



Quantitative genetics theory for non-inbred populations in linkage disequilibrium

José Marcelo Soriano Viana

Universidade Federal de Viçosa, Departamento de Biologia Geral, Viçosa, MG, Brazil.

Abstract

Although linkage disequilibrium, epistasis and inbreeding are common phenomena in genetic systems that control quantitative traits, theory development and analysis are very complex, especially when they are considered together. The objective of this study is to offer additional quantitative genetics theory to define and analyze, in relation to non-inbred cross-pollinating populations, components of genotypic variance, heritabilities and predicted gains, assuming linkage disequilibrium and absence of epistasis. The genotypic variance and its components, additive and due to dominance genetic variances, are invariant over the generations only in regard to completely linked genes and to those in equilibrium. When the population is structured in half-sib families, the additive variance in the parents' generation and the genotypic variance in the population can be estimated. When the population is structured in full-sib families, none of the components of genotypic variance can be estimated. The narrow sense heritability at plant level can be estimated from the parent-offspring or mid parent-offspring regression. When there is dominance, the narrow sense heritability estimate in the F_2 is biased due to linkage disequilibrium when estimated by the Warner method, but not when estimated by means of the plant F_2 -family F_3 regression. The bias is proportional to the number of pairs of linked genes, without independent assortment, and to the degree of dominance, and tends to be positive when genes in the coupling phase predominate or negative and of higher value when genes in the repulsion phase predominate. Linkage disequilibrium is also cause of bias in estimates of the narrow sense heritabilities at full-sib family mean and at plant within half-sib and full-sib families levels. Generally, the magnitude of the bias is proportional to the number of pairs of genes in disequilibrium and to the frequency of recombining gametes.

Key words: components of genotypic variance, heritabilities, half-sib families, full-sib families.

Received: August 26, 2003; Accepted: June 25, 2004.

Introduction

Linkage disequilibrium, epistasis and inbreeding are phenomena that when considered separately or especially together, make the characterization and analysis of genetic systems responsible for quantitative traits very complex. Although the fundamental theory for each was established many years ago (Cockerham, 1954, 1956; Kempthorne, 1955, 1957; Schnell, 1961, 1963; Cockerham and Weir, 1977), discussion on linkage disequilibrium and epistasis related to breeding is surprisingly uncommon in statistical genetics (Kempthorne, 1973), quantitative genetics (Wricke and Weber, 1986; Hallauer and Miranda Filho, 1988; Comstock, 1996; Falconer and Mackay, 1996; Lynch and Walsh, 1998) and biometrical genetics (Mather and Jinks, 1974; Kearsley and Pooni, 1996) books, compared to other subjects relevant to breeders. Linkage disequilibrium and absence of epistasis are compulsorily assumed in almost all the methodologies developed to analyze quantita-

tive traits. The consequence, clearly, is biased estimates of genetic parameters and predicted gains, as linkage and genetic interaction are the rule and not the exception.

Although relevant, knowledge of the effects of linkage disequilibrium on the coefficients of the components of genotypic variance, including epistasis, on the covariance between relatives (Cockerham, 1956; Schnell, 1961, 1963; Weir *et al.*, 1980) does not represent all the theory of quantitative genetics for polygenic systems with genes in disequilibrium. The objective of this study was to supply additional theoretical information, characterizing and analyzing the components of genotypic variance, the heritabilities and the expected gains from selection, in relation to non-inbred cross-pollinating populations, assuming absence of epistasis.

Theory and Discussion

Components of genotypic variance

In the gametic pool of a non-inbred population (generation n), the probabilities of the gametes AB, Ab, aB and ab are (Kempthorne, 1973)

Send correspondence to José Marcelo Soriano Viana. Universidade Federal de Viçosa, Departamento de Biologia Geral, 36.570-000 Viçosa, MG, Brazil. E-mail: jmsviana@ufv.br.

$$p_{11}^{(n)} = p_a p_b + \Delta^{(n)}$$

$$p_{10}^{(n)} = p_a q_b - \Delta^{(n)}$$

$$p_{01}^{(n)} = q_a p_b - \Delta^{(n)}$$

$$p_{00}^{(n)} = q_a q_b + \Delta^{(n)}$$

where p is the probability of the gene that increases the trait expression, q is the probability of the allelic gene that decreases it and $\Delta^{(n)} = p_{11}^{(n)} p_{00}^{(n)} - p_{10}^{(n)} p_{01}^{(n)}$ is the measure of linkage disequilibrium. The genotype probabilities in generation 0 are

$$f_{22}^{(0)} = p_a^2 p_b^2 + 2p_a p_b \Delta^{(-1)} + [\Delta^{(-1)}]^2$$

$$f_{21}^{(0)} = 2p_a^2 p_b q_b + 2p_a (q_b - p_b) \Delta^{(-1)} - 2[\Delta^{(-1)}]^2$$

$$f_{20}^{(0)} = p_a^2 q_b^2 - 2p_a q_b \Delta^{(-1)} + [\Delta^{(-1)}]^2$$

$$f_{12}^{(0)} = 2p_a q_a p_b^2 + 2(q_a - p_a) p_b \Delta^{(-1)} - 2[\Delta^{(-1)}]^2$$

$$f_{11c}^{(0)} = 2p_a q_a p_b q_b + 2p_a p_b \Delta^{(-1)} + 2q_a q_b \Delta^{(-1)} + 2[\Delta^{(-1)}]^2$$

$$f_{11r}^{(0)} = 2p_a q_a p_b q_b - 2p_a q_b \Delta^{(-1)} - 2q_a p_b \Delta^{(-1)} + 2[\Delta^{(-1)}]^2$$

$$f_{10}^{(0)} = 2p_a q_a q_b^2 - 2(q_a - p_a) q_b \Delta^{(-1)} - 2[\Delta^{(-1)}]^2$$

$$f_{02}^{(0)} = q_a^2 p_b^2 - 2q_a p_b \Delta^{(-1)} + [\Delta^{(-1)}]^2$$

$$f_{01}^{(0)} = 2q_a^2 p_b q_b - 2q_a (q_b - p_b) \Delta^{(-1)} - 2[\Delta^{(-1)}]^2$$

$$f_{00}^{(0)} = q_a^2 q_b^2 + 2q_a q_b \Delta^{(-1)} + [\Delta^{(-1)}]^2$$

where f_{ij} is the probability of carrier of i and j copies of the genes of loci A and B that increase the trait expression, respectively. The c and r indexes identify the double-heterozygotes in the coupling and repulsion phases.

The genotypic values are

$$G_{22} = (m_a + a_a) + (m_b + a_b) = (M_a + A_{AA} + D_{AA}) + (M_b + A_{BB} + D_{BB})$$

$$G_{12} = (m_a + d_a) + (m_b + a_b) = (M_a + A_{Aa} + D_{Aa}) + (M_b + A_{BB} + D_{BB})$$

$$G_{02} = (m_a - a_a) + (m_b + a_b) = (M_a + A_{aa} + D_{aa}) + (M_b + A_{BB} + D_{BB})$$

$$G_{21} = (m_a + a_a) + (m_b + d_b) = (M_a + A_{AA} + D_{AA}) + (M_b + A_{Bb} + D_{Bb})$$

$$G_{11} = (m_a + d_a) + (m_b + d_b) = (M_a + A_{Aa} + D_{Aa}) + (M_b + A_{Bb} + D_{Bb})$$

$$G_{01} = (m_a - a_a) + (m_b + d_b) = (M_a + A_{aa} + D_{aa}) + (M_b + A_{Bb} + D_{Bb})$$

$$G_{20} = (m_a + a_a) + (m_b - a_b) = (M_a + A_{AA} + D_{AA}) + (M_b + A_{bb} + D_{bb})$$

$$G_{10} = (m_a + d_a) + (m_b - a_b) = (M_a + A_{Aa} + D_{Aa}) + (M_b + A_{bb} + D_{bb})$$

$$G_{00} = (m_a - a_a) + (m_b - a_b) = (M_a + A_{aa} + D_{aa}) + (M_b + A_{bb} + D_{bb})$$

where, for each gene, m is the mean of the genotypic values of the homozygotes, a is the deviation between the genotypic value of the homozygote of greatest expression and m , d is the deviation due to dominance, $M = m + (p - q)a + 2pqd$ is the population mean, A is additive genetic value and D is genetic value due to dominance. Regarding one gene, $A = 2q\alpha$ and $D = -2q^2d$, if the individual is homozygous for the gene that increases the trait expression, or $A = (q - p)\alpha$ and $D = 2pqd$, if the individual is heterozygous, or

$A = -2p\alpha$ and $D = -2p^2d$, if the individual is homozygous for the gene that reduces the trait expression, where $\alpha = a + (q - p)d$ is the average effect of a gene substitution (Falconer and Mackay, 1996).

The genotypic mean of generation 0 is

$$E(G) = f_{22}^{(0)}(G_{22}) + \dots + f_{00}^{(0)}(G_{00}) = m_a + (p_a - q_a)a_a + 2p_a q_a d_a + m_b + (p_b - q_b)a_b + 2p_b q_b d_b = M$$

The genotypic variance is

$$\begin{aligned} \sigma_G^{2(0)} &= f_{22}^{(0)}(G_{22})^2 + \dots + f_{00}^{(0)}(G_{00})^2 - M^2 = \\ &= (2p_a q_a \alpha_a^2 + 2p_b q_b \alpha_b^2 + 4\Delta^{(-1)} \alpha_a \alpha_b) + \\ &= (4p_a^2 q_a^2 d_a^2 + 4p_b^2 q_b^2 d_b^2 + 8[\Delta^{(-1)}]^2 d_a d_b) = \\ &= \sigma_A^{2(0)} + \sigma_D^{2(0)} \end{aligned}$$

where:

$$\text{Cov}(A_a, A_b)^{(0)} = 2\Delta^{(-1)} \alpha_a \alpha_b$$

$$\text{Cov}(A_a, D_b)^{(0)} = \text{Cov}(D_a, A_b)^{(0)} = 0$$

$$\text{Cov}(D_a, D_b)^{(0)} = 4[\Delta^{(-1)}]^2 d_a d_b$$

Considering k genes,

$$\sigma_A^{2(0)} = 2 \sum_{i=1}^k p_i q_i \alpha_i^2 + 4 \sum_{i=1}^k \sum_{j=1}^k \Delta_{ij}^{(-1)} \alpha_i \alpha_j$$

$$\sigma_D^{2(0)} = 4 \sum_{i=1}^k p_i^2 q_i^2 d_i^2 + 8 \sum_{i=1}^k \sum_{j=1}^k [\Delta_{ij}^{(-1)}]^2 d_i d_j$$

In spite of the probabilities of the genotypes altering from generation to generation, because $\Delta_{ij}^{(n)} = (1 - r_{ij}) \Delta_{ij}^{(n-1)} = (1 - r_{ij})^n \Delta_{ij}^{(0)}$ (Kempthorne, 1973), where r_{ij} is the frequency of recombining gametes, the population mean is constant. The genotypic variance changes as the population approaches equilibrium. The genotypic variance is invariant only for completely linked genes ($r_{ij} = 0$) and those in equilibrium ($\Delta_{ij}^{(n)} = 0$).

If generation 0 is the F_2 derived from the cross of two pure lines, $p_i = 0, 1/2$ or 1 , and

$$\Delta_{ij}^{(-1)} = \left(\frac{1 - r_{ij}}{2} \right)^2 - \left(\frac{r_{ij}}{2} \right)^2, \text{ for coupling genes, and}$$

$$\Delta_{ij}^{(-1)} = \left(\frac{r_{ij}}{2} \right)^2 - \left(\frac{1 - r_{ij}}{2} \right)^2, \text{ for repulsion genes}$$

Among and within family genotypic variances

When the generation 0 is structured in half-sib families, the genotypic means of the progenies are

$$\begin{aligned} \bar{G}_{AABB}^{(1)} &= [p_a p_b + \Delta^{(0)}] (M + A_{AA} + A_{BB} + D_{AA} + D_{BB}) + \\ &= [p_a q_b - \Delta^{(0)}] (M + A_{AA} + A_{Bb} + D_{AA} + D_{Bb}) + \end{aligned}$$

$$\begin{aligned} & \left[q_a p_b - \Delta^{(0)} \right] (M + A_{Aa} + A_{BB} + D_{Aa} + D_{BB}) + \\ & \left[q_a q_b + \Delta^{(0)} \right] (M + A_{Aa} + A_{Bb} + D_{Aa} + D_{Bb}) = \\ & M + \frac{1}{2} A_{AA} + \frac{1}{2} A_{BB} = M + \frac{1}{2} A_{AABB} \\ \bar{G}_{AABB}^{(1)} &= M + \frac{1}{2} A_{AA} + \frac{1}{2} A_{BB} = M + \frac{1}{2} A_{AABB} \\ \dots \\ \bar{G}_{aabb}^{(1)} &= M + \frac{1}{2} A_{aa} + \frac{1}{2} A_{bb} = M + \frac{1}{2} A_{aabb} \end{aligned}$$

Therefore, the genotypic mean of a half-sib family is also equal to the mean of the population plus half the additive genetic value of the common parent. The among and within genotypic variances are

$$\sigma_{\text{GaHSF}}^{2(1)} = f_{22}^{(0)} (\bar{G}_{AABB}^{(1)})^2 + \dots + f_{00}^{(0)} (\bar{G}_{aabb}^{(1)})^2 - M^2 = \frac{1}{4} \sigma_A^{2(0)}$$

$$\sigma_{\text{GwHSF}}^{2(1)} = \sigma_A^{2(1)} + \sigma_D^{2(1)} - \frac{1}{4} \sigma_A^{2(0)}$$

Or further,

$$\sigma_{\text{GaHSF}}^{2(1)} = \frac{1}{4} \sigma_A^{2(1)} + \sum_{i=1}^k \sum_{j=1}^k \left(\frac{r_{ij}}{1-r_{ij}} \right) \Delta_{ij}^{(0)} \alpha_i \alpha_j$$

$$\sigma_{\text{GwHSF}}^{2(1)} = \frac{3}{4} \sigma_A^{2(1)} + \sigma_D^{2(1)} - \sum_{i=1}^k \sum_{j=1}^k \left(\frac{r_{ij}}{1-r_{ij}} \right) \Delta_{ij}^{(0)} \alpha_i \alpha_j$$

Thus $\sigma_A^{2(0)}$ and, depending on the sample data, $\sigma_G^{2(1)}$ are estimable.

The genotypic mean of a full-sib family is, generically,

$$\bar{G}_{\text{FSF}} = \frac{G+G'}{2} + u_a d_a + u_b d_b \quad (u_i = 0, 1 \text{ or } -1/2)$$

where G and G' are the genotypic values of the parents. Thus,

$$\begin{aligned} \sigma_{\text{GaFSF}}^{2(1)} &= \frac{1}{2} \sigma_G^{2(0)} + d_a^2 V(u_a) + d_b^2 V(u_b) + \\ & d_a \text{Cov}(G+G', u_a) + d_b \text{Cov}(G+G', u_b) + \\ & 2d_a d_b \text{Cov}(u_a, u_b) \end{aligned}$$

But,

$$E(u_i) = 0 \quad (i = 1, \dots, k)$$

$$E[(u_i)^2] = 3p_i^2 q_i^2 \quad (i = 1, \dots, k)$$

$$\text{Cov}(G+G', u_a) = -4p_a^2 q_a^2 d_a - 4[\Delta^{(-1)}]^2 d_b$$

$$\text{Cov}(G+G', u_b) = -4p_b^2 q_b^2 d_b - 4[\Delta^{(-1)}]^2 d_a$$

$$\text{Cov}(u_a, u_b) = 3[\Delta^{(-1)}]^2$$

Thus,

$$\sigma_{\text{GaFSF}}^{2(1)} = \frac{1}{2} \sigma_A^{2(0)} + \frac{1}{4} \sigma_D^{2(0)}$$

$$\sigma_{\text{GwFSF}}^{2(1)} = \sigma_A^{2(1)} + \sigma_D^{2(1)} - \frac{1}{2} \sigma_A^{2(0)} - \frac{1}{4} \sigma_D^{2(0)}$$

Or further,

$$\sigma_{\text{GaFSF}}^{2(1)} = \frac{1}{2} \sigma_A^{2(1)} + \frac{1}{4} \sigma_D^{2(1)} + 2 \sum_{i=1}^k \sum_{j=1}^k \left(\frac{r_{ij}}{1-r_{ij}} \right) \Delta_{ij}^{(0)} \alpha_i \alpha_j +$$

$$2 \sum_{i=1}^k \sum_{j=1}^k \left[\left(\frac{1}{1-r_{ij}} \right)^2 - 1 \right] [\Delta_{ij}^{(0)}]^2 d_i d_j$$

$$\sigma_{\text{GwFSF}}^{2(1)} = \frac{1}{2} \sigma_A^{2(1)} + \frac{3}{4} \sigma_D^{2(1)} - 2 \sum_{i=1}^k \sum_{j=1}^k \left(\frac{r_{ij}}{1-r_{ij}} \right) \Delta_{ij}^{(0)} \alpha_i \alpha_j -$$

$$2 \sum_{i=1}^k \sum_{j=1}^k \left[\left(\frac{1}{1-r_{ij}} \right)^2 - 1 \right] [\Delta_{ij}^{(0)}]^2 d_i d_j$$

In this case, no component of the genotypic variance is estimable.

Heritabilities and genetic gains

The narrow sense heritability at individual level in the generation 0 is the square of the correlation between the phenotypic and additive genetic values (Falconer and MacKay, 1996; Viana, 2002), given by

$$h^2 = \frac{\sigma_A^{2(0)}}{\sigma_P^{2(0)}}$$

Its estimation, although rare, can be done by the regression of the mean phenotypic value of non-inbred progeny (generation 1) as a function of the phenotypic value of a parent or of the mean phenotypic value of the parents (generation 0) (Hallauer and Miranda Filho, 1988). Assuming that the genotypic value and environmental effect are independent variables, the parent-offspring covariance is

$$\text{Cov}(P^{(0)}, O^{(1)}) = f_{22}^{(0)} (A_{22} + D_{22}) \left(\frac{1}{2} A_{22} \right) + \dots +$$

$$f_{00}^{(0)} (A_{00} + D_{00}) \left(\frac{1}{2} A_{00} \right) - 0.0 = \frac{1}{2} \sigma_A^{2(0)}$$

The parent-grand offspring covariance is

$$\text{Cov}(P^{(0)}, O^{(2)}) = f_{22}^{(0)} (A_{22} + D_{22}) \left(\frac{1}{4} A_{22} \right) + \dots +$$

$$f_{00}^{(0)} (A_{00} + D_{00}) \left(\frac{1}{4} A_{00} \right) - 0.0 = \frac{1}{4} \sigma_A^{2(0)}$$

Thus, the ancestor-descendent covariance is

$$\text{Cov}(P^{(0)}, O^{(n)}) = \left(\frac{1}{2}\right)^n \sigma_A^{2(0)}$$

The covariance between S_0 plant and S_1 progeny is

$$\text{Cov}(S_0^{(0)}, S_1^{(1)}) = \text{Cov}(G, G + u_a d_a + u_b d_b) = \sigma_G^{2(0)} + d_a \text{Cov}(G, u_a) + d_b \text{Cov}(G, u_b) \quad (u_i = 0 \text{ or } -1/2)$$

But

$$E(u_i) = -p_i q_i \quad (i = 1, \dots, k)$$

$$E[(u_i)^2] = \frac{1}{2} p_i q_i \quad (i = 1, \dots, k)$$

$$\text{Cov}(G, u_a) = -p_a q_a [(q_b - p_a) \alpha_a + 2p_a q_a d_a] - (q_a - p_a) \Delta^{(-1)} \alpha_b - 2[\Delta^{(-1)}]^2 d_b$$

$$\text{Cov}(G, u_b) = -p_b q_b [(q_b - p_b) \alpha_b + 2p_b q_b d_b] - (q_b - p_b) \Delta^{(-1)} \alpha_a - 2[\Delta^{(-1)}]^2 d_a$$

$$\text{Cov}(u_a, u_b) = \frac{1}{2} (q_a - p_a)(q_b - p_b) \Delta^{(-1)} + [\Delta^{(-1)}]^2$$

Thus, considering k genes,

$$\text{Cov}(S_0^{(0)}, S_1^{(1)}) = \sigma_A^{2(0)} + \frac{1}{2} \sigma_D^{2(0)} + \sum_{i=1}^k [p_i q_i (p_i - q_i) \alpha_i d_i] + \sum_{i=1}^k \sum_{j=1}^k \Delta_{ij}^{(-1)} [(p_i - q_i) d_i \alpha_j + (p_j - q_j) d_j \alpha_i]$$

In the presence of dominance, the estimate of narrow sense heritability at F_2 plant level is biased because of linkage disequilibrium when the additive variance is estimated by the method of Warner (1952) but not when estimated by parent-offspring regression (Frey and Horner, 1957), because

$$\sigma_A^{2(0)} = 2\text{Cov}(S_0^{(0)}, S_1^{(1)}) - \sigma_G^{2(0)}$$

The genotypic variances in the backcross generations of F_1 plants with the parents are

$$\sigma_{G_{BC1}}^2 = \sum_{i=1}^k p_i q_i \alpha_i^2 + 4 \sum_{i=1}^k p_i^2 q_i^2 d_i^2 + 2 \sum_{i=1}^k \sum_{j=1}^k \Delta_{ij}^{(-1)} (\alpha_i \alpha_j + \theta_{ij} d_i d_j - \theta_{ij} \alpha_i d_j - d_i \alpha_j) - \frac{1}{2} \sum_{i=1}^{k'} \theta_i \alpha_i d_i$$

$$\sigma_{G_{BC2}}^2 = \sum_{i=1}^k p_i q_i \alpha_i^2 + 4 \sum_{i=1}^k p_i^2 q_i^2 d_i^2 + 2 \sum_{i=1}^k \sum_{j=1}^k \Delta_{ij}^{(-1)} (\alpha_i \alpha_j + \theta_{ij} d_i d_j + \theta_{ij} \alpha_i d_j + d_i \alpha_j) + \frac{1}{2} \sum_{i=1}^{k'} \theta_i \alpha_i d_i$$

where p_i is equal to 0, 1/2 or 1, θ_{ij} is equal to 1, in the case of coupling genes, or -1, in the case of repulsion genes, θ_i is equal to 1 or -1, and k' is the number of non-fixed genes.

Thus,

$$2\sigma_G^{2(0)} - \sigma_{G_{BC1}}^2 - \sigma_{G_{BC2}}^2 = \sigma_A^{2(0)} + 8 \sum_{i=1}^k \sum_{j=1}^k r_{ij} \Delta_{ij}^{(-1)} d_i d_j$$

Table 1 - Bias (%) in the estimate of the narrow sense heritability at F_2 plant level, obtained by the Warner method, considering k non-fixed genes (equal allelic frequencies), k_d genes in disequilibrium, in coupling (c) and repulsion (r), and different degrees of dominance (d/a) and recombinant gamete frequencies (r_{ab}).

$ d/a $	r_{ab}	$\Delta^{(-1)}, c$	$k = 2, k_d = 2, c$	$k = 20, k_d = 2, c$	$k = 20, k_d = 4, c$	$\Delta^{(-1)}, r$	$k = 2, k_d = 2, r$	$k = 20, k_d = 2, r$	$k = 20, k_d = 4, r$
2	0	0.25	0.0	0.0	0.0	-0.25	0.0	0.0	0.0
2	0.1	0.2	35.6	5.9	25.9	-0.2	-320.0	-7.0	-73.8
2	0.2	0.15	60.0	9.1	42.3	-0.15	-240.0	-10.2	-90.0
2	0.3	0.1	68.6	9.2	46.4	-0.1	-160.0	-10.0	-75.8
2	0.4	0.05	53.3	6.3	34.3	-0.05	-80.0	-6.5	-43.6
2	0.5	0	0.0	0.0	0.0	0	0.0	0.0	0.0
1	0	0.25	0.0	0.0	0.0	-0.25	0.0	0.0	0.0
1	0.1	0.2	8.9	1.5	6.5	-0.2	-80.0	-1.7	-18.5
1	0.2	0.15	15.0	2.3	10.6	-0.15	-60.0	-2.5	-22.5
1	0.3	0.1	17.1	2.3	11.6	-0.1	-40.0	-2.5	-18.9
1	0.4	0.05	13.3	1.6	8.6	-0.05	-20.0	-1.6	-10.9
1	0.5	0	0.0	0.0	0.0	0	0.0	0.0	0.0
0.5	0	0.25	0.0	0.0	0.0	-0.25	0.0	0.0	0.0
0.5	0.1	0.2	2.2	0.4	1.6	-0.2	-20.0	-0.4	-4.6
0.5	0.2	0.15	3.8	0.6	2.6	-0.15	-15.0	-0.6	-5.6
0.5	0.3	0.1	4.3	0.6	2.9	-0.1	-10.0	-0.6	-4.7

Table 1 (cont)

d / a	r _{ab}	Δ ⁽⁻¹⁾ , c	k = 2,	k = 20,	k = 20,	Δ ⁽⁻¹⁾ , r	k = 2,	k = 20,	k = 20,
			k _d =2, c	k _d =2, c	k _d =4, c		k _d =2, r	k _d =2, r	k _d =4, r
0.5	0.4	0.05	3.3	0.4	2.1	-0.05	-5.0	-0.4	-2.7
0.5	0.5	0	0.0	0.0	0.0	0	0.0	0.0	0.0
0	0	0.25	0.0	0.0	0.0	-0.25	0.0	0.0	0.0
0	0.1	0.2	0.0	0.0	0.0	-0.2	0.0	0.0	0.0
0	0.2	0.15	0.0	0.0	0.0	-0.15	0.0	0.0	0.0
0	0.3	0.1	0.0	0.0	0.0	-0.1	0.0	0.0	0.0
0	0.4	0.05	0.0	0.0	0.0	-0.05	0.0	0.0	0.0
0	0.5	0	0.0	0.0	0.0	0	0.0	0.0	0.0

The magnitudes of the bias in relation to polygenic systems with two and 20 genes, with groups of two and four in disequilibrium (Table 1), reveal that the heritability estimates in F₂, obtained by the Warner method and assuming linkage equilibrium, can be very biased. Generally, the bias is proportional to the number of pairs of genes in disequilibrium (linked without independent assortment) and to the degree of dominance, tending to be positive when genes in the coupling phase predominate and negative when genes in the repulsion phase predominate. In this case, the bias should have greater magnitude.

Therefore, linkage disequilibrium is not the cause of bias in predicting genetic gains from mass selection in cross-pollinating populations, with estimation of the additive variance by the parent-offspring or mid parent-offspring regression. In the presence of dominance, the prediction of gains with selection in F₂ is biased due to linkage disequilibrium when the additive variance is estimated by the Warner method, but not when estimated by the parent-offspring regression.

The narrow sense heritability at half-sib family mean level is the square of the correlation between the mean phenotypic value of the progeny and the additive genetic value of the common parent (Viana, 2002), given by

$$h^2 = \frac{(1/4)\sigma_A^{2(0)}}{\sigma_{PaHSF}^{2(1)}}$$

The narrow sense heritability at plant within family level, which corresponds to the square of the mean correlation, considering the different progenies, between the phenotypic and additive genetic values of the same individual is

$$h^2 = \frac{(3/4)\sigma_A^{2(0)} - 4 \sum_{i=1}^k \sum_{j=1}^k r_{ij} \Delta_{ij}^{(-1)} \alpha_i \alpha_j}{\sigma_{PwHSF}^{2(1)}}$$

Linkage disequilibrium is, then, cause of bias only in predicting gain from within selection.

With full-sib families, the narrow sense heritabilities at progeny mean level, that is the square of the correlation between the progeny mean phenotypic value and the mean additive genetic value of the parents (Viana, 2002), and at plant within family level are

$$h^2 = \frac{(1/2)\sigma_A^{2(0)}}{\sigma_{PaFSF}^{2(1)}}$$

$$h^2 = \frac{(1/2)\sigma_A^{2(1)} - 2 \sum_{i=1}^k \sum_{j=1}^k \left(\frac{r_{ij}}{1-r_{ij}} \right) \Delta_{ij}^{(0)} \alpha_i \alpha_j}{\sigma_{PwFSF}^{2(1)}}$$

Assuming linkage equilibrium, the estimated heritabilities are

$$h^2 = \frac{(1/2)\sigma_A^{2(0)} + 2 \sum_{i=1}^k \sum_{j=1}^k r_{ij} \Delta_{ij}^{(-1)} \alpha_i \alpha_j}{\sigma_{PaFSF}^{2(1)}} + \frac{4 \sum_{i=1}^k \sum_{j=1}^k [1 - (1-r_{ij})^2] [\Delta_{ij}^{(-1)}]^2 d_i d_j}{\sigma_{PaFSF}^{2(1)}}$$

$$h^2 = \frac{(1/2)\sigma_A^{2(1)} + 4 \sum_{i=1}^k \sum_{j=1}^k \left(\frac{r_{ij}}{1-r_{ij}} \right) \Delta_{ij}^{(0)} \alpha_i \alpha_j}{\sigma_{PwFSF}^{2(1)}} + \frac{4 \sum_{i=1}^k \sum_{j=1}^k \left[\left(\frac{1}{1-r_{ij}} \right)^2 - 1 \right] [\Delta_{ij}^{(0)}]^2 d_i d_j}{\sigma_{PwFSF}^{2(1)}}$$

Therefore, the predicted gains with among and within selection are biased due to linkage disequilibrium.

The analysis of the bias due to linkage disequilibrium in the estimates of genetic population parameters is complex because of the infinite combinations of genotype and gene frequencies, degree of dominance, frequency of re-

combining gametes, number of genes in the polygenic system and number of pairs of genes in disequilibrium. Considering two and 20 genes in the genic system, with groups of two and four in disequilibrium (Tables 2, 3 and 4), it can be seen that the bias in the narrow sense heritability estimates at plant within half-sib family level, at full-sib family level and at plant within full-sib progeny

level is generally reduced but may have a very high magnitude, depending on the number of genes and the number of pairs of genes in disequilibrium. The bias is proportional to the number of genes, to the number of pairs of genes in disequilibrium and to the frequency of recombining gametes. The magnitude of the bias is minimized when the frequency of the dominant genes is intermediate. When the dominant

Table 2 - Bias (%) in the estimates of the narrow sense heritabilities at plant within half-sib family (h^2_{wHSF}), at full-sib family mean (h^2_{FSF}) and at plant within full-sib progeny levels (h^2_{wFSF}), considering k non-fixed genes, k_d genes in disequilibrium, dominant genes at low frequencies (0.1 to 0.2), and different degrees of dominance (d/a) and recombining gamete frequencies (r_{ab}).

d / a	r_{ab}	$\Delta^{(-1)}$	h^2_{wHSF}			h^2_{FSF}			h^2_{wFSF}		
			k = 2, $k_d=2$	k = 20, $k_d=2$	k = 20, $k_d=4$	k = 2, $k_d=2$	k = 20, $k_d=2$	k = 20, $k_d=4$	k = 2, $k_d=2$	k = 20, $k_d=2$	k = 20, $k_d=4$
2	0	0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.1	0.025	2.1	0.2	1.4	1.7	0.2	1.1	5.0	0.6	3.1
2	0.2	0.025	4.4	0.5	2.8	3.4	0.4	2.2	10.5	1.2	6.5
2	0.3	0.026	6.9	0.8	4.3	5.2	0.6	3.3	16.5	1.8	10.2
2	0.4	0.026	9.6	1.0	5.9	7.0	0.8	4.5	23.3	2.4	14.2
2	0.5	0.027	12.6	1.3	7.6	8.9	1.0	5.7	30.8	3.1	18.5
1	0	0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.1	0.025	2.1	0.2	1.3	1.6	0.2	1.0	4.9	0.6	3.1
1	0.2	0.025	4.4	0.5	2.8	3.3	0.4	2.1	10.3	1.1	6.4
1	0.3	0.026	6.9	0.8	4.3	5.0	0.6	3.2	16.2	1.7	10.0
1	0.4	0.026	9.6	1.0	5.9	6.8	0.8	4.3	22.9	2.4	13.9
1	0.5	0.027	12.5	1.3	7.6	8.6	1.0	5.5	30.2	3.0	18.1
0.5	0	0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.5	0.1	0.025	2.1	0.2	1.3	1.6	0.2	1.0	4.8	0.5	3.0
0.5	0.2	0.025	4.4	0.5	2.7	3.2	0.4	2.0	10.1	1.1	6.3
0.5	0.3	0.026	6.8	0.8	4.2	4.9	0.6	3.1	16.0	1.7	9.9
0.5	0.4	0.026	9.5	1.0	5.8	6.6	0.8	4.2	22.5	2.3	13.7
0.5	0.5	0.027	12.4	1.3	7.5	8.3	1.0	5.4	29.8	3.0	17.9
0	0	0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0	0.1	0.025	2.1	0.2	1.3	1.5	0.2	1.0	4.7	0.5	3.0
0	0.2	0.025	4.3	0.5	2.7	3.1	0.4	2.0	9.9	1.1	6.2
0	0.3	0.026	6.7	0.7	4.2	4.7	0.5	3.0	15.6	1.7	9.6
0	0.4	0.026	9.3	1.0	5.7	6.4	0.7	4.1	22.0	2.3	13.4
0	0.5	0.027	12.1	1.3	7.3	8.1	0.9	5.2	29.0	2.9	17.4

Table 3 - Bias (%) in the estimates of the narrow sense heritabilities at plant within half-sib family (h^2_{wHSF}), at full-sib family mean (h^2_{FSF}) and at plant within full-sib progeny levels (h^2_{wFSF}), considering k non-fixed genes, k_d genes in disequilibrium, dominant genes at intermediate frequencies (0.4 to 0.6), and different degrees of dominance (d/a) and recombining gamete frequencies (r_{ab}).

d / a	r_{ab}	$\Delta^{(-1)}$	h^2_{wHSF}			h^2_{FSF}			h^2_{wFSF}		
			k = 2, $k_d=2$	k = 20, $k_d=2$	k = 20, $k_d=4$	k = 2, $k_d=2$	k = 20, $k_d=2$	k = 20, $k_d=4$	k = 2, $k_d=2$	k = 20, $k_d=2$	k = 20, $k_d=4$
2	0	-0.021	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.1	-0.019	-1.0	-0.1	-0.6	-0.5	-0.0	-0.3	-1.9	-0.2	-1.1
2	0.2	-0.017	-1.7	-0.2	-1.0	-0.9	-0.1	-0.5	-3.5	-0.3	-2.0
2	0.3	-0.015	-2.2	-0.2	-1.3	-1.3	-0.1	-0.8	-4.6	-0.4	-2.7
2	0.4	-0.013	-2.5	-0.2	-1.5	-1.6	-0.2	-0.9	-5.3	-0.5	-3.1

Table 3 (cont.)

d / a	r _{ab}	Δ ⁽⁻¹⁾	h ² _{wHSF}			h ² _{FSF}			h ² _{wFSF}		
			k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4	k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4	k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4
2	0.5	-0.011	-2.6	-0.3	-1.6	-1.7	-0.2	-1.0	-5.6	-0.6	-3.3
1	0	-0.021	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.1	-0.019	-1.0	-0.1	-0.6	-0.7	-0.1	-0.4	-2.3	-0.2	-1.3
1	0.2	-0.017	-1.8	-0.2	-1.1	-1.3	-0.1	-0.8	-4.0	-0.4	-2.4
1	0.3	-0.015	-2.4	-0.2	-1.4	-1.7	-0.2	-1.0	-5.2	-0.5	-3.1
1	0.4	-0.013	-2.7	-0.3	-1.6	-2.0	-0.2	-1.2	-5.9	-0.6	-3.5
1	0.5	-0.011	-2.8	-0.3	-1.7	-2.1	-0.2	-1.2	-6.2	-0.6	-3.7
0.5	0	-0.021	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.5	0.1	-0.019	-1.1	-0.1	-0.6	-0.8	-0.1	-0.5	-2.4	-0.2	-1.4
0.5	0.2	-0.017	-1.9	-0.2	-1.1	-1.4	-0.1	-0.8	-4.1	-0.4	-2.4
0.5	0.3	-0.015	-2.4	-0.2	-1.4	-1.8	-0.2	-1.1	-5.4	-0.5	-3.2
0.5	0.4	-0.013	-2.8	-0.3	-1.6	-2.1	-0.2	-1.2	-6.1	-0.6	-3.6
0.5	0.5	-0.011	-2.9	-0.3	-1.7	-2.2	-0.2	-1.3	-6.3	-0.6	-3.8
0	0	-0.021	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0	0.1	-0.019	-1.1	-0.1	-0.6	-0.8	-0.1	-0.5	-2.4	-0.2	-1.4
0	0.2	-0.017	-1.9	-0.2	-1.1	-1.4	-0.1	-0.8	-4.2	-0.4	-2.5
0	0.3	-0.015	-2.4	-0.2	-1.4	-1.9	-0.2	-1.1	-5.4	-0.5	-3.2
0	0.4	-0.013	-2.8	-0.3	-1.7	-2.1	-0.2	-1.3	-6.1	-0.6	-3.7
0	0.5	-0.011	-2.9	-0.3	-1.7	-2.2	-0.2	-1.3	-6.4	-0.6	-3.8

Table 4 - Bias (%) in the estimates of the narrow sense heritabilities at plant within half-sib family (h²_{wHSF}), at full-sib family mean (h²_{FSF}) and at plant within full-sib progeny levels (h²_{wFSF}), considering k non-fixed genes, k_d genes in disequilibrium, dominant genes at high frequencies (0.8 to 0.9), and different degrees of dominance (d/a) and recombining gamete frequencies (r_{ab}).

d / a	r _{ab}	Δ ⁽⁻¹⁾	h ² _{wHSF}			h ² _{FSF}			h ² _{wFSF}		
			k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4	k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4	k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4
2	0	-0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.1	-0.024	-2.7	-0.2	-1.5	3.9	0.3	2.2	-0.3	-0.0	-0.2
2	0.2	-0.024	-5.3	-0.5	-3.0	7.2	0.6	4.0	-1.1	-0.1	-0.6
2	0.3	-0.024	-7.8	-0.7	-4.5	9.9	0.8	5.5	-2.5	-0.2	-1.4
2	0.4	-0.024	-10.1	-0.9	-5.9	11.9	1.0	6.6	-4.3	-0.4	-2.5
2	0.5	-0.024	-12.3	-1.2	-7.3	13.3	1.1	7.3	-6.5	-0.6	-3.9
1	0	-0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.1	-0.024	-2.7	-0.2	-1.5	-0.2	-0.0	-0.1	-4.2	-0.4	-2.4
1	0.2	-0.024	-5.3	-0.5	-3.0	-0.5	-0.0	-0.3	-8.2	-0.7	-4.7
1	0.3	-0.024	-7.8	-0.7	-4.5	-1.1	-0.1	-0.6	-12.2	-1.1	-7.1
1	0.4	-0.024	-10.1	-0.9	-5.9	-1.9	-0.2	-1.0	-16.0	-1.5	-9.5
1	0.5	-0.024	-12.3	-1.2	-7.3	-2.9	-0.2	-1.6	-19.7	-2.0	-11.8
0.5	0	-0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.5	0.1	-0.024	-2.9	-0.2	-1.6	-2.1	-0.2	-1.1	-6.2	-0.5	-3.5
0.5	0.2	-0.024	-5.5	-0.5	-3.2	-4.2	-0.3	-2.3	-11.9	-1.1	-6.8
0.5	0.3	-0.024	-8.1	-0.7	-4.7	-6.3	-0.5	-3.5	-17.2	-1.6	-10.0
0.5	0.4	-0.024	-10.5	-1.0	-6.1	-8.4	-0.7	-4.6	-22.1	-2.1	-13.1

Table 4 (cont.)

d / a	r_{ab}	$\Delta^{(-1)}$	h_{wHSF}^2			h_{FSF}^2			h_{wFSF}^2		
			k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4	k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4	k = 2, k _d = 2	k = 20, k _d = 2	k = 20, k _d = 4
0.5	0.5	-0.024	-12.8	-1.2	-7.6	-10.5	-0.9	-5.8	-26.6	-2.7	-16.0
0	0	-0.024	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0	0.1	-0.024	-2.9	-0.2	-1.6	-2.2	-0.2	-1.2	-6.4	-0.6	-3.6
0	0.2	-0.024	-5.6	-0.5	-3.2	-4.4	-0.4	-2.4	-12.2	-1.1	-7.0
0	0.3	-0.024	-8.1	-0.7	-4.7	-6.6	-0.6	-3.7	-17.6	-1.6	-10.2
0	0.4	-0.024	-10.6	-1.0	-6.2	-8.9	-0.7	-4.9	-22.6	-2.2	-13.3
0	0.5	-0.024	-12.9	-1.2	-7.6	-11.1	-0.9	-6.1	-27.2	-2.7	-16.3

genes are at reduced frequencies the bias is positive. If in the population the frequency of the dominant genes is intermediate to high, the bias tends to be negative. Except for overdominance, and in populations with high frequencies of dominant genes, the bias is less affected by the degree of dominance, with a tendency to be proportional to the degree of dominance when the dominant genes are at reduced frequencies, but inversely proportional when the frequency of the dominant genes is intermediate to high. With the same exception, the bias in the estimate of heritability at plant within full-sib family level is greater than that of the estimate of heritability at progeny mean level.

References

- 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 and Weir BS (1977) Two-locus theory in quantitative genetics. In: Pollak E, Kempthorne O and Bailey Jr TB (eds) *Proceedings of the International Conference on Quantitative Genetics*. Iowa State University Press, Ames, pp 247-269.
- Comstock RE (1996) *Quantitative Genetics with Special Reference to Plant and Animal Breeding*. Ames, Iowa State University Press, 436 pp.
- Falconer DS and Mackay TFC (1996) *Introduction to Quantitative Genetics*. 4th edition. Longman, New York, 464 pp.
- Frey KJ and Horner T (1957) Heritability in standard units. *Agronomy Journal* 49:59-62.
- Hallauer AR and Miranda Filho JB (1988) *Quantitative Genetics in Maize Breeding*. 2nd edition. Iowa State University Press, Ames, 468 pp.
- Kearsey MJ and Pooni HS (1996) *The Genetical Analysis of Quantitative Traits*. Chapman & Hall, London, 352 pp.
- Kempthorne O (1973) *An Introduction to Genetic Statistics*. Iowa State University Press, Ames, 545 pp.
- Kempthorne O (1955) The theoretical values of correlations between relatives in random mating populations. *Genetics* 40:153-167.
- Lynch M and Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, Massachusetts, 980 pp.
- Mather K and Jinks JL (1974) *Biometrical Genetics*. 2nd edition. Cornell University Press, Ithaca, New York, 382 pp.
- Schnell FW (1961) Some general formulations of linkage effects in inbreeding. *Genetics* 46:947-957.
- Schnell FW (1963) The covariance between relatives in the presence of linkage. In: Hanson WD and Robinson HF (eds) *Statistical Genetics and Plant Breeding*. National Academy of Sciences - National Research Council, Washington, pp 468-483.
- Viana JMS (2002) Heritability at family mean level. *Revista Árvore* 26:271-278.
- Warner JN (1952) A method for estimating heritability. *Agronomy Journal* 44:427-430.
- Weir BS, Cockerham CC and Reynolds J (1980) The effects of linkage and linkage disequilibrium on the covariances of noninbred relatives. *Heredity* 45:351-359.
- Wricke G and Weber WE (1986) *Quantitative Genetics and Selection in Plant Breeding*. Walter de Gruyter, Berlin, 406 pp.

Associate Editor: Louis Bernard Klaczko