Theor Appl Genet (1996) 93:102–109

R. L. Wu

Detecting epistatic genetic variance with a clonally replicated design: models for low- vs high-order nonallelic interaction

Received: 3 January 1996 / Accepted: 26 January 1996

Abstract A quantitative genetic model, that uses known family structure with clonal replicates to separate genetic variance into its additive, dominance and epistatic components, is available in the current literature. Making use of offspring testing, this model is based on the theory that components of variance from the linear model of an experimental design may be expressed in terms of expected covariances among relatives. However, if interactions between a pair of quantitative trait loci (QTLs) explain a large proportion of the total epistasis, it will seriously overestimate the additive and dominance variances but underestimate the epistatic variance. In the present paper, a new model is developed to manipulate this problem by combining parental and offspring material into the same test. Under the condition described above, the new model can provide an accurate estimate for additive × additive variances. Also, its accuracy in estimating dominance and total epistatic variances is much greater than the accuracy of the previous model. However, if there is obvious evidence showing the major contribution of high-order interactions, especially among ≥ 4QTLs, to the total epistasis, the previous model is more appropriate to partition the genetic variance for a quantitative trait. The re-analysis of an example from a factorial mating design in poplar shows large differences in estimating variance components between the new and previous models when two different assumptions (low- vs high-order epistatic interactions) are used. The new model will be an alternative to estimating the mode of quantitative inheritance for species, especially for long-lived, predominantly outcrossing forest trees, that can be clonally replicated.

Key words Additive variance · Clone
Dominant variance · Two-locus interaction

Communicated by J. MacKey

R. L. Wu
Division of Ecosystem Science and Conservation, University of Washington, Seattle, WA 98195-2100, USA

Introduction

Epistasis, or the genetic interaction among alleles at different but functionally related loci, results from a nonlinear relationship between genes and characters (Falconer 1989). A considerable body of research on biochemistry, physiology, and biometrical genetics strongly suggests that interlocus interaction is ubiquitous in quantitative traits (Hayman and Mather 1955; Hayman 1958; Wright 1980; Mather and Jinks 1982). One of the greatest contributions of Sewall Wright to evolutionary theory was to show how epistasis plays a central role in phenotypic evolution and speciation (Wright 1932, 1980; Templeton 1979, 1980; Carson and Templeton 1984; Provine 1986; Wade 1992). Theoretical work shows that quantitative variation in wild populations may harbor potentially strong epistatic components (Crow and Kimura 1970) because the small phenotypic effect of these components often hide a larger molecular effect. Thus, the cryptic epistatic variation may become important when genotypic frequencies are substantially disturbed by selection or by a population bottleneck (Carson and Templeton 1984; Bryant et al. 1986; Goodnight 1987, 1988; Tachida and Cockerham 1989; Bryant and Meffert 1992; Wade 1992).

Epistasis may also be an important cause of heterosis. By incorporating different components of epistasis into their genetic models, Jinks and Jones (1958) observed that there was a correlation between the presence or absence of heterosis and the presence or absence of nonallelic interactions in an annual plant. In a two-locus diallelic model, Minvielle (1987) showed that heterosis might stem from multiplicative epistatic interactions rather than dominance effects. The contribution of multiplicative accumulation to heterosis in a complex trait has been investigated both theoretically and empirically (Wright 1922; Dempster 1942; Ricky 1942; Williams 1959; Griffing 1990; Schnell and Cockerham 1992).

Although epistasis plays a pronounced role in evolution and breeding, empirical investigations on this

phenomenon are surprisingly few. One of the reasons for this may be due to the fact that segregating generations to detect epistasis cannot be obtained in many species (Mather and Jinks 1982). A methodological limitation to deal with this issue exists even in current molecular-marker-based mapping (Tanksley 1993). The partitioning of genetic variance into its additive, dominance, and epistatic components was initially introduced by Fisher (1981). Cockerham (1954) further subdivided the epistatic component into additive × additive, additive × dominant and dominant × dominant interactions. All these interactions can be estimated on the basis of generation means (Hayman and Mather 1955). Foster and Shaw (1988) took advantage of cloning to isolate epistasis from the total genetic variance by expressing components of variance from the linear model of an experimental design in terms of expected covariances among relatives (Cockerham 1963). Owing to its robust nature, Foster-Shaw's method has been applied to many forest tree species (e.g., Foster and Shaw 1988; Foster 1990; Mullin et al. 1992; Mullin and Part 1994; Rönnberg-Wästljung et al. 1994).

Unfortunately, Foster-Shaw's method cannot provide accurate estimates of genetic variance components. On the one hand, the additive and dominance variances estimated from the covariances among half-sibs or full-sibs are contaminated by a portion of the epistasis, especially low-order interactions (Cockerham 1963). The estimate of epistatic variance, on the other hand, is only a fraction of the epistasis contained by within-family variance rather than the total epistatic variance. Although the biased estimates could be relaxed somewhat by assuming strong high-order interactions among multiple loci (at least ≥ 3) (Mullin and Park 1992), they would be misleading in the case where two-locus interactions exist. When lowest-order interactions form a portion of the epistasis, the estimated additive and dominance variances will be seriously contaminated and the estimated epistatic variance will be only about one-quarter of the total epistasis. So far, the question of how epistasis can be accurately ruled out based on different interaction types has not been explored.

In this paper, I attempt to detect more accurate estimators of epistasis under the assumption of digenic epistasis by combining parental and offspring information in a mating design. From accuracy analysis and an example in *Populus*, the new procedure can be used as an alternative to partition the genetic variance of a quantitative trait into its causal componenets for species that can be cloned.

Statistical model

A number of mating designs can be used to generate relatives and to estimate genetic parameters, depending on the biological properties of a species. Here, a factorial mating design (NCII; Cockerham 1963) is considered. Assume f unrelated females and m unrelated males to be randomly sampled from a panmictic population. With regular Mendelian diploid behavior at meiosis and no cytoplasmic or maternal effects, the f females are factorially crossed to the m males to produce fm full-sib families each with c individuals. Each individual is then cloned to produce rn identical ramets.

A randomized complete block design is used to partition phenotypic variance into genetic and environmental components. The design has a total of fmc genotypes (clones) each with r replicates and n-individual plots. A general linear statistical model for the mating design can be expressed as

$$\begin{aligned} y_{ijkpq} = \mu + \mathrm{F}_i + \mathrm{M}_j + (\mathrm{F}\times\mathrm{M})_{ij} + [\mathrm{C}(\mathrm{F}\times\mathrm{M})]_{k(ij)} + \mathrm{R}_p + (\mathrm{F}\times\mathrm{R})_{ip} \\ + (\mathrm{M}\times\mathrm{R})_{jp} + (\mathrm{F}\times\mathrm{M}\times\mathrm{R})_{ijp} \\ + [\mathrm{C}(\mathrm{F}\times\mathrm{M})\times\mathrm{R}]_{k(ij)p} + \mathrm{E}_{ijkpq} \end{aligned} \quad (1)$$

for $i = 1, 2, \ldots, f$; $j = 1, 2, \ldots, m$; $k = 1, 2, \ldots, c$; $p = 1, 2, \ldots, r$;

$q = 1, 2 \ldots, n$

where y_{ijkpq} is the observed value of the qth individual of the kth clone within the ijth full-sib family in the pth replicate; μ is the overall mean; F_i and M_j are the general combining ability (GCA) effects of the ith female and jth male parents, respectively; $(\mathrm{F}\times\mathrm{M})_{ij}$ is the specific combing ability (SCA) effect of the ith female and jth male parent; $[\mathrm{C}(\mathrm{F}\times\mathrm{M})]_{k(ij)}$ is the kth clonal effect within the ijth full-sib family; R_p is the pth replicate effect; $(\mathrm{F}\times\mathrm{R})_{ip}$ and $(\mathrm{M}\times\mathrm{R})_{jp}$ are the interaction effects between the ith female or jth male GCA and the pth replicate, respectively; $(\mathrm{F}\times\mathrm{M}\times\mathrm{R})_{ijp}$ is the interaction effect between the SCA

Table 1 Analysis of variance and expected mean squares for a clonally replicated genetic trial based on a linear model of Eq. 1

Source	Degree of freedom	MS	EMS[a]
Replicate(R)	$r-1$	$\mathrm{MS_R}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\,\mathrm{V_{FMR}} + ncf\,\mathrm{V_{MR}} + ncm\mathrm{V_{FR}} + ncmf\,\mathrm{V_R}$
Female(F)	$f-1$	$\mathrm{MS_F}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\,\mathrm{V_{FMR}} + ncm\mathrm{V_{FR}} + nr\mathrm{V_{C(FM)}} + ncr\,\mathrm{V_{FM}} + ncrm\mathrm{V_F}$
Male(M)	$m-1$	$\mathrm{MS_M}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\,\mathrm{V_{FMR}} + ncf\,\mathrm{V_{MR}} + nr\mathrm{V_{C(FM)}} + ncr\,\mathrm{V_{FM}} + ncrf\,\mathrm{V_M}$
F × M	$(f-1)(m-1)$	$\mathrm{MS_{FM}}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\mathrm{V_{FMR}} + nr\mathrm{V_{C(FM)}} + ncr\mathrm{V_{FM}}$
Clone(C)/(F × M)	$fm(c-1)$	$\mathrm{MS_{C(FM)}}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nr\mathrm{V_{C(FM)}}$
F × R	$(f-1)(r-1)$	$\mathrm{MS_{FR}}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\mathrm{V_{FMR}} + ncm\mathrm{V_{FR}}$
M × R	$(m-1)(r-1)$	$\mathrm{MS_{MR}}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\mathrm{V_{FMR}} + ncf\,\mathrm{V_{MR}}$
F × M × R	$(f-1)(m-1)(r-1)$	$\mathrm{MS_{FMR}}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}} + nc\ \mathrm{V_{FMR}}$
C/(F × M) × R	$fm(c-1)(r-1)$	$\mathrm{MS_{C(FM)R}}$	$\mathrm{V_E} + n\mathrm{V_{C(FM)R}}$
Error	$fmcr(n-1)$	$\mathrm{MS_E}$	$\mathrm{V_E}$

[a] $\mathrm{V_E}$, variance due to the error effect; $\mathrm{V_{C(FM)R}}$, variance due to the interaction effect between clone-within-family and replicate; $\mathrm{V_{FMR}}$, variance due to the interaction effect between the SCA and replicate; $\mathrm{V_{MR}}$, variance due to the interaction effect between the male GCA and replicate; $\mathrm{V_{FR}}$, variance due to the interaction effect between the female GCA and replicate; $\mathrm{V_{C(FM)}}$, variance due to the clone-within-family effect; $\mathrm{V_{FM}}$, variance due to the SCA effect; $\mathrm{V_M}$, variance due to the male GCA effect; $\mathrm{V_F}$, variance due to the female GCA effect; $\mathrm{V_R}$, variance due to the replicate effect

of the ith female and jth male parents and the pth replicate; $[C(F \times M) \times R]_{k(ij)p}$ is the interaction effect between the kth clone with the ijth family and the pth replicate; E_{ijkpq} is the residual effect.

If the data are balanced and all of the model effects are assumed to be random, the analysis of variance is carried out with the degrees of freedom and expected mean squares as given in Table 1. The components of variance due to the female (V_F) and male GCA (V_M), SCA (V_{FM}), and clone-within-family [$V_{C(FM)}$] effects are estimated using mean squares and the expected mean squares. The sampling variances of these components are, respectively, given as follows (Hallauer and Miranda 1988):

$$\mathrm{Var}(V_F) = \frac{2}{(ncrm)^2}\left[\frac{MS_F^2}{f+1} + \frac{MS_{FM}^2}{(f-1)(m-1)+2} + \frac{MS_{FR}^2}{(f-1)(r-1)+2} + \frac{MS_{FMR}^2}{(f-1)(m-1)(r-1)+2}\right],$$

$$\mathrm{Var}(V_M) = \frac{2}{(ncrf)^2}\left[\frac{MS_M^2}{m+1} + \frac{MS_{FM}^2}{(f-1)(m-1)+2} + \frac{MS_{FR}^2}{(m-1)(r-1)+2} + \frac{MS_{FMR}^2}{(f-1)(m-1)(r-1)+2}\right],$$

$$\mathrm{Var}(V_{FM}) = \frac{2}{(ncr)^2}\left[\frac{MS_{FM}^2}{(f-1)(m-1)+2} + \frac{MS_{C(FM)}^2}{fm(c-1)+2} + \frac{MS_{FMR}^2}{(f-1)(m-1)(r-1)+2} + \frac{MS_{C(FM)R}^2}{fm(c-1)(r-1)+2}\right],$$

$$\mathrm{Var}\left[V_{C(FM)}\right] = \frac{2}{(nr)^2}\left[\frac{MS_{C(FM)}^2}{fm(c-1)+2} + \frac{MS_{C(FM)R}^2}{fm(c-1)(r-1)+2}\right].$$

The variance of full-sib family means (V_{FS}) is the summation of the female and male GCA variances and the SCA variance. The covariance between the means of the offspring family and the means of both parents [Cov(OP)] may be estimated from the analysis of covariance model (Table 2). Supposing that the offspring is a trait (X) and the corresponding mid-parent value is another trait (Y), the sources of covariance between these two traits include offspring family, replicate, offspring family × replicate, and error effects. The mean cross-product, based on the fm offspring families, is equated to its expected mean-product to obtain an unbiased estimate [$\mathrm{Cov}_O(X, Y)$] of the covariance of the offspring means and mid-parent values. The sampling variance of $\mathrm{Cov}_O(X, Y)$ is given as

$$\mathrm{Var}[\mathrm{Cov}_O(X, Y)] = \frac{2}{(nr)^2}\left[\frac{MCP_O^2}{fm+2} + \frac{MCP_{OR}^2}{(fm-1)(r-1)+2}\right].$$

Estimation of genetic variance components

In genetic experiments, groups of relatives provide sources of experimental variances that can be controlled and analyzed. Genetic effects may be studied by measuring the similarity or covariance (Cov) of group members (Cockerman 1963). Because family groups are assumed to be produced by crossing unrelated diploid parents from a random mating population, the among-group variation is a measure of the family covariance. Thus, the variance components of full-sib family, female and male GCAs, SCA, and clone-within-family, derived from the model of Eq. 1, are expressed in terms of expected covariances among relatives (Becker 1984; Foster and Shaw 1988):

$$V_{FS} = \mathrm{Cov}(\text{full-sibs}) \tag{2a}$$

$$V_F = V_M = \mathrm{Cov}(\text{half-sibs}) \tag{2b}$$

$$V_{FM} = \mathrm{Cov}(\text{full-sibs}) - 2\mathrm{Cov}(\text{half-sibs}) \tag{2c}$$

$$V_{C(FM)} = V_G - \mathrm{Cov}(\text{full-sibs}) \tag{2d}$$

where V_G is the total genetic variance of a quantitative trait, including additive (V_A), dominance (V_D), and epistatic variances (V_I). Assuming Mendelian behavior and linkage equilibrium at loci, and given the trait to be affected by a total of n QTLs, the covariance among relatives can be expressed as a function of genetic variance component (Fisher 1918):

$$\mathrm{Cov}(\text{relatives}) = \alpha V_A + \beta V_D + \alpha^2 V_{AA} + \alpha\beta V_{AD} + \beta^2 V_{DD} + \ldots$$

where α is the genetic correlation among relatives (Wright's coefficient of relationship), β is the coefficient of double coancestry (the probability that both alleles in two individuals are identical by descent), and V_{AA}, V_{AD}, and V_{DD} are the epistatic variances due to the first-order additive × additive effect, the first-order additive × dominance effect, and the first-order dominance × dominance effect. Based on the above assumptions and relationships, the covariances among full-sibs, half-sibs, and between offspring and the corresponding mid-parent may be, respectively, written (Cockerham 1954; Hallauer and Miranda 1988) as:

$$\mathrm{Cov}(\text{full-sibs}) = \frac{1}{2}V_A + \frac{1}{4}V_D + \frac{1}{4}V_{AA} + \frac{1}{8}V_{AD} + \frac{1}{16}V_{DD} + \sum_{i=3}^{n}\frac{1}{2^i}V_{AA\ldots A} + \sum_{i=3}^{n}\frac{1}{4^i}V_{DD\ldots D} + \sum_{i=3}^{n}\sum_{k=1}^{i-1}\frac{1}{2^{2i-k}}V_{AA\ldots ADD\ldots D} \tag{3a}$$

$$\mathrm{Cov}(\text{half-sibs}) = \frac{1}{4}V_A + \frac{1}{16}V_{AA} + \sum_{i=3}^{n}\frac{1}{4^i}V_{AA\ldots A} \tag{3b}$$

$$\mathrm{Cov}(\mathrm{OP}) = \frac{1}{2}V_A + \frac{1}{4}V_{AA} + \sum_{i=3}^{n}\frac{1}{2^i}V_{AA\ldots A} \tag{3c}$$

where i is the number of QTLs displaying inter-locus interactions, k of which are purely additive in effect; $V_{AA\ldots A}$, $V_{DD\ldots D}$, and $V_{AA\ldots ADD\ldots D}$ are the epistatic variances due to high-order additive by additive, dominance by dominance, and additve by dominance interactions among QTLs (≥ 3), respectively. According to Eqs. 2a–2d and 3a–3c, observed experimental variance or covariance components can be

Table 2 Analysis of covariance of the mean of offspring family (X) and the mean of parents (Y)

Source	Degree of freedom	MCP	EMCP[a]
Replicate (R)	$r-1$	MCP_R	$\mathrm{Cov}_E(X, Y) + n\mathrm{Cov}_{OR}(X, Y) + nfm\mathrm{Cov}_R(X, Y)$
Offspring (O)	$fm-1$	MCP_O	$\mathrm{Cov}_E(X, Y) + n\mathrm{Cov}_{OR}(X, Y) + nr\mathrm{Cov}_O(X, Y)$
O × R	$(fm-1)(r-1)$	MCP_{OR}	$\mathrm{Cov}_E(X, Y) + n\mathrm{Cov}_{OR}(X, Y)$
Error	$fmr(n-1)$	MCP_E	$\mathrm{Cov}_E(X, Y)$

[a] $\mathrm{Cov}_E(X, Y)$, covariance of X and Y due to the error effect; $\mathrm{Cov}_{OR}(X, Y)$, covariance of X and Y due to the interaction effect between offspring family and replicate; $\mathrm{Cov}_O(X, Y)$, covariance of X and Y due to the offspring family effect; $\mathrm{Cov}_R(X, Y)$, covariance of X and Y due to the replicate effect

written in terms of causal components of genetic variance as:

$$V_F = V_M = \frac{1}{4}V_A + \frac{1}{16}V_{AA} + \sum_{i=3}^{n}\frac{1}{4^i}V_{AA\ldots A} \tag{4a}$$

$$V_{FM} = \frac{1}{4}V_D + \frac{1}{8}V_{AA} + \frac{1}{8}V_{AD} + \frac{1}{16}V_{DD} + \sum_{i=3}^{n}\left[\frac{1}{2^i} - \frac{1}{2^{2i-1}}\right]V_{AA\ldots A} + \sum_{i=3}^{n}\frac{1}{4^i}V_{DD\ldots D} + \sum_{i=3}^{n}\sum_{k=1}^{i-1}\frac{1}{2^{2i-k}}V_{AA\ldots ADD\ldots D} \tag{4b}$$

$$V_{C(FM)} = \frac{1}{2}V_A + \frac{3}{4}V_D + \frac{3}{4}V_{AA} + \frac{7}{8}V_{AD} + \frac{15}{16}V_{DD} + \sum_{i=3}^{n}\left[1 - \frac{1}{2^i}\right]V_{AA\ldots A} + \sum_{i=3}^{n}\left[1 - \frac{1}{4^i}\right]V_{DD\ldots D} + \sum_{i=3}^{n}\sum_{k=1}^{i-1}\left[1 - \frac{1}{2^{2i-k}}\right]V_{AA\ldots ADD\ldots D} \tag{4c}$$

$$Cov_O(X,Y) = \frac{1}{2}V_A + \frac{1}{4}V_{AA} + \sum_{i=3}^{n}\frac{1}{2^i}V_{AA\ldots A}. \tag{4d}$$

Unknown genetic components, i.e., additive variance (V_A), dominance variance (V_D), the lowest-order epistatic variances (V_{AA}, V_{AD}, and V_{DD}), and higher-order epistatic variances ($V_{AA\ldots A}$, $V_{DD\ldots D}$, and $V_{AA\ldots DD\ldots D}$), cannot be determined because they are expressed in only the four equations above. However, an approximate estimate (V_A^*) of additive genetic variance can be obtained from Eqs. 4a and 4d:

$$V_A^* \approx 4(V_F + V_M) - 2Cov_O(X,Y) \tag{5a}$$

$$= V_A + \sum_{i=3}^{n}\left[\frac{1}{2^{2i-3}} - \frac{1}{2^{i-1}}\right]V_{AA\ldots A} \tag{5b}$$

where * denotes an approximate estimator. Two properties of the estimator, V_A^*, are established from the above equations: (1) V_A^* can provide unbiased estimate for V_A when the total epistasis is due only to digenic interactions; and (2) if high-order interactions involving groups of QTLs ≥ 3 exist, V_A^* will underestimate V_A since the second term in Eq. 5b has negative coefficients. However, the underestimated bias is only a minute portion ($1/2^{2i-3} - 1/2^{i-1}$) of the high-order interaction variance. Similarly, V_{AA} can also be estimated from Eqs. 4a and 4d:

$$V_{AA}^* \approx 8[Cov_O(X,Y) - (V_F + V_M)] \tag{6a}$$

$$= V_{AA} + 8\sum_{i=3}^{n}\left[\frac{1}{2^i} - \frac{1}{2^{2i-1}}\right]V_{AA\ldots A}. \tag{6b}$$

The estimator, V_{AA}^*, will be unbiased for V_{AA} when the lowest-order interactions account for the epistasis. However, if there exist high-order interactions among additive QTLs, V_{AA}^* is contaminated by $8(1/2^i - 1/2^{2i-1})$ of the high-order interaction variances. This portion is as large as 3/4 for the second-order interaction. As a result, V_{AA}^* is generally suggested to include the first- plus second-order additive × additive interaction variances.

After V_{AA} is estimated, the remaining question is how to estimate the rest of the epistasis, i.e., $V_{AD} + V_{DD}$. No procedure is available to provide an exact estimate for the summation. However, its two approximate expressions, $V_{AD} + 3/2\,V_{DD}$ and $2/3\,V_{AD} + V_{DD}$, can be derived by substracting Eq. 4b, multiplied by 3, from Eq. 4c. These two expressions are an overestimate and underestimate of $V_{AD} + V_{DD}$, respectively. Since the relative magnitude of V_{AD} and V_{DD} is unknown, a more accurate estimate for $V_{AD} + V_{DD}$ may be obtained from the average of the two expressions, i.e.,

$$V_{AD}^* + V_{DD}^* \approx \frac{5}{6}V_{AD}^* + \frac{5}{4}V_{DD}^*$$

$$\approx \frac{5}{3}V_F + \frac{5}{3}V_M - 5V_{FM} + \frac{5}{3}V_{C(FM)} - \frac{10}{3}Cov_O(X,Y) \tag{7a}$$

$$= \frac{5}{6}V_{AD} + \frac{5}{4}V_{DD} + \sum_{i=3}^{n}\left[\frac{5}{3} + \frac{5}{3\cdot 2^{2i-3}} - \frac{5}{2^{i-1}}\right]V_{AA\ldots A} + \sum_{i=3}^{n}\left[\frac{5}{3} - \frac{5}{3^{2(i-1)}}\right]V_{DD\ldots D} + \sum_{i=3}^{n}\sum_{k=1}^{i-1}\left[\frac{5}{3} - \frac{5}{2^{2i-k-2}}\right]V_{AA\ldots ADD\ldots DD}. \tag{7b}$$

The estimate of the total first-order interaction epistatic variance is approximately obtained by summing Eqs. 6a and 7a:

$$V_I^* = V_{AA}^* + V_{AD}^* + V_{DD}^* \approx -\frac{19}{3}V_F - \frac{19}{3}V_M - 5V_{FM} + \frac{5}{3}V_{C(FM)} + \frac{14}{3}Cov_O(X,Y). \tag{8}$$

Combining Eqs. 6b and 7b, the properties of V_I^* can be established: (1) in the case with only digenic interactions, V_I^* is a satisfactory estimate for the total epistasis. Eq. 8 can provide a good estimate typically when $V_{AD} = 3/2\,V_{DD} \approx V_{DD}$. However, if $V_{AD} \gg V_{DD}$, then V_I can be well estimated via multiplying Eq. 7a by 6/5 to obtain a more approximate estimate for $V_{AD} + V_{DD}$. If $V_{AD} \ll V_{DD}$, then a good estimate of V_I is obtained via multiplying Eq. 7a by 4/5; and (2) the estimator, described by Eq. 8, also includes high-order interaction epistatic terms if the latter exist. These high-order terms are overestimated since their coefficients are greater than 1.

Dominance variance, V_D, is estimated by summing Eqs. 4b and 4c and then replacing $V_{AD} + V_{DD}$ with Eq. 7a:

$$V_D \approx \frac{10}{3}(V_F + V_M) + 6V_{FM} - \frac{2}{3}V_{C(FM)} - \frac{8}{3}Cov_O(X,Y) \tag{9a}$$

$$= V_D + \frac{1}{6}V_{DD} - \frac{1}{4}V_{AD} - \sum_{i=3}^{n}\left[\frac{2}{3} - \frac{1}{2^{i-2}} + \frac{1}{3\cdot 2^{2(i-2)}}\right]V_{AA\ldots A} - \sum_{i=3}^{n}\left[\frac{2}{3} - \frac{5}{3\cdot 2^{2(i-1)}}\right]V_{DD\ldots D} - \sum_{i=3}^{n}\sum_{k=1}^{i-1}\left[\frac{2}{3} - \frac{5}{3\cdot 2^{2i-k-2}}\right]V_{AA\ldots ADD\ldots D}. \tag{9b}$$

Analogous to the total first-order epistasis, the accuracy for estimating V_D relies on the accuracy with which the summation $V_{AD} + V_{DD}$ is estimated and on the magnitude of higher-order interaction variances.

Assuming that V_F, V_M, V_{FM}, $V_{C(FM)}$, and $Cov_O(X,Y)$ are independent from each other, the sampling variances of additive, dominance, and epistatic variances are expressed as:

$$Var(V_A) \approx 16[Var(V_F) + Var(V_M)] + 4Var[Cov_O(X,Y)]$$

$$Var(V_D) \approx \frac{100}{9}[Var(V_F) + Var(V_M)] + 36Var(V_{FM}) + \frac{4}{9}Var[V_{C(FM)}] + \frac{64}{9}Var[Cov_O(X,Y)]$$

$$Var(V_I) \approx \frac{36}{9}[Var(V_F) + Var(V_M)] + 25Var(V_{FM}) + \frac{25}{9}Var[V_{C(FM)}] + \frac{196}{9}Var[Cov_O(X,Y)]$$

where the sampling variances of all high-order interaction variances are ignored.

Application

Accuracy analysis

If the total epistasis for a quantitative trait is only derived from digenic interactions, the present method is more accurate for partitioning genetic variance into the causal components than previous methods (Cockerham 1963; Foster and Shaw 1988). Additive (V_A) and additive × additive variances (V_{AA}) can be exactly estimated in this case based on Eqs. 5a,b and 6a,b. The estimator for V_A, suggested in Foster and Shaw (1988), is contaminated by 1/4 of the additive × additive variance, and it is impossible to obtain the estimator for V_{AA} from Foster-Shaw's method (Table 3). The estimators of dominance and epistatic variances, described by Eqs. 11a and 8, are also closer to their actual values than the corresponding estimators by Cockerham (1963) and Foster and Shaw (1988). Cockerham's method overestimates the dominance variance by $1/2V_{AA} + 1/2V_{AD} + 1/4V_{DD}$, while the present method only has the deviation of $1/6V_{AD} + 1/4V_{DD}$ (Table 3). It can be seen that this deviation is inversely included in the estimation of the total epistatic variance. The epistatic estimate from Foster-Shaw's method is downwardly biased by $3/4V_{AA} + 1/2V_{AD} + 1/4V_{DD}$ (Table 3).

The estimators for the components of additive and dominance variance, suggested in this paper, will be contaminated by a portion of the epistatic components when higher-order interactions are included in the epistasis (Table 3). If the higher-order interactions, involving groups of QTLs ≥ 4, contribute to the major portion of the epistasis, the present method will lose its advantages in accuracy, as compared to Foster-Shaw's method.

Numerical example

Foster and Shaw (1988) reported a study of estimating the types of gene action based on a 3 × 4 factorial mating design with clonal replicates for a poplar (*Poplus deltoides* Bartr.). Table 4 lists the variances due to female and male GCAs, SCA, and clone-within-family and the corrsponding standard errors in 3-year stem diameter at 1.4 m height. In their study, Foster and Shaw assumed that the epistasis in radial growth mainly results from high-order interactions among multiple QTLs. Here, the alternative assumption that the epistasis for this trait is due to the lowest-order interactions will be considered. As will be clear, the analytical results are quite different when based on different assumptions. Since the covariance of the means of offspring family and the means of both parents was not estaimated in Foster and Shaw's paper, a simulation will be performed based on its relationship to Cov(half-sibs) in terms of genetic variance components. It can be seen from Eqs. 3b and 3c that the value of Cov(OP) should range from 2Cov(half-sibs) to 4Cov(half-sibs). Based on this relationship, let $\gamma = 2\text{Cov(half-sibs)}/\text{Cov(OP)}$ $(1/2 \leq \gamma \leq 1)$. The ratio reflects the magnitude of additive × additive epistatic variance of all orders. (However, only the lowest-order epistasis is assumed here.) If γ approaches 1, this indicates that the additive × additive variance is close to 0 as compared to the additive variance (V_A). If γ approaches 1/2, the additive × additive variance will tend to be infinite relative to V_A.

Under different γ values, the components of additive, dominance, and epistatic variances were simulated by the present method (Fig. 1). When $\gamma = 1.0$, the estimate of additive variance components from the present method were identical to that from Foster-Shaw's method (Fig. 1 A; Table 4). However, so long as $\gamma < 1.0$, i.e., there is additive × additive epistatic variance in 3-year stem diameter of *P. deltoides*, Foster-Shaw's es-

Table 3 Coefficients of components of two-, three-, and four-locus interaction variances included in the estimates of additive (V_A^*), dominance (V_D^*), and epistatic variances (V_I^*), from the present (1) and Foster-Shaw's (2) methods. Positive (or negative) denotes the three estimators above to be contaminated (or discounted) by the corresponding epistasis terms

Epistatic term		V_A^*		V_D^*		V_I^*	
		1	2	1	2	1	2
Two-locus interaction	V_{AA}	0	1/4	0	1/2	0	−1/4
	V_{AD}	0	0	1/6	1/2	−1/6	−1/2
	V_{DD}	0	0	−1/4	1/4	1/4	−1/4
Three-locus interaction	V_{AAA}	−1/8	1/16	−1/4	3/8	3/8	−7/16
	V_{AAD}	0	0	−3/8	1/4	3/8	−1/4
	V_{ADD}	0	0	−1/4	1/8	1/4	−1/8
	V_{DDD}	0	0	−1/4	1/16	1/4	−1/16
Four-locus interaction	V_{AAAA}	−3/32	1/64	−7/16	7/32	17/32	−15/64
	V_{AAAD}	0	0	−59/96	1/8	59/96	−1/8
	V_{AADD}	0	0	−27/48	1/16	27/48	−1/16
	V_{ADDD}	0	0	−11/24	1/32	11/24	−1/32
	V_{DDDD}	0	0	−41/64	1/64	41/64	−1/64

Table 4 Variance components due to female and male GCAs, SCA, and clone-within-family and Foster-Shaw's estimators of additive (V_A), dominance (V_D), and epistatic variance components (V_I) for 3-year stem diameter at 1.4 m height in *P. deltoides* (Foster and Shaw 1988)

Source	Symbol	Value ± SE
Experimental variance		
Female (F)	V_F	0.0525 ± 0.0930
Male (M)	V_M	0.3684 ± 0.3057
F × M	V_{FM}	0.0887 ± 0.1064
Clone (C)/(FM)	$V_{C(FM)}$	0.5431 ± 0.1081
Causal component		
Additive	V_A	0.8418
Dominant	V_D	0.3548
Epistatic	V_I	0.0000

timator will be upwardly biased. Their estimator also overestimated dominance variances (Fig. 1 B, Table 4) when only digenic interactions exist. According to Foster-Shaw's assumption, the total epistasis is absent (since its calculated value is negative, which may result from large sampling variances) (Table 4). However, considerable additive × additive variance is found under the two-locus interaction assumption (Fig. 1 C), despite the rest of the epistasis failing to be detected (Fig. 1 D).

Discussion

Estimation of epistatic variance components is a long standing problem. Despite its obvious importance in breeding and evolution, we are still painfully short of a reliable method to obtain such an estimate. Although use of clones has provided possibilities to isolate epistasis from the total genetic variance (Foster and Shaw 1988), the estimates of the epistatic component are largely frustrated by the assumption that the total epistasis is primarily due to high-order inter-allelic interactions. At present, even QTL mapping cannot provide a reasonable solution to this problem since too many parameters are required to determine the types of epistasis (Tanksley 1993). The present paper attempts to propose a new procedure that is totally independent of Foster-Shaw's (1988) assumption. If the epistasis for a quantitative trait is limited to interactions between a pair of QTLs, the new procedure will have two significant advantages over Foster-Shaw's method. They are (1) the partitioning of the genetic variance is more accurate, and (2) the additive × additive component can be isolated from the total epistasis.

Under the condition of only digenic interactions, the new procedure provides an exact estimate for additive and additive × additive variances. When this condition

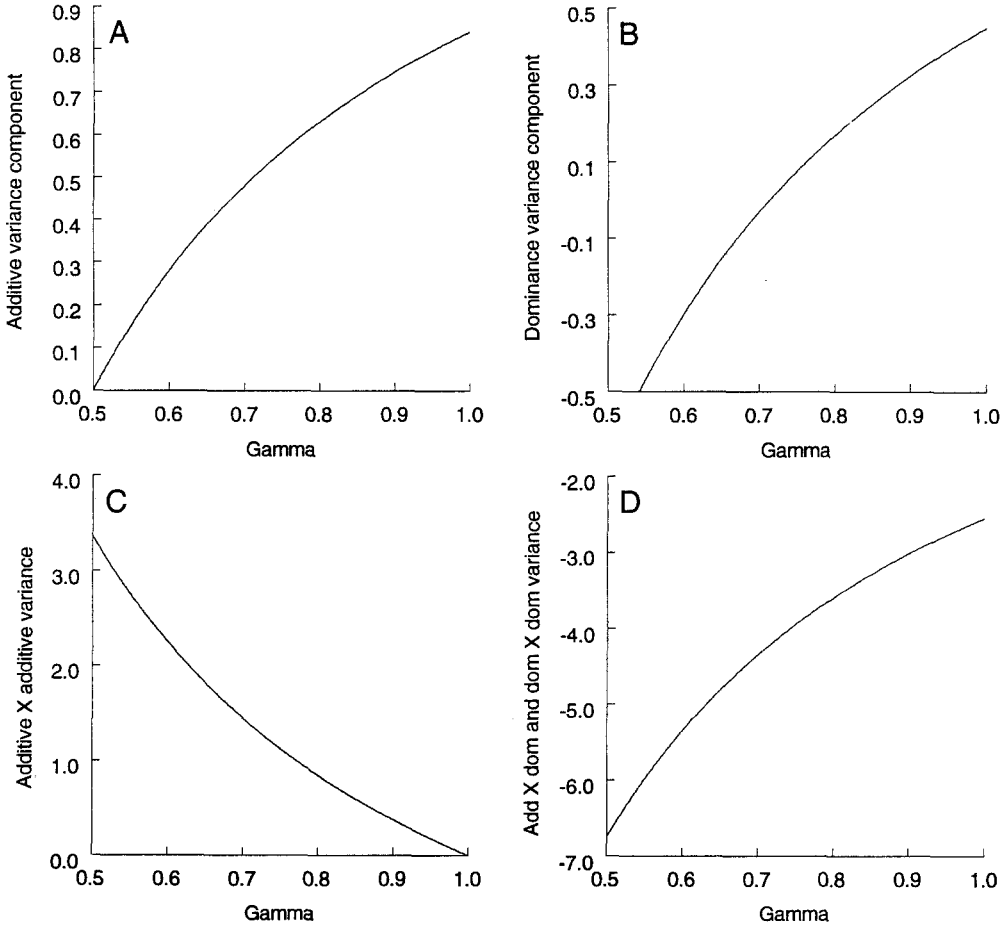

Fig 1 The estimates of additive (**A**), dominance (**B**), additive × additive (**C**), and additive × dominance plus dominance × dominance epistatic variances (**D**) in 3-year stem diameter of *P. deltoides* from the present method under different γ values (data from Foster and Shaw 1990)

is not met, the estimates of these two variances will be biased. However, if digenic interactions contribute to a major portion of the total epistasis, the estimated additive and adddive × additive variances by the covariances among half-sibs and between the offspring and the mid-parent will be quite reasonable.

The exact estimates for dominance, additive × dominance and dominance × dominance variances cannot be obtained without additional information about gene action, allele frequencies, and the number of QTLs involved, even if only two-locus interactions are assumed. However, the present estimators for dominance and epistatic variance components are much better than Foster-Shaw's estimators under the two-locus interaction model. Once a major portion of the total epistasis is due to higher-order interactions involving groups of QTLs ≥ 4, the present procedure will not be comparable to Foster-Shaw's method.

In practice, the applicability of Foster-Shaw's method may be limited, since there is little evidence for high-order epistatic variances despite the fact that the matter has not been pursued extensively. Traditional statistical genetics generally assumes the existence of only two-locus interactions. Although this assumption is made due to an increasing algebraic difficulty in manipulating high-order interactions (Mather and Jinks 1982), it seems to have also shown some biological basis, as revealed by current molecular marker-aided QTL analyses. Based on these molecular analyses, a significant proportion of the genetic variance in a number of quantitative traits can be explained by only a few QTLs with large effects (e.g., Doebley and Stec 1991, 1993; Paterson et al. 1991; Stuber et al. 1992; deVicente and Tanksley 1993; Damerval et al. 1994; Bradshaw and Steller 1995). Thus, for these traits, it is less possible that interactions involving groups of numerous QTLs form a major portion of the epistasis. However, if there is a difficulty in determining the relative importance of low- vs high-order interactions for the total epistasis, both the new method and Foster-Shaw's method are generally recommended for use. Researchers are encouraged to choose a better result based on other knowledge. As shown by the *Populus* example, the two methods would lead to very different conclusions.

Except for the assumption of two-locus interactions, the present estimates for the types of gene action of a quantitative trait from a mating and experimental design require the same other assumptions as Foster-Shaw's method (see Comstock et al. 1958; Foster and Shaw 1988). These assumptions have been described above and are summarized as follows: (1) the parents are randomly sampled from a random mating population; (2) there is a regular Mendelian diploid behavior at meisos; (3) cytoplasmic or maternal effects are absent; and (4) linkage equilibrium exists such that QTLs affecting an observed trait segregate independently or, where linkage does exist, the distribution of genotypes is as expected in the absence of linkage (Cockerham 1956). Violations of the above assumptions will result in biased estimates for genetic variance components. Wu (1995) suggested a genetic model for the resolution of quantitative variance in triploid progeny and further discussed the possible accuracy of his estimators. The effects of relaxing each of the other assumptions on the estimates requires further study.

In order to make it possible to estimate genetic variance components from the present method, both parents and their offspring should be planted side by side in a clonally replicated plantation. Although there are many biological techniques available to clone tree species (see Ahuja and Libby 1993), C effects by cloning have been documented in several species (Libby and Jund 1962; Cannell et al. 1988; Farmer et al. 1989). Since C effects may bias estimates of variance components (Libby and Jund 1962), appropriate measures should be taken to minimize these non-genetic effects before the replicated trial is established (Frampton and Foster 1993).

Acknowledgements The author is grateful to Dr. H. D. Bradshaw, Jr and Professor J. MacKey for helpful discussions regarding this work. This study is supported by grants from Washington Technology Centers (09-1022) and U.S. Department of Energy (19X-43382C).

References

Ahuja MR, Libby WJ (1993) Clonal forestry I. Genetics and biotechnology. Springer-Verlag, Berlin Heidelberg

Becker WA (1984) Manual of Quantitative genetices, 4th edn. Academic Enterprises. Pullman, Washington

Bradshaw HD Jr, Stettler RF (1995) Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, form, and phenology traits in forest tree. Genetics 139:963–973

Bryant EH, Meffert LM (1992) The effect of serial founderflush cycles on quantitative genetic variation in the housefly. Heredity 70:122–129

Bryant EH, McCommas SA, Combs LM (1986) The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. Genetics 114:1191–1211

Cannell MGR, Sheppard LJ, Cahalan C (1988) C effects and second generation clone performance in *Picea sitchensis* and *Pinus contorta*. Silvae Genet 37:15–19

Carson HL, Templeton AR (1984) Genetic resolutions in relation to speciation phenomena: the founding of new populations. Annu Rev Ecol Syst 15:97–131

Cockerham CC (1954) An extension of the concept of partitioning hereditary variance for analysis of covariance among relatives when epistasis is present. Genetics 39:859–882

Cockerham CC (1956) Effects of linkage on the covariances between relatives. Genetics 41:138–141

Cockerham CC (1963) Estimation of genetic variances. In: Hanson WD, Robinson HF (eds) Statistical genetics and plant breeding. Proc Natl Acad Sci Natl Res Council USA 982:53–94

Comstock RE, Kelleher T, Morrow EB (1958) Genetic variation in an asexual species, the garden strawberry. Genetics 43:634–646

Crow JF, Kimura M (1970) An introduction to population genetics theory. Harper and Row, New York

Damerval C, Maurice A, Josse JM, deVienne D (1994) Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of genome expression. Genetics 137:289–301

Dempster ER (1942) "Mock dominance". Science 97:464–465

deViente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596
Doebley J, Stec A (1991) Genetic analysis of the morphological differences between maize and teosinte. Genetics 129:285–295
Doebley J, Stec A (1993) Inheritance of the morphological differences between maize and teosinte: comparison of results for two F_2 populations. Genetics 134:559–570
Falconer DS (1989) Introduction to quantitative genetics, 3rd edn. Longman, Harlow, England
Farmer RE Jr, Freitag M, Garlick K (1989) Genetic variance and "C" effects in balsam poplar rooting. Silvae Genet 38:62–65
Fisher RA (1918) The correlations between relatives on the supposition of Mendelian inheritance. Trans Roy Soc Edin 52:399–433
Foster GS (1990) Genetic control of rooting ability of stem cuttings from lobolly pine. Can J For Res 20:1361–1368
Foster GS, Shaw DV (1988) Using clonal replicates to explore genetic variation in a perennial plant species. Theor Appl Genet 76:788–794
Frampton LJ Jr, Foster GS (1993) Field testing vegetative propagules. In: Ahuja MR, Libby WJ (eds) Clonal forestry I. Genetics and biotechnology. Springer-Verlag, Berlin Heidelberg, pp 110–134
Goodnight CJ (1987) On the effect of founder events on epistatic genetic variance. Evolution 41:80–91
Goodnight CJ (1988) Epistasis and the effect of founder events on the additive genetic variance. Evolution 42:399–403
Griffing B (1990) Use of a controlled-nutrient experiment to test heterosis hypotheses. Genetics 126:753–767
Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding, 2nd edn. Iowa State University Press, Ames, Iowa
Hayman BI (1958) The separation of epistatic from additive and dominance variation in generation means. Heredity 12:371–390
Hayman BI, Mather K (1955) The description of genic interaction in continunous variation. Biometrics 10:69–82
Jinks JL, Jones ZM (1958) Estimation of the components of heterosis. Genetics 43:223–234
Libby WJ, Jund E (1962) Variance associated with cloning. Heredity 17:533–540
Mather K (1949) Biometrical genetics. Methuen, London
Mather K, Jinks JL (1982) Biometrical genetics, 3rd edn. Chapman and Hall, London
Minvielle F (1987) Dominance is not necessary for heterosis: a two-locus model. Genet Res 49:245–247
Mullin TJ, Park JS (1992) Estimating genetic gains from alternative breeding strategies for clonal forestry. Can J For Res 22:14–23
Mullin TJ, Park JS (1994) Genetic parameters and age-age correlations in a clonally replicated test of black spruce after 10 years. Can J For Res 24:2330–2341
Mullin TJ, Morgenstern EK, Park JS, Fowler DP (1992) Genetic parameters from a clonally replicated test of black spruce (*Picea mariana*). Can J For Res 22:24–36
Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander S, Tanksley D (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics 127:181–197
Provine WB (1986) Sewall Wright and evolutionary biology. University of Chicago Press, Chicago
Richey FD (1942) Mock-dominance and hybrid vigor. Science 96:280–281
Rönnberg-Wästljung AC, Gullberg U, Nilsson C (1994) Genetic parameters of growth characters in *Salix viminalis* growth in Sweden. Can J For Res 24:1960–1969
Schnell FW, Cockerham CC (1992) Multiplicative vs arbitrary gene action in heterosis. Genetics 131:461–469
Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–829
Tachida H, Cockerham CC (1989) A building block model for quantitative genetics. Genetics 121:839–844
Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205–233
Templeton AR (1979) The unit of selection in *Drosophila mercatorum*. II. Genetic revolution and the origin of coadapted genomes in parthenogenetic strains. Genetics 92:1265–1282
Templeton AR (1980) The theory of speciation via the founder principle. Genetics 94:1011–1038
Wade MJ (1992) Sewall Wright: gene interaction and the shifting balance theory. Oxf Surv Evol Biol 8:33–62
Williams W (1959) Heterosis and the genetics of complex characters. Nature 184:527–530
Wright S (1922) The effect of inbreeding and crossbreeding on guinea pigs. III. Crosses between highly inbred families. Tech Bull US Dept Agric 1121
Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proc 6th Int Congr Genet 1:356–366
Wright S (1980) Genic and organismic selection. Evolution 34:825–843
Wu R (1995) A quantitative genetic model of mixed diploid and triploid progenies in tree breeding and evolution. Theor Appl Genet 90:683–690