

Epistasis of Quantitative Trait Loci Under Different Gene Action Models

Rong-Cai Yang¹

Alberta Agriculture, Food and Rural Development, Edmonton, Alberta T6H 5T6, Canada and Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta T6G 2P5, Canada

Manuscript received July 15, 2003

Accepted for publication April 9, 2004

ABSTRACT

Modeling and detecting nonallelic (epistatic) effects at multiple quantitative trait loci (QTL) often assume that the study population is in zygotic equilibrium (*i.e.*, genotypic frequencies at different loci are products of corresponding single-locus genotypic frequencies). However, zygotic associations can arise from physical linkages between different loci or from many evolutionary and demographic processes even for unlinked loci. We describe a new model that partitions the two-locus genotypic values in a zygotic disequilibrium population into equilibrium and residual portions. The residual portion is of course due to the presence of zygotic associations. The equilibrium portion has eight components including epistatic effects that can be defined under three commonly used equilibrium models, Cockerham's model, F_2 -metric, and F_{∞} -metric models. We evaluate our model along with these equilibrium models theoretically and empirically. While all the equilibrium models require zygotic equilibrium, Cockerham's model is the most general, allowing for Hardy-Weinberg disequilibrium and arbitrary gene frequencies at individual loci whereas F_2 -metric and F_{∞} -metric models require gene frequencies of one-half in a Hardy-Weinberg equilibrium population. In an F_2 population with two unlinked loci, Cockerham's model is reduced to the F_2 -metric model and thus both have a desirable property of orthogonality among the genic effects; the genic effects under the F_{∞} -metric model are not orthogonal but they can be easily translated into those under the F_2 -metric model through a simple relation. Our model is reduced to these equilibrium models in the absence of zygotic associations. The results from our empirical analysis suggest that the residual genetic variance arising from zygotic associations can be substantial and may be an important source of bias in QTL mapping studies.

WITH increasing availability of fine-scale and highly dense genetic maps for different organisms, it is now possible to simultaneously identify and map several quantitative trait loci (QTL) controlling a trait of economic and/or adaptive significance in plant and animal species. Unlike earlier efforts to map individual QTL "one QTL at a time" (see LYNCH and WALSH 1998 for review), this new approach requires that gene action models specify both allelic (additive and dominance) effects at individual loci and nonallelic interactions (epistatic effects) between loci. However, these genic effects must be defined in a "reference" population because they depend not only on genotypic values but also on genotypic frequencies of the reference population. The most obvious choice of reference population has been an "ideal" random mating population where Hardy-Weinberg and linkage equilibria are assumed but the gene frequencies may be arbitrary at different loci (KEMP THORNE 1957; CROW and KIMURA 1970; LYNCH and WALSH 1998). COCKERHAM (1954) considered a more general population where departures from Hardy-Weinberg equilibrium at individual loci such as inbreeding could be accommodated but the genotypic

frequencies between loci would be uncorrelated. On the basis of this reference population, COCKERHAM (1954) then developed a set of orthogonal scales (contrasts) so that the two-locus genotypic values could be partitioned into independent components due to additive, dominance, and epistatic effects. Cockerham's model has been a standard for conventional quantitative genetic analysis and recently for modeling multiple QTL (KAO and ZENG 2002).

For populations derived from a cross between two inbred lines such as F_2 and subsequent populations, two simpler models, F_2 -metric and F_{∞} -metric, have often been used. While translation of the genic effects defined under the F_2 -metric and F_{∞} -metric models can be easily done (VAN DER VEEN 1959), each model has its own characteristics. The F_2 -metric model is actually a special case of Cockerham's model where the gene frequencies are one-half and the genotypic frequencies are those expected under Hardy-Weinberg equilibrium at each of two uncorrelated loci. Thus the F_2 -metric model possesses the same property of orthogonality among the genic effects as Cockerham's model and has been widely used (*e.g.*, HAYMAN and MATHER 1955; KEMP THORNE 1957; COCKERHAM and ZENG 1996; GOODNIGHT 2000; KAO and ZENG 2002). On the other hand, the genic effects defined under the F_{∞} -metric model are not orthogonal, but this model has been equally popular (*e.g.*, CROW

¹Address for correspondence: Department of Agricultural, Food and Nutritional Science, 410 Agriculture/Forestry Centre, University of Alberta, Edmonton, AB T6G 2P5, Canada.
E-mail: rong-cai.yang@ualberta.ca

TABLE 1

Joint frequencies of the nine genotypes for two loci, each with two alleles, *A* and *a* at locus *A* and *B* and *b* at locus *B*. These genotypic frequencies are expressed in terms of their single-locus genotypic frequencies and zygotic associations

Locus <i>A</i>	Locus <i>B</i>			Total
	<i>BB</i>	<i>Bb</i>	<i>bb</i>	
<i>AA</i>	$P_{AB}^{AB} = P_A^A P_B^B + \omega_{AB}^{AB}$	$P_{Ab}^{AB} = P_A^A P_B^b + \omega_{Ab}^{AB}$	$P_{aB}^{AB} = P_A^a P_B^B + \omega_{aB}^{AB}$	$P_A^A = p_A^2 + D_A$
<i>Aa</i>	$P_{AB}^{Ab} = P_A^A P_B^b + \omega_{AB}^{Ab}$	$P_{Ab}^{Ab} + P_{aB}^{Ab} = P_A^A P_B^b + \omega_{ab}^{Ab}$	$P_{aB}^{Ab} = P_A^a P_B^B + \omega_{aB}^{Ab}$	$P_A^a = 2p_A p_a - 2D_A$
<i>Aa</i>	$P_{aB}^{aB} = P_A^a P_B^B + \omega_{aB}^{aB}$	$P_{ab}^{aB} = P_A^a P_B^b + \omega_{ab}^{aB}$	$P_{ab}^{ab} = P_A^a P_B^b + \omega_{ab}^{ab}$	$P_a^a = p_a^2 + D_A$
Total	$P_B^B = p_B^2 + D_B$	$P_{-b}^B = 2p_B p_b - 2D_B$	$P_{-b}^b = p_b^2 + D_B$	1

The possible values of zygotic associations are constrained by the single-locus frequencies such that only four of the nine zygotic associations need to be defined and the remaining five are entirely expressed in terms of the four defined zygotic associations. For example, if the zygotic associations for four double homozygotes, ω_{AB}^{AB} , ω_{Ab}^{Ab} , ω_{aB}^{aB} and ω_{ab}^{ab} are defined, then the zygotic associations for the remaining five genotypes are expressed as follows: $\omega_{AB}^{Ab} = -(\omega_{AB}^{AB} + \omega_{Ab}^{Ab})$, $\omega_{aB}^{Ab} = -(\omega_{aB}^{aB} + \omega_{ab}^{ab})$, $\omega_{aB}^{aB} = -(\omega_{aB}^{aB} + \omega_{ab}^{ab})$, $\omega_{ab}^{aB} = -(\omega_{ab}^{ab} + \omega_{aB}^{aB})$, and $\omega_{ab}^{ab} = \omega_{AB}^{AB} + \omega_{Ab}^{Ab} + \omega_{aB}^{aB} + \omega_{ab}^{ab}$.

and KIMURA 1970; MATHER and JINKS 1982; YANG and BAKER 1990; HALEY and KNOTT 1992; CARLBORG *et al.* 2000; YI and XU 2002) probably because of its easy and clear genetic interpretation as noted by VAN DER VEEN (1959) and MATHER and JINKS (1982).

The key assumption with all these gene action models is that the genotypic frequencies at different loci are products of the appropriate single-locus genotypic frequencies; in other words, the population is in zygotic equilibrium. However, zygotic associations can arise from physical linkages between different loci or from many evolutionary and demographic processes even for unlinked loci (see YANG 2002 for review). In the presence of zygotic associations, a large number of genic disequilibria including Hardy-Weinberg and linkage disequilibria are required for a complete characterization of non-random associations at different loci (COCKERHAM and WEIR 1973; WEIR 1996). These disequilibria have been considered in some general gene action models (*e.g.*, GALLAIS 1974; WEIR and COCKERHAM 1977; KAO and ZENG 2002) but a prohibitively large number of genic effects arising from such models have prevented them from being estimated in experimental quantitative genetics and QTL mapping studies. In this article, we first develop a new model that partitions the two-locus genotypic values in a zygotic disequilibrium population into equilibrium and residual portions. Such partitioning is made possible by the decomposition of the two-locus genotypic frequencies into the expected frequencies and deviations (zygotic associations) as given in YANG (2000). The residual portion is of course due to the presence of zygotic associations. We also examine the partitioning of the equilibrium portion into different genic effects

at the two loci using Cockerham's model and the F_2 -metric and F_∞ -metric models.

COCKERHAM'S MODEL WITH ZYGOTIC ASSOCIATIONS

Two-locus genotypic frequencies and zygotic associations: For two loci, each with two alleles, *A* and *a* at locus *A* and *B* and *b* at locus *B*, there are 9 possible genotypes (10 if the coupling and repulsion double heterozygotes are distinguishable). Following YANG (2000), we write frequencies of these genotypes as, $P_{yz}^{uy} = P_{yz}^{vz}$, which result from union of gametes *uy* and *vz* with *u*, *v* = *A* or *a*, and *y*, *z* = *B* or *b* (Table 1). The genotypic frequencies at individual loci are the marginal totals of the appropriate two-locus genotypic frequencies. For example, the frequency of genotype *AA* is

$$P_A^A = P_{AB}^{AB} + P_{Ab}^{AB} + P_{aB}^{Ab}$$

With the genotypic frequencies at locus *A*, the frequency of allele *A* is

$$p_A = P_A^A + \frac{1}{2} P_A^a$$

and that of allele *a* is $p_a = 1 - p_A$.

Departures from Hardy-Weinberg equilibrium at locus *A* are

$$D_A = P_A^A - p_A^2 = P_a^a - p_a^2 = -[\frac{1}{2} P_A^a - p_A p_a] \quad (1a)$$

and those at locus *B* are

$$D_B = P_B^B - p_B^2 = P_b^b - p_b^2 = -[\frac{1}{2} P_b^b - p_B p_b]. \quad (1b)$$

In a Hardy-Weinberg equilibrium population, $D_A = D_B =$

0. In inbred populations, D_A and D_B depend on the level of inbreeding as measured by the inbreeding coefficients at loci A (f_A) and B (f_B). For example, the Hardy-Weinberg disequilibrium at locus A is $D_A = f_A p_A p_a$, where

$$f_A = \frac{P_A^A - p_A^2}{p_A p_a} = \frac{P_a^a - p_a^2}{p_A p_a} = 1 - \frac{P_a^A}{2p_A p_a}.$$

YANG (2000) defined a zygotic association as a deviation of the two-locus genotypic frequency from the product of single-locus frequencies,

$$\omega_{yz}^{uy} = P_{yz}^{uy} - P_v^u P_z^y, \quad (2)$$

which has the range of

$$-P_v^u P_z^y \leq \omega_{yz}^{uy} \leq \min[P_v^u (1 - P_z^y), (1 - P_v^u) P_z^y].$$

Because these zygotic associations are constrained by the single-locus genotypic frequencies,

$$\sum_{u,v} \omega_{yz}^{uy} = \sum_{y,z} \omega_{yz}^{uy} = 0,$$

only 4 of the 9 zygotic associations can be defined freely and the remaining 5 are expressed entirely in terms of the four defined zygotic associations (Table 1).

For compact and clear presentation of subsequent developments, we describe the relations of two-locus genotypic frequencies with their expectations and zygotic associations (Table 1) in matrix form,

$$\mathbf{P} = \mathbf{\Psi} + \mathbf{\Omega}, \quad (3)$$

where

$$\mathbf{P} = \text{diag}\{P_{AB}^{AB} P_{Ab}^{AB} P_{aB}^{AB} P_{ab}^{AB} P_{AB}^{AB} + P_a^a P_{Ab}^{Ab} P_{aB}^{Ab} P_{ab}^{Ab} P_{ab}^{ab}\},$$

$$\mathbf{\Psi} = \text{diag}\{P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B P_A^A P_B^B\}$$

$$\mathbf{\Omega} = \text{diag}\{\omega_{AB}^{AB} \omega_{Ab}^{AB} \omega_{aB}^{AB} \omega_{ab}^{AB} \omega_{AB}^{AB} \omega_{Ab}^{Ab} \omega_{aB}^{Ab} \omega_{ab}^{Ab} \omega_{AB}^{Ab} \omega_{Ab}^{Ab} \omega_{aB}^{Ab} \omega_{ab}^{Ab} \omega_{AB}^{ab} \omega_{Ab}^{ab} \omega_{aB}^{ab} \omega_{ab}^{ab}\}$$

with $\text{diag}\{\}$ signaling a diagonal matrix.

Cockerham's model: In the absence of zygotic associations ($\mathbf{\Omega} = \mathbf{0}$ and $\mathbf{P} = \mathbf{\Psi}$), COCKERHAM (1954) developed a regression model to partition the two-locus genotypic value into eight orthogonal scales,

$$\mathbf{G} = \mathbf{1}\mu + \sum_{t=1}^8 \mathbf{w}_t \beta_t, \quad (4)$$

where $\mathbf{G} = [G_{AB}^{AB} G_{Ab}^{AB} G_{aB}^{AB} G_{ab}^{AB} G_{AB}^{Ab} G_{Ab}^{Ab} G_{aB}^{Ab} G_{ab}^{Ab} G_{AB}^{ab} G_{Ab}^{ab} G_{aB}^{ab} G_{ab}^{ab}]'$ is the vector of genotypic values corresponding to genotypes $[AABB AABb Aabb AaBB AaBb Aabb aaBB aaBb aabb]$, $\mathbf{1}$ is a vector of ones, $\mu = (\mathbf{G}' \mathbf{\Psi} \mathbf{1})$ is the population mean, \mathbf{w}_t is the t th column vector of matrix \mathbf{W} ,

$$\mathbf{W} = \begin{bmatrix} 2p_a & \frac{-1}{8P_A^A} & 2p_b & \frac{-1}{8P_B^B} & 4p_a p_b \\ 2p_a & \frac{-1}{8P_A^A} & p_b - p_B & \frac{1}{4P_B^B} & 2p_a(p_b - p_B) \\ 2p_a & \frac{-1}{8P_A^A} & -2p_B & \frac{-1}{8P_B^B} & -4p_a p_B \\ p_a - p_A & \frac{1}{4P_A^A} & 2p_b & \frac{-1}{8P_B^B} & 2(p_a - p_A)p_b \\ p_a - p_A & \frac{1}{4P_A^A} & p_b - p_B & \frac{1}{4P_B^B} & (p_a - p_A)(p_b - p_B) \\ p_a - p_A & \frac{1}{4P_A^A} & -2p_B & \frac{-1}{8P_B^B} & -2(p_a - p_A)p_B \\ -2p_A & \frac{-1}{8P_A^A} & 2p_b & \frac{-1}{8P_B^B} & -4p_A p_b \\ -2p_A & \frac{-1}{8P_A^A} & p_b - p_B & \frac{1}{4P_B^B} & -2p_A(p_b - p_B) \\ -2p_A & \frac{-1}{8P_A^A} & -2p_B & \frac{-1}{8P_B^B} & 4p_A p_B \end{bmatrix}$$

$$\begin{bmatrix} \frac{-p_a}{4P_B^B} & \frac{-p_b}{4P_A^A} & \frac{1}{64P_A^A P_B^B} \\ \frac{p_a}{2P_B^B} & \frac{p_b - p_B}{8P_A^A} & \frac{-1}{32P_A^A P_B^B} \\ \frac{-p_a}{4P_B^B} & \frac{p_B}{4P_A^A} & \frac{1}{64P_A^A P_B^B} \\ \frac{p_A - p_a}{8P_B^B} & \frac{p_b}{2P_A^A} & \frac{-1}{32P_A^A P_B^B} \\ \frac{(p_a - p_A)}{4P_B^B} & \frac{(p_b - p_B)}{4P_A^A} & \frac{1}{16P_A^A P_B^B} \\ \frac{p_A - p_a}{8P_B^B} & \frac{-p_B}{2P_A^A} & \frac{-1}{32P_A^A P_B^B} \\ \frac{p_A}{4P_B^B} & \frac{-p_b}{4P_A^A} & \frac{1}{64P_A^A P_B^B} \\ \frac{-p_A}{2P_B^B} & \frac{p_B - p_B}{8P_A^A} & \frac{-1}{32P_A^A P_B^B} \\ \frac{p_A}{4P_B^B} & \frac{p_B}{8P_A^A} & \frac{1}{64P_A^A P_B^B} \end{bmatrix}$$

and β_t is the coefficient of partial regression of \mathbf{G} on \mathbf{w}_t ,

$$\beta_t = (\mathbf{w}_t' \mathbf{\Psi} \mathbf{G}) / (\mathbf{w}_t' \mathbf{\Psi} \mathbf{w}_t). \quad (5)$$

It should be noted that the orthogonal scales given in the \mathbf{W} matrix are slight modifications of those by COCKERHAM (1954) to be consistent with more recent uses of these scales (COCKERHAM and ZENG 1996; KAO and ZENG 2002). Like the usual orthogonal contrasts, the orthogonal scales satisfy the two basic requirements: (i) $\mathbf{w}_t' \mathbf{\Psi} \mathbf{1} = 0$ and (ii) $\mathbf{w}_t' \mathbf{\Psi} \mathbf{w}_{t'} = 0$ for $t \neq t'$. In addition, these eight orthogonal scales correspond to the additive

and dominance effects at locus *A* (\mathbf{w}_1 and \mathbf{w}_2), additive and dominance effects at locus *B* (\mathbf{w}_3 and \mathbf{w}_4), and four epistatic effects between the two loci, \mathbf{w}_5 ($= \mathbf{w}_1 \times \mathbf{w}_3$), \mathbf{w}_6 ($= \mathbf{w}_1 \times \mathbf{w}_4$), \mathbf{w}_7 ($= \mathbf{w}_2 \times \mathbf{w}_3$), and \mathbf{w}_8 ($= \mathbf{w}_2 \times \mathbf{w}_4$).

The orthogonal scales in the \mathbf{W} matrix can be used to partition the total genetic variance in a zygotic equilibrium population (σ_G^2) into eight independent components,

$$\sigma_G^2 = \mathbf{G}'\Psi\mathbf{G} - \mu^2 = \sum_{i=1}^8 (\mathbf{w}'_i\Psi\mathbf{w}_i)\beta_i^2 = \sum_{i=1}^8 \sigma_i^2, \quad (6a)$$

where

$$\sigma_i^2 = (\mathbf{w}'_i\Psi\mathbf{G})^2 / (\mathbf{w}'_i\Psi\mathbf{w}_i). \quad (6b)$$

COCKERHAM (1954) gave detailed expressions for σ_i^2 's in an inbred population where Hardy-Weinberg disequilibrium is measured in terms of inbreeding coefficient. He also noted that in the absence of inbreeding ($f_A = f_B = 0$ and $D_A = D_B = 0$), these variance components became those in a Hardy-Weinberg equilibrium population (KEMPTHORNE 1957; CROW and KIMURA 1970; LYNCH and WALSH 1998). Because the different scales are orthogonal (*i.e.*, $\mathbf{w}'_t\Psi\mathbf{w}_{t'} = 0$ for all $t \neq t'$), there are no covariances between them.

When $p_A = p_B = 1/2$ and $D_A = D_B = 0$ (*i.e.*, the gene and genotypic frequencies are those in an F_2 population), Cockerham's model is reduced to the F_2 -metric model. In this particular case where the F_2 population is in Hardy-Weinberg and zygotic equilibria [$\mathbf{\Omega} = \mathbf{0}$ and $\mathbf{P} = \Psi = (1/16)\text{diag}\{1\ 2\ 1\ 2\ 4\ 2\ 1\ 2\ 1\}$], the expressions for partial regression coefficients in Cockerham's model (β_i 's in Equation 5) are greatly simplified and are identified with specific genic effects,

$$\begin{aligned} a_1 = \beta_1 &= (\bar{G}_A^+ - \bar{G}_A^-) / 2 \\ d_A = \beta_2 &= (2\bar{G}_A^+ - \bar{G}_A^+ - \bar{G}_A^-) / 2 \\ a_B = \beta_3 &= (\bar{G}_B^+ - \bar{G}_B^-) / 2 \\ d_B = \beta_4 &= (2\bar{G}_B^+ - \bar{G}_B^+ - \bar{G}_B^-) / 2 \\ aa = \beta_5 &= (G_{AB}^{AB} - G_{Ab}^{Ab} - G_{aB}^{aB} + G_{ab}^{ab}) / 4 \\ ad = \beta_6 &= (-G_{AB}^{AB} + 2G_{Ab}^{Ab} - G_{aB}^{aB} + G_{ab}^{ab} + 2G_{ab}^{ab} + G_{ab}^{ab}) / 4 \\ da = \beta_7 &= (-G_{AB}^{AB} + 2G_{Ab}^{Ab} - G_{aB}^{aB} + G_{ab}^{ab} + 2G_{ab}^{ab} + G_{ab}^{ab}) / 4 \\ dd = \beta_8 &= (G_{AB}^{AB} - 2G_{Ab}^{Ab} + G_{aB}^{aB} - 2G_{ab}^{ab} + 4G_{ab}^{ab} - 2G_{ab}^{ab} + G_{ab}^{ab} - 2G_{ab}^{ab} + G_{ab}^{ab}) / 4 \end{aligned} \quad (7)$$

(COCKERHAM and ZENG 1996; CHEVERUD 2000; KAO and ZENG 2002), where $\bar{G}_{v^u}^y = \sum_{y,z} P_{vz}^{uy} G_{vz}^{uy}$ and $\bar{G}_z^y = \sum_{u,v} P_{vz}^{uy} G_{vz}^{uy}$ are the average values of genotypes uv (AA , Aa , or aa) at locus *A* and yz (BB , Bb , or bb) at locus *B*, respectively. This set of genic effects along with the mean (μ) can be expressed in terms of another set under the F_∞ -metric model through a simple translation as given in VAN DER VEEN (1959). A more detailed comparison of the F_2 -metric model with the F_∞ -metric model is made later in F_2 -METRIC *vs.* F_∞ -METRIC MODELS.

Departure from Cockerham's model: In the presence of zygotic associations ($\mathbf{\Omega} \neq \mathbf{0}$ and $\mathbf{P} = \Psi + \mathbf{\Omega}$), however, a residual term (ϵ) needs to be added to Cockerham's regression model and Equation 4 becomes

$$\mathbf{G} = \mathbf{1}\mu' + \sum_{i=1}^8 \mathbf{w}_i\beta_i + \mathbf{1}\epsilon, \quad (8)$$

where μ' is the mean of the zygotic disequilibrium population, $\mu' = \mathbf{G}'\mathbf{P}\mathbf{1} = \mu + \mathbf{G}'\mathbf{\Omega}\mathbf{1}$. Thus, the total genetic variance in the zygotic disequilibrium population is

$$\begin{aligned} \sigma_G^{\prime 2} &= \mathbf{G}'\mathbf{P}\mathbf{G} - (\mu')^2 = \mathbf{G}'\Psi\mathbf{G} + \mathbf{G}'\mathbf{\Omega}\mathbf{G} - (\mu')^2 \\ &= \sigma_G^2 + \sigma_\epsilon^2 - 2\mu(\mu' - \mu), \end{aligned} \quad (9)$$

where σ_G^2 is the total genetic variance under Cockerham's model as given in Equation 6a and σ_ϵ^2 is the residual variance arising from zygotic associations. When the means and genetic variances in equilibrium and nonequilibrium populations are known, σ_ϵ^2 can be obtained from Equation 9,

$$\begin{aligned} \sigma_\epsilon^2 &= \sigma_G^{\prime 2} - \sigma_G^2 + 2\mu(\mu' - \mu) \\ &= \sigma_G^{\prime 2} - \sum_{i=1}^8 \sigma_i^2 + 2\mu(\mu' - \mu). \end{aligned} \quad (10a)$$

Alternatively, it can be calculated directly from genotypic values and zygotic associations,

$$\sigma_\epsilon^2 = \mathbf{G}'\mathbf{\Omega}\mathbf{G} - (\mathbf{G}'\mathbf{\Omega}\mathbf{1})^2. \quad (10b)$$

Using the relations among zygotic associations given in Table 1, we have

$$\begin{aligned} \sigma_\epsilon^2 &= \omega_{AB}^{AB}E_1 + \omega_{Ab}^{Ab}E_2 + \omega_{aB}^{aB}E_3 + \omega_{ab}^{ab}E_4 \\ &\quad - [\omega_{AB}^{AB}e_1 + \omega_{Ab}^{Ab}e_2 + \omega_{aB}^{aB}e_3 + \omega_{ab}^{ab}e_4]^2, \end{aligned} \quad (10c)$$

where

$$\begin{aligned} E_1 &= (G_{AB}^{AB})^2 - (G_{Ab}^{Ab})^2 - (G_{aB}^{aB})^2 + (G_{ab}^{ab})^2, \\ E_2 &= (G_{Ab}^{Ab})^2 - (G_{Ab}^{Ab})^2 - (G_{aB}^{aB})^2 + (G_{ab}^{ab})^2, \\ E_3 &= (G_{aB}^{aB})^2 - (G_{aB}^{aB})^2 - (G_{ab}^{ab})^2 + (G_{ab}^{ab})^2, \\ E_4 &= (G_{ab}^{ab})^2 - (G_{ab}^{ab})^2 - (G_{ab}^{ab})^2 + (G_{ab}^{ab})^2, \end{aligned}$$

and

$$\begin{aligned} e_1 &= G_{AB}^{AB} - G_{Ab}^{Ab} - G_{aB}^{aB} + G_{ab}^{ab}, \\ e_2 &= G_{Ab}^{Ab} - G_{Ab}^{Ab} - G_{aB}^{aB} + G_{ab}^{ab}, \\ e_3 &= G_{aB}^{aB} - G_{aB}^{aB} - G_{ab}^{ab} + G_{ab}^{ab}, \\ e_4 &= G_{ab}^{ab} - G_{ab}^{ab} - G_{ab}^{ab} + G_{ab}^{ab}. \end{aligned}$$

It should be noted that while σ_ϵ^2 is called the residual "variance" it is not necessarily positive because the zygotic associations in $\mathbf{\Omega}$ that are the deviations can be either positive or negative.

Application: DOEBLEY *et al.* (1995) identified two QTL, UMC107 (designated as locus *A*) and BV302 (designated as locus *B*), controlling differences in plant and inflorescence architecture between cultivated maize (*Zea mays* ssp. *mays*) and teosinte (*Z. mays* ssp. *parviglumis*) in the BC₃F₂ population (Teosinte-MIL × Teosinte-M3L) derived from an original cross of Reventador maize × *parviglumis* teosinte. A total of nine morphological and inflorescence traits were measured. Two of those traits, average length of vegetative internodes in the primary lateral branch (LBIL) and percentage of cupules lacking the pedicellate spikelet (PEDS), have been reanalyzed to examine the magnitude of epistasis between loci *A* and *B* (LYNCH and WALSH 1998; GOODNIGHT 2000; KAO and ZENG 2002). However, all these analyses have assumed $p_A = p_B = \frac{1}{2}$ and $D_A = D_B = 0$ as in an F₂ population with two uncorrelated loci. We test to determine if such assumption holds for Doebley *et al.*'s data. We then use the actual genotypic frequencies as given in Table 7 of KAO and ZENG (2002) and the trait values in Table 8 of DOEBLEY *et al.* (1995) to illustrate the application of our new model developed above.

The observed and expected genotypic frequencies and their differences (disequilibria) at loci UMC107 and BV302 are given in Table 2. A sample of 183 individuals from the BC₃F₂ population was recorded in DOEBLEY *et al.* (1995) but as KAO and ZENG (2002) pointed out, 21 individuals had a missing value for one or more traits measured. Thus, a sample size of 161 individuals is used for our analysis. The expected values are derived assuming Cockerham's reference population (CRP) and the F₂ population. For a single locus, CRP may not be in Hardy-Weinberg equilibrium with arbitrary gene frequencies whereas the F₂ population is in Hardy-Weinberg equilibrium with the gene frequencies being constrained to the fixed value of one-half. Thus, the deviations of observed from expected single-locus genotypic frequencies in CRP are simply Hardy-Weinberg disequilibria but such deviations in the F₂ population are Hardy-Weinberg disequilibria plus the differences between expected values in the two reference populations. For example, for genotype *AA* at locus *A*, the difference between observed and expected genotypic frequencies in CRP is $0.2050 - 0.2228 = -0.0178$ but the corresponding difference under the F₂ population is $0.2050 - 0.2500 = -0.0450$, which equals to $-0.0178 + (0.2228 - 0.2500)$. The expected two-locus genotypic frequencies in CRP are the products of observed single-locus genotypic frequencies. For example, the expected frequency of genotype *AABb* is $(0.2050 \times 0.5280) = 0.1082$. In contrast, the expected two-locus genotypic frequencies in the F₂ population are the products of expected single-locus genotypic frequencies. According to the chi-square tests, neither Hardy-Weinberg disequilibria at individual loci nor zygotic associations between the two loci are significant. The observed genotypic frequencies fit well

to those expected in CRP and the F₂ population. Because of small numbers of individuals for some genotypes (*e.g.*, *AAbb*), we also carry out Monte Carlo exact tests (WEIR 1996). The "exact" *P* values are very similar to the asymptotic *P* values from the chi-square tests. Thus, the assumption of F₂ gene and genotypic frequencies used for analyzing Doebley *et al.*'s data (LYNCH and WALSH 1998; GOODNIGHT 2000; KAO and ZENG 2002) is probably justified.

Despite insignificant departures from Hardy-Weinberg and zygotic equilibria, we retain these disequilibria for illustrating the use of our new model. The mean values of nine morphological and inflorescence traits for each of the nine genotypic classes as given in DOEBLEY *et al.*'s (1995) Table 8 are used for our analysis. We partition the genotypic values into eight genic effects defined under Cockerham's model and the F₂-metric model plus a residual effect (Table 3). The estimated genic effects under the F₂-metric model for LBIL are identical to those in LYNCH and WALSH (1998, pp. 88–90) and in KAO and ZENG's (2002) Table 9 (under the "Cockerham" column), barring rounding errors. In most cases, the values under Cockerham's model and the F₂-metric model are similar as would be expected from insignificant Hardy-Weinberg and zygotic disequilibria (*cf.* Table 2). However, there are a few exceptions, particularly with the residual effects for LBIL and percentage of staminate spikelets in primary lateral inflorescence (STAM), where the two models differ most. These differences are likely attributable to the deviations of gene frequencies from one-half under Cockerham's model.

Table 4 presents the partitioning of total genetic variance (σ_c^2) using (9) for each of the nine traits in the presence of zygotic associations. Two different partitions are obtained, depending on whether the equilibrium portion of total variance (σ_c^2) is calculated using Cockerham's model or the F₂-metric model. For example, for LBIL, $\sigma_c^2 = \sigma_c^2 + \sigma_e^2 - 2\mu(\mu' - \mu) = 354.395 - 29.588 - 2(56.005)(56.684 - 56.005) = 360.790$ under Cockerham's model but $\sigma_c^2 = 381.209 - 157.364 - 2(56.888)(56.684 - 56.888) = 360.790$ under the F₂-metric model. In other words, since the two models are based on different reference populations, they have different population means and genetic variances under zygotic equilibrium (μ and σ_c^2). Like the patterns of estimated genic effects in Table 3, contributions (percentages) due to individual variance components of σ_c^2 under Cockerham's model *vs.* the F₂-metric model are similar in most cases; the residual variances are similar as well with the two notable exceptions for LBIL and STAM. Further inspection on contributions due to individual variance components of σ_c^2 under both models shows that additive effects are predominant components for all traits except for PEDS whereas dominance and epistatic effects are minor. Epistasis is most important for PEDS with its effects accounting for nearly 40% of the total equilibrium variance under both models.

TABLE 2
One-locus and two-locus genotypic frequencies and chi-square tests in a population
derived from a cross of Reventador maize × *parviglumis* teosinte

Genotype	Observed	CRP		F ₂ population	
		Expected	Disequilibrium	Expected	Disequilibrium
Locus A					
AA	0.2050	0.2228	-0.0178	0.2500	-0.0450
Aa	0.5342	0.4984	0.0358	0.5000	0.0342
aa	0.2609	0.2788	-0.0179	0.2500	0.0109
	Test for HWD at locus A:	$\chi^2 = 0.83$; d.f. = 2; $P = 0.6608$		$\chi^2 = 1.76$; d.f. = 2; $P = 0.4152$	
Locus B					
BB	0.2422	0.2562	-0.0140	0.2500	-0.0078
Bb	0.5280	0.5000	0.0280	0.5000	0.0280
bb	0.2298	0.2438	-0.0140	0.2500	-0.0202
	Test for HWD at locus B:	$\chi^2 = 0.50$; d.f. = 2; $P = 0.7775$		$\chi^2 = 0.55$; d.f. = 2; $P = 0.7585$	
Loci A-B					
AABB	0.0497	0.0497	0.0000	0.0625	-0.0128
AABb	0.1367	0.1082	0.0284	0.1250	0.0116
AAbb	0.0186	0.0471	-0.0285	0.0625	-0.0439
AaBB	0.1242	0.1294	-0.0052	0.1250	-0.0008
AaBb	0.2609	0.2820	-0.0211	0.2500	0.0109
Aabb	0.1491	0.1228	0.0263	0.1250	0.0241
aaBB	0.0683	0.0632	0.0051	0.0625	0.0058
aaBb	0.1304	0.1377	-0.0073	0.1250	0.0054
aabb	0.0621	0.0600	0.0022	0.0625	-0.0004
	Test for ZA for loci A-B:	$\chi^2 = 5.31$; d.f. = 4; $P = 0.2568$		$\chi^2 = 4.19$; d.f. = 4; $P = 0.3805$	

Observed and expected genotypic frequencies and their differences (disequilibria) at two diallelic loci, UMC107 (designated as locus A) and BV302 (designated as locus B), in the BC₃F₂ population (Teosinte-MIL × Teosinte-M3L) derived from an original cross of Reventador maize × *parviglumis* teosinte as described by DOEBLEY *et al.* (1995) and KAO and ZENG (2002) are shown. The expected values are derived under Cockerham's reference population (CRP) and the F₂ population. Chi-square tests are carried out to determine the presence of Hardy-Weinberg disequilibria (HWD) at individual loci and zygotic associations (ZA) between the two loci. Sample size is $n = 161$.

It is evident from Table 4 that the residual portion of the total genetic variance (σ_e^2) is generally small compared to the equilibrium portion (σ_c^2). This is of course due largely to insignificant zygotic disequilibrium between the two interacting loci, UMC107 and BV302 (Table 2). To gauge the impact of zygotic disequilibrium on the magnitude of σ_e^2 , we let the Teosinte-MIL × Teosinte-M3L population be at the maximum level of zygotic disequilibrium that is obtainable if double homozygotes and heterozygotes are equally frequent in the absence of single heterozygotes at both loci (YANG 2003). One of such possibilities is that the required P matrix would be estimated by $(1/161)\text{diag}\{20\ 0\ 20\ 0\ 81\ 0\ 20\ 0\ 20\}$. With this particular genotypic distribution, the magnitudes of σ_e^2 are larger for all nine traits and for some traits are as much as or even greater than those of σ_c^2 (results not presented). For example, $\sigma_e^2 = -425.448$ for LBIL under the F₂-metric model with the

maximum zygotic disequilibrium, compared to $\sigma_e^2 = -157.364$ with insignificant zygotic disequilibrium as shown in Table 4; in both cases, $\sigma_c^2 = 381.209$. Evidently, the magnitude of zygotic association affects the partitioning of the total genetic variance. Further research is needed to examine detailed relationships between nonextreme values of zygotic disequilibrium and the residual variance.

F₂-METRIC VS. F_∞-METRIC MODELS

Conceptual distinction: As noted above, Cockerham's model is reduced to the F₂-metric model for F₂ and other populations derived from a cross between two inbred lines (*cf.* Equation 6). However, since the 1950s, the F₂-metric model has coexisted with another popular model of gene action, the F_∞-metric model. In essence, genic effects under Cockerham's model or the F₂-metric

TABLE 3
Partitioning the two-locus genotypic values in a population derived from a cross of
Reventador maize \times *parviglumis* teosinte

Trait ^a	Genic effects								
	Locus UMC107		Locus BV302		Epistasis between UMC107-BV302				ϵ
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	
CUPL	1.333	-0.004	1.479	0.175	0.119	0.230	0.350	-0.295	-0.007
	1.325	0.000	1.475	0.200	0.100	0.250	0.350	-0.300	0.044
CUPR	-1.756	0.229	-1.390	-0.019	0.246	0.066	-0.053	0.246	-0.004
	-1.775	0.225	-1.375	-0.025	0.250	0.050	-0.050	0.250	-0.094
DISA	-1.040	-0.058	-0.880	0.146	0.164	-0.010	-0.176	0.221	-0.004
	-1.038	-0.063	-0.863	0.138	0.175	-0.025	-0.175	0.225	-0.053
INNO	-0.246	-0.078	-0.271	0.082	0.015	-0.100	-0.178	0.369	0.006
	-0.238	-0.088	-0.263	0.063	0.025	-0.125	-0.175	0.375	-0.003
LBIL	14.339	-4.329	19.525	-5.200	3.213	-19.397	3.951	-17.764	0.321
	15.100	-3.925	19.475	-5.625	2.700	-18.300	3.750	-18.050	1.204
LIBN	2.158	-0.669	-0.536	0.326	-0.447	-0.571	-0.856	1.181	0.032
	2.225	-0.700	-0.525	0.250	-0.400	-0.650	-0.850	1.200	0.178
PEDS	-1.030	-0.684	-1.138	-1.389	1.406	1.791	1.214	1.476	-0.024
	-1.050	-0.750	-1.125	-1.350	1.350	1.700	1.250	1.500	-0.004
STAM	32.910	-0.141	2.303	5.026	3.868	-8.462	-7.815	11.564	0.226
	33.125	-0.375	2.875	4.125	4.250	-9.250	-7.750	11.750	1.920
YOKE	-1.024	-0.103	-0.913	-0.152	0.477	0.050	-0.346	0.000	-0.026
	-1.025	-0.100	-0.875	-0.150	0.500	0.050	-0.350	0.000	-0.064

Partitioning the genotypic values for two interacting loci, UMC107 and BV302, affecting nine morphological traits in the BC₃F₂ population (Teosinte-MIL \times Teosinte-M3L) derived from an original cross of Reventador maize \times *parviglumis* teosinte as described by DOEBLEY *et al.* (1995) and KAO and ZENG (2002) is shown. The partitions for each trait include eight genic effects ($\beta_1 - \beta_8$) under Cockerham's model (top) and the F₂-metric model (bottom) plus a residual (ϵ) arising from zygotic associations.

^a CUPL, average length of cupules (internodes) in the inflorescence (in millimeters); CUPR, number of cupules in a single rank of the ear; DISA, tendency of ear to shatter (1-5 scale), 1 = 100% disarticulating, 5 = 0% disarticulating; INNO, number of vegetative internodes in the lateral branch; LBIL, average length of vegetative internodes in the primary lateral branch (in millimeters); LIBN, number of branches in primary lateral inflorescence; PEDS, percentage of cupules lacking the pedicellate spikelet; STAM, percentage of staminate spikelets in primary lateral inflorescence; and YOKE, degree to which the fruitcases are in yoked pairs (1-5 scale). 1, 0% yoked fruitcases; 5, 100% yoked fruitcases.

model are deviations from the mean of a noninbred equilibrium population whereas genic effects under the F₂-metric model are contrasts among genotypes without reference to any population. For example, the additive effect at locus A under Cockerham's model or the F₂-metric model (β_1) is a weighted average of the two differences ($\bar{G}_A^A - \bar{G}_a^A$) and ($\bar{G}_a^A - \bar{G}_a^a$) corresponding to the comparisons of genotypes, AA - Aa and Aa - aa, with each difference representing an effect of replacing a by A. The dominance effect (β_2) would be present, *i.e.*, $\beta_2 = (\bar{G}_A^A - \bar{G}_a^A) - (\bar{G}_a^A - \bar{G}_a^a) \neq 0$, unless the two differences are exactly the same. On the other hand, the two genic effects at locus A under the F₂-metric model are one-half the difference between the two homozygotes and the deviation of the heterozygote from the homozygote mean, respectively (*e.g.*, MATHER and JINKS 1982). The F₂-metric model is also referred to as the homozygote-based model (WRIGHT 1987).

VAN DER VEEN (1959) gave three reasons why the F₂-metric model would be preferred. First, nonzero coef-

ficients of the genic effects defined under the F₂-metric model are uniquely associated with homozygosity and heterozygosity at different loci. Second, the expressions for describing heterosis and other genetic phenomena are simpler and more surveyable under the F₂-metric model. Finally the F₂-metric model leads to simpler and symmetrical conditions and descriptions of relationships identified for classical F₂ segregation ratios with epistasis. In contrast, KAO and ZENG (2002) recently argued against the use of the F₂-metric model particularly in QTL mapping studies because there is possible bias in estimating allelic effects in the presence of non-allelic (epistatic) effects due to the fact that the genic effects defined under the F₂-metric model in an F₂ population are not orthogonal. In what follows, we clarify the relationships and differences between the two models through theoretical analysis and numerical evaluation.

Parameter relations: With two uncorrelated loci in an F₂ population ($p_A = p_B = 1/2$ and $D_A = D_B = 0$), the values of nine possible genotypes can be expressed as

TABLE 4

Partitioning the two-locus genetic variance in a population derived from a cross of Reventador maize \times *parviglumis* teosinte

Trait ^a	μ'	μ	Percentage of σ_G^2										σ_G^2	σ_ϵ^2	$\sigma_G'^2$
			Locus UMC107		Locus BV302		Epistasis between UMC107-BV302								
			σ_1^2	σ_2^2	σ_3^2	σ_4^2	σ_5^2	σ_6^2	σ_7^2	σ_8^2					
CUPL	9.031	9.024	43.47	0.00	54.57	0.41	0.16	0.32	0.77	0.29	1.891	0.258	2.026		
	9.031	9.075	43.74	0.00	54.20	0.50	0.12	0.39	0.76	0.28	2.007	-0.778	2.026		
CUPR	9.131	9.127	60.18	0.56	38.50	0.00	0.56	0.02	0.01	0.16	2.370	0.225	2.514		
	9.131	9.038	61.69	0.50	37.02	0.01	0.61	0.01	0.01	0.15	2.554	1.651	2.514		
DISA	2.860	2.855	56.54	0.10	41.33	0.60	0.66	0.00	0.42	0.35	0.885	0.058	0.919		
	2.860	2.806	57.83	0.10	39.97	0.51	0.82	0.01	0.41	0.34	0.931	0.288	0.919		
INNO	2.922	2.928	35.16	1.95	43.60	2.12	0.06	1.47	4.76	10.88	0.080	-0.035	0.082		
	2.922	2.919	35.13	2.38	42.92	1.22	0.19	2.43	4.77	10.95	0.080	0.019	0.082		
LBIL	55.684	56.005	26.85	1.34	50.76	1.91	0.64	12.32	0.53	5.65	354.395	-29.588	360.790		
	55.684	56.888	29.91	1.01	49.75	2.08	0.48	10.98	0.46	5.34	381.209	-157.364	360.790		
LIBN	2.147	2.179	80.17	4.21	5.05	0.99	1.62	1.41	3.26	3.29	2.688	-0.401	2.427		
	2.147	2.325	81.85	4.05	4.56	0.52	1.32	1.75	2.99	2.98	3.024	-1.424	2.427		
PEDS	0.879	0.855	17.40	4.19	21.66	17.14	15.28	13.18	6.24	4.90	2.823	0.145	2.927		
	0.879	0.875	18.79	4.79	21.57	15.53	15.53	12.32	6.66	4.79	2.933	0.000	2.927		
STAM	42.267	42.493	93.26	0.00	0.47	1.18	0.61	1.55	1.36	1.58	537.363	-17.111	539.488		
	42.267	44.188	93.24	0.01	0.70	0.72	0.77	1.82	1.28	1.47	588.402	-218.632	539.488		
YOKE	2.189	2.163	51.02	0.28	41.34	0.61	5.22	0.03	1.50	0.00	0.951	0.153	0.993		
	2.189	2.125	52.83	0.25	38.50	0.57	6.29	0.03	1.54	0.00	0.994	0.270	0.993		

Partitioning the total genetic variance ($\sigma_G'^2$) for two interacting loci, UMC107 and BV302, affecting nine morphological traits in the BC₃F₂ population (Teosinte-M1L \times Teosinte-M3L) derived from an original cross of Reventador maize \times *parviglumis* teosinte as described by DOEBLEY *et al.* (1995) and KAO and ZENG (2002) is shown. The genetic variance with zygotic equilibrium (σ_G^2) for each trait is partitioned into eight independent components ($\sigma_1^2 - \sigma_8^2$) under Cockerham's model (top) and the F₂-metric model (bottom).

^aDetailed descriptions of these traits are given in Table 3.

$$\mathbf{G} = \mathbf{M}_\delta \mathbf{E}_\delta, \tag{11}$$

where δ indicates whether the genic effects are defined under the F₂-metric model ($\delta = 0$) or F_∞-metric model ($\delta = 1$); $\mathbf{E}_\delta = [\mu_{(\delta)} \ a_{A(\delta)} \ d_{A(\delta)} \ a_{B(\delta)} \ d_{B(\delta)} \ aa_{(\delta)} \ ad_{(\delta)} \ da_{(\delta)} \ dd_{(\delta)}]'$ contains the parameters defined in (6) for $\delta = 0$ and

$$\mathbf{M}_\delta = \begin{bmatrix} 1 & 1 & \frac{\delta-1}{2} & 1 & \frac{\delta-1}{2} & 1 & \frac{\delta-1}{2} & \frac{\delta-1}{2} & \frac{(\delta-1)^2}{4} \\ 1 & 1 & \frac{\delta-1}{2} & 0 & \frac{\delta+1}{2} & 0 & \frac{\delta+1}{2} & 0 & \frac{\delta^2-1}{4} \\ 1 & 1 & \frac{\delta-1}{2} & -1 & \frac{\delta-1}{2} & -1 & \frac{\delta-1}{2} & -\frac{\delta-1}{2} & \frac{(\delta-1)^2}{4} \\ 1 & 0 & \frac{\delta+1}{2} & 1 & \frac{\delta-1}{2} & 0 & 0 & \frac{\delta+1}{2} & \frac{\delta^2-1}{4} \\ 1 & 0 & \frac{\delta+1}{2} & 0 & \frac{\delta+1}{2} & 0 & 0 & 0 & \frac{(\delta+1)^2}{4} \\ 1 & 0 & \frac{\delta+1}{2} & -1 & \frac{\delta-1}{2} & 0 & 0 & -\frac{\delta+1}{2} & \frac{\delta^2-1}{4} \\ 1 & -1 & \frac{\delta-1}{2} & 1 & \frac{\delta-1}{2} & -1 & -\frac{\delta-1}{2} & \frac{\delta-1}{2} & \frac{(\delta-1)^2}{4} \\ 1 & -1 & \frac{\delta-1}{2} & 0 & \frac{\delta+1}{2} & 0 & -\frac{\delta+1}{2} & 0 & \frac{\delta^2-1}{4} \\ 1 & -1 & \frac{\delta-1}{2} & -1 & \frac{\delta-1}{2} & 1 & -\frac{\delta-1}{2} & -\frac{\delta-1}{2} & \frac{(\delta-1)^2}{4} \end{bmatrix} \tag{12}$$

Obviously, $\mathbf{G} = \mathbf{M}_0 \mathbf{E}_0$ and $\mathbf{G} = \mathbf{M}_1 \mathbf{E}_1$ describe the partitioning of genotypic values under F₂-metric and F_∞-metric models, respectively. Since the F₂-metric model is a

special case of Cockerham's model, columns 2–9 in the \mathbf{M}_0 matrix are identical to the \mathbf{W} matrix when $p_A = p_B = \frac{1}{2}$ and $D_A = D_B = 0$. The translation of parameters from F₂-metric into F_∞-metric can be made through the relation $\mathbf{E}_0 = \mathbf{T} \mathbf{E}_1$ (VAN DER VEEN 1959), where

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{13}$$

Thus, the same genotypic values can be partitioned into the mean and eight genic effects defined under either the F₂-metric or the F_∞-metric model so long as appropriate translations are identified,

$$\mathbf{G} = \mathbf{M}_0 \mathbf{E}_0 = \mathbf{M}_0 \mathbf{T} \mathbf{E}_1 = \mathbf{M}_1 \mathbf{E}_1 = \mathbf{M}_1 \mathbf{T}^{-1} \mathbf{E}_0. \tag{14}$$

In addition, it is quite obvious from (14) that $\mathbf{M}_0 = \mathbf{M}_1 \mathbf{T}^{-1}$ and $\mathbf{M}_1 = \mathbf{M}_0 \mathbf{T}$.

TABLE 5
Coded genotypic values and F₂ segregation ratios at two unlinked loci (A and B)
for 10 nonepistatic models and 10 epistatic models

Model no.	Genotype									F ₂ segregation ratio
	AABB	AABb	AAbb	AaBB	AaBb	Aabb	aaBB	aaBb	aabb	
1	4x	3x	2x	3x	2x	x	2x	x	0	1:4:6:4:1
2	4x	4x	2x	4x	4x	2x	2x	2x	0	9:6:1
3	4x	2x	2x	2x	0	0	2x	0	0	1:6:9
4	4x	5x	2x	5x	6x	3x	2x	3x	0	4:4:1:4:2:1
5	4x	x	2x	x	-2x	-x	2x	-x	0	1:2:4:1:4:4
6	6x	4x	2x	5x	3x	x	4x	2x	0	1:2:3:4:3:2:1
7	6x	6x	2x	6x	6x	2x	4x	4x	0	9:3:3:1
8	6x	2x	2x	4x	0	0	4x	0	0	1:3:3:9
9	6x	8x	2x	7x	9x	3x	4x	6x	0	4:2:2:3:1:2:1:1
10	6x	0	2x	3x	-3x	-x	4x	-2x	0	1:1:2:1:3:2:2:4
11	x	x	0	x	x	0	0	0	0	9:7
12	x	wx	0	wx	wx	0	0	0	0	1:8:7
13	x	x	x	x	x	x	x	x	0	15:1
14	x	x	0	x	x	0	x	x	x	13:3
15	2x	2x	0	2x	2x	0	x	x	x	9:3:4
16	2x	2x	2x	2x	2x	2x	x	x	0	12:3:1
17	2x	x	0	2x	x	0	2x	2x	2x	7:6:3
18	2x	3x	0	2x	3x	0	x	x	0	6:3:3:4
19	x	3x	2x	x	3x	2x	3x	0	2x	7:4:3:2
20	9x	9x	-3x	9x	9x	-3x	-3x	-3x	x	9:1:6

With F₂ gene and genotypic frequencies [*i.e.*, $\mathbf{\Omega} = \mathbf{0}$ and $\mathbf{P} = \mathbf{\Psi} = (\frac{1}{16})\text{diag}\{1\ 2\ 1\ 2\ 4\ 2\ 1\ 2\ 1\}$] the population mean is

$$\mathbf{G}'\mathbf{\Psi}\mathbf{1} = \begin{cases} \mu_{(0)}, & \text{if } F_2\text{-metric} \\ \mu_{(1)} + \frac{1}{2}d_{A(1)} + \frac{1}{2}d_{B(1)} + \frac{1}{4}dd_{(1)}, & \text{if } F_\infty\text{-metric} \end{cases} \quad (15)$$

The total equilibrium genetic variance can be partitioned under the F₂-metric model,

$$\begin{aligned} \sigma_G^2 &= \mathbf{G}'\mathbf{\Psi}\mathbf{G} - \mu_0^2 = \mathbf{E}'_0 \mathbf{M}'_0 \mathbf{\Psi} \mathbf{M}_0 \mathbf{E}_0 - \mu_{(0)}^2 \\ &= \frac{a_{A(0)}^2}{2} + \frac{d_{A(0)}^2}{4} + \frac{a_{B(0)}^2}{2} + \frac{d_{B(0)}^2}{4} + \frac{aa_{(0)}^2}{4} + \frac{ad_{(0)}^2}{8} + \frac{da_{(0)}^2}{8} + \frac{dd_{(0)}^2}{16} \end{aligned} \quad (16a)$$

and under F_∞-metric model,

$$\begin{aligned} \sigma_G^2 &= \mathbf{E}'_1 \mathbf{M}'_1 \mathbf{\Psi} \mathbf{M}_1 \mathbf{E}_1 - [\mu_{(1)} + \frac{1}{2}d_{A(1)} + \frac{1}{2}d_{B(1)} + \frac{1}{4}dd_{(1)}]^2 \\ &= \frac{[a_{A(1)} + \frac{1}{2}ad_{(1)}]^2}{2} + \frac{[d_{A(1)} + \frac{1}{2}dd_{(1)}]^2}{4} + \frac{[a_{B(1)} + \frac{1}{2}da_{(1)}]^2}{2} + \frac{[d_{B(1)} + \frac{1}{2}dd_{(1)}]^2}{4} \\ &\quad + \frac{aa_{(1)}^2}{4} + \frac{ad_{(1)}^2}{8} + \frac{da_{(1)}^2}{8} + \frac{dd_{(1)}^2}{16} \end{aligned} \quad (16b)$$

It is evident from comparing (16a) and (16b) that the two models differ in partitioning the total genetic variance σ_G^2 in an F₂ population. Because, with F₂ gene and genotypic frequencies, the genic effects under the F_∞-metric model are not orthogonal the additive and dominance variances include covariances between allelic and nonallelic effects plus a portion of appropriate epistatic variances. It is also evident that the eight independent

components of σ_G^2 as identified in Cockerham's model have simpler forms under the F₂-metric and F_∞-metric models,

$$\begin{aligned} \sigma_1^2 &= \frac{a_{A(0)}^2}{2} = \frac{[a_{A(1)} + \frac{1}{2}ad_{(1)}]^2}{2}, & \sigma_5^2 &= \frac{aa_{(0)}^2}{4} = \frac{aa_{(1)}^2}{4}, \\ \sigma_2^2 &= \frac{d_{A(0)}^2}{4} = \frac{[d_{A(1)} + \frac{1}{2}dd_{(1)}]^2}{4}, & \sigma_6^2 &= \frac{ad_{(0)}^2}{8} = \frac{ad_{(1)}^2}{8}, \\ \sigma_3^2 &= \frac{a_{B(0)}^2}{2} = \frac{[a_{B(1)} + \frac{1}{2}da_{(1)}]^2}{2}, & \sigma_7^2 &= \frac{da_{(0)}^2}{8} = \frac{da_{(1)}^2}{8}, \\ \sigma_4^2 &= \frac{d_{B(0)}^2}{4} = \frac{[d_{B(1)} + \frac{1}{2}dd_{(1)}]^2}{4}, & \sigma_8^2 &= \frac{dd_{(0)}^2}{16} = \frac{dd_{(1)}^2}{16} \end{aligned} \quad (17)$$

Numerical evaluation: We evaluate 20 genetic models exhibiting different segregating ratios for two independent loci in F₂ populations (Table 5). These models are chosen to represent varying levels of additive, dominance, and epistatic effects under the F₂-metric and F_∞-metric models. YANG and BAKER (1990) also examined these genetic models, but only with the F_∞-metric model. The first 10 models involve no epistasis and they are variants of the Mendelian F₂ ratio of 9:3:3:1 for two independently segregating loci (model 7). These models vary from those of pure additive genic effects (models 1 and 6) to those of strong dominance (*e.g.*, models 9 and 10). Models 11–19 are those with classical epistatic ratios that can be found in F₂ populations (STRICKBERGER

TABLE 6

Coefficients required for equality among eight genic effects derived from the two-locus F_2 -metric ($\delta = 0$) and F_{∞} -metric ($\delta = 1$) for each of 20 genetic models

Model no. ^a	δ	$a_{A(\delta)}$	$d_{A(\delta)}$	$a_{B(\delta)}$	$d_{B(\delta)}$	$aa_{(\delta)}$	$ad_{(\delta)}$	$da_{(\delta)}$	$dd_{(\delta)}$
1	0	1	0	1	0	0	0	0	0
	1	1	0	1	0	0	0	0	0
2	0	1	1	1	1	0	0	0	0
	1	1	1	1	1	0	0	0	0
3	0	1	-1	1	-1	0	0	0	0
	1	1	-1	1	-1	0	0	0	0
4	0	2	1	2	1	0	0	0	0
	1	2	1	2	1	0	0	0	0
5	0	2	-1	2	-1	0	0	0	0
	1	2	-1	2	-1	0	0	0	0
6	0	2	0	1	0	0	0	0	0
	1	2	0	1	0	0	0	0	0
7	0	2	2	1	1	0	0	0	0
	1	2	2	1	1	0	0	0	0
8	0	2	-2	1	-1	0	0	0	0
	1	2	-2	1	-1	0	0	0	0
9	0	4	2	2	1	0	0	0	0
	1	4	2	2	1	0	0	0	0
10	0	4	-2	2	-1	0	0	0	0
	1	4	-2	2	-1	0	0	0	0
11	0	2	2	2	2	3	3	3	3
	1	1	1	1	1	1	1	1	1
12	0	1	2	1	2	1	0	0	1
	1	1	0	1	0	1	0	0	1
13	0	2	2	2	2	-1	-1	-1	-1
	1	1	1	1	1	-1	-1	-1	-1
14	0	-3	-3	1	1	2	2	2	2
	1	-1	-1	1	1	1	1	1	1
15	0	3	3	1	1	2	2	2	2
	1	0	0	1	1	1	1	1	1
16	0	2	2	10	10	-5	-5	-5	-5
	1	1	1	3	3	-3	-3	-3	-3
17	0	-3	-3	2	0	2	0	2	0
	1	-1	-1	1	0	1	0	1	0
18	0	1	1	$\frac{5}{7}$	$\frac{5}{13}$	$\frac{5}{2}$	$\frac{5}{6}$	$\frac{5}{2}$	$\frac{5}{6}$
	1	1	1	$\frac{1}{3}$	$\frac{1}{5}$	1	$\frac{1}{3}$	1	$\frac{1}{3}$
19	0	1	1	-2	1	-1	$\frac{1}{4}$	-1	$\frac{1}{4}$
	1	1	1	0	1	1	$-\frac{1}{4}$	1	$-\frac{1}{4}$
20	0	1	1	1	1	1	1	1	1
	1	1	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

For example, the equality among the eight genic effects for model 4 is $2a_{A(\delta)} = 2a_{B(\delta)} = d_{A(\delta)} = d_{B(\delta)}$; $aa_{(\delta)} = ad_{(\delta)} = da_{(\delta)} = dd_{(\delta)} = 0$ under both F_2 -metric ($\delta = 0$) and F_{∞} -metric ($\delta = 1$).

^a Genetic models are described in Table 5.

1976). Model 12 is the partial complementary model, a modification of model 11 with $0 \leq w \leq 1$. Model 20 is hypothetical with the relation of $a_{A(0)} = a_{B(0)} = d_{A(0)} = d_{B(0)} = aa_{(0)} = ad_{(0)} = da_{(0)} = dd_{(0)}$ under the F_2 -metric model. The multiplier x in each of the genotypic values in Table 5 can be any nonzero integer. For convenience, we take $x = 1$ and $w = 0.5$ (for model 12) for subsequent calculations.

We partition each genotypic value into the mean and eight genic effects under either F_2 -metric or F_{∞} -

metric, thereby allowing for identification of relationships among the genic effects for each of the 20 genetic models (Table 6). The relationships identified under F_2 -metric and F_{∞} -metric are the same for nonepistatic models (models 1–10) but different for epistatic models (models 11–20). For example, the genic relationships for model 4 are $2a_{A(\delta)} = 2a_{B(\delta)} = d_{A(\delta)} = d_{B(\delta)}$, $aa_{(\delta)} = ad_{(\delta)} = da_{(\delta)} = dd_{(\delta)} = 0$ under either F_2 -metric ($\delta = 0$) or F_{∞} -metric ($\delta = 1$), whereas the genic relationships for model 16 are $2a_{A(0)} = 10a_{B(0)} = 2d_{A(0)} = 10d_{B(0)} = -5aa_{(0)} = -5ad_{(0)} = -5da_{(0)} = -5dd_{(0)}$ under F_2 -metric ($\delta = 0$) but $a_{A(1)} = 3a_{B(1)} = d_{A(1)} = 3d_{B(1)} = -3aa_{(1)} = -3ad_{(1)} = -3da_{(1)} = -3dd_{(1)}$ under F_{∞} -metric ($\delta = 1$). It is evident from Equation 13 that while the conditions of $a_{A(0)} = a_{A(1)}$ and $a_{B(0)} = a_{B(1)}$ always hold regardless of whether or not epistasis is present, those of $d_{A(0)} = d_{A(1)}$ and $d_{B(0)} = d_{B(1)}$ are true only in the absence of epistasis. It should be emphasized that because these genic relationships are derived from the coded genotypic values with x being set to 1 (*cf.* Table 5), they will be multiplied by a constant if x is set to be any other nonzero integer.

The contributions due to individual genic effects are calculated using (17) (Table 7). The nonepistatic models (models 1–10) differ only in the sizes of additive genic effects and levels of dominance at each of the two loci. The 10 epistatic models (models 11–20) show varying levels of epistasis as evident from inspecting percentages of the contributions of individual epistatic effects to the total variance. Model 13 (duplicate epistasis) and model 19 (partial dominant epistasis) were the most epistatic of the models considered. Epistatic effects in model 19 account for almost 75% of the total variance whereas those in model 13 account for ~61% of the total variance, although the constitutions of the four epistatic variance components in both F_2 -metric and F_{∞} -metric are quite different.

DISCUSSION

The models of gene actions including epistasis between different loci developed for conventional quantitative genetics and recent QTL mapping studies do not have the same level of applicability because the genic effects in these models are defined with reference to different types of populations. These models range from the ones for general experimental and natural populations where gene frequencies are arbitrary and Hardy-Weinberg disequilibrium is often present (*e.g.*, Cockerham's model by COCKERHAM 1954) to the ones for special populations derived from a cross between two inbred lines with gene frequencies of one-half (*e.g.*, F_2 -metric model by HAYMAN and MATHER 1955 or F_{∞} -metric model by MATHER and JINKS 1982). The key assumption involved in all the models is that the study population is in zygotic equilibrium (*i.e.*, genotypic frequencies at different loci are products of corresponding single-locus genotypic frequencies). However, zygotic associations

TABLE 7
Percentages of total variance in an equilibrium F₂ population accounted for by
eight individual variance components for 20 genetic models

Model no. ^a	Percentage of total variance in F ₂							
	Locus A		Locus B		Epistasis between loci A and B			
	σ ₁ ²	σ ₂ ²	σ ₃ ²	σ ₄ ²	σ ₅ ²	σ ₆ ²	σ ₇ ²	σ ₈ ²
1	50.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00
2	33.33	16.67	33.33	16.67	0.00	0.00	0.00	0.00
3	33.33	16.67	33.33	16.67	0.00	0.00	0.00	0.00
4	16.67	33.33	16.67	33.33	0.00	0.00	0.00	0.00
5	16.67	33.33	16.67	33.33	0.00	0.00	0.00	0.00
6	20.00	0.00	80.00	0.00	0.00	0.00	0.00	0.00
7	13.33	6.67	53.33	26.67	0.00	0.00	0.00	0.00
8	13.33	6.67	53.33	26.67	0.00	0.00	0.00	0.00
9	6.67	13.33	26.67	53.33	0.00	0.00	0.00	0.00
10	6.67	13.33	26.67	53.33	0.00	0.00	0.00	0.00
11	28.57	14.29	28.57	14.29	6.35	3.17	3.17	1.59
12	34.78	4.35	34.78	4.35	17.39	0.00	0.00	4.35
13	13.33	6.67	13.33	6.67	26.67	13.33	13.33	6.67
14	5.13	2.56	46.15	23.08	10.26	5.13	5.13	2.56
15	5.13	2.56	46.15	23.08	10.26	5.13	5.13	2.56
16	57.47	28.74	2.30	1.15	4.60	2.30	2.30	1.15
17	22.22	11.11	50.00	0.00	11.11	0.00	5.56	0.00
18	13.33	6.67	26.13	45.07	1.07	4.80	0.53	2.40
19	11.11	5.56	2.78	5.56	5.56	44.44	2.78	22.22
20	24.24	12.12	24.24	12.12	12.12	6.06	6.06	3.03

^a Genetic models are described in Table 5.

can arise from physical linkages between different loci or from many evolutionary and demographic processes even for unlinked loci (YANG 2002). Zygotic associations are considered in some general gene action models (e.g., GALLAIS 1974; WEIR and COCKERHAM 1977) but a prohibitively large number of genic effects arising from such models have prevented them from being estimated in experimental quantitative genetics and QTL mapping studies. This article develops a new model that partitions the two-locus genotypic values in a zygotic disequilibrium population into equilibrium and residual portions. The equilibrium portion has eight genic effects that can be defined under those equilibrium models (Cockerham’s model and the F₂-metric and F_∞-metric models). The residual portion is of course due to the presence of zygotic associations and can be a substantial part of the total genetic variance (Table 4). Thus, bias will occur when an equilibrium gene action model is used in modeling and detecting epistatic QTL for a zygotic disequilibrium population.

KAO and ZENG (2002) also considered associations of genotypic frequencies between different loci in modeling epistatic effects but our model differs from theirs in two ways. First our model separates the genotypic values and genetic variance into a portion arising from zero interlocus association (Cockerham’s model) and the other portion arising from the presence of inter-

locus associations; in contrast, Kao and Zeng’s model incorporated interlocus associations into Cockerham’s model (F₂-metric model) by defining genetic and statistical parameters for the same genic effects. Second, Kao and Zeng’s model considered linkage disequilibrium in a Hardy-Weinberg equilibrium F₂ population, which would be only one of several genic disequilibria required to completely characterize zygotic associations; the residual variance (σ_e²) in our model essentially measures the relative importance of genetic variance due to zygotic associations (Equation 10c and see Table 4 for examples). While our model does not expound the relationship between those genic disequilibria and zygotic associations, it can be easily achieved using COCKERHAM and WEIR’s (1973) disequilibrium functions (YANG 2002). For example, the zygotic association for genotype AABB (ω_{AB}^{AB}) can be expressed as

$$\begin{aligned} \omega_{AB}^{AB} &= P_{AB}^{AB} - P_A^A \cdot P_B^B \\ &= 2p_A D_{..B}^{AB} + 2p_B D_{.A.}^{AB} + 2p_A p_B D_{..}^{AB} + 2p_A p_B D_{.B}^A + (D_{..}^{AB})^2 \\ &\quad + (D_{.B}^A)^2 + D_{AB}^{AB}, \end{aligned}$$

where each genic disequilibrium (*D*) is the deviation of a frequency from that based on random association of genes and accounting for any lower-order disequilibria. For example, linkage disequilibrium (*D*_{..}^{AB}) is the deviation of frequency of gamete AB from the product of

frequencies of allele A at locus A and allele B at locus B , $D_{..}^{AB} = P_{..}^{AB} - p_A p_B$, with

$$P_{..}^{AB} = P_{AB}^{AB} + P_{Ab}^{AB} + P_{aB}^{AB} + P_{ab}^{AB}.$$

When zygotes result from random union of gametes as assumed in KAO and ZENG (2002), all nongametic disequilibria including Hardy-Weinberg disequilibrium disappear (*e.g.*, $D_A^A = D_B^A = D_B^B = D_{AB}^{AB} = 0$). In this case, the zygotic association for genotype $AABB$ (ω_{AB}^{AB}) reduces to

$$\omega_{AB}^{AB} = 2p_A p_B D_{..}^{AB} + (D_{..}^{AB})^2,$$

which agrees with the frequency of the same genotype given in Table 6 of KAO and ZENG (2002). The same argument is true for the remaining eight genotypes.

Our theoretical and empirical (Table 4) analyses clearly show substantial differences between Cockerham’s model and the F_2 -metric model. KAO and ZENG (2002) considered the F_2 population where gene frequencies are one-half and genotypic frequencies are those expected under Hardy-Weinberg equilibrium at each of the two loci. In this case, Cockerham’s model is the same as the F_2 -metric. In all other cases, however, the two models would be different. In analyzing the PEDS data of DOEBLEY *et al.* (1995), GOODNIGHT (2000) used Cockerham’s model to calculate eight variance components of the total genetic variance using three gene frequencies at both UMC107 and BV302: $p_A = p_B = 0.25, 0.5$, and 0.75 , but no Hardy-Weinberg disequilibrium was considered at either locus. His case of $p_A = p_B = 0.5$ corresponds to the F_2 -metric as evident from our analysis of the same trait (Table 4). Nevertheless, it is evident from his Table 4 that changes in gene frequencies greatly affect the distributions of variance components due to the genic effects, which is certainly consistent with our finding.

We distinguish the genic effects defined under the F_2 -metric model ($\delta = 0$) from those under the F_∞ -metric model ($\delta = 1$). Such distinction should help to reduce the confusion arising from the use of the same notations for both models. Such distinction also enables us to clearly identify the differences in parameter relations between the two sets of parameters for different genetic models describing F_2 segregation patterns (Table 6). Until recently, there was a lack of appreciation that different gene action models would lead to different partitions of genic effects and subsequently different variance components due to the genic effects. Because the F_∞ -metric model has clear interpretation advantages as noted by VAN DER VEEN (1959) and MATHER and JINKS (1982), it has often been used for modeling multiple QTL (*e.g.*, HALEY and KNOTT 1992; CARLBORG *et al.* 2000; YI and XU 2002). However, because the genic effects defined under the F_∞ -metric model with the F_2 gene and genotypic frequencies are not orthogonal, bias will occur in estimating allelic effects when nonallelic (epistatic) effects are present; in contrast, the genic effects under the F_2 -metric model are orthogonal and

thus unbiased estimates of allelic effects can be obtained regardless of whether or not epistasis is present (CHEVERUD 2000; KAO and ZENG 2002). Since the translation between the F_2 -metric and F_∞ -metric models can be readily done through the simple relation (13), we recommend that in QTL mapping studies the F_2 -metric model should be used to first estimate the genic effects and then the estimated genic effects should be converted into those under the F_∞ -metric model to capture its interpretation advantages.

Our comparative assessment of different gene action models suggests the need for caution when reading the standard labels of “additive genetic variance,” “dominance variance,” etc., in the literature. As shown in this article and elsewhere (*e.g.*, COCKERHAM 1963; COCKERHAM and TACHIDA 1988), genic effects and their variances change with the population of reference. For example, in the CRP, the additive variance at locus A can be obtained from (6b) and (7),

$$\sigma_1^2(\text{CRP}) = 2p_A p_a \left[\frac{P_A^a}{p_A} (\bar{G}_A^a - \bar{G}_a^a) + \frac{P_a^a}{p_a} (\bar{G}_a^a - \bar{G}_a^a) \right]^2 \left/ \left(\frac{P_A^a}{p_A} + \frac{P_a^a}{p_a} \right) \right.$$

If CRP is an inbred population (IP), this additive variance becomes

$$\sigma_1^2(\text{IP}) = 2p_A p_a (1 + f_A) \left[a_{A(0)} + \frac{1 - f_A}{1 + f_A} (p_a - p_A) 2d_{A(0)} \right]^2,$$

which is reduced to the more familiar form in a random mating population ($f_A = 0$),

$$\sigma_1^2(f_A = 0) = 2p_A p_a [a_{A(0)} + (p_a - p_A) 2d_{A(0)}]^2.$$

Clearly, the so-called additive variance is “contaminated” with the dominance effect unless the gene frequencies are one-half as obtained under F_2 -metric model (Equation 17) or the dominance effect is absent. On the other hand, as shown in Equation 17, the additive variance under the F_∞ -metric model is contaminated with the epistatic effect. With zygotic associations (ZA) as in most natural populations, the additive variance takes an even more complicated form,

$$\sigma_A^2(\text{ZA}) = \sigma_1^2(\text{CRP}) + \sigma_3^2(\text{CRP}) + \text{part of } \sigma_\epsilon^2,$$

whose amount can be quantified in terms of various genic disequilibria. For example, part of the dominance and additive \times additive variances within finite populations undergoing bottleneck, drift, or population subdivision has been found to behave as additive variance (COCKERHAM and TACHIDA 1988; GOODNIGHT 1988; WHITLOCK *et al.* 1993). Such nonadditive genetic variances contribute to immediate and permanent response to selection in finite populations in an intricate fashion. These discussions serve to emphasize that the additive variance is not free from the influence of nonadditive genic effects and thus does not have simplicity and “clean” meaning unless they arise from the F_2 -metric model. Fortunately, most of the current QTL mapping

activities particularly with plant and laboratory species have focused on the use of planned mating designs corresponding to the F_2 -metric model (LYNCH and WALSH 1998).

The extension of our two-locus model for nonequilibrium populations to three or more loci is straightforward. For example, for three uncorrelated loci each with two alleles, the frequencies of 27 possible genotypes are the appropriate products of three single-locus genotypic frequencies. These three-locus genotypic frequencies are required for partitioning the three-locus genotypic values into 26 genic effects as expected under Cockerham's model or the F_2 -metric model: 3 additive, 3 dominance, 12 two-locus epistatic, and 8 three-locus epistatic effects. In the presence of two-locus and/or three-locus zygotic associations, an additional residual genic effect arises from nonzero deviations of the observed from expected three-locus genotypic frequencies. Such deviations are actually functions of 12 two-locus zygotic associations and 8 three-locus zygotic associations. It remains to be investigated what the exact forms of these functions are. Perhaps the functions of multilocus heterozygosity developed by YANG (2002) may be modified to accommodate individual genotypes as required here. Our model can also be extended to the case of multiple alleles per locus. For example, with three alleles at each of two independent loci, there are 6 possible genotypes at each locus and 36 genotypes at the two loci. The genotypic values are partitioned into the mean and 35 genic effects under Cockerham's model or the F_2 -metric model: 10 (5 at each locus) allelic effects and 25 nonallelic (epistatic) effects. With zygotic associations between the two loci, the residual genic effect is due to the presence of 25 zygotic associations that can be freely defined.

I am grateful to Marjorie Asmussen for her help and regret her recent and sudden death while this manuscript was still under review. I thank Yun-Xin Fu for stepping in to complete the review and two reviewers for valuable comments. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada grant OGP0183983.

LITERATURE CITED

- CARLBORG, O., L. ANDERSSON and B. KINGHORN, 2000 The use of a genetic algorithm for simultaneous mapping of multiple interactive quantitative trait loci. *Genetics* **155**: 2003–2010.
- CHEVERUD, J. M., 2000 Detecting epistasis among quantitative trait loci, pp. 58–81 in *Epistasis and the Evolutionary Process*, edited by J. B. WOLF, E. D. BRODIE III and M. J. WADE. Oxford University Press, New York.
- COCKERHAM, C. C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859–882.
- COCKERHAM, C. C., 1963 Estimation of genetic variances. *Statistical genetics and plant breeding*. NAS-NRC **982**: 53–94.
- COCKERHAM, C. C., and H. TACHIDA, 1988 Permanency of response to selection for quantitative characters in finite populations. *Proc. Natl. Acad. Sci. USA* **85**: 1563–1565.
- COCKERHAM, C. C., and B. S. WEIR, 1973 Descent measures for two loci with some applications. *Theor. Popul. Biol.* **4**: 300–330.
- COCKERHAM, C. C., and Z.-B. ZENG, 1996 Design III with marker loci. *Genetics* **143**: 1437–1456.
- CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetic Theory*. Harper & Row, New York.
- DOEBLEY, J., A. STEC and C. GUSTUS, 1995 *teosinte branched1* and origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* **141**: 333–346.
- GALLAIS, A., 1974 Covariances between arbitrary relatives with linkage and epistasis in the case of linkage disequilibrium. *Biometrics* **30**: 429–446.
- GOODNIGHT, C. J., 1988 Epistasis and the effect of founder events on the additive genetic variance. *Evolution* **42**: 441–454.
- GOODNIGHT, C. J., 2000 Quantitative trait loci and gene interaction: the quantitative genetics of metapopulation. *Heredity* **84**: 587–598.
- HALEY, C. S., and S. A. KNOTT, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315–324.
- HAYMAN, B. I., and K. MATHER, 1955 The description of genetic interactions in continuous variation. *Biometrics* **11**: 69–82.
- KAO, C.-H., and Z.-B. ZENG, 2002 Modeling epistasis of quantitative trait loci using Cockerham's model. *Genetics* **160**: 1243–1261.
- KEMPTHORNE, O., 1957 *An Introduction to Genetic Statistics*. John Wiley & Sons, New York.
- LYNCH, M., and B. WALSH, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
- MATHER, K., and J. L. JINKS, 1982 *Biometrical Genetics*, Ed. 3. Chapman & Hall, London.
- STRICKBERGER, M. W., 1976 *Genetics*, Ed. 2. Macmillan, New York.
- VAN DER VEEN, J. H., 1959 Tests of non-allelic interaction and linkage for quantitative characters in generations derived from two diploid pure lines. *Genetica* **30**: 201–232.
- WEIR, B. S., 1996 *Genetic Data Analysis II*. Sinauer Associates, Sunderland, MA.
- WEIR, B. S., and C. C. COCKERHAM, 1977 Two-locus theory in quantitative genetics, pp. 247–269 in *Proceedings of the International Conference on Quantitative Genetics*, edited by E. POLLAK, O. KEMPTHORNE and T. B. BAILEY, JR. Iowa State University Press, Ames, IA.
- WHITLOCK, M. C., P. C. PHILLIPS and M. J. WADE, 1993 Gene interaction affects the additive genetic variance in subdivided populations with migration and extinction. *Evolution* **47**: 1758–1769.
- WRIGHT, A. J., 1987 Covariance of inbred relatives with special reference to selfing. *Genetics* **115**: 545–552.
- YANG, R.-C., 2000 Zygotic associations and multilocus statistics in a nonequilibrium diploid population. *Genetics* **155**: 1449–1458.
- YANG, R.-C., 2002 Analysis of multilocus zygotic associations. *Genetics* **161**: 435–445.
- YANG, R.-C., 2003 Gametic and zygotic associations. *Genetics* **165**: 447–450.
- YANG, R.-C., and R. J. BAKER, 1990 Effects of epistasis on tests for linkage in self-pollinated species. *Theor. Appl. Genet.* **79**: 219–224.
- YI, N., and S. XU, 2002 Mapping quantitative trait loci with epistatic effects. *Genet. Res.* **79**: 185–198.

Communicating editor: Y.-X. Fu

