Theor Appl Genet (1994) 89:160–166

J. Zhu · B. S. Weir

Analysis of cytoplasmic and maternal effects. II. Genetic models for triploid endosperms

Received: 28 July 1993 / Accepted: 27 September 1993

Abstract Genetic models for quantitative traits of triploid endosperms are proposed for the analysis of direct gene effects, cytoplasmic effects, and maternal gene effects. The maternal effect is partitioned into maternal additive and dominance components. In the full genetic model, the direct effect is partitioned into direct additive and dominance components and high-order dominance component, which are the cumulative effects of three-allele interactions. If the high-order dominance effects are of no importance, a reduced genetic model can be used. Monte Carlo simulations were conducted in this study for demonstrating unbiasedness of estimated variance and covariance components from the MINQUE (0/1) procedure, which is a minimum norm quadratic unbiased estimation (MINQUE) method setting 0 for all the prior covariances and 1 for all the prior variances. Robustness of estimating variance and covariance components for the genetic models was tested by simulations. Both full and reduced genetic models are shown to be robust for estimating variance and covariance components under several situations of no specific effects. Efficiency of predicting random genetic effects for the genetic models by the MINQUE (0/1) procedure was compared with the best linear unbiased prediction (BLUP). A worked example is given to illustrate the use of the reduced genetic model for kernel growth characteristics in corn (*Zea mays* L.).

Key words Monte Carlo simulation · Endospermic traits · Cytoplasmic and maternal effects · Variance and covariance components · Genetic prediction

Communicated by A. L. Kahler

J. Zhu
Department of Agronomy, Zheijiang Agricultural University, Hangzhou, Zheijiang, China

B. S. Weir (✉)
Program in Statistical Genetics, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA

Introduction

Endosperms are the major storage organ of cereal grains, which provide more than half of the food energy and protein consumed on earth. An understanding of the inheritance of characteristics and nutrition content of endosperms is important to cereal breeding with respect to improving yield potential and seed quality. Triploid endosperm is supplied with carbohydrate reserves by the maternal plant. Although some cereal crops are polyploid, the genetic behavior of most cereal crops, including wheat, is like that of diploids.

There have been several triploid models proposed for the analysis of quantitative traits of endosperm (Gale 1975; Bogyo et al. 1988; Mo 1988). All of these triploid models, under the assumption of no maternal and cytoplasmic effects, include only the direct effects of nuclear genes in endosperm cells.

Maternal effects and the cytoplasmic inheritance of nutrition content and kernel characteristics have been widely studied in cereal crops. Maternal effects for oleic and linoleic acids have been found in corn (Garwood et al. 1970; Poneleit and Bauman 1970), and strong maternal effects were exhibited for the fatty acid compositions of triglycerides and phospholipids in corn (Weber 1983). By means of diallel analysis of F_1s and F_2s from all combinations of four inbred lines, Poneleit and Egli (1983) observed maternal effects but not cytoplasmic effects for kernel growth characteristics in corn. In wheat (*Triticum aestivum* L.) Dhaliwal (1977) found maternal effects for seed proteins. Maternal effects are also involved in the inheritance of protein and lysine content in barley (*Hordeum vulgare* L.) (Ullrich and Eslick 1978). Cytoplasmic influences on yield and grain quality have been found in maize (Rao and Fleming 1978) and in wheat (Kofoid and Maan 1982). Since maternal plants supply the assimilates for grain filling and since the photosynthetic activity of the maternal plant is determined by the maternal nuclear genes and cytoplasmic genes, a real genetic model with biological

accuracy should consist of maternal genetic effects and cytoplasmic effects along with direct genetic effects.

Foolad and Jones (1992) recently presented genetic models for analyzing quantitative seed characters. The model includes testa, cytoplasm, and embryo effects as well as endosperm effects. Diallel analysis for reciprocal crosses provides a way for studying maternal or cytoplasmic effects (Henderson 1948; Griffing 1956; Cockerham and Weir 1977; Eisen et al. 1983; Beavis et al. 1987). Zhu and Weir (1994) have proposed a model for quantitative traits controlled by diploid nuclear genes, cytoplasmic genes, and maternal nuclear genes. A minimum norm quadratic unbiased estimation (MINQUE) procedure for the estimation of variance and covariance components for one trait and of covariance components for two traits was also developed for analyzing genetic models with correlations between direct and maternal genetic effects (Zhu and Weir 1994). A linear unbiased prediction (LUP) method proved to be efficient for the prediction of random genetic effects (Zhu and Weir 1994).

The objective of the present study is to introduce genetic models for quantitative traits of triploid endosperms and to evaluate the unbiasedness and efficiency of estimating variances and covariances and of predicting genetic effects by the MINQUE procedure for these genetic models. An example is given with data from kernel growth characteristics in corn (Poneleit and Egli 1983) to illustrate the use of the models.

Models and methodology

When genetic entries are assigned at random in complete blocks, the genetic model including the direct genetic effect (G_o), cytoplasmic effect (C) and maternal genetic effect (G_m) is

$$y = \mu + G_o + C + G_m + B + \varepsilon$$
$$= \mu + G + B + \varepsilon$$

where μ is the population mean, the total genetic effect G is $G_o + C + G_m$, B is the effect of randomized complete block, and ε is the residual error.

When a set of completely inbred lines of cereal crops are used for conducting modified diallel crosses, the direct genetic effect (G_o) for triploid endosperms can be partitioned as

$$G_o = \sum_i \tau_i A_i + \sum_{i \le j} \delta_{ij} D_{ij} + \sum_{i,i' \le j} k_{ii'j} H_{ii'j}$$

where A_i is the cumulative additive effect of direct genes from line i, $A_i \sim (0, \sigma_A^2)$; D_{ij} is the cumulative dominance effect of direct genes from line i and line j, $D \sim (0, \sigma_D^2)$; and $H_{ii'j}$ is the cumulative high-order dominance effect of direct genes from line i, line i' and line j, $H_{ii'j} \sim (0, \sigma_H^2)$. The cytoplasmic effect (C) is a function of $\Sigma_i \gamma_i C_i$ where $C_i \sim (0, \sigma_C^2)$. The maternal genetic effect (G_m) for a diploid maternal plant can be expressed as

$$G_m = \sum_i \tau_{m_i} A_{m_i} + \sum_{i \le j} \delta_{m_{ij}} D_{m_{ij}}$$

where A_{m_i} is the cumulative additive effect of maternal genes from line i, $A_{m_i} \sim (0, \sigma_{A_m}^2)$; and $D_{m_{ij}}$ is the cumulative dominance effect of maternal genes from line i and line j, $D_{m_{ij}} \sim (0, \sigma_{D_m}^2)$. Direct effects are correlated with maternal effects, with covariance components $\mathrm{Cov}(A_i, A_{m_i}) = \sigma_{A.A_m}$ and $\mathrm{Cov}(D_{ij}, D_{m_{ij}}) = \sigma_{D.D_m}$. Under the same genetic assumptions as those for diploid seed model (Zhu and Weir 1994), the genetic model can be rewritten as a linear model for the mean observation in the lth block of the kth type of genetic entry from lines i, j:

$$y_{ijkl} = \mu + G_{ijk} + B_l + \varepsilon_{ijkl}$$

where μ is the constant population mean, B_l is the effect of the lth randomized complete block, ε_{ijkl} is the residual error, and the total genetic effect G_{ijk} depends on the specific genetic entry of endosperms. For inbred line $P_i (i = j)$ and $F_{1_{ij}} (i \neq j)$ from maternal $P_i \times$ paternal $P_j (k = 1)$

$$G_{ij1} = 2A_i + A_j + D_{ii} + 2D_{ij} + H_{iij} + C_i + 2A_{m_i} + D_{m_{ij}};$$

for $F_{2_{ij}} (k = 2)$:

$$G_{ij2} = 1.5A_i + 1.5A_j + D_{ii} + D_{jj} + D_{ij} + 0.25H_{iii} + 0.25H_{jjj} + 0.25H_{iij} + 0.25H_{ijj} + C_i + A_{m_i} + A_{m_j} + D_{m_{ij}};$$

and for backcross BC_j from maternal $F_{1_{ij}} \times$ paternal $P_j (k = 3)$:

$$G_{ij3} = A_i + 2A_j + 0.5D_{ii} + 1.5D_{jj} + D_{ij} + 0.5H_{jjj} + 0.5H_{iij} + C_i + A_{m_i} + A_{m_j} + D_{m_{ij}}.$$

The genetic model can be written as a matrix form

$$y = \mathbf{1}\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_H\mathbf{e}_H + \mathbf{U}_C\mathbf{e}_C + \mathbf{U}_{A_m}\mathbf{e}_{A_m} + \mathbf{U}_{D_m}\mathbf{e}_{D_m} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon$$
$$= \mathbf{1}\mu + \sum_{u=1}^{\infty} \mathbf{U}_u\mathbf{e}_u \qquad (1)$$

with the variance-covariance matrix $\mathrm{Var}(\mathbf{y})$:

$$\mathrm{Var}(\mathbf{y}) = \sigma_A^2\mathbf{U}_A\mathbf{U}_A' + \sigma_D^2\mathbf{U}_D\mathbf{U}_D' + \sigma_H^2\mathbf{U}_H\mathbf{U}_H' + \sigma_C^2\mathbf{U}_C\mathbf{U}_C' + \sigma_{A_m}^2\mathbf{U}_{A_m}\mathbf{U}_{A_m}' + \sigma_{D_m}^2\mathbf{U}_{D_m}\mathbf{U}_{D_m}' + \sigma_B^2\mathbf{U}_B\mathbf{U}_B' + \sigma_{A.A_m}(\mathbf{U}_A\mathbf{U}_{A_m}' + \mathbf{U}_{A_m}\mathbf{U}_A') + \sigma_{D.D_m}(\mathbf{U}_D\mathbf{U}_{D_m}' + \mathbf{U}_{D_m}\mathbf{U}_D') + \sigma_\varepsilon^2\mathbf{I}$$
$$= \sum_{v=1}^{10} \theta_u\mathbf{V}_u$$
$$= \mathbf{V}_{(\theta)}$$

where $\mathbf{U}_u$ is the known incidence matrix relating to the random vector $\mathbf{e}_u \sim (\mathbf{o}, \sigma_u^2\mathbf{I})$ for $u = 1, 2, \ldots, 8$; $\mathbf{U}_u'$ is the transpose of $\mathbf{U}_u$, $\mathbf{U}_8 = \mathbf{I}$ is an identity matrix; $\mathbf{V}_u = \mathbf{U}_u\mathbf{U}_u'$ for $u = 1, 2, \ldots, 7$; $\mathbf{V}_8 = (\mathbf{U}_1\mathbf{U}_5' + \mathbf{U}_5\mathbf{U}_1')$, $\mathbf{V}_9 = (\mathbf{U}_2\mathbf{U}_6 + \mathbf{U}_6\mathbf{U}_2')$, and $\mathbf{V}_{10} = \mathbf{I}$.

Since the high-order dominance effect $H_{ii'j}$ is the cumulative effect of a three-allele interaction at each locus, it may not be of as much importance as the two-allele interaction for general dominance effect D_{ij}. If the high-order dominance effects are negligible or not of much concern in practice, a reduced genetic model can be employed by excluding them. The reduced genetic model can be written in a matrix form for all entries involving endosperms of inbred lines, F_1s and F_2s $(k = 1, 2)$,

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_C\mathbf{e}_C + \mathbf{U}_{A_m}\mathbf{e}_{A_m} + \mathbf{U}_{D_m}\mathbf{e}_{D_m} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon$$
$$= 1\mu + \sum_{u=1}^{7} \mathbf{U}_u\mathbf{e}_u \qquad (2)$$

with the variance-covariance matrix Var(y):

$$\begin{aligned}\mathrm{Var}(\mathbf{y}) &= \sigma_A^2\mathbf{U}_A\mathbf{U}_A' + \sigma_D^2\mathbf{U}_D\mathbf{U}_D' + \sigma_C^2\mathbf{U}_C\mathbf{U}_C' \\ &\quad + \sigma_{A_m}^2\mathbf{U}_{A_m}\mathbf{U}_{A_m} + \sigma_{D_m}^2\mathbf{U}_{D_m}\mathbf{U}_{D_m} + \sigma_B^2\mathbf{U}_B\mathbf{U}_B' \\ &\quad + \sigma_{A.A_m}(\mathbf{U}_A\mathbf{U}_{A_m} + \mathbf{U}_{A_m}\mathbf{U}_A') \\ &\quad + \sigma_{D.D_m}(\mathbf{U}_D\mathbf{U}_{D_m} + \mathbf{U}_{D_m}\mathbf{U}_D') + \sigma_e^2\mathbf{I} \\ &= \sum_{u=1}^{9}\theta_u\mathbf{V}_u \\ &= \mathbf{V}_{(\theta)}\end{aligned}$$

where $\mathbf{U}_7 = \mathbf{I}$ is an identity matrix; $\mathbf{V}_u = \mathbf{U}_u\mathbf{U}_u'$ for $u = 1, 2, \ldots, 6$; $\mathbf{V}_7 = (\mathbf{U}_1\mathbf{U}_4' + \mathbf{U}_4\mathbf{U}_1')$, $\mathbf{V}_8 = (\mathbf{U}_2\mathbf{U}_5' + \mathbf{U}_5\mathbf{U}_2')$, and $\mathbf{V}_9 = \mathbf{I}$.

Variance and covariance components for one trait and covariance components for two traits can be estimated by MINQUE(0/1) with the jackknife procedure (Zhu and Weir 1994). For the full model of triploid endosperm, phenotypic variance V_p can be partitioned as

$$\begin{aligned}V_P &= V_{G_o} + V_C + V_{G_m} + 2C_{G_o.G_m} + V_e \\ &= (V_A + V_D + V_H) + V_C + (V_{A_m} + V_{D_m}) + 2(C_{A.A_m} + C_{D.D_m}) + V_e\end{aligned}$$

where direct genetic variance V_{G_o} = (additive variance V_A + dominance variance V_D); V_C is cytoplasmic variance; maternal genetic variance V_{G_m} = (maternal additive variance V_{A_m} + maternal dominance variance V_{D_m}); genetic covariance $C_{G_o.G_m}$ = (additive covariance $C_{A.A_m}$ + dominance covariance $C_{D_o.D_m}$); V_e is residual variance.

Variance and covariance for direct and maternal genetic effects are different for generations F_1, F_2, and BC_j. Therefore they should be estimated separately for each generation. For F_1 seeds on plants of inbred lines,

$$\hat{V}_A = 5\hat{\sigma}_A^2 \quad \hat{V}_D = 5\hat{\sigma}_D^2 \quad \hat{V}_H = \hat{\sigma}_H^2$$
$$\hat{V}_{A_m} = 4\hat{\sigma}_{A_m}^2 \quad \hat{V}_{D_m} = 4\hat{\sigma}_{D_m}^2$$
$$\hat{C}_{A.A_m} = 4\hat{\sigma}_{A.A_m} \quad \hat{C}_{D.D_m} = \hat{\sigma}_{D.D_m}$$

For F_2 seeds on F_1 plants,

$$\hat{V}_A = 4\tfrac{1}{2}\hat{\sigma}_A^2 \quad \hat{V}_D = 3\hat{\sigma}_D^2 \quad \hat{V}_H = \tfrac{1}{4}\hat{\sigma}_H^2$$

For BC_j seeds on F_1 plants,

$$\hat{V}_A = 5\hat{\sigma}_A^2 \quad \hat{V}_D = 3\tfrac{1}{2}\hat{\sigma}_D^2 \quad \hat{V}_H = \tfrac{1}{2}\hat{\sigma}_H^2$$

For both F_2 and BC_j seeds,

$$\hat{V}_{A_m} = 2\hat{\sigma}_{A_m}^2 \quad \hat{V}_{D_m} = \hat{\sigma}_{D_m}^2$$
$$\hat{C}_{A.A_m} = 3\hat{\sigma}_{A.A_m} \quad \hat{C}_{D.D_m} = \hat{\sigma}_{D.D_m}$$

Cytoplasmic variance ($\hat{V}_C = \hat{\sigma}_C^2$) and residual variance ($\hat{V}_e = \hat{\sigma}_e^2$) can be estimated for all the generations.

Phenotypic covariance C_P between two traits can also be partitioned as

$$\begin{aligned}C_P &= C_{G_o} + C_C + C_{G_m} + 2C_{G_o/G_m} + C_e \\ &= (C_A + C_D + C_H) + C_C + (C_{A_m} + C_{D_m}) + 2(C_{A/A_m} + C_{D/D_m}) + C_e\end{aligned}$$

where direct genetic covariance C_{G_o} = (direct additive covariance C_A + direct dominance covariance C_D); C_C is cytoplasmic covariance; maternal genetic covariance C_{G_m} = (maternal additive covariance C_{A_m} + maternal dominance covariance C_{D_m}); genetic covariance between nuclear and maternal gene effects for two traits $C_{G_o.G_m}$ = (additive covariance C_{A/A_m} + dominance covariance C_{D/D_m}); C_e is residual covariance. Covariance components of each generation can be estimated as the same ways as for variance components. Under the assumption of no high-order dominance effects, variance and covariance of high-order dominance effects should be dropped from the phenotypic variance and covariance for the reduced model. Genetic effects in the full or reduced genetic models can be predicted by LUP with jackknife procedure (Zhu and Weir 1994).

Monte Carlo simulation results

Two mating designs of modified diallel crosses from six inbred lines were employed in this study. For the full genetic model, modified diallel crosses with F_1s and reciprocal F_1s, F_2s and reciprocal F_2s, and backcrosses BCs were conducted with a total of 90 genetic entries in each block. For the reduced genetic model, modified diallel crosses with F_1s and reciprocal F_1s, F_2s and reciprocal F_2s, and six inbred lines were used with total 66 genetic entries in each block. Monte Carlo simulations were conducted by the same ways as in Zhu and Weir (1994). Mean Squared Error (MSE) is usually used as the main criterion for comparing efficiency of estimation methods. Since MSE is a function of parameter and its estimate, efficiency of estimation cannot be evaluated by MSE for different parameter values. Coefficients of efficiency (C.E.) defined as C.E. = $(\mathrm{MSE})^{1/2}/(|\theta| + |\mathrm{bias}|)$

Table 1 Bias, C.E. and power value from simulations by MINQUE(0/1) with the jackknife procedure for modified diallel crosses

Parameter	True value	Full model			Reduced model[a]		
		Bias	C.E.	Power[b]	Bias	C.E.	Power[b]
One trait							
σ_A^2	20	−0.82	0.87	0.63	−0.23	0.81	0.60
σ_D^2	10	0.04	0.63	0.58	0.19	0.41	0.97
σ_H^2	10	1.67[c]	1.37	0.12	–	–	–
σ_C^2	10	0.56	0.73	0.52	−0.42	0.68	0.41
$\sigma_{A_m}^2$	45	0.51	0.79	0.71	1.49	0.68	0.75
$\sigma_{D_m}^2$	40	−0.65	0.42	0.96	0.95	0.41	0.97
$\sigma_{A.A_m}$	15	1.33	1.17	0.50	0.61	1.15	0.54
$\sigma_{D.D_m}$	10	0.12	0.71	0.49	0.43	0.63	0.57
σ_e^2	20	−0.01	0.09	1.00	−0.10	0.12	1.00
Two traits							
$\sigma_{A/A}$	10	0.99	1.34	0.35	0.75	1.21	0.31
$\sigma_{D/D}$	5	−0.05	0.92	0.20	−0.09	0.61	0.48
$\sigma_{H/H}$	5	0.05	2.36	0.06	–	–	–
$\sigma_{C/C}$	5	−0.25	1.14	0.15	−0.27	1.06	0.13
σ_{A_m/A_m}	22.5	−1.56	1.08	0.26	0.01	1.13	0.33
σ_{D_m/D_m}	20	−0.25	0.62	0.43	0.67	0.62	0.53
σ_{A/A_m}	15	−1.13	1.10	0.38	−0.30	1.16	0.42
σ_{D/D_m}	10	−0.08	0.70	0.47	0.31	0.62	0.43
$\sigma_{e/e}$	10	0.01	0.15	0.83	0.06	0.19	0.63

[a] Reduced model with $\sigma_H^2 = 0$ for one trait or $\sigma_{H/H} = 0$ for two traits

[b] Probability of rejecting the null hypothesis of no variation by the t test with $\alpha = 0.05$

[c] Bias > 10% of the true value

were used in this study for evaluating estimation efficiency for variance and covariance components with different parameter values. The smaller the C.E. value, the more efficient is the estimation.

Simulation results of bias, C.E. and power value for one-trait variance and covariance components are listed in Table 1. For the full model, genetic variance components other than σ_H^2 are well estimated with small biases and high power values (> 50%). The high-order dominance variance component σ_H^2 tends to be overestimated by six-parent modified diallel crosses with F_1s, F_2s, and BCs. The power value for detecting significance of σ_H^2 is about 12%. It is indicated that more genetic entries should be included for unbiased estimation of σ_H^2 with reasonable power. Under the assumption of no high-order dominance effects ($\sigma_H^2 = 0$), all of the variance components can be estimated with negligible bias by the reduced model of six-parent modified diallel crosses.

Robustness of estimating one-trait variance and covariance components was tested by simulations under the conditions of no specific variation. When the true value is zero for a specific parameter, the conclusion of non-significance can be drawn with a probability around 95% by the *t*-test. For the full genetic model, other variance and covariance components can be well estimated with decreased C.E. values and increased power values if there is no high-order dominance effects ($\sigma_H^2 = 0$). For the reduced genetic model, even though there do exist high-order dominance effects ($\sigma_H^2 = 10$), direct additive variance and maternal genetic variance components along with cytoplasmic variance component can still be well estimated; slight biases are observed only for σ_D^2 and $\sigma_{D.D_m}^2$. It is indicated that no matter whether or not the high-order dominance effects exist correct conclusions can be obtained for maternal genetic effects, cytoplasmic effects, and direct additive effects of endospermic traits by the reduced genetic model.

For both full and reduced genetic models, additive variance and covariance components and cytoplasmic variance component will be better estimated if there are no direct and maternal dominance effects. Without cytoplasmic effects ($\sigma_C^2 = 0$), other variance and covariance components can be estimated with similar bias, C.E. and power value as the cytoplasmic effects are present. No cytoplasmic and maternal genetic effects ($\sigma_C^2 = \sigma_{A_m}^2 = \sigma_{D_m}^2 = \sigma_{A.A_m}^2 = \sigma_{D.D_m}^2 = 0$) will result in estimates with negligible biases for variance components of direct genetic effects. The variance for residual error can always be well estimated with negligible bias and high power value in any situation.

Simulation results for covariance components of two traits are also presented in Table 1. For the full genetic model, all of the covariance components are well estimated. Except for error covariance $\sigma_{e/e}$, power values are low for detecting significant covariance components. The power value for testing the high-order dominance variation is extremely low. For the reduced genetic model, all of the covariance components can be estimated with negligible biases. The power values for testing genetic covariance components are relatively lower than that of error covariance $\sigma_{e/e}$. More genetic entries are needed for the modified diallel mating designs in order to gain reasonable power for detecting significant covariance components.

When comparing efficiency of estimation for the full genetic model, estimation of maternal genetic variances and covariances is more efficient than that of direct genetic variances and covariances, and estimation for the high-order dominance effects is of the least efficiency. For the reduced genetic model, equally efficient estimations are obtained for variances and covariances of direct and maternal dominance effects and for those of cytoplasmic and maternal additive effects. Dominance variances and covariances are always more efficiently estimable than additive variances and covariances in two models. Since C.E. values for variance components of one trait are much smaller than those for correspondent covariance component of two traits, the estimation of covariance components is of less efficiency than that of variance components. But estimation efficiencies for σ_{A/A_m} vs $\sigma_{A.A_m}$ and σ_{D/D_m} vs $\sigma_{D.D_m}$ are about equal. In this study, the modified diallel mating design for the reduced genetic model did not include endosperms of backcrosses *BC*'s and the number of genetic entries were about three-quarters of those for the full genetic model. However, the reduced genetic model tends to give smaller C.E. for most variances and covariances than the full genetic model. It is concluded that the reduced genetic model is more efficient for estimating variances and covariances of endosperm traits than the full genetic model provided that the high-order dominance effects are not important.

Table 2 Prediction of genetic effects by LUP and BLUP methods for modified diallel crosses[a]

Parameter	Full model				Reduced model[b]			
	Dist. of $\hat{\mathbf{e}}_{u(0/1)}$		Dist. of $\hat{\mathbf{e}}_{u(\theta)}$		Dist. of $\hat{\mathbf{e}}_{u(0/1)}$		Dist. of $\hat{\mathbf{e}}_{u(\theta)}$	
	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
A	6.46	5.03	6.37	4.97	6.27	4.75	6.15	4.62
D	10.50**	4.19	9.62	2.90	8.02**	3.44	7.63	3.13
H	17.74**	4.27	17.12	3.98	–	–	–	–
C	3.98	2.73	3.82	2.69	4.00	2.15	3.92	2.02
A_m	8.80	10.32	8.56	10.32	8.26	11.59	8.23	11.59
D_m	16.95*	16.29	16.41	15.94	15.75	14.44	15.52	14.75

[a] Absolute bias of mean prediction of genetic effects is $10^{-4} \sim 10^{-7}$ for $\hat{\mathbf{e}}_{i(0/1)}$, and $10^{-2} \sim 10^{-4}$ for $\hat{\mathbf{e}}_{i(\theta)}$

[b] Reduced model with $\sigma_H^2 = 0$

*, ** Significantly different from the BLUP at the 0.05 and 0.01 level, respectively.

Simulation results for prediction of genetic effects are listed in Table 2. Parameter values used by simulations are the same as those in Table 1 for one trait. The MINQUE (0/1) procedure gives linear unbiased prediction (LUP) for random genetic effects just as the BLUP does. For the full genetic model, the distance of $\hat{\mathbf{e}}_{u(0/1)}$ to $\tilde{\mathbf{e}}_u$ approaches that of the BLUP $\hat{\mathbf{e}}_{u(\theta)}$ for direct and maternal additive effects and cytoplasmic effects, but the distance of $\hat{\mathbf{e}}_{u(0/1)}$ is significantly larger than that of the BLUP $\hat{\mathbf{e}}_{u(\theta)}$ for direct and maternal dominance effects and for high-order dominance effects. Therefore, direct and maternal additive effects as well as cytoplasmic effects could be predicted by the MINQUE (0/1) procedure almost as efficiently as by the BLUP procedure for the full genetic model. If additive effects are more important than dominance effects, the predicted additive effects could serve as major criteria for evaluating breeding materials.

Except for direct dominance effects, all of the other genetic effects can be predicted by the MINQUE(0/1) procedure as efficiently as by the BLUP procedure for the reduced genetic model. If endosperm traits are mainly under the influence of maternal genetic effects and the high-order dominance effects are negligible, the genetic merits of breeding materials can be predicted quite well by the MINQUE(0/1) procedure for the reduced genetic model.

Example: kernel growth characteristics of corn

Data for kernel growth characteristics of corn (*Zea mays* L.) from Table 5 of (Poneleit and Egli 1983) were used as an example for the demonstration of estimating variance and covariance components and of predicting genetic merits. The kernel growth characteristics studied are effective filling period (EFP) and kernel growth rate (KGR). These two traits are referred to mainly as endospermic characteristics. Cell means of four-parent modified diallel mating with kernels for inbred lines, F_1s and reciprocal F_1s, and F_2s and reciprocal F_2s were first analyzed by the full genetic model without backcrosses to test high-order dominance effects. Negative estimates for σ_H^2 indicated that high-order dominance effects may not be important for both of these kernel growth characteristics. The reduced genetic model was then used for estimation and prediction. The experimental size of four-parent modified diallel crosses may be not big enough for efficiently detecting variation of genetic effects.

Jackknife estimates and their standard errors were obtained by resampling the cell mean of each genetic entry with 27 degrees of freedom. Estimates of results of variance and covariance components are summarized in Table 3. Significances were detected for maternal dominance variance of EFP and KGR. Positive estimates were observed for EFP additive variances of direct and maternal gene effects, but these variances were not significantly different from zero. As shown by simulation (Table 1), estimation for additive variances is less efficient than that for dominance variance. There might have been additive variances of direct and maternal gene effects for EFP if the experimental size had been larger than a 4 × 4 diallel cross. Cytoplasm variance was negligible for EFP. Estimates of direct additive, dominance variances and cytoplasm variance were negative for KGR. There were no significant covariances between

Table 3 Estimates of variance and covariance components by the reduced model for kernel growth characteristics of corn

Variance	EFP Estimate ± SE	KGR Estimate ± SE	Covariance	EFP vs KGR Estimate ± SE
σ_A^2	3.015 ± 2.706	−0.208 ± 0.299	$\sigma_{A/A}$	−0.201 ± 0.641
σ_D^2	−0.953 ± 1.440	−0.067 ± 0.075	$\sigma_{D/D}$	−0.055 ± 0.092
σ_C^2	0.516 ± 0.960	−0.039 ± 0.164	$\sigma_{C/C}$	0.054 ± 0.303
σ_{Am}^2	4.247 ± 4.136	0.284 ± 0.699	$\sigma_{Am/Am}$	−1.739 ± 1.512
σ_{Dm}^2	7.369* ± 3.763	1.121* ± 0.525	$\sigma_{Dm/Dm}$	2.181* ± 0.936
$\sigma_{A.Am}$	−3.843 ± 3.332	0.002 ± 0.415	$\sigma_{A/Am}$	0.630 ± 1.086
$\sigma_{D.Dm}$	0.883 ± 1.079	0.309 ± 0.188	$\sigma_{D/Dm}$	0.224 ± 0.554
σ_e^2	0.853 ± 0.862	0.345* ± 0.199	$\sigma_{e/e}$	−0.101* ± 0.287

* $P < 0.05$

Table 4 Predicted genetic effects and standard errors from cell mean data by the reduced model for kernel growth characteristics of corn

	$i=1$	$i=2$	$i=3$	$i=4$
EFP				
$D_{m_{i1}}$	−3.229** ± 0.813			
$D_{m_{i2}}$	4.354** ± 1.115	−2.683** ± 0.465		
$D_{m_{i3}}$	1.848 ± 0.933	−0.146 ± 0.344	−1.899** ± 0.488	
$D_{m_{i4}}$	3.903** ± 1.115	−0.411 ± 0.481	0.489 ± 0.454	−2.228** ± 0.397
KGR				
$D_{m_{i1}}$	−1.394** ± 0.258			
$D_{m_{i2}}$	0.317 ± 0.261	−1.114** ± 0.217		
$D_{m_{i3}}$	1.045* ± 0.396	1.131** ± 0.355	−1.638** ± 0.288	
$D_{m_{i4}}$	0.468 ± 0.344	0.324 ± 0.319	1.414** ± 0.497	−0.552 ± 0.277

* $P < 0.05$. ** $P < 0.01$

direct and maternal gene effects. It was concluded that maternal gene effects, especially maternal dominance effects, were the most important contribution to the genetic variation of EFP and KGR. Maternal dominance covariance was positive and significant between EFP and KGR. There might be some negative covariance component of maternal additive effects between these two traits.

The predicted maternal dominance effects by MINQUE(0/1) with jackknife procedure are listed in Table 4. The highest predicted value for maternal dominance effects for EFP was $\hat{D}_{m_{12}}$, which could be the genetic mechanism for the highest EFP value observed in F_2 kernels of crosses 1×2 and 2×1. The highest $\hat{D}_{m_{34}}$ for KGR indicated that line 3 and line 4 might be good candidates for parents of a hybrid with high kernel growth rate. The homozygous maternal dominance effects were negative for both EFP and KGR. Heterosis for kernel growth characteristics can be measured by the average of the homozygous dominance effects. The negative values of $(\Sigma_i \hat{D}_{m_{ii}})/4$ were observed for both EFP (-1.875 ± 0.517) and KGR (-1.009 ± 0.201). Heterosis for kernel growth characteristics was then expected for the F_2 kernels on F_1 plants. This provides an explanation for the phenomenon that significant increases in EFP and KGR were observed for the F_2 kernels as compared to homozygous and some F_1 kernels.

Discussion

The genetic models proposed in this paper are for quantitative traits of the endosperm, which is a triploid organ nursed by the diploid maternal plant. These models can also be used for analyzing cereal seedling traits upon which maternal effects are often observed. Some important cereal crops, e.g., wheat and oats, are allopolyploid plants, but their meiotic divisions behave like those of diploid plants. The models are also applicable for these allopolyploid crops.

The genetic models are based on the assumption that no previous information is available for the cytoplasm resources. If p inbred lines for the modified diallel crosses belong to q cytoplasm resources the ith line is known to have the hth type of cytoplasm, the cytoplasm effect for G_{ijk} is then defined as C_h for $h = 1, 2, \ldots, q \leq p$. If all of the p inbred lines have the same type of cytoplasm, the cytoplasm effect can simply be omitted. One of the assumptions for the genetic models is that of no inter-action of genetic effect and environmental effect. If the assumption is true, data for quantitative traits of endosperms can be collected from one-environment experiments. When there exists some kind of gene-by-environment interaction, an extended genetic model with all kinds of interaction terms should be used for analysis of interaction variance and covariance components, and experiments should be conducted in several environments.

Both full and reduced genetic models are shown to be robust for estimating variance and covariance components under several situations of no specific effects. The genetic models partition the total genetic effect into direct genetic effects, cytoplasmic effects, and maternal genetic effects. Even if there are no cytoplasmic effects or maternal genetic effects, direct additive and dominance variance components can still be well estimated. The triploid models for direct genetic effects (Gale 1975; Bogyo et al. 1988; Mo 1988) would result in correct conclusions for genetic variations of endospermic traits only under the situations of no cytoplasmic effects and maternal genetic effects. If maternal genetic effects and/or cytoplasmic effects are really important for quantitative traits of endosperms, triploid models would give biased estimates. The models proposed by Foolad and Jones (1992) will provide unbiased estimation for maternal, cytoplasm and endosperm effects. By those models, observation on individual seeds is needed for more than ten generations to estimate variance components. And estimation of genetic covariance components between different traits is not available.

Since seeds are the offspring of maternal plants, experiments for the genetic analysis of endosperm traits have to be conducted with special caution. The maternal plants of genetic entries are the experimental units. When experiments for the genetic models are conducted, plants of F_1s and inbred lines are assigned at random within complete blocks for producing BC and F_1 seeds and/or self-pollinated inbred line seeds and F_2 seeds later on. Artificial emasculation and pollination will have apparent impacts of mechanical damage on kernel growth. Even for the reduced genetic model with only inbred lines, F_1s and F_2s, kernels for inbred lines and F_2s should be produced by the same procedures of emasculation and pollination as those for F_1s.

For many cereal crops artificial emasculation and pollination are very labor intensive. It will be very hard to conduct experiments for a modified diallel mating design with all combinations of F_1s, F_2s, and BCs from more than six inbred lines. Since cytoplasmic male-sterile lines have been found and used for producing hybrid seeds in major cereal crops, a specific unbalanced mating design of the reduced genetic model can be used for the estimation of variance and covariance components and for the prediction of genetic effects. The unbalanced mating involves three types of lines, A lines as male-sterile lines, B lines as maintainer lines, and R lines as fertility restorer lines. A set of A lines with the same cytoplasmic resource are pollinated by a set of R lines for producing F_1 seeds. F_2 seeds are produced by natural self-pollination from F_1 plants of A lines $\times$ R lines. Inbred line seeds are from B lines and R lines. There are no cytoplasmic effects for this mating design if all of the A lines have the same cytoplasmic resource. The mating design is unbalanced because of no reciprocal F_1s and F_2s involved. By this mating

design all of the seeds can be produced without artificial emasculation. Monte Carlo simulations for the unbalanced mating designs from six *A* lines and five *R* lines with three replications showed that variance and covariance components can be well and efficiently estimated (bias $\leq 6.47\%$ of the parameter value, and C.E. ≤ 1.13).

Acknowledgments This research was supported in part by NIH Grant GM32518.

References

Beavis WD, Pollak E, Frey KJ (1987) A theoretical model for quantitatively inherited traits influenced by nuclear-cytoplasmic interactions. Theor Appl Genet 74:571–578

Bogyo TP, Lance RCM, Chevalier P, Nilan RA (1988) Genetic models for quantitatively inherited endosperm characters. Heredity 60:61–67

Cockerham CC, Weir BS (1977) Quadratic analyses of reciprocal crosses. Biometrics 33:187–203

Dhaliwal HS (1977) Genetic control of seed proteins in wheat. Theor Appl Genet 50:235–239

Eisen EJ, Hórstgen-Schwark G, Saxton AM, Bandy TR (1983) Genetic interpretation and analysis of diallel crosses with animals. Theor Appl Genet 65:17–23

Foolad MR, Jones RA (1992) Models to estimate maternally controlled genetic variation in quantitative seed characters. Theor Appl Genet 83:360–366

Gale MD (1975) High-amylase breeding and genetical aspects of the problem. Cereal Res Commun 4: 231–243

Garwood DL, Weber RJ, Lamber RJ, Alexander DE (1970) Effect of different cytoplasm on oil, fatty acids, plant height, and ear height in maize (*Zea mays* L.). Crop Sci 10:39–41

Griffing B (1956) Concept of general and specific combining ability in relation to diallel crossing systems. Aust J Biol Sci 9:463–493

Henderson CR (1948) Estimation of general, specific and maternal combining abilities in crosses among inbred lines of swine. PhD thesis, Iowa State University, Ames, Iowa

Kofoid KD, Maan SS (1982) Agronomic and bread-making performance of fertile alloplasmic wheat. Crop Sci 22:725–729

Mo H (1988) Genetic expression for endosperm traits. In: Weir BS, Eisen EJ, Goodman MM, Namkoong GN (eds) Proc 2nd Int Conf Quant Genet. Sinauer Assoc, Boston Mass., pp 478–487

Poneleit CG, Bauman LF (1970) Diallel analysis of fatty acids in corn (*Zea mays* L.) oil. Crop Sci 10:338–341

Poneleit CG, Egli DB (1983) Differences between reciprocal crosses of maize for kernel growth characteristics. Crop Sci 23: 871–875

Rao AP, Fleming AA (1978) Cytoplasmic-genotypic effects in the GT 112 maize inbred with four cytoplasms. Crop Sci 18: 935–937

Ullrich SE, Eslick RF (1978) Protein and lysine maternal inheritance effects in high protein and high lysine barley mutant Hiproly. Barley Genet Newsl 8:108–109

Weber EJ (1983) Variation in corn (*Zea mays* L.) for fatty acid compositions of triglycerides and phospholipids. Biochemical Genet 21:1–13

Zhu J, Weir BS (1994) Analysis of cytoplasmic and maternal effects. I. A genetic model for diploid plant seeds and animals. Theor Appl Genet (in press)