIN et al. (1999) extended the "deterministic sampling" approach to analyze two linked QTLs for a single or two quantitative traits. LEBRETON et al. (1998) used a nonparametric bootstrap method to test close linkage vs. pleiotropy of coincident QTLs.

An important progress in multi-trait mapping is its extension to map QTLs across different environments and developmental stages. By modeling the functional relationships between QTL expression and environment, KOROL et al. (1998a) proposed a maximum likelihood method to examine QTL x environment interactions when the number of environments is large. QTLs affecting environmental and developmental variation have been poorly understood, despite their particular importance in fundamental biology and applied animal and plant breeding. In developmental genetics, epigenetic inheritance is an important topic and its occurrence is dependent on the developmental environment of the organism. Through the stimulation of some signals, the organism may activate some "sleeping" genes to alter some biochemical pathways or morphological structures to better adapt to the environmental changes. New QTLs induced by a change in development have been mapped in *Populus* (WU, 1999).

Epistasis, coined by W. BATESON (1909), has been a major focus in the study of trait heredity and variation for several decades. Its actual role in evolution, speciation and breeding is unclear. One of the reasons is that it is difficult to quantify this complicated genetic phenomenon. Statistical methods for mapping epistatic QTLs have been now available and, thus, allows for a better resolution to epistasis than any time before.

The current statistical methods for mapping QTLs associated with complex developmental processes are still based on some simplified assumptions. For example, we assume only one or two QTLs affecting correlated traits. However, if we consider the traits as an integrated organism, such QTL numbers will not be adequate to describe the genetic mechanisms underlying phenotypic integration. Also, we consider these developmental processes separately rather comprehensively. For example, the simultaneous consideration of multiple traits is important, yet the integrated organism may respond to environmental changes (CHAPIN, 1991). It will be very interesting to map the QTLs behind such an integrated response.

Acknowledgments

The senior author thanks R. R. SEDEROFF and other members of the Forest Biotechnology Group at North Carolina State University for encouragement and support on writing this review. We are grateful to three anonymous referees for constructive comments on this manuscript. This work is partly supported by a grant from the Ministry of Forestry, China, and by the North Carolina State University Forest Biotechnology Associates.

Literature

ATCHLEY, W. R. and HALL, B. K.: A model for development and evolution of complex morphological structures. Biol. Rev. **66**: 101–157 (1991). — ATCHLEY, W. R., XU, S. and VOGL, C.: Developmental quantitative genetic models of evolutionary change. Dev. Genet. **15**: 92–103 (1994). — BATESON, W.: MENDEL's principles of heredity: a defense. The University Press, Cambridge (1909). — BRADSHAW, H. D. and STETTLER, R. F.: Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, form and phenology in a forest tree. Genetics **139**: 963–973 (1995). — BROMAN, K. W.: Identifying quantitative trait loci in experimental crosses. Ph.D thesis, University of California, Berkeley (1997). — CHAPIN, F. S.: Integrated responses of plants to stress. BioScience **41**: 29–36 (1991). — CHEVERUD, J. M., ROUNTMAN, E. J., DURANTE, F. A. M., VAN SWINDEREN, B., COTHRAN, K. and PEREL, C.: Quantitative trait loci for murine growth. Genetics **142**: 1305–1319 (1996). — CHURCHILL, G. A. and DOERGE, R. W.: Empirical threshold values for quantitative trait mapping. Genetics **138**: 963–971 (1994). — COWLEY, D. E. and ATCHLEY, W. R.: Quantitative genetic models for development, epigenetic selection and phenotypic evolution. Evolution **46**: 495–518 (1992). — DOEBLEY, J., STEC, A. and GUSTUS, C.: *Teosinte* branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics **141**: 333–346 (1995). — FALCONER, D. S.: Introduction to quantitative genetics. 3rd Ed. Longman Sci. and Tech., Harlow, UK (1989). — GRATTAPAGLIA, D., BERTOLUCCI, F. L. G., PENCHEL, R. and SEDEROFF, R. R.: Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. Genetics **144**: 1205–1214 (1996). — GRATTAPAGLIA, D., BERTOLUCCI, F. L. G. and SEDEROFF, R. R.: Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. Genetics **144**: 1205–1214 (1995). — GRIGNOLA, F. E., ZHANG, Q. and HOESCHELE, I.: Mapping linked quantitative trait loci via residual maximum likelihood. Genet. Sel. Evol. **29**: 529–544 (1997). — GROOVER, A., DEVY, M., FIDDLER, T., LEE, J., MEGRAW, T., MITCHELL-OLDS, T., SHERMAN, B., VUJCIC, C., WILLIAMS, C. and NEALE, D.: Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics **138**: 1293–1300 (1994). — JACOB, H. J., LINDPAINTER, K., LINCOLN, S. E., KUSUMI, K., BUNKER, R. K., MAO, Y.-P., GANTEN, D., DZAU, V. J. and LANDER, E. S.: Genetic mapping of a gene causing hypertension in the stroke-prone rat. Cell **67**: 213–224 (1991). — JANSEN, R. C.: Interval mapping of multiple quantitative trait loci. Genetics **135**: 205–211 (1993). — JANSEN, R. C.: Controlling the type I and type II errors in mapping quantitative trait loci. Genetics **138**: 871–881 (1994). — JANSEN, R. C. and STAM, P.: High resolution of quantitative traits into multiple loci via interval mapping. Genetics **136**: 1447–1455 (1994). — JIANG, C. and ZENG, Z.-B.: Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics **140**: 1111–1127 (1995). — KAO, C.-H.: Statistical Methods for Locating Pastions and Analyzing Epistasis of Multiple Quantitative Trait Genes Using Molecular Marker Information. Ph. D. thesis. North Carolina State University, Raleigh, NC (1995). — KAO, C.-H. and ZENG, Z.-B.: General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics **53**: 653–665 (1997). — KOROL, A. B., RONIN, Y. I. and KIRZHNER, V. M.: Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics **140**: 1137–1147 (1995). — KOROL, A. B., RONIN, Y. I. and NEVO, E.: Approximate analysis of QTL-environment interaction with no limits on the number of environments. Genetics **148**: 2015–2028 (1998a). — KOROL, A. B., RONIN, Y. I. and NEVO, E. et al.: Multi-interval mapping of correlated trait complexes. Heredity **80**: 273–284 (1998b). — LANDER, E. S. and BOTSTEIN, D.: Mapping MENDELian factors underlying quantitative traits using RFLP linkage maps. Genetics **121**: 185–199 (1989). — LANDER, E. S. and KRUGLYAK, L.: Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. Nat. Genet. **11**: 241–247 (1995). — LEBRETON, C. M., VISSCHER, P. M., HALEY, C.S., SEMIKHODSKII and QUARRIE, S. A.: A nonparametric bootstrap method for testing close linkage vs. pleiotropy of coincident quantitative trait loci. Genetics **150**: 931–943 (1998). — MANGIN, B., THOQUET, P. and GRIMSLEY, N.: Pleiotropic QTL analysis. Biometrics **54**: 88–99 (1998). — MENG, X.-L. and RUBIN, D. B.: Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika **80**: 267–278 (1993).

— NUZHDIN, S. V., PASYUKOVA, E. G., DILDA, C.L., ZENG, Z.-B. and MACKAY, T. F. C.: Sex-specific quantitative trait loci affecting longevity in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. **94**: 9734–9739 (1997). — PATERSON, A. H., DAMON, S., HEWITT, J. D., ZAMIR, D., RABINOWITCH, LINCOLN, S. E., LANDER, E. S. and TANKSLEY, S. D.: MENDELian factors underlying quantitative traits in tomato: comparisons across species, generations and environments. Genetics **127**: 181–197 (1991). — PHILLIPS, P. C.: The language of gene interaction. Genetics **149**: 1167–1171 (1998). — RONIN, Y. I., KIRZHNER, V. M. and KOROL, A. B.: Linkage between loci of quantitative traits and marker loci – multi-trait analysis with a single marker. Theor. Appl. Genet. **90**: 776–786 (1995). — RONIN, Y. I., KOROL, A. B. and NEVO, E.: Single- and multiple-trait mapping analysis of linked quantitative traits loci: Some asymptotic analytical approximations. Genetics **151**: 387–396 (1999). — STUBER, C. W., LINCOLN, S. E., WOLFF, D. W., HELENTJARIS, T. and LANDER, E. S.: Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics **132**: 823–839 (1992). — VOGL, C., ATCHLEY, W. R., COWLEY, D. E., CRENSHAW, P., MURRAY, J. D. et al.: The epigenetic influence of growth hormone on skeletal development. Growth Dev. Aging **57**: 163–182 (1993). — WU, R. L.: Genetic mapping of QTLs affecting tree growth and architecture in *Populus:* Implications for ideotype breeding. Theor. Appl. Genet. **96**: 447–457 (1998a). — WU, R. L.: The detection of plasticity genes in heterogeneous environments. Evolution **52**: 967–977 (1998b). — WU, R. L.: Mapping epigenetic QTLs altering a developmental trajectory: Theory and application. Dev. Genet., in press (1999). — WU, R., BRADSHAW jr., H. D. and STETTLER, R. F.: Molecular genetics of growth and development in *Populus* (Salicaceae). V. Mapping quantitative trait loci affecting leaf variation. Am. J. Bot. **84**: 143–153 (1997). — WU, R. L., LIU, H. X. and HAN, Y. F.: Statistical methods for mapping quantitative trait loci in forest trees. Scientia Silvae Sinica **35**(2): 100–117 (1999). — WU, R. and STETTLER, R. F.: Quantitative genetics of growth and development in Populus. II. The partitioning of genotype x environment interaction in stem growth. Heredity **78**: 124–134 (1997). — YAN, J. Q., ZHU, J., HE, C. X., BENMOUSSA, M. and WU, P.: Molecular dissection of developmental behavior of plant height in rice *(Oryza sativa* L.). Genetics **150**: 1257–1265 (1998). — YATES, F. E.: Self-organization systems. In: BOYD, C. A. R. and NOBLE, D. (eds): The logic of life – The challenge of integrative physiology. Oxford University Press, Oxford. pp. 189–1994 (1993). — ZENG, Z.-B.: Precision mapping of quantitative trait loci. Genetics **136**: 1457–1568 (1994). — ZHANG, H. and FORDE, B. G.: An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. Science **279**: 407–409 (1998).

Genetic Analysis of Needle Proteins in Maritime Pine

2. Variation of Protein Accumulation

By P. COSTA and C. PLOMION[1])

INRA, Laboratoire de Génétique et Amélioration des Arbres Forestiers, BP45, F-33610 Cestas, France

(Received 29th December 1998)

Abstract

Experimental variation and genetic background effects in protein accumulation were studied in a three-generation F_2 inbred pedigree of maritime pine *(Pinus pinaster* AIT.). Proteins extracted from needles were revealed by high-resolution silver-stained two-dimensional polyacrylamide gel electrophoresis and analysed with a computer-assisted system for single spot quantification. The integrated intensity of 77% of the studied spots showed a linear relationship with the total amount of protein loaded into the gel. A significant difference of integrated intensity was found among both parents and their hybrid for 31% of the studied proteins, from which 78% followed a non-additive mode of inheritance. The extent of the observed non-additivity is discussed and compared with results found in similar experiments in pea, maize and wheat. Finally, QTL mapping allowed the detection of PQL (Protein Quantity Loci) that explained part of the quantitative variation of protein accumulation.

Key words: Pinus pinaster, two-dimensional electrophoresis, proteins, inheritance, mapping, QTL, additivity.

[1]) Corresponding author: CHRISTOPHE PLOMION, INRA, Laboratoire de Génétique et Amélioration des Arbres Forestiers, BP45, F-33610 Cestas, France. Tel. 33 5 57 97 90 76, Fax 33 5 57 97 90 88
email. plomion @pierroton.inra.fr

Introduction

Qualitative variations of proteins revealed by two-dimensional polyacrylamide gel electrophoresis (2-DE) (O'FARRELL, 1975) were used for genetic studies in maritime pine *(Pinus pinaster* AIT.). Protein markers have been used both in genetic mapping (BAHRMAN and DAMERVAL, 1989; GERBER et al., 1993; PLOMION et al., 1995, 1997) and in population genetic studies (BAHRMAN et al., 1994; PETIT et al., 1995).

To analyse the genetic basis of quantitative variation of proteins separated by 2-DE, DAMERVAL et al. (1994) measured the quantity of each protein in a F_2 progeny of maize. The protein quantity was assessed through integrated optical density of each single spot using an automatic image-analysis system. Their study was based on QTL (Quantitative Trait Loci) mapping procedure (reviewed by TANKSLEY, 1993; KEARSEY, 1998). They used a linkage map to locate the regulatory factors or "PQL" (Protein Quantity Loci) that would explain part of the spot intensity variation. They concluded that multifactorial control of protein quantity variation was a general feature of genome expression. Recently, DE VIENNE et al. (1999) combined the PQL methodology with a traditional QTL mapping experiment, to characterise QTLs of economically important traits. They showed that three PQLs controlling the quantity of a single leaf protein and three QTLs of height growth in maize were co-located.

Given the large genome size of conifers (OHRI and KOSHOO, 1986; WAKAMIYA et al., 1993), the PQL approach seems to be a